2011 KDI Journal of Economic Policy Conference

Globalization, Human Capital and Inequality

Hosted by

KDI KAEA

KDI

2011 KDI Journal of Economic Policy Conference

Globalization, Human Capital and Inequality

Hosted by
K D I
KAEA

Korea Development Institute
47 Hoegiro, Dongdaemun-gu
Seoul 130-740, Korea

ISBN 978-89-8063-669-3

Contents

CHAPTER 1
Integrating Personality Psychology into Economics
(*James Heckman*) 1
1. A Brief History of Personality Psychology 2
2. Cognition: “g”– a single factor that is claimed to represent intelligence 3
3. Personality Traits 4
4. The Person-Situation Debate: A Strong Influence on Behavioral Economics 6
5. Personality Psychology After the Person-Situation Debate 6
6. The Predictive Power of Personality Traits 7
7. Difficulties in Synthesizing Studies of the Effects of Personality 7
8. Main Findings from Predictive Analyses 7
9. Conceptualizing Personality Within an Economic Model 13
10. An Economic Framework for Conceptualizing and Measuring Personality and Personality Traits 14
11. A Definition of Personality 19
12. Stability and Change in Personality Traits and Preferences 23
13. Personality and Preference Parameters 31
14. Summary and Conclusions 32

References 34

CHAPTER 2
‘학교교육 수준 및 실태 분석 연구: 중학교’ 자료를 이용한 사교육비 지출의 성적 향상 효과 분석(*강창희*) 38
제1절 서 론 39

제2절 계량분석모형 43
제3절 분석자료: 학교교육 수준 및 실태 분석 연구 51
제4절 실증분석 결과 56
제5절 결 론 66

참고문헌 68

CHAPTER 3
Education and Globalization on Income Inequality:
Cross-Country Evidence(*Kang H. Park*) 74
1. Introduction 75
2. Review of Literature 77
3. Model, Data and Variables 81
4. Emprirical Results 84
5. Conclusions 89

References 91

CHAPTER 4
Should I Marry Again?(*Jinyoung Kim*) 95
1. Introduction 96
2. The Model 100
3. Empirical Implementation 111
4. Empirical Findings 116
5. Concluding Remarks 122

References 125
Appendix 128

CHAPTER 5
Stereotypes and Inequality: A Structure of Identity Choice
(*Young Chul Kim, Glenn C. Loury*) 137
1. Introduction 138
2. Model with the Identity Choice 142

3. Stereotyping Equilibrium 145
4. Affective Stereotyping Equilibria with Multiple PSE 153
5. Affective Stereotyping Equilibria with Unique PSE 161
6. Further Discussions 166
7. Conclusion 167
8. Appendix: Proofs 168

References 174

CHAPTER 6
The Dynamics of Income Inequality in Korea(*Jongsung Kim*) 175
1. Introduction 176
2. Income Inequality in Korea 178
3. Trends in Income Inequality in Korea 182
4. New Evidence of Income Inequality and its Decomposition 190
5. Discussion and Conclusions 195

References 199

CHAPTER 7
The Impact of Mortgage Securitization on the Housing Bubble and Subprime Mortgage Crisis: A Self-organization Perspective (*Hong Y. Park, Hyejung Chang, Kaustav Misra*) 201
1. Introduction 202
2. Securitization of Mortgages: Its Complexity and Impact 204
3. Securitization of Mortgages, Mortgage Rates and Housing Bubble 209
4. Model and Empirical Results 214
5. Empirical Analysis and Discussion 216
6. Solutions to the Subprime Mortgage Crisis 222
7. Summary and Conclusion 229

References 231
Appendix 233

CHAPTER 8
Real Estate Investors, the Leverage Cycle and the Housing Market Crisis(*Andrew Haughwout, Donghoon Lee, Joseph Tracy, Wilbert van der Klaauw*) 235
1. Related Literature 237
2. Investors and the Leverage Cycle 249
3. Results 254
4. Mortgage Products and Leverage 264
5. Delinquencies and de-leveraging during the bust 269
6. Conclusion 279

References 281

CHAPTER 9
Optimal Provision of Loans and Insurance against Unemployment from Lifetime Perspective (*Joseph Stiglitz, Jungyoll Yun*) 283
1. Introduction 284
2. The Model and Baseline Optimum 288
3. Presence of Private Loan Market 302
4. Welfare Effects of Loan-Cosigning in the Provision of Loans 305
5. Conclusion 311

References 312
Appendix 314

CHAPTER 10
The Effects of Free Trade Agreements and Fair Trade for All (*Junsoo Lee*) 319
1. Introduction 320
2. Econometric Issues 322
3. Estimation Methods and Empirical Results 325
4. Concluding Remarks 336

References 338
Appendix 339

CHAPTER 11
Floating Exchange Rates and Macroeconomic Independence
(*Yoonbai Kim, Lian An, Yu You*) 341
1. Introduction 342
2. Methodology 346
3. Empirical Results 350
4. Conclusion 365

References 366

CHAPTER 12
Uncovered Interest Parity Puzzle: Asymmetric Responses
(*Byung-Joo Lee*) 370
1. Introduction 371
2. The forward premium puzzle 372
3. Empirical analysis 375
4. Conclusion 391

References 393

List of Tables

CHAPTER 1
Table 1 The Big Five Domains and Their Facets 5
Table 2 Incentives and Performance on Intelligence Tests 17
Table 3 Standard Preference Parameters and Conceptually Similar Measures in the Psychology Literature 31
Table 4 Empirical Studies of the Links Between Preferences and Traits 32

CHAPTER 2
〈표 1〉 분석자료에 대한 기술통계량 55
〈표 2〉 OLS 및 도구변수법 추정 결과: 종속변수는 세 과목 평균 점수 57
〈표 3〉 OLS 및 도구변수법 추정 결과: 종속변수는 개별 과목 점수 60
〈표 4〉 사교육비 효과에 대한 구간추정법의 추정 결과 63
〈표 5〉 처치수준별 성적의 평균, $\hat{E}[y(t)]$ 64

CHAPTER 3
Table 1 a. Regression of Income Inequality on Income (Year 1990) 84
b. Regression of Income Inequality on Income (Year 2000) 85
Table 2 a. Regression of Income Inequality on Income and Education Variables (Year 1990) 86
b. Regression of Income Inequality on Income and Education Variables (Year 2000) 86
Table 3 a. Regression of Income Inequality on Income, Education and Globalization Variables (Year 1990) 87
b. Regression of Income Inequality on Income, Education and Globalization Variables (Year 2000) 88

CHAPTER 4
Table 1 Variable Descriptions and Summary Statistics 115
Table 2 Human Capital Investment 118
Table 3 Childbearing with a New Partner: Binary Choice Model 120

CHAPTER 6

Table 1 Changes of income in the 2009 first quarter from the previous year (%) 186
Table 2 Income Decomposition in selected waves 193

CHAPTER 7

Table 1 Estimation Results 216

CHAPTER 8

Table 1 Distribution of Mortgage Loan Documentation Level, by Year 243
Table 2 Mortgage Leverage During the Housing Boom 245
Table 3 Average balance per mortgage, by investor status 258
Table 4 Combined Loan to Value ratios at Origination, Securitized Non-prime Mortgages 266
Table 5 Annual Investor Effects on 90DPD Delinquency Rate 276
Table 6 Annual Investor Effects on 90DPD Delinquency Rate . Panel Definition of Investor 278

CHAPTER 10

Table 1 Estimation Results of Cross-sectional Regression 327
Table 2 Panel Estimation Results with Bilateral and Country-time FE 330
Table 3 Panel IV Estimation Results 332
Table 4 Dynamic Panel Hsiao-Anderson and IV Estimation 333
Table 5 Dynamic Panel Results Using GMM Estimation 334

CHAPTER 11

Table 1 Variance Decompositions 352
Table 2 Unit-root Tests 355
Table 3 The Johansen Tests of Cointegration 357
Table 4 Long-Run Exclusion, Stationarity and Weak Exogeneity 358
Table 5 Tests of Hypotheses 359
Table 6 The Moving Average Representation 362

CHAPTER 12

Table 1	UIP Slope Estimates for Developed Countries: U.S. Dollar rate	377
Table 2	UIP Slope Estimates for Developing Countries: U.S. Dollar rate	378
Table 3	Rejection of UIP test for each currency: U.S. Dollar rate	379
Table 4	Summary of all slope estimates	379
Table 5	Bilateral UIP slope estimates	381
Table 6	Regression of UIP slope on key currencies	383
Table 7	Asymmetric UIP relationship: F-test	387
Table 8	Asymmetric UIP slope estimates	388
Table 9	Paired t test for the means of + and - forward premium	389

List of Figures

CHAPTER 1

Figure 1 An Hierarchical Scheme of General Intelligence and Its Components 4
Figure 2 Association of the Big Five and Intelligence with Years of Schooling in GSOEP 8
Figure 3 Distribution of Cognitive and Non-Cognitive Skills by Education Group 9
Figure 4 Distribution of Cognitive and Non-Cognitive Skills by Education Group 9
Figure 5 Probability of Being a 4-year-college Graduate or Higher at Age 30, Males 10
Figure 6 Probability of Being a 4-year-college Graduate or Higher at Age 30, Males 10
Figure 7 Correlations of the Big Five and Intelligence with Course Grades 11
Figure 8 Associations with Job Performance 12
Figure 9 Correlations of Mortality with Personality, IQ, and Socioeconomic Status (SES) 12
Figure 10 Roberts's Model of Personality 13
Figure 11 AFQT Score Decomposed by IQ, Rosenberg, and Rotter 18
Figure 12 DAT scores and GPA decomposed by IQ and Personality 19
Figure 13 Cumulative Mean-Level Changes in Personality Across the Life Cycle 23
Figure 14 Cumulative Mean-Level Changes in Personality Across the Life Cycle 24
Figure 15 Cumulative Mean-Level Changes in Personality Across the Life Cycle 24
Figure 16 Cumulative Mean-Level Changes in Personality Across the Life Cycle 25
Figure 17 A Life Cycle Framework for Organizing Studies and Integrating Evidence: Period Life Cycle 27
Figure 18 Causal Effect of Schooling on ASVAB Measures of Cognition 27

Figure 19 Causal Effect of Schooling on ASVAB Measures of Cognition 28
Figure 20 Causal Effect of Schooling on ASVAB Measures of Cognition 28
Figure 21 Causal Effect of Schooling on Two Measures of Personality 29
Figure 22 Perry Preschool Program: IQ, by Age and Treatment Group 30

CHAPTER 3

Figure 1 Three Income Inequalities by Bourguignon and Morrisson 76

CHAPTER 5

Figure 1 Phenotypic Stereotyping Equilibria 146
Figure 2 Human Capital Investment with Affective Discrimination Absent 147
Figure 3 Human Capital Investment and Affective Behavior 150
Figure 4 Human Capital Investment Rate 152
Figure 5 ASE given Multiple PSE: Case with $\overline{\pi} > \pi_h$ 154
Figure 6 ASE given Multiple PSE: Case with $\overline{\pi} < \pi_h$ 156
Figure 7 ASE given Unique PSE: Case with $\overline{\pi} < \pi_u$ 162
Figure 8 ASE given Unique PSE: Case with $\overline{\pi} < \pi_u$ 164

CHAPTER 6

Figure 1 Gini Index 183
Figure 2 Decile Ratio (P90/P10) 184
Figure 3 Quintile Ratio (P80/P20) 185
Figure 4 Proportion of household income of the lower-income households 187
Figure 5 Pattern of Relative Poverty Rate 188
Figure 6 Proportion of Household Income of Each Quintile Group 188

CHAPTER 7

Figure 1 Securitization and Mortage Prices 209
Figure 2 A Scatter Plot of Nominal Housing Prices and Mortgage Securitization 220

CHAPTER 8

Figure 1 House Price Dynamics - US and AZ, CA, FL and NV 240
Figure 2 Maximum House Bids and LTV 247
Figure 3 Investor shares in new mortgage borrowing(cont) 255
Figure 4 Investor share in aggregate first mortgage balance 257
Figure 5 Balance share of 2+ first mortgage holders by state 257
Figure 6 Average quarterly transition rates(t-1to t) in # of first mortgages, by year 260
Figure 7 Investor share of purchase origination balances, LP and CCP definitions 262
Figure 8 Share securitized non-prime in aggregate mortgage balance, by investor type 265
Figure 9 Share of aggregate mortg age balance in individual accounts, By investor type 267
Figure 10 Proportion of mortgage pay offs that were originated within The past three years, by investor type 268
Figure 11 Investor share of 90+ DPD first-mortgage balances, by investor type 270
Figure 12-1 Aggregate 90+DPD securitized non-prime mortgage balances($B), by investor type 271
Figure 12-2 Investor share 90+DPD securitized non-prime balances, LP and CCPdefinitions 272
Figure 13 Average quarterly transition rates by investor type 272
Figure 14 Investor effect on quarterly flow into 90+ delinquency 275

CHAPTER 9

Figure 1 Model 291
Figure 2 Optimal Mix (r_1^*, R_1^*) 317
Figure 3 Optimal Mix (r_1^0, R_1^*) 318

CHAPTER 11

Figure 1 Impulse Responses 353

CHAPTER 12

Figure 1 % change of spot rate, forward premium and risk premium 374
Figure 2 One month Eurorate for British Pound and Japanese Yen 387

Figure 3	Canadian Dollar-U.S. Dollar 1 year change & 1 year forward permium.	390
Figure 4	Norwegian Krone-U.K. Pound 1 year change & 1 year forward premium.	390

CHAPTER 1

Integrating Personality Psychology into Economics*

by
James Heckman
(University of Chicago, American Bar Foundation,
University College Dublin)

Abstract

This paper reviews the problems and potential benefits of integrating personality psychology into economics. Economists have much to learn from and contribute to personality psychology.

What can economists learn from and contribute to personality psychology? What do we learn from personality psychology? Personality traits predict many behaviors—sometimes with the same or greater strength as conventional cognitive traits. Personality psychology considers a wider array of actions than are usually considered by economists and enlarges the economist's way to describe and model the

---

* This research was supported by grants from NIH R01-HD054702, R01-HD065072, and K01-AG033182; the University of Chicago; The Institute for New Economic Thinking (INET); A New Science of Virtues: A Project of the University of Chicago; the American Bar Foundation; a conference series from the Spencer Foundation; the JB & MK Pritzker Family Foundation; the Buffett Early Childhood Fund; the Geary Institute, University College Dublin, Ireland; and an anonymous funder. This paper draws on Almlund *et al.*(2011). I thank Angela Duckworth and Tim Kautz for helpful comments.

** 본 논문은 2011 년 8 월 9 일 KDI 에서 개최된 2011 KDI Journal of Economic Policy Conference 에서 발표된 것임.

*** 제임스 헤크만: (e-mail) jjh@uchicago.edu, (address) 1126 E 59th St. Chicago, IL 60637, USA

world. Personality traits are not set in stone. They change over the life cycle. They are a possible avenue for policy intervention.

Personality psychologists lack precise models. Economics provides a clear framework for recasting the field. Economics now plays an important role in clarifying the concepts and empirical content of psychology. More precise models reveal basic identification problems that plague measurement in psychology. At an empirical level, “cognitive” and “noncognitive” traits are not easily separated.

Moreover, personality psychologists typically present correlations and not causal relationships. Many contemporaneously measured relationships suffer from the problem of reverse causality. Economists can apply their tools to define and estimate causal mechanisms. In addition, psychological measures have substantial measurement error. Econometric tools account for measurement error, and doing so makes a difference. Economists formulate and estimate mechanisms of investment—how traits can be changed for the better.

There are major challenges in integrating personality psychology and economics. Economists need to link the traits of psychology with the preferences, constraints and expectation mechanisms of economics. We need to develop rigorous methods for analyzing causal relationships in both fields. We also need to develop a common language and a common framework to promote interdisciplinary exchange.

There is a danger in assuming that basic questions of content and identification have been answered by psychologists at the level required for rigorous economic analysis. In explaining outcomes, how important is the person? How important is the situation? How important is their interaction? I address these issues in this paper.

1. A Brief History of Personality Psychology

Alfred Binet, architect of the first modern intelligence test that became the Stanford-Binet IQ test, noted that performance in school

"...admits of other things than intelligence; to succeed in his studies, one must have qualities which depend on attention, will, and character; for example a certain docility, a regularity of habits, and especially continuity of effort. A child, even if intelligent, will learn little in class if he never listens, if he spends his time in playing tricks, in giggling, is playing truant."

-Binet (1916, p.254)

All later pioneers have made similar statements. Many feature the Big Five trait "Conscientiousness" as a main determinant of success.[1] Before considering the Big Five traits, it is useful to briefly examine the modern concept of cognition by way of contrast.

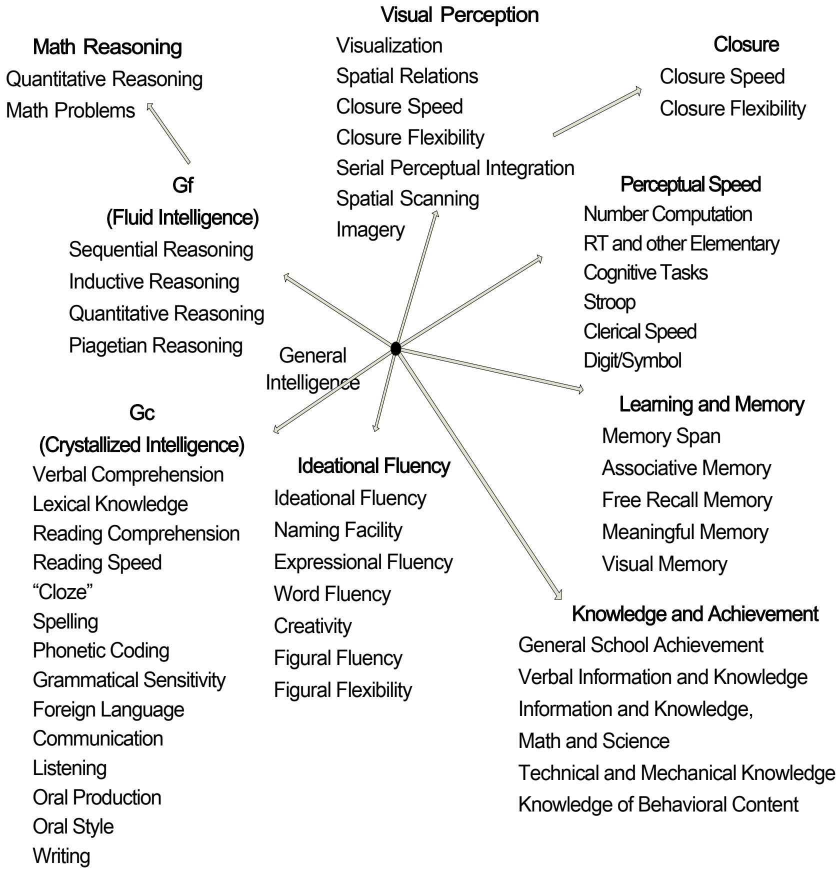
2. Cognition: "g"— a single factor that is claimed to represent intelligence

Traditional "g" is a product of early Twentieth Century psychology. The concept of "g" has been broadened even beyond the traditional subcomponents of "fluid" and "crystallized" intelligence. Figure 1 summarizes current thinking where "g" or general intelligence is at the top of a large pyramid of cognitive traits.

---

1 See Almlund *et al*. (2011).

| Figure 1 | An Hierarchical Scheme of General Intelligence and Its Components



Source: Recreated from Ackerman and Heggestad (1997), based on Carroll (1993).

3. Personality Traits

Early pioneers used a lexical approach to define personality. They classified words that are used to describe people. This practice culminated in the "Big Five" derived from factor analysis of measurements of personality extracted from a variety of measures— observer reports, tests and measured productivity on the job (Costa and McCrae, 1992; Goldberg, 1993). No single "g_p" explains all traits. There are strong correlations within clusters but weak correlations across clusters.

❙ Table 1 ❙ The Big Five Domains and Their Facets

Big Five Personality Factor	American Psychology Association Dictionary description	Facets (and correlated trait adjective)	Related Traits	Childhood Temperament Traits
Conscientiousness	"the tendency to be organized, responsible, and hardworking"	Competence (efficient) Order (organized) Dutifulness (not careless) Achievement striving (ambitious) Self-discipline (not lazy) Deliberation (not impulsive)	Grit Perseverance Delay of gratification Impulse control Achievement striving Ambition Work ethic	Attention/(lack of) distractibility Effortful control Impulse control/delay of gratification Persistence Activity*
Openness to Experience	"the tendency to be open to new aesthetic, cultural, or intellectual experiences"	Fantasy (imaginative) Aesthetic (artistic) Feelings (excitable) Actions (wide interests) Ideas (curious) Values (unconventional)	—	Sensory sensitivity Pleasure in low- intensity activities Curiosity
Extraversion	"an orientation of one's interests and energies toward the outer world of people and things rather than the inner world of subjective experience; characterized by positive affect and sociability"	Warmth (friendly) Gregariousness (sociable) Assertiveness (self-confident) Activity (energetic) Excitement seeking (adventurous) Positive emotions (enthusiastic)	—	Surgency Social dominance Social vitality Sensation seeking Shyness* Activity* Positive emotionality Sociability/affiliation
Agreeableness	"the tendency to act in a cooperative, unselfish manner"	Trust (forgiving) Straight-forwardness (not demanding) Altruism (warm) Compliance (not stubborn) Modesty (not show-off) Tender-mindedness (sympathetic)	Empathy Perspective taking Cooperation Competitiveness	Irritability* Aggressiveness Willfulness
Neuroticism/ Emotional Stability	Emotional stability is "predictability and consistency in emotional reactions, with absence of rapid mood changes." Neuroticism is "a chronic level of emotional instability and proneness to psychological distress."	Anxiety (worrying) Hostility (irritable) Depression (not contented) Self-consciousness (shy) Impulsiveness (moody) Vulnerability to stress (not self-confident)	Internal vs. External Locus of control Core self-evaluation Self-esteem Self-efficacy Optimism Axis I psychopathologies (mental disorders) including depression and anxiety disorders	Fearfulness/behavioral inhibition Shyness* Irritability* Frustration (Lack of) soothability Sadness

Notes: Facets specified by the NEO-PI-R personality inventory (Costa and McCrae [1992]). Trait adjectives in parentheses from the Adjective Check List (Gough and Heilbrun [1983]). *These temperament traits may be related to two Big Five factors.

Source: Table adapted from John and Srivastava (1999).

The Big Five predict many outcomes. The Big Five are defined without reference to any context (i.e., situation). This practice gives rise to an identification problem that I discuss below.

4. The Person-Situation Debate: A Strong Influence on Behavioral Economics

Is variation across people in behavior a consequence of personal traits or of situations? Economists are still badly divided over this question. The modern origins of the debate start with the works of psychologist Walter Mischel:

> *"...with the possible exception of intelligence, highly generalized behavioral consistencies have not been demonstrated, and the concept of personality traits as broad dispositions is thus untenable"*
> -Mischel (1968, p.146)

Many behavioral economists hold a similar view and appeal to Mischel as a guiding influence.

> *"The great contribution to psychology by Walter Mischel [. . .] is to show that there is no such thing as a stable personality trait."*
> -Thaler (2008)

The accumulated evidence speaks strongly against the claims of Mischel and the behavioral economists.[2]

5. Personality Psychology After the Person-Situation Debate

Correlational evidence shows that for many outcomes, measured personality traits are as predictive, and are sometimes more predictive, than standard measures of cognition. Traits are stable across situations.

2 See Almlund *et al*. (2011).

Situations also matter. Behavioral genetics show that personality traits are as heritable as cognitive traits. Alterations in brain structure and function through accidents, disease and by experiments affect measured personality.[3]

6. The Predictive Power of Personality Traits

A growing body of evidence suggests that personality measures–especially those related to Conscientiousness, and, to a lesser extent, Neuroticism–predict a wide range of outcomes. The predictive power of any particular personality measure tends to be less than the predictive power of IQ but in some cases rivals or exceeds it.

7. Difficulties in Synthesizing Studies of the Effects of Personality

Measures of personality and cognition differ among studies. Different studies use different measures of predictive power. Many studies do not address the question of causality, i.e., does the measured trait cause (rather than just predict) the outcome?

Few economists or psychologists working on the relationship between personality and outcomes address the issue of causality, and when they do so, it is usually by employing early measures of cognition and personality to predict later outcomes. This practice trades an endogeneity problem with an errors in variables problem. Almlund *et al*. (2011) discuss alternative approaches to causality building on the analysis of Hansen *et al*. (2004).

8. Main Findings from Predictive Analyses

The predictive power of "g" decreases with the level of job complexity.

---

3 See Almlund *et al*. (2011).

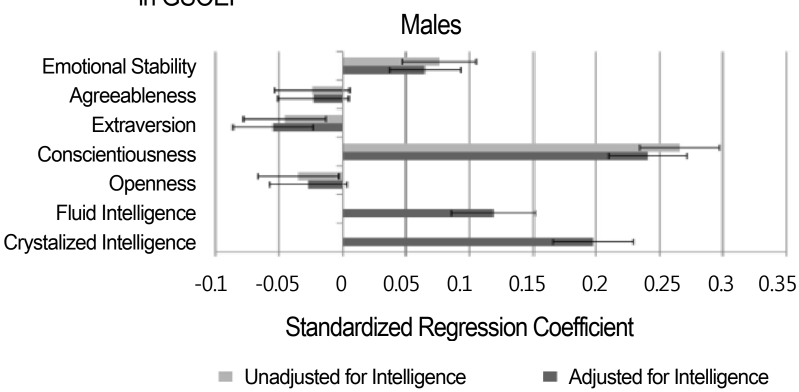
Personality traits are predictive at all levels of job complexity. Conscientiousness is the most predictive Big Five trait across many outcomes such as educational attainment, grades, job performance across a range of occupational categories, longevity and criminality. Neuroticism (and related Locus of Control) predicts schooling outcomes and labor market search. Other traits play roles at finer levels. I now present examples of the power of personality traits.

8. 1. Educational Attainment and Achievement

In explaining educational attainment, Conscientiousness plays a powerful role. See Figure 2.

Another example is the GED in America. GEDs are high school dropouts who exam certify to be high school equivalents. They have the same cognitive skills as high school graduates but much lower noncognitive skills. See Figures 3 and 4.

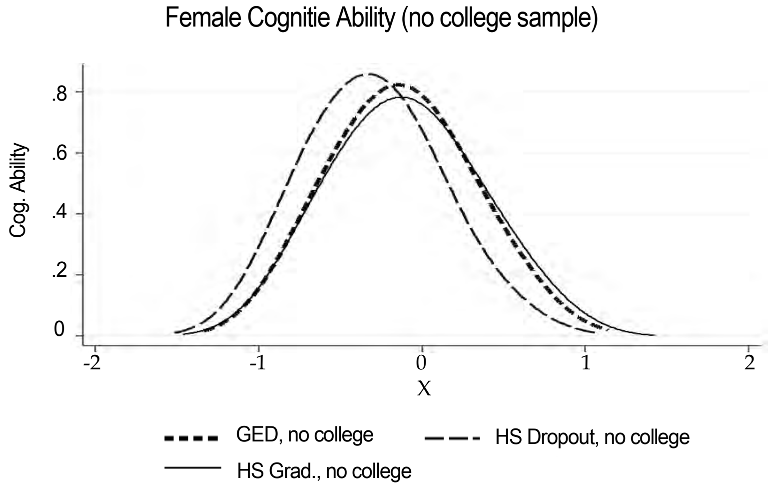
Figure 2 Association of the Big Five and Intelligence with Years of Schooling in GSOEP



Note: The figure displays standardized regression coefficients from multivariate of years of school attended on the Big Five and intelligence, controlling for age and age-squared. The bars represent standard errors. The Big Five coefficients are corrected for attenuation bias. The Big Five were measured in 2005. Years of schooling were measured in 2008. Intelligence was measured in 2006. The measures of intelligence were based on components of the Wechsler Adult Intelligence Scale (WAIS). The data is a representative sample of German adults between the ages of 21 and 94.

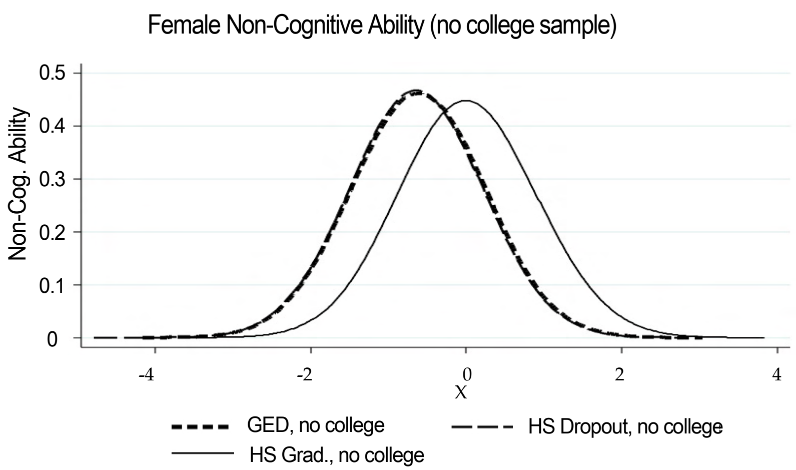
Source: German Socio-Economic Panel (GSOEP), waves 2004-2008, calculations performed by Pia Pinger (See Almlund *et al.*, 2011.).

| Figure 3 | Distribution of Cognitive and Non-Cognitive Skills by Education Group



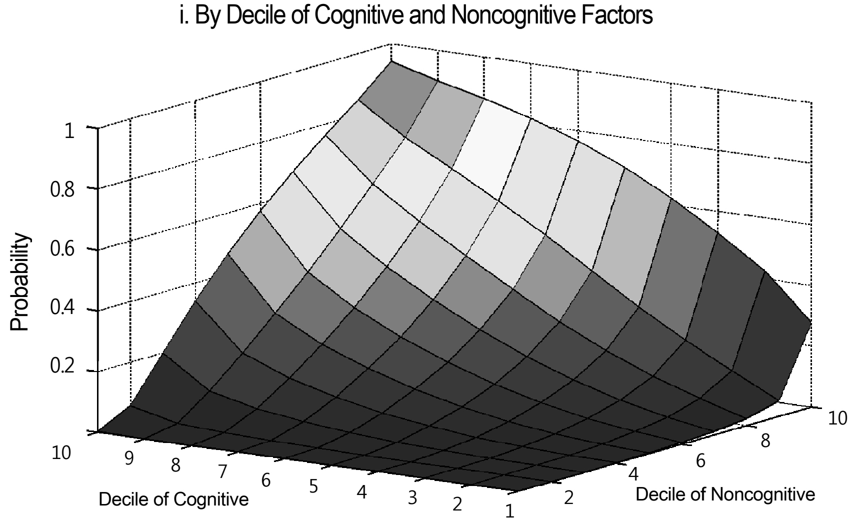
Source: Heckman *et al.* (2011).

| Figure 4 | Distribution of Cognitive and Non-Cognitive Skills by Education Group



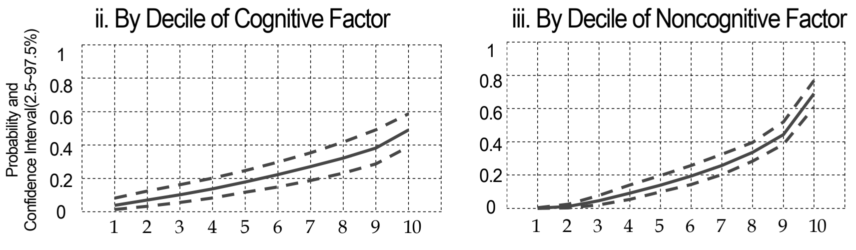
Source: Heckman *et al.* (2011).

| Figure 5 | Probability of Being a 4-year-college Graduate or Higher at Age 30, Males



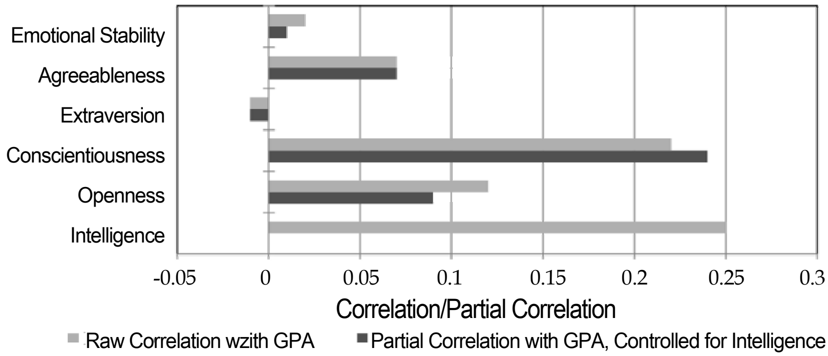
Notes: The data are simulated from the estimates of the model and the NLSY79 sample. Higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5~97.5% confidence intervals. The upper curve is the joint density. The two marginal curves (ii) and (iii) are evaluated at the mean of the trait not being varied.
Source: Heckman *et al*. (2006, Figure 21).

| Figure 6 | Probability of Being a 4-year-college Graduate or Higher at Age 30, Males



Notes: The data are simulated from the estimates of the model and the NLSY79 sample. Higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5~97.5% confidence intervals. The upper curve is the joint density. The two marginal curves (ii) and (iii) are evaluated at the mean of the trait not being varied.
Source: Heckman *et al*. (2006, Figure 21).

| Figure 7 | Correlations of the Big Five and Intelligence with Course Grades



Notes: All correlations are significant at the 1% level. The correlations are corrected for scale reliability and come from a meta analysis representing a collection of studies representing samples of between N=31,955 to N=70,926, depending on the trait. The meta-analysis did not clearly specify when personality was measured relative to course grades.
Source: Poropat (2009).

GEDs earn at the rate of dropouts. Their lower levels of noncognitive skill leads to lower wages than ordinary high school graduates even though they have the same level of cognitive skills.

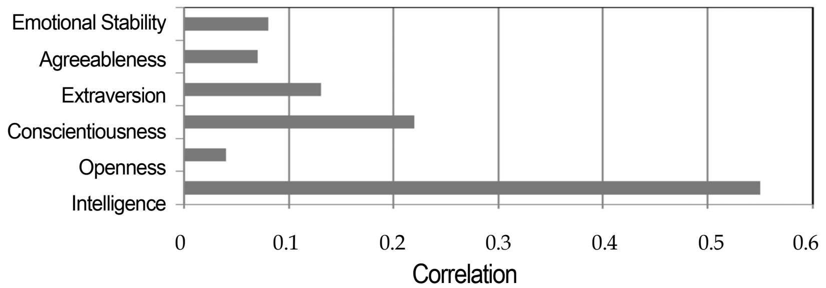
Cognitive and noncognitive skills are both important in explaining college graduation. See Figures 5 and 6. Persons with low levels of noncognitive skills are unlikely to graduate college, as are persons with low levels of cognitive skills.

Similar results hold for course grades. See Figure 7. Indeed, course grades are a good measure of conscientiousness (See Almlund *et al*., 2011; Borghans *et al*., 2011.).

8. 2. Labor Market Outcomes

Intelligence is the greatest single predictor of job performance, especially in complex tasks, but noncognitive skills are also important predictors. See Figure 8.

| Figure 8 | Associations with Job Performance



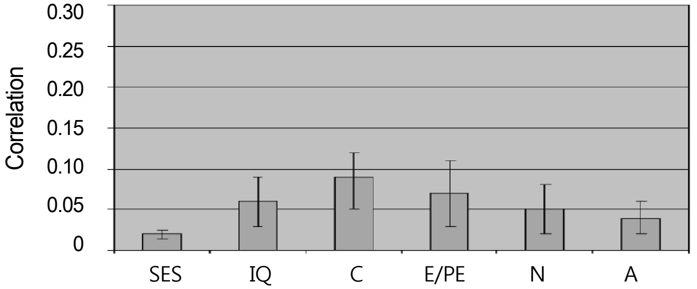
Notes: The values for personality are correlations that were corrected for sampling error, censoring, and measurement error. Job performance was based on performance ratings, productivity data and training proficiency. The authors do report the timing of the measurements of personality relative to job performance. Of the Big Five, the coefficient on Conscientiousness is the only one that is statistically significant with a lower bound on the 90credibility value of 0.10. The value for IQ is a raw correlation.

Sources: The correlations reported for personality traits come from a meta-analysis conducted by Barrick and Mount (1991). The correlation reported for IQ and job performance come from Schmidt and Hunter (2004).

8. 3. Longevity

Personality traits also predict longevity. In particular, Conscientiousness is a better predictor than IQ. See Figure 9.

| Figure 9 | Correlations of Mortality with Personality, IQ, and Socioeconomic Status (SES)



Notes: The figure represents results from a meta-analysis of 34 studies. Average effects (in the correlation metric) of low socioeconomic status (SES), low IQ, low Conscientiousness (C), low Extraversion /Positive Emotion (E/PE), Neuroticism (N), and low Agreeableness (A) on mortality. Error bars represent standard error. The lengths of the studies represented vary from 1 year to 71 years.

Source: Roberts *et al.* (2007).

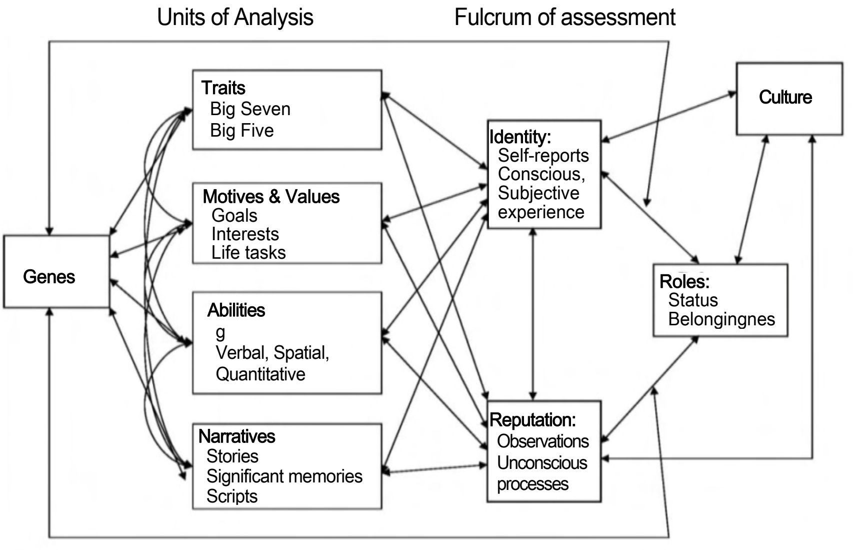
9. Conceptualizing Personality Within an Economic Model

How should one conceptualize these correlations and establish a causal basis for them? Recent work (Almlund *et al.*, 2011) develops economic models of personality and their implications for measurement of personality and preference. They place the concept of personality within an economic framework. Personality is defined as an emergent property of a system. Economic models frame and solve a central identification problem in empirical psychology: How to go from measurements of personality to personality traits.

It is important to distinguish **personality traits** from **measured personality**. One definition of personality by a leading **psychologist** is:

> *"Personality traits are the relatively enduring patterns of thoughts, feelings, and behaviors that reflect the tendency to respond in certain ways under certain circumstances."*
>
> -Roberts (2009, p.140)

| Figure 10 | Roberts's Model of Personality



Source: Roberts (2006).

His conceptual framework for personality is presented in Figure 10. Personality is a property of a system. This type of analysis is typical of the models used in personality psychology.

10. An Economic Framework for Conceptualizing and Measuring Personality and Personality Traits

How can we interpret personality within economic models? Through preferences (the standard approach), constraints (Borghans *et al.*, 2008) or through expectations? Or does it operate through all three?

10. 1. Personality Affects Productivity

Almlund *et al.* (2011) develop models in which productivity in task j depends on the traits of agents represented by trait vector θ, and the "effort" they expend on the task, e_j:

$$P_j = \phi_j(\theta, e_j), \quad j \in \mathcal{J} = \{1, \ldots, J\}, e_j \in \mathcal{E}, \theta \in \Theta \tag{1}$$

Traits θ are endowments, like a public good. $\sum_{j=1}^{J} e_j = \bar{e}$. $\bar{e}$ is endowment.

$\phi_j(\theta, e_j)$ is concave and increasing in e_j; $\frac{\partial^2 \phi_j}{\partial\theta\partial e_i'} \geq 0, \forall j$. R_j is the reward per unit task output. The agent in assumed to maximize

$$\sum_{j=1}^{J} R_j \phi_j\,(\theta, e_j) \tag{2}$$

with respect to $\{e_j\}_{j=1}^{J}$ subject to the constraint $\sum_{j=1}^{J} e_j = \bar{e}$. In general as $R_j \uparrow e_j \uparrow$.

Effort in one task might diminish effort in another. If tasks are mutually exclusive, we obtain the Roy model (Heckman and Honoré, 1990; Heckman and Sedlacek, 1985).

10. 2. Identifying Personality Traits From Measured Performance on Tasks

I next consider a basic identification problem. Some tasks may require only a single trait or only a subset of all of the traits. Divide θ into "mental" (μ) and "personality" (π) traits, θ_μ and θ_π. To use performance on a task (or on multiple measures of the task) to identify a trait requires that performance on certain tasks (performance on a test, performance in an interpersonal situation, etc.) depends exclusively on one component of θ, say $\theta_{1,j,}$ as well as on the effort used in the task. Thus measurement assumes task j output is generated by the following relationship:

$$P_j = \phi_j\,(\underbrace{\theta_{1,j}}_{\substack{\text{single}\\ \text{trait}\\ \text{used in}\\ \text{trait } j}},\ e_j)$$

We need to standardize for effort at a benchmark level, say e^*, to use P_j to identify a measure of the trait $\theta_{1,j}$.

The activity of picking a task (or a collection of tasks) that measure a particular trait ($\theta_{1,j}$ in our example) is called **operationalization** in psychology. Demonstrating that a measure successfully operationalizes a trait is called **construct validity**. Note, however, that we need to standardize for effort to measure the trait. Otherwise variation in effort produces variation in the measured trait across situations with different incentives.

10. 3. A Fundamental Identification Problem

Operationalization and construct validation require heroic assumptions. Even if one adjusts for effort in a task, measured productivity may depend on multiple traits. Thus two components of θ (say $\theta_{1,\mu}, \theta_{1,\pi}$) may determine productivity in j. Without further information, one cannot infer which of the two traits produces the productivity in j. In general, even having two (or more) measures of productivity that depend on $(\theta_{1,\mu}, \theta_{1,\pi})$ is not enough to identify the separate components.

Consider the following case of two productivity measures for the two tasks j and j':

$$P_j = \phi_j(\theta_{1,\mu}, \theta_{1,\pi}, e_j)$$

$$P_{j'} = \phi_{j'}(\theta_{1,\mu}, \theta_{1,\pi}, e_{j'}), \qquad j \neq j'.$$

Standardize measurements at a common level of effort $e_j = e_{j'} = e^*$. Note that if the supports of e_j and $e_{j'}$ are disjoint, no $(\theta_{1,\mu}, \theta_{1,\pi})$ exists. Assume that the $\phi_k()$ are known. If the system of equations satisfies a local rank condition, then one can solve for the pair $(\theta_{1,\mu}, \theta_{1,\pi})$ at e^*. **Only the pair is identified**. One cannot (without further information) determine which component of the pair the $\theta_{1,\mu}$ or $\theta_{1,\pi}$.

In the absence of **dedicated constructs** (constructs that are generated by only one component of θ), there is an intrinsic identification problem that arises in using measures of productivity in tasks to infer traits. Analysts have to make one normalization in order to identify the traits. However, we need only one such construct joined with patterned structures on how θ enters other task to identify the vector θ (e.g. one example is a recursive, triangular structure). See the discussion in Almlund *et al*. (2011).

10. 4. Examples of Nonidentification

IQ and achievement test scores reect incentives and efforts, and capture both cognitive and personality traits. Table 2 summarizes the evidence that paying disadvantaged students for correct answers on IQ tests substantially raises measured IQ. Almlund *et al*. (2011) summarize many other studies

A considerable fraction of the variance in achievement tests is explained by personality traits. See Figure 11. Grades are explained more by the Big Five traits than by IQ. See Figure 12.

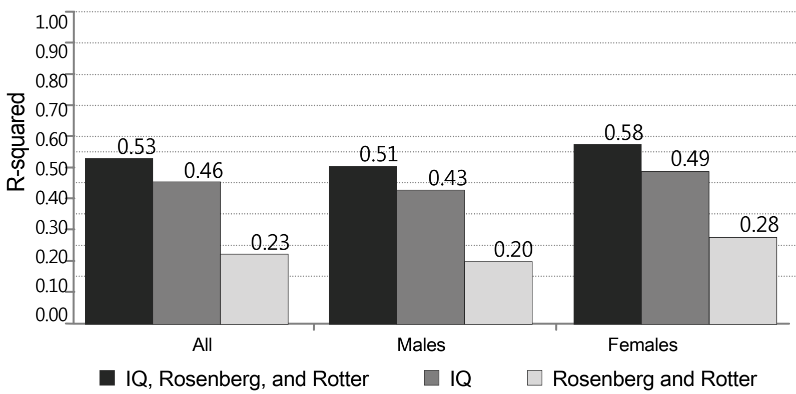
Table 2 Incentives and Performance on Intelligence Tests

Study	Sample and Study Design	Experimental Group	Effect size of incentive (in standard deviations)	Summary
Edlund (1972)	Between subjects study. 11 matched pairs of low SES children; children were about one standard deviation below average in IQ at baseline	M&M candies given for each right answer	Experimental group scored 12 points higher than control group during a second testing on an alternative form of the Stanford Binet (about 0.8 standard deviations)	"…a carefully chosen consequence, candy, given contingent on each occurrence of correct responses to an IQ test, can result in a significantly higher IQ score." (p.319)
Breuning and Zella (1978)	Within and between subjects study of 485 special education high school students all took IQ tests, then were randomly assigned to control or incentive groups to retake tests. Subjects were below-average in IQ.	Incentives such as record albums, radios (<$25) given for improvement in test performance	Scores increased by about 17 points. Results were consistent across the Otis-Lennon, WISC-R, and Lorge-Thorndike tests.	"In summary, the promise of individualized incentives contingent on an increase in IQ test performance (as compared with pretest performance) resulted in an approximate 17-point increase in IQ test scores. These increases were equally spread across subtests… The incentive condition effects were much less pronounced for students having pretest IQs between 98 and 120 and did not occur for students having pretest IQs between 121 and 140." (p.225)

10. 5. Measures of Personality in Psychology Based on Linear Factor Analysis

Such measures account for measurement error, and identify factors that can be interpreted as traits. Cunha *et al*. (2010) develop nonlinear factor models (nonlinear and nonparameteric). Using these models they establish that measurement error is quantitatively important. The share of error variance for proxies of cognition, personality and investment ranges from 1~90%. Not accounting for measurement error produces downward-biased estimates of self-productivity effects and perverse estimates of investment effects.

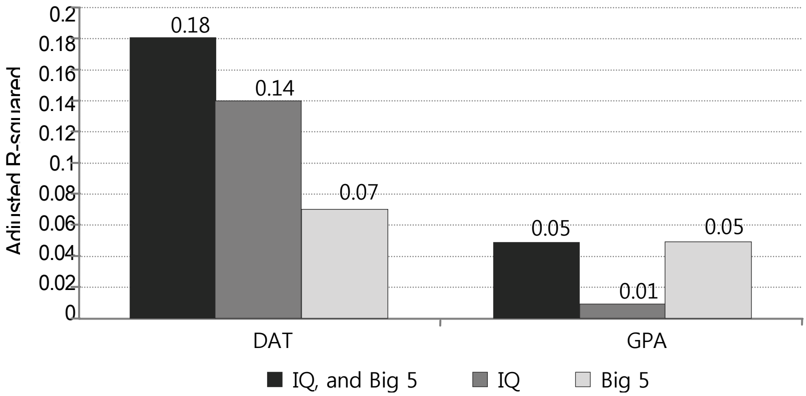
| Figure 11 | AFQT Score Decomposed by IQ, Rosenberg, and Rotter



Notes: The data come from the NLSY. Rosenberg, and Rotter were administered in 1979. The ASVAB was administered in 1980. To account for varying levels of schooling at the time of the test, scores have been adjusted for schooling at the time of the test conditional on final schooling using the method developed in Hansen *et al*. (2004). AFQT is constructed from the Arithmetic Reasoning, Word Knowledge, Numeric Operations, and Paragraph Comprehension ASVAB subtests. DAT and DAT percentile, IQ, and GPA are from high school transcript data. IQ is pooled across several IQ tests using IQ percentiles. GPA is the individual's core-subject GPA from each year of school. Sample excludes the military over-sample. Background variables include mother's highest grade completed, father's highest grade completed, southern residence at age 14, urban residence at age 14, living in a broken home at age 14, receiving newspapers in the household at age 14, receiving magazines in the household at age 14, and the household having a library card at age 14.

Source: Borghans *et al*. (2011).

| Figure 12 | DAT scores and GPA decomposed by IQ and Personality



Notes: Data is from Stella Maris, a high school in the Netherlands. Students were administered part of a Raven's IQ test and personality questions based on the Big 5. DAT and GPA are from high school records.

Source: Borghans *et al.* (2011).

11. A Definition of Personality

I now add preferences and goals to the analysis. Preferences and goals also shape effort. They are personality traits broadly defined. Income is the return to productivity:

$$\text{Income} = \sum_{j=1}^{J} R_j \, P_j$$

Preferences are defined over final consumption goods X, productivity P and effort e:

$$U(X, P, e \mid \psi), \psi \in \Psi. \tag{3}$$

Agents have preferences over goods, agents may value the output of tasks in their own right and agents may value the effort devoted to tasks. The agents maximize (3) with respect

$$\underbrace{Y}_{\substack{\text{exogenous}\\ \text{flow of}\\ \text{income}}} + R'P = \underbrace{W'}_{\substack{\text{prices}\\ \text{of}\\ \text{goods}}} X \tag{4}$$

11. 1. Adding Uncertainty

Let $\mathcal{I}$ be the information possessed by an agent. "E" denotes the expectation operator. The agent can be interpreted as making decisions based on

$$E[U\,(X, P, e \mid \psi)|\mathcal{I}]. \tag{5}$$

11. 2. Personality Traits

Personality traits are the components of e, θ and ψ that affect behavior. We observe **measured personality**—behaviors generated by incentives, goals, and traits.

11. 3. Actions

Actions are styles of behavior that affect how tasks are accomplished. They are aspects of behavior that go beyond effort. Smiling, cajoling, etc. are examples. Tasks are accomplished by taking actions. The i^{th} possible action to perform task j is denoted $a_{i,j}, i \in \{1, \dots, K_j\}$. Array actions in a vector $a_j = \left(a_{1,j}, \dots, a_{K_j,j}\right) \in \mathcal{A}$. Actions may be the same or different across the tasks. The productivity of the agent in task j depends on the actions taken in that task:

$$P_j = \tau_j\left(a_{1,j}, a_{2,j}, \dots, a_{k_j,j}\right). \tag{6}$$

The actions themselves depend on traits θ and "effort" $e_{i,j}$:

$$a_{i,j} = \nu_{i,j}\,(\theta, e_{i,j}) \tag{7}$$

where

$$\sum_{i=1}^{K_j} e_{i,j} = e_j \text{ and } \sum_{j=1}^{J} e_j = \bar{e}.$$

Actions generalize the notion of effort to a broader class of behaviors. Let $\mathcal{M}$ be the set of actions, including actions that do not directly contribute to productivity. Let M be the index set of items in $\mathcal{M}$.

$$a_{i,m} = \nu_{i,m}\left(\theta, e_{i,m}\right), m \in M, \mathcal{A} \subseteq \mathcal{M}.$$

The agent solves

$$\max E\left[U\left(a, X, P, e \mid \psi\right) \mid \mathcal{I}\right]$$

with respect to X and e given the stated constraints.

We can introduce situations indexed by $h \in \mathcal{H}$. For a person with traits θ and effort vector e_j with action $a_{i,j}$, using the specification (7), the action function can be expanded to be dependent on situation h:

$$a_{i,j,h} = \nu_{i,j}\left(\theta, e_{i,j,h}, h\right). \tag{8}$$

11. 4. A Definition of Personality

Let $T \in \mathcal{T}$ be a vector of traits $(\theta, \psi, \bar{e})$. Personality is a response function.

Personality: $a = a(R, W, T, h, Y, \mathcal{I}).$ (9)

The behavior that constitutes personality is defined as a pattern of actions in response to the constraints, endowments, and incentives facing agents given their goals and preferences.

Actions—not traits—constitute the data used to identify the traits. Personality psychologists use actions (e.g., "dispositions") to infer traits. Identification issues similar to those previously discussed apply to this broader set of measurements of behaviors.

11. 5. Personality as Enduring Actions

Many personality psychologists define personality as "*enduring patterns of thoughts, feelings and behaviors*" that reflect tendencies of persons to respond in certain ways under certain circumstances (See Cervone and Pervin [2009]). What are enduring patterns of actions? "Enduring actions" are the average of the a functions for a person with a given trait vector $T = t$ over situations and efforts.

11. 6. Average Actions

Consider task j and trait vector $T = (\theta, \psi, \bar{e})$. Define the average action for information set $\mathcal{I}$:

$$\bar{a}_{T,j,\mathcal{I}} = \int_{S_{T,\mathcal{I}}(h,e_{i,j})} v_{i,j}\left(\theta, e_{i,j}, h\right) g\left(h, e_{i,j} \mid T = (\theta, \psi, \bar{e}), \mathcal{I}\right) dh \, de_{i,j}$$

where $\mathcal{S}_{T,\mathcal{I}}(h, e_{i,j})$ is the support $(h, e_{i,j})$ of given T and $\mathcal{I}$. $g\left(h, e_{i,j} | T = (\theta, \psi, \bar{e}), \mathcal{I}\right)$ is the density of $(h, e_{i,j})$ given $T = (\theta, \psi, \bar{e})$ and information set $\mathcal{I}$. $\bar{a}_{T,i,\mathcal{I}}$ is the "enduring action" of agents across situations in task j with information $\mathcal{I}$, i. e., the **average personality**. Only if $v_{i,j}$ is separable in the T, the marginal effect of personality trait vector θ is the same in all situations.

One can define the "enduring traits" in a variety of ways, say by averaging over tasks, j, situations, h, or both. Only under separability in T will one obtain the same marginal effect of θ. Epstein (1979) and a subsequent literature present evidence against nonseparability but in favor of an "enduring trait" that is common across situations. He argues strongly against the extreme form of situational specificity assumed in modern behavioral economics.

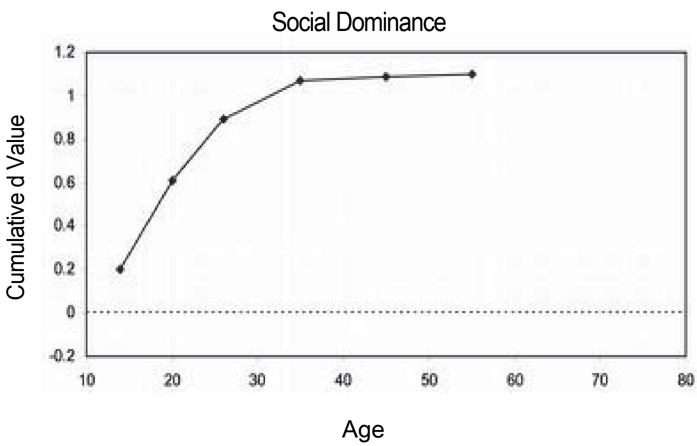
12. Stability and Change in Personality Traits and Preferences

While it is commonly thought that personality traits are stable, at least in adult life, in fact traits change over the life cycle. See Figures 13~16.

12. 1. Processes of Development Discussed in the Literature

There are many hypothesized mechanisms of change. Two common processes discussed in the literature are ontogeny (programmed developmental processes common to all persons) and sociogeny (shared socialization processes). Personality also changes through external forces above and beyond common ontogenic and sociogenic processes. Such changes operate through alterations in normal biology, such as brain lesions and chemical interventions. A channel that receives a lot of attention in economics is investment: educational interventions and parental investment that affect personality throughout the life cycle.

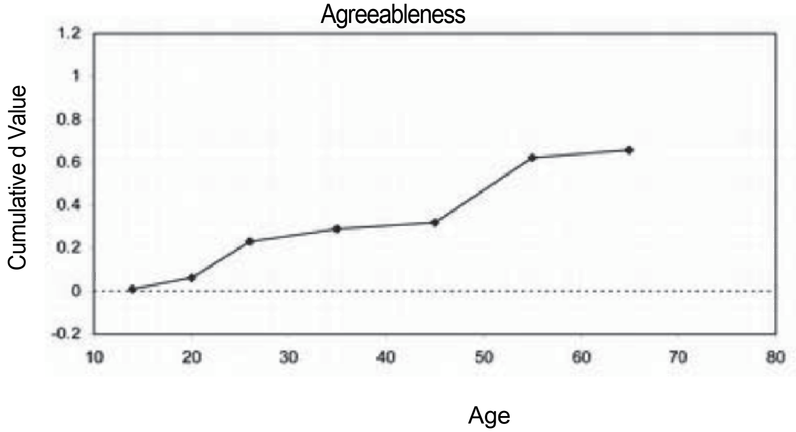
| Figure 13 | Cumulative Mean-Level Changes in Personality Across the Life Cycle



Note: Social vitality and social dominance are aspects of Big Five Extraversion. Cumulatived values represent total lifetime change in units of standard deviations ("effect sizes").

Source: Figure taken from Roberts *et al.* (2006) and Roberts and Mroczek (2008). Reprinted with permission of the authors.

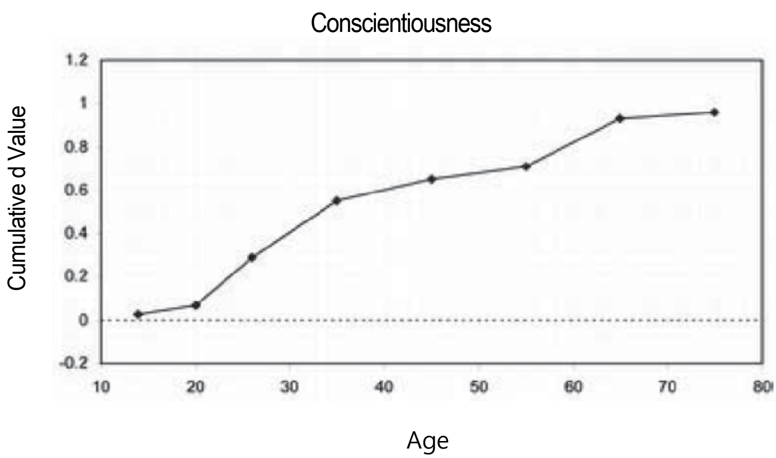
| **Figure 14** | Cumulative Mean-Level Changes in Personality Across the Life Cycle



Note: Social vitality and social dominance are aspects of Big Five Extraversion. Cumulative d values represent total lifetime change in units of standard deviations ("effect sizes").

Source: Figure taken from Roberts *et al.* (2006) and Roberts and Mroczek (2008). Reprinted with permission of the authors.

| **Figure 15** | Cumulative Mean-Level Changes in Personality Across the Life Cycle



Note: Social vitality and social dominance are aspects of Big Five Extraversion. Cumulative d values represent total lifetime change in units of standard deviations ("effect sizes").

Source: Figure taken from Roberts *et al.* (2006) and Roberts and Mroczek (2008). Reprinted with permission of the authors.

| Figure 16 | Cumulative Mean-Level Changes in Personality Across the Life Cycle

Note: Social vitality and social dominance are aspects of Big Five Extraversion. Cumulative d values represent total lifetime change in units of standard deviations ("effect sizes").

Source: Figure taken from Roberts *et al.* (2006) and Roberts and Mroczek (2008). Reprinted with permission of the authors.

12. 2. Life Cycle Dynamics

Let T^{υ} be traits at age $\upsilon, \upsilon \in \{1, \dots, V\} \in \mathcal{V}$. Information $\mathcal{I}^{\upsilon}$ may be updated through various channels of learning. The technology of skill formation (Cunha and Heckman, 2007, 2009) postulates the following equation of motion:

$$T^{\upsilon+1} = \eta^{\upsilon}\left(\underbrace{T^{\upsilon}}_{\text{self-productivity}}, \underbrace{IN^{\upsilon}}_{\text{investment}}, h^{\upsilon}\right), \upsilon = 0, \dots, V-1. \tag{10}$$

Functions can be nonautonomous (υ-dependent). Situations may change over time as a function of past actions, past situations, investment, information, and the like:

$$h^{\upsilon+1} = \chi^{\upsilon}(h^{\upsilon}, IN^{\upsilon}, a^{\upsilon}). \tag{11}$$

Information $\mathcal{I}^{\upsilon}$ may also change over the life cycle through experimentation and learning:

$$\mathcal{J}^{\upsilon+1} = \rho^{\upsilon}(\mathcal{J}^{\upsilon},\ a^{\upsilon},\ T^{\upsilon}, IN^{\upsilon}, h^{\upsilon}). \tag{12}$$

Figure 17 summarizes the dynamics of skill formation as formulated in Cunha and Heckman (2007, 2009).

Cunha *et al*. (2010) estimate technology (10) using longitudinal data on the development of children with rich measures of parental investment and of child traits. Self-productivity becomes stronger as children become older, for both cognitive and noncognitive capability formation. The elasticity of substitution for cognitive inputs is smaller in the adolescent years, so that it is more difficult to compensate for the effects of adverse environments on cognitive endowments at later ages than it is at earlier ages.

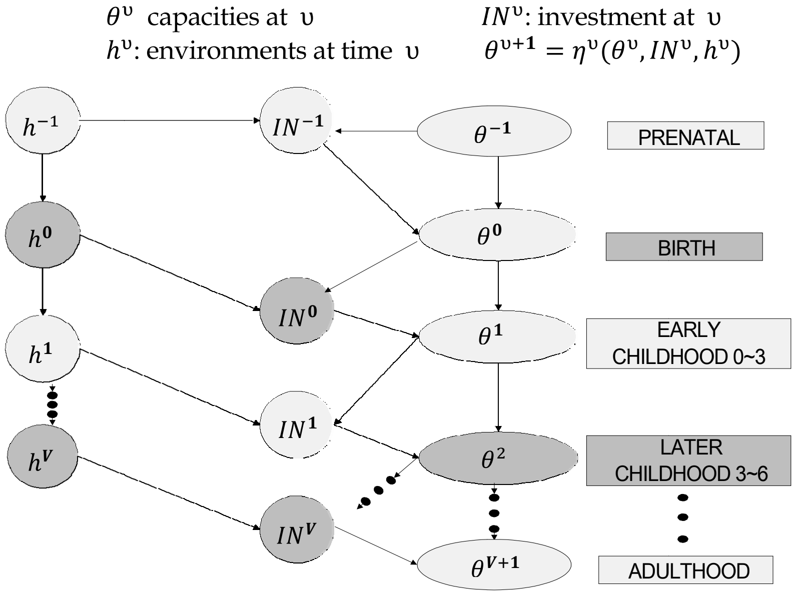
This finding explains the evidence on ineffective cognitive remediation strategies for disadvantaged adolescents. Personality traits foster the development of cognition but not vice versa. Cunha *et al*. (2010) show that it is equally easy to substitute for deficits in personality traits at both early and late stages for socioemotional skills over the life cycle.

Overall, 16% of the variation in educational attainment is explained by factors extracted from adolescent cognitive traits, 12% is due to factors extracted from adolescent personality (socioemotional traits), and 15% is due to factors extracted from measured parental investments.

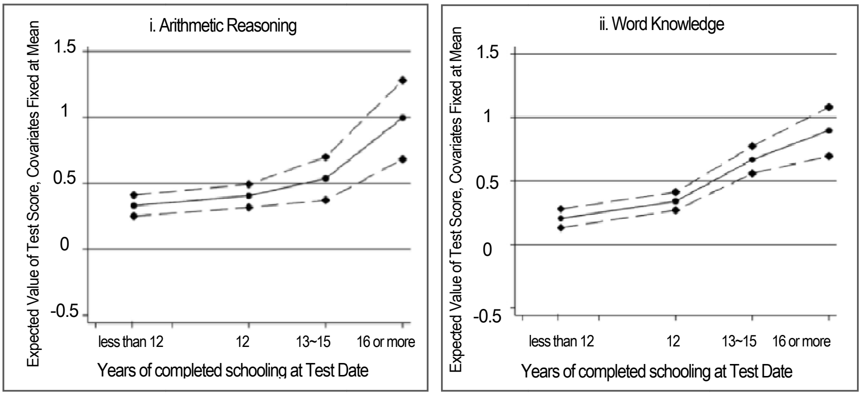
12. 3. The Causal Effects of Schooling on Cognitive and Personality Traits

Using the methodology of Hansen *et al*. (2004), it is possible to estimate the causal effect of schooling on cognitive and noncognitive measurements. See Figures 18~21. Schooling has substantial effects on both types of traits.

| Figure 17 | A Life Cycle Framework for Organizing Studies and Integrating Evidence: Period Life Cycle



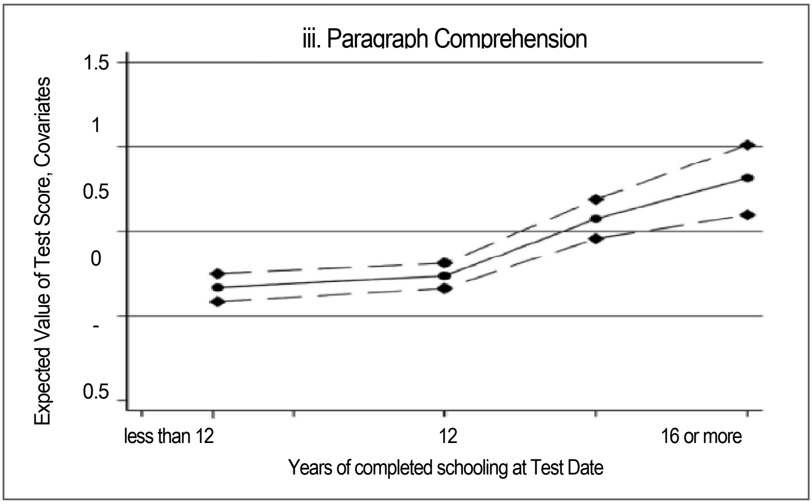
| Figure 18 | Causal Effect of Schooling on ASVAB Measures of Cognition



Notes: Effect of schooling on components of the ASVAB. The first four components are averaged to create male's with average ability. We standardize the test scores to have within-sample mean zero, variance one. The model is estimated using the NLSY79 sample. Solid lines depict average test scores, and dashed lines, confidence intervals.

Source: Heckman *et al.* (2006, Figure 4).

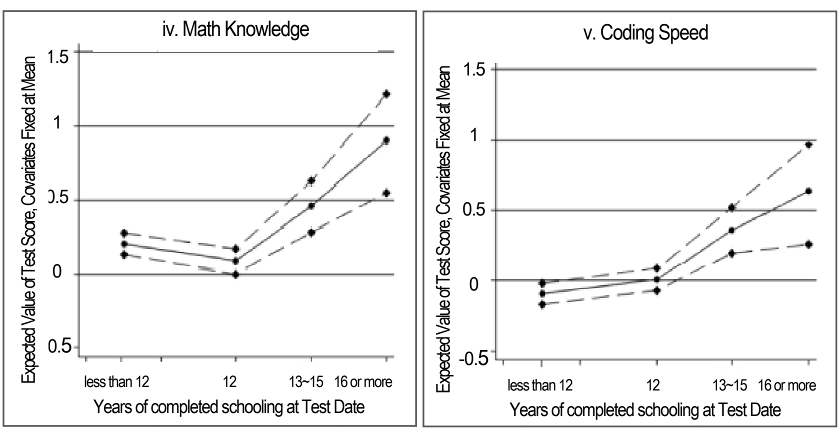
| Figure 19 | Causal Effect of Schooling on ASVAB Measures of Cognition



Notes: Effect of schooling on components of the ASVAB. The first four components are averaged to create male's with average ability. We standardize the test scores to have within-sample mean zero, variance one. The model is estimated using the NLSY79 sample. Solid lines depict average test scores, and dashed lines, confidence intervals.

Source: Heckman *et al*. (2006, Figure 4).

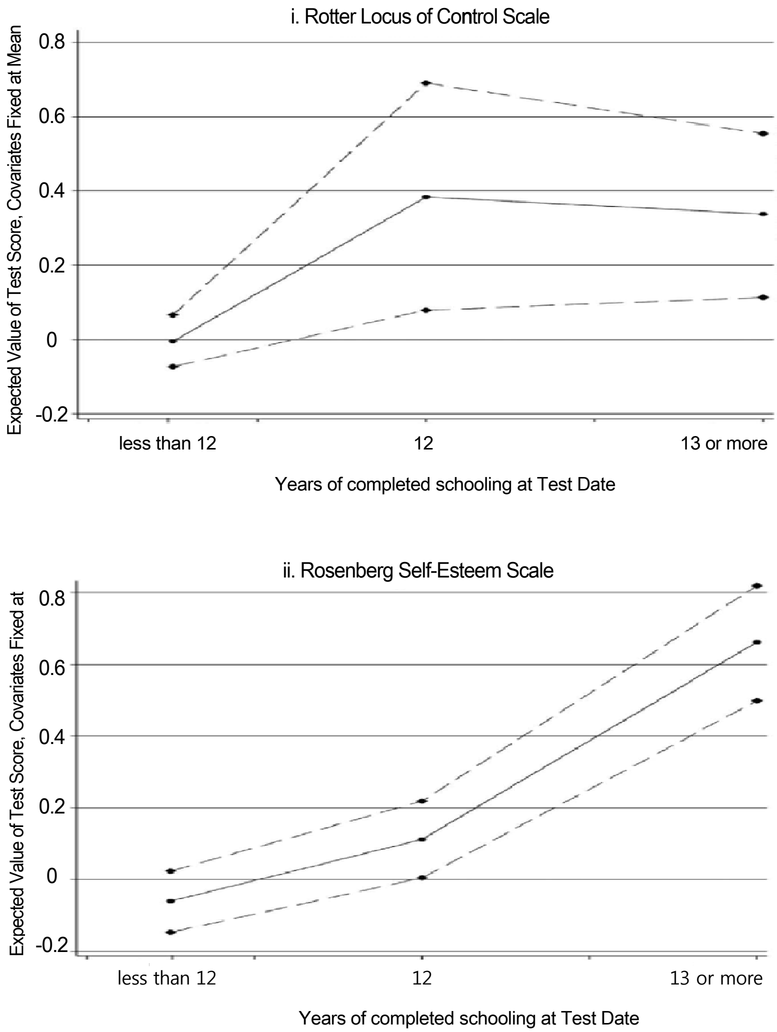
| Figure 20 | Causal Effect of Schooling on ASVAB Measures of Cognition



Notes: Effect of schooling on components of the ASVAB. The first four components are averaged to create male's with average ability. We standardize the test scores to have within-sample mean zero, variance one. The model is estimated using the NLSY79 sample. Solid lines depict average test scores, and dashed lines, confidence intervals.

Source: Heckman *et al*. (2006, Figure 4).

| Figure 21 | Causal Effect of Schooling on Two Measures of Personality



Notes: The figures show the causal effect of schooling on two measures of personality traits. We standardize the test scores to have within-sample mean zero, variance one. The model is estimated using the NLSY79 sample. Solid lines depict average test scores, and dashed lines, confidence intervals.

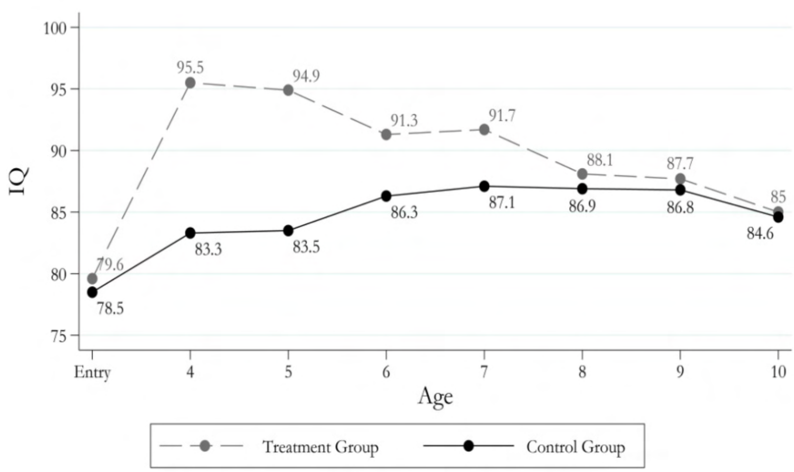
Source: Heckman *et al*. (2006, Figure 4).

12. 4. The Evidence from Interventions

The Perry Preschool program intervened early in the lives of disadvantaged children. It has a 7~10% rate of return per annum (See Heckman *et al*., 2010.). The Perry Preschool Program did not have a lasting improvement on cognitive ability, but it did improve important later-life outcomes through changes in personality (Heckman *et al*., 2011).

The Perry Preschool Program worked primarily through socioemotional channels. It raised scores on achievement tests but not IQ tests. As previously noted, socioemotional factors and cognitive factors both explain performance on achievement tests (Duckworth, 2007; Borghans *et al*., 2008; Borghans *et al*., 2009).

| Figure 22 | Perry Preschool Program: IQ, by Age and Treatment Group



Notes: IQ measured on the Stanford-Binet Intelligence Scale (Terman and Merrill, 1960). Test was administered at program entry and each of the ages indicated.
Source: Cunha *et al*. (2006) and Heckman and Masterov (2007) based on data provided by the High Scope Foundation.

13. Personality and Preference Parameters

Measures of personality predict a wide range of life outcomes that economists study. Personality psychologists define traits as relatively stable, person-specific determinants of behavior. Preferences are the natural counterpart of these traits in economics. However, the exact link between personality and preferences is unclear. Table 3 shows one possible correspondence between conventional economic preference parameters and personality measures.

An empirical Literature is emerging that attempts to make thes correspondence. See Table 4.

Table 3 Standard Preference Parameters and Conceptually Similar Measures in the Psychology Literature

Preference parameter	Personality measures
Time Preference	Conscientiousness Self-control Affective mindfulness Consideration of future consequences Elaboration of consequences Time preference
Risk Aversion	Impulsive sensation seeking Balloon Analogue Risk Task
Leisure Preference	Achievement Striving Endurance Industriousness
Social Preference	Warmth Gregariousness Trust Altruism Tender-mindedness Hostility

Table 4 Empirical Studies of the Links Between Preferences and Traits

Preferences	Personality measures	Empirical study
Time Preference	Conscientiousness, Self-contrel, Affective mindfulness, Elaboration of consequences, Consideration of future consequences.	Daly, Delaney and Harmen [2009]
	Extraversion Time preference	Dohmen, Falk, Huffman *et al.* [2009]
Risk Aversion	Sensation Seeking	Zuckerman [1994], Eckel and Grossman [2002]
	Openness	Dohmen, Falk, Huffman *et al.* [2010]
	Neuroticism, ambition, Agreeableness	Borghans, Golsteyn, Heckman *et al.* [2009]
	Balloon Analogue Risk Task	Lejuez, Aklin, Zvolensky *et al.* [2003]
Social Preferences		
Altruism	Neuroticism, Agreeableness	Ashton, Paunonen, Helmes *et al.* [1998],Osiński [2009] , Bekkers [2006]
Reciprocity	Neuroticism, Agreeableness, Conscientiousness	Dohmen, Falk, Huffman *et al.* [2008]
Trust	Neuroticism, Agreeableness, Openness, Conscientiousness	Dohmen, Falk, Huffman *et al.* [2008]

14. Summary and Conclusions

What can economists take from and contribute to personality psychology? What do we learn from personality psychology? Personality traits predict many behaviors sometimes with the same strength as conventional cognitive traits. Personality psychology considers a wider array of actions than are usually considered by economists. It enlarges the economist's way to describe and model the world. Cognition is one aspect of personality broadly defined.

Personality traits are not set in stone. They change over the life cycle. They are a possible avenue for intervention and policy.

Personality psychologists lack precise models. Economics provides a framework for recasting the field. More precise models reveal basic identification problems that plague measurement in psychology. Such analyses show that, at an empirical level, "cognitive" and "noncognitive" traits are not easily separated.

Personality psychologists typically present correlations—not causal relationships. Many contemporaneously measured relationships suffer from the problem of reverse causality. Econometric tools can be used to define and estimate causal mechanisms and to understand the causes of effects. Psychological measures have substantial measurement error. Econometric tools account for measurement error, and doing so makes a difference. Economists can formulate and estimate mechanisms of investment—how traits can be changed for the better.

There are major challenges in linking the traits of psychology with the preferences, constraints and expectation mechanisms of economics. Developing rigorous methods for analyzing causal relationships in both fields remains to be done. Developing a common language and framework to promote interdisciplinary exchange is required. There is a danger in assuming that basic questions of content and identification have been answered by psychologists at the level required for rigorous economic analysis.

References

Ackerman, P. L. and E. D. Heggestad, "Intelligence, Personality, and Interests: Evidence for Overlapping Traits," *Psychological Bulletin* 121, 1997, pp.219~245.

Almlund, M., A. Duckworth, J. J. Heckman, and T. Kautz, "Personality Psychology and Economics," In E. A. Hanushek, S. Machin, and L. Wößmann (eds.), *Handbook of the Economics of Education*, Volume 4. Amsterdam: Elsevier, Forthcoming, 2011.

Ashton, M. C., S. V. Paunonen, E. Helmes, and D. N. Jackson, Kin Altruism, Reciprocal Altruism, and the Big Five Personality Factors, *Evolution and Human Behavior* 19 (1), 1998, pp.243~255.

Barrick, M. R. and M. K. Mount, "The Big Five Personality Dimensions and Job Performance: A Meta-analysis," *Personnel Psychology* 44 (1), 1991, pp.1~26.

Bekkers, R., "Traditional and Health-related Philanthropy: The Role of Resources and Personality," *Social Psychology Quarterly* 69 (4), 2006, pp.349~366.

Binet, A. and T. Simon, *The Development of Intelligence in Children* (The Binet-Simon Scale), Psychological Science, Baltimore, MD: Williams & Wilkins Co, 1916.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. ter Weel, "The Economics and Psychology of Personality Traits," *Journal of Human Resources* 43 (4), Fall 2008, pp.972~1059.

Borghans, L., B. H. Golsteyn, J. J. Heckman, and H. Meijers, "Gender Differences in Risk Aversion and Ambiguity Aversion," *Journal of the European Economic Association* 7 (2-3), April 2009, pp.649~658.

Borghans, L., B. H. H. Golsteyn, J. J. Heckman, and J. E. Humphries, "IQ, Achievement, and Personality," Unpublished Manuscript, University of Maastricht and University of Chicago (revised from the 2009 version), 2011.

Borghans, L., H. Meijers, and B. ter Weel, "The role of Noncognitive Skills in Explaining Cognitive Test Scores," *Economic Inquiry* 46 (1), January 2008, pp.2~12.

Carroll, J. B., Human Cognitive Abilities: A Survey of Factor-Analytic Studies. New York: Cambridge University Press, 1993.

Cervone, D. and L. A. Pervin, *Personality: Theory and Research* (11th ed.), Hoboken:

John Wiley and Sons, Inc, 2009.

Costa, P. T. and R. R. McCrae, Revised NEO Personality Inventory (NEO PI-R) and the NEO Five-Factor Inventory (NEO-FFI) Professional Manual, Odessa, FL: Psychological Assessment Resources, 1992.

Cunha, F. and J. J. Heckman, “The Technology of Skill Formation,” *American Economic Review* 97 (2), May 2007, pp.31~47.

Cunha, F. and J. J. Heckman, “The Economics and Psychology of Inequality and Human Development,” *Journal of the European Economic Association* 7 (2-3), April 2009, pp.320~364. Presented as the Marshall Lecture, European Economics Association, Milan, Italy, August 29, 2008.

Cunha, F., J. J. Heckman, L. J. Lochner, and D. V. Masterov, “Interpreting the Evidence on Life Cycle Skill Formation,” In E. A. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, Chapter 12, Amsterdam: North-Holland, 2006, pp.697~812.

Cunha, F., J. J. Heckman, and S. M. Schennach, “Estimating the Technology of Cognitive and Noncognitive Skill Formation,” *Econometrica* 78 (3), May 2010, pp.883~931.

Daly, M., L. Delaney, and C. P. Harmon, “Psychological and Biological Foundations of Time Preferences,” *Journal of the European Economic Association* 7 (2-3), 2009, pp.659~669.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde, “Representative Trust and Reciprocity: Prevalence and Determinants,” *Economic Inquiry* 46 (1), 2008, pp.84~90.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde, “Are Risk Aversion and Impatience Related to Cognitive Ability?” *American Economic Review* 100 (3), 2010, pp.1238~1260.

Duckworth, A. L., Unpublished Dataset, University of Pennsylvania, Department of Psychology, 2007.

Eckel, C. C. and P. J. Grossman, “Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk,” *Evolution and Human Behavior* 23 (4), 2002, pp.281~295.

Epstein, S., “The Stability of Behavior: I. On Predicting most of the People much of the Time,” *Journal of Personality and Social Psychology* 37 (7), 1979 , pp.1097~1126.

Goldberg, L. R., “The Structure of Phenotypic Personality Traits,” *American Psychologist* 48 (1), 1993, pp.26~34.

Gough, H. G. and A. B. Heilbrun, “The Adjective Check List Manual,” Palo Alto, CA: Consulting Psychologists Press, 1983.

Hansen, K. T., J. J. Heckman, and K. J. Mullen, “The Effect of Schooling and Ability on

Achievement Test Scores," *Journal of Econometrics* 121 (1-2), July–August 2004, pp.39~98.

Heckman, J. J. and B. E. Honoré, "The Empirical Content of the Roy Model," *Econometrica* 58 (5), September 1990, pp.1121~1149.

Heckman, J. J., J. E. Humphries, S. Urzúa, and G. Veramendi, "The Effects of Educational Choices on Labor Market, Health, and Social Outcomes," Unpublished Manuscript, University of Chicago, *Department of Economics*, 2011.

Heckman, J. J., L. Malofeeva, R. Pinto, and P. A. Savelyev, "Understanding the Mechanisms Through which an Influential Early Childhood Program Boosted Adult Outcomes," Unpublished Manuscript, University of Chicago, Department of Economics (first draft, 2008). Under revision, *American Economic Review*, 2011.

Heckman, J. J. and D. V. Masterov, "The Productivity Argument for Investing in Young Children," *Review of Agricultural Economics* 29 (3), 2007, pp. 446~493.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Q. Yavitz, "The Rate of Return to the HighScope Perry Preschool Program," *Journal of Public Economics* 94 (1-2), February 2010, pp.114~128.

Heckman, J. J. and G. L. Sedlacek, "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-selection in the Labor Market," *Journal of Political Economy* 93 (6), December 1985, pp.1077~1125.

Heckman, J. J., J. Stixrud, and S. Urzúa, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics* 24 (3), July 2006, pp.411~482.

John, O. P. and S. Srivastava, "The Big Five Trait Taxonomy: History, Measurement and Theoretical Perspectives," In L. A. Pervin and O. P. John (eds.), *Handbook of Personality: Theory and Research*, Chapter 4, New York: The Guilford Press, 1999, pp.102~138.

Lejuez, C. W., W. M. Aklin, M. J. Zvolensky, and C. M. Pedulla, "Evaluation of the Balloon Analogue Risk Task (BART) as a Predictor of Adolescent Real-world Risk-taking Behaviours," *Journal of Adolescence* 26 (4), 2003, pp.475~479.

Mischel, W., *Personality and Assessment*, New York: Wiley, 1968.

Osinski, J., "Kin Altruism, Reciprocal Altruism and Social Discounting," *Personality and Individual Differences* 47 (4), 2009, pp.374~378.

Poropat, A. E., "A Meta-analysis of the Five-factor Model of Personality and Academic Performance," *Psychological Bulletin* 135 (2), 2009, pp.322~338.

Roberts, B. W., "Personality Development and Organizational Behavior," In B. Staw (ed.), *Research on Organizational Behavior*, Volume 27, Chapter 1, Oxford, UK: Elsevier Science/JAT Press, 2006, pp.1~41.

Roberts, B. W., "Back to the Future: Personality and Assessment and Personality Development," *Journal of Research in Personality* 43 (2), 2009 , pp.137~145.

Roberts, B. W., N. R. Kuncel, R. L. Shiner, A. Caspi, and L. R. Goldberg, "The Power of Personality: The Comparative Validity of Personality Traits," Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes, *Perspectives in Psychological Science* 2 (4), December 2007, pp.313~345.

Roberts, B. W. and D. Mroczek, "Personality Trait Change in Adulthood," *Current Directions in Psychological Science* 17 (1), 2008, pp.31~35.

Roberts, B. W., K. E. Walton, and W. Viechtbauer, "Patterns of Mean-level Change in Personality Traits Across the Life Course: A Meta-analysis of Longitudinal Studies," *Psychological Bulletin* 132 (1), 2006, pp.1~25.

Schmidt, F. L. and J. Hunter, "General Mental Ability in the World of Work: Occupational Attainment and Job Performance," *Journal of Personality and Social Psychology* 86 (1), 2004, pp.162~173.

Terman, L. M. and M. A. Merrill, Stanford-Binet Intelligence Scale: Manual for the Third Revision Form L-M. Boston: Houghton Mifflin, 1960.

Thaler, R. H., "A Short Course in Behavioral Economics," Edge Master Class, Sonoma, CA, July 25-27, July 2008.

Zuckerman, M., Behavioral Expressions and Biosocial Bases of Sensation Seeking, New York: Cam- Bridge University Press, 1994.

CHAPTER 2

'학교교육 수준 및 실태 분석 연구: 중학교' 자료를 이용한 사교육비 지출의 성적 향상 효과 분석*

강 창 희
(중앙대학교 경영경제대학 경제학부 부교수)

<초 록>

본 논문은 한국교육개발원의 '학교교육 수준 및 실태 분석 연구: 중학교' 자료를 이용하여 중학교 3학년 학생에 대한 사교육비 지출이 어느 정도의 성적 향상효과가 있는지를 추정한다. 사교육비의 내생성을 통제하기 위해 본 논문은 도구변수법과 비모수 구간추정법(nonparametric bounds analysis)을 사용한다. 분석 결과, 두 방법에서 공통적으로 사교육비 지출의 증가가 유의미한 정도의 성적 향상으로 연결된다는 확실한 증거가 발견되지 않는다. 도구변수법의 결과에 의하면, 10% 높은 사교육비 지출은 국어, 영어, 수학 성적을 각각 약 1.24%, 1.28%, 0.75% 향상시킨다. 구간추정법에서는 국어, 영어, 수학 과목에서 모두 10% 증가된 사교육비 지출의 효과가 0보다 크다는 증거를 찾기 어렵다. 본 논문의 실증 결과는 내생성을 통제한 여타의 연구들과 비슷한 결과를 보여준다.

* 강창희: (e-mail) ckang@cau.ac.kr, (address) Department of Economics, Chung-Ang University, 221 Heukseok-Dong Dongjak-gu, Seoul 156-756, Korea.

제1절 서 론

최근 국내외 교육경제학 연구에서 흔히 등장하는 하나의 주제는 학교 밖 사교육(private tutoring)이 학생의 학업성적에 미치는 인과효과의 크기를 추정하는 문제이다. 예를 들어 Dang(2007)과 Dang and Rogers(2008)는 베트남의 자료를 이용하여, 그리고 Ono(2007)는 일본의 자료를 이용하여 사교육의 성적 향상효과를 추정하였다. 이들의 연구 결과에 의하면, 사교육은 학생의 성적을 향상시키는 강력한 효과가 있다. 반면에, 미국 자료를 이용한 Briggs(2001), 터키의 자료를 이용한 Gurun and Millimet(2008) 및 우리나라의 '한국교육고용패널'을 이용한 Kang(2007)의 연구와 '한국교육종단연구' 자료를 이용한 강창희・이삼호(2010) 등에 의하면, 사교육은 학생의 교육성과에 그다지 큰 긍정적인 영향을 미치지 못한다.

이상의 연구들은 사교육의 효과 연구를 선도했다는 측면에서 기여한 바가 크지만, 일부 연구의 실증 결과들은 사교육의 인과효과라고 결론 내리기에는 많은 한계들이 존재한다. 예를 들어 Dang(2007)은 통계모형에서 사교육의 내생성을 통제하기 위하여 학생이 거주하는 지역(commune)의 학교들에서 부과하는 튜터링 가격(tutoring fees)을 도구변수로 사용한다. 그러나 저자가 스스로 인정하듯이, 이 도구변수는 학생 거주지역의 생활수준과 연관이 있을 수 있고, 이로 인하여 사용된 도구변수가 외생적이라고 단정하기 힘들다. Ono(2007)는 일본의 로닌(ronin, 일본에서 관찰되는 대학입시 재수현상)이 진학하는 학교의 수준에 미치는 영향을 분석하였다. 그는 로닌의 도구변수로서 학생의 출신지역 내에 있는 대학들의 평균적인 수준을 사용하였다. Dang(2007)에서와 같이 이 도구변수 또한 학생 본인 및 그 가족의 관측되지 않는 특성을 통하여 학생이 진학하는 대학의 수준과 연관될 수 있으므로 외생적이라고 인정하기 어렵다. Briggs(2001)는 미국의 자료를 이용하여 코칭(coaching)이 대입시험(SAT 또는 ACT)의 성적에 미치는 영향을 추정하였다. 그는 코칭의 내생성을 통제하기 위하여 Heckman의 선택편의 교정모형을 사용하고 있으나, 효과를 식별하기 위하여 단지 함수의 비선형성에 의존할 뿐 특정한 도구변수를 사용하고 있지는 않다.

우리나라에서 사교육에 대한 연구는 다른 나라들보다도 앞서 2000년대 초반 이후 본격화되기 시작하였다. 우리나라의 학부모들이 전통적으로 높은 사교육비를 지출하여 왔음을 생각할 때 이는 그리 놀라운 일이 아니다. 우리나라에서 사교육의 효과에 관한 연구는 다름 아닌 교육학, 교육심리학 및 교육사회학 분야의 연구자들에 의해 시작되었다(한대동 외[2001]; 오영수·윤정식[2003]). 사교육의 인과효과를 추정하는 문제가 정책효과나 처치효과 분석의 하나의 응용분야라는 인식하에서 최근 경제학자들이 사교육의 효과 추정 연구에 참여하고 있다(Kang[2007]; 강창희·이삼호[2010]; 김진영[2007]; 남기곤[2008]; 최형재[2008]).

지난 10여 년 동안 교육학 및 경제학 연구자들이 우리나라에서 사교육이 성적에 어떠한 효과를 미치는지를 추정하기 위해 많은 노력을 기울여 왔지만, 아직 그 효과의 크기에 대하여 합의된 견해가 형성되어 있지는 않다. 그 이유는 무엇보다도 대부분의 기존 실증연구들에서 사교육활동이나 사교육비가 가지는 내생성(endogeneity)이 적절하게 통제되지 못했기 때문이다. 예를 들면 조혜영·이경상(2005)과 김진영(2007)의 연구에서는 학생이 받은 사교육의 양을 표시하는 변수로서 사교육시간이, 그리고 이수정·임현정(2009)에서는 사교육비용이 사용되었다. 사교육이 교육성과에 미치는 영향을 추정하기 위하여 이들 연구에서는 사교육변수에 대하여 단순회귀분석방법(Ordinary Least Squares: OLS)을 적용하였다. 하지만 무작위 또는 자연실험을 통하지 않은 경우 단순회귀분석의 추정치는 인과관계(causation)가 아니라 상관관계(correlation)만을 표시한다는 점은 이미 잘 알려진 사실이다. 일부 교육학 연구자들은 단순회귀분석법 대신에 위계선형모형(Hierarchical Linear Model: HLM)을 통계 분석에 활용한다(임천순 외[2004]; 박현정 외[2008]). 하지만 HLM 또한 우도함수를 도출하는 과정에서 사교육비에 대한 외생성을 가정하기 때문에 OLS법과 동일한 한계에 직면한다.

사교육의 효과를 추정하는 대부분의 연구들이 사교육비의 내생성을 무시하거나 제한적으로만 고려한 데 비하여, 아래의 세 연구들은 내생성을 추정과정에서 명시적으로 다룬다. 먼저 Kang(2007)에서는 사교육비 지출액의 도구변수로서 학생의 출생순위가 첫째아(first-born)인지의 여부가 사용되

었다. 이는 Black *et al.*(2005)에서 이미 제안된 도구변수로서, 개인의 출생 순위는 출생과 동시에 외생적으로 결정되지만, 첫째아는 다른 형제들에 비하여 더 많은 교육투자를 받는다는 사실에 근거한다. 위의 도구변수 추정방법을 이용하여 한국직업능력개발원의 ‘교육고용패널’ 자료를 분석한 Kang(2007)의 결과에 따르면, 사교육은 학생의 성적에 양의 영향을 미치기는 하지만 그 크기는 그리 크지 않다. 사교육비 지출이 10% 증가할 때 학생의 성적은 평균 0.4% 정도 상승한다. 둘째, 최형재(2008)는 종속변수로서 진학한 대학의 질적 수준을 사용하였다. Kang(2007)에서와 마찬가지로, 이 연구에서 사용된 사교육비의 도구변수는 학생의 출생순위이다. 그의 결과에 의하면, 10% 높은 사교육비 지출은 학생이 상위 31개 대학이나 의학과에 입학할 확률을 약 0.6~0.7% 정도 상승시킨다. 마지막으로 김지하·김정은(2009)에서는 내생성 통제를 위하여 성향점수 매칭법이 사용되었다. 이 연구는 내생성에 대한 인식 측면에서 과거의 연구들보다 진일보하였지만 사용된 사교육변수가 사교육비(tutoring expenditures)가 아니라 사교육 참여 여부여서 사교육의 효과를 정밀하게 탐구하는 데에는 한계가 있다. 물론 성향점수 매칭법이 인과효과의 추정에 어느 정도 효과적인지에 대한 실증방법론상의 의문도 여전히 남아 있다.

본 논문에서 우리는 사교육비의 내생성을 명시적으로 고려함과 동시에 그동안 사교육비의 효과 연구에서 사용되지 않았던 새로운 자료를 사용한다. 먼저 통계방법의 측면에서, 사교육의 내생성을 통제하는 방법으로 우리는 첫째아 여부를 사교육비 지출액의 도구변수로 활용하는 2단계 추정법을 적용한다(Kang[2007]; 최형재[2008]). 첫째아 여부는 사교육비와는 양의 연관관계를 가지고 있기는 하지만, 이 변수가 종속변수인 학업성적과 영의 상관관계를 보이는 진정한 외생변수(exogenous variable)라고는 보기 어렵다. 우리는 아래에서 첫째아 여부 도구변수는 사교육비의 효과를 과대추정(overstate)할 가능성이 있음을 지적한다.

본 논문에서 사용되는 두 번째의 통계방법은 실증경제학에서 최근 소개되기 시작한 비모수 구간추정법(nonparametric bounds methods)이다. 비모수 구간추정법은 Manski(1990)에 의해 최초로 경제학에 소개되었고, 이후

Manski(1997)와 Manski and Pepper(2000)에 의하여 그 방법이 보다 구체적으로 개발되었다. 이 방법이 실제로 적용된 사례로는 Blundell *et al.*(2007), Gerfin and Schellhorn(2006), Gonzalez(2005), Kreider and Pepper(2007), Lechner(1999), Manski and Nagin(1998), Pepper(2000) 및 강창희・이삼호(2010) 등이 있다. 비모수 구간추정법의 기본 아이디어는 외생성보다는 완화된 가정을 적용하여 통상적인 점추정치(point estimates) 대신에 인과효과의 범위, 즉 그 하한(lower bound)과 상한(upper bound)을 구하는 방법이다. 이와 같이 추정된 인과효과의 범위가 충분히 작다면 우리는 인과효과의 실제 크기가 이 범위 내에 존재하는 것으로 해석할 수 있다.

우리는 도구변수법과 비모수 구간추정법을 한국교육개발원에서 2004년 중학교 3학년 학생 약 14,000명을 대상으로 구축한 '학교교육 수준 및 실태 분석 연구: 중학교' 자료(이하 '중학교 실태 연구 자료')에 적용한다. 본 자료를 이용한 실증분석 결과에 의하면, 사교육비 지출은 학생의 학업성적에 다소 긍정적인 영향을 미치기는 하지만 그 영향이 그리 크지는 않다. 도구변수법의 결과에 의하면, 10% 높은 사교육비 지출은 국어성적을 약 1.24%, 영어성적을 약 1.28%, 수학성적을 약 0.75%, 그리고 세 과목 전체의 평균성적을 약 0.74% 정도 향상시킨다. 구간추정법의 결과에 의하면, 10% 높은 사교육비 지출은 국어, 영어, 수학 성적을 각각 최대 0.77~0.84%, 2.16~2.64%, 2.04~2.84% 정도까지 향상시킨다. 그러나 모든 과목에서 공통적으로 효과의 최솟값이 0과 다르다는 실증적인 증거가 부족하다. 이를 보수적으로 해석하면, 국어, 영어, 수학 과목에서 모두 높은 사교육비 지출이 유의미한 정도로 성적을 향상시킨다는 증거를 발견하기 어렵다. 이상의 결과는 동일한 통계방법을 고등학교 3학년의 대입수능시험 성적자료에 적용한 Kang(2007)이나, 중학교 2~3학년 학생의 학업성취도에 적용한 강창희・이삼호(2010)의 연구 결과와 일맥상통한다. 이는 본 논문에서 추정한 사교육비 성적 향상효과의 크기가 신뢰할 만한 수치임을 암시한다.

본 논문은 다음과 같이 구성된다. 제Ⅱ장에서는 계량분석모형이, 제Ⅲ장에서는 사용된 자료에 대한 설명이 다루어진다. 제Ⅳ장에서는 분석 결과가 기술되고, 제Ⅴ장 결론으로 본 논문은 마무리된다.

제2절 계량분석모형

횡단면 자료를 사용하는 본 논문의 실증분석에 적용되는 통계모형은 다음과 같다.[1]

$$y_i = \beta_0 + s_i\beta_1 + X_i\beta_2 + \epsilon_i \tag{1}$$

이 식에서 y_i는 사교육비의 효과 측정에 사용되는 종속변수로서 학생 i의 시험성적(개별 과목 또는 복수 과목의 평균)을 표시한다. 각 과목마다 시험의 난이도나 시험에 참가하는 학생들의 구성이 조금씩 다를 수 있기 때문에 서로 다른 시험의 성적을 표준화할 필요가 있다. 이를 위하여 우리는 각 과목의 원점수를 그 과목 전체 샘플의 평균과 표준편차를 이용하여 표준화한(정규화한) Z-점수를 아래의 통계 분석에서 사용한다. s_i는 자료에서 측정된 월평균 사교육비 지출액(S_i)의 자연 로그값을 표시하고, X_i는 i의 개인 배경, 가정 배경 및 학교 배경을 표시하는 벡터이다. ϵ_i는 모형의 오차항이다.

사교육비의 효과에 관한 통상적인 연구들에서는 식 (1)을 단순회귀법(OLS) 또는 위계선형모형(HLM; Raudenbush and Bryk[2002]) 추정법을 이용하여 추정하였다. OLS와 HLM의 추정량이 β_1에 대한 일치추정량(consistent estimates)이 되기 위해 필요한 조건은 '$Cov(s_i, \epsilon_i) = 0$'이다. OLS나 HLM에서는 공통적으로 오차항에 대하여 '$\epsilon_i \sim d(0, \sigma^2)$'을 가정한다. 여기에서 d는 평균이 0이고 분산이 σ^2인 임의의 확률분포(예를 들어 정규분포)를 표시한다. 이때 ϵ_i의 평균 0은 식 (1) 우변의 설명변수들, 특히 s_i와 아무런 관련이 없는 임의의 상수이기 때문에 '$\epsilon_i \sim d(0, \sigma^2)$' 가정은 결국 '$Cov(s_i, \epsilon_i) = 0$'을 함축한다. 즉, 분석자료에서 '$Cov(s_i, \epsilon_i) = 0$'의 관계가 성립한다는 가정하에서 OLS와 HLM 추정법은 사교육비 지출의 진정한 인과효과(β_1)에 대한 일치추정량을 도출한다. 그러나 만약 분석자료에서 '$Cov(s_i, \epsilon_i) = 0$'의 가정이 성립되지 않는다면 추정치 $\widehat{\beta_1}$는 β_1에 대하여 부

1 본 장에서 소개되는 실증분석방법은 강창희・이삼호(2010)의 한국교육개발원 연구보고서 제Ⅲ장에 보다 자세히 설명된 사교육비의 인과효과 추정방법 중 일부를 차용하여 본 논문의 목적에 맞게 약간 수정하였다.

정확한 정보를 제공한다. '$Cov(s_i, \epsilon_i) > 0$'이 성립하는 경우 $\widehat{\beta_1}$는 β_1에 대한 과대추정치(overestimates)가 되고, '$Cov(s_i, \epsilon_i) < 0$'이 성립하는 경우 $\widehat{\beta_1}$는 β_1의 과소추정치(underestimates)가 된다.

예를 들어 학생의 부모가 교육열이나 소득이 높으면 그 학생에 대한 사교육비 지출이 높다고 예상할 수 있다. 이 상황에서 교육열이나 소득이 높은 부모들이 그것이 낮은 부모들에 비하여 지적인 능력 또한 평균적으로 우수하다면, 우리는 '$Cov(s_i, \epsilon_i) > 0$'의 관계가 성립한다고 추측할 수 있다. 만약 부모의 교육열 또는 부모의 지적인 능력 등과 같이 사교육비 지출과 학생의 학업성적에 동시에 영향을 미칠 가능성이 있는 변수들이 식 (1)의 우변에서 적절히 통제되지 않는 경우 OLS나 HLM을 통해 구한 $\widehat{\beta_1}$은 사교육비의 진정한 인과효과를 과장할 위험이 있다.

'$Cov(s_i, \epsilon_i) < 0$'가 성립하는 반대의 가능성도 상상해 볼 수 있다. 예를 들어 학부모의 교육열이나 소득이 동일하더라도 학생의 사교육 이전의 성적 또는 동기부여 수준(motivation)이 낮으면 부모는 자녀의 성적 향상을 위해 많은 사교육비를 지출할 가능성이 있다. 이 경우 '$Cov(s_i, \epsilon_i) < 0$'이 성립하고 이로 인하여 $\widehat{\beta_1}$는 사교육비의 진정한 효과를 과소평가할 위험이 있다. 결국 이론적인 다양한 가능성들은 '$Cov(s_i, \epsilon_i) = 0$'이 성립할 확률보다는 그것이 성립하지 않을 확률이 더 높음을 암시한다. 결국 사교육비와 성적 간의 인과관계에 관한 통계 분석의 핵심은 '$Cov(s_i, \epsilon_i) \neq 0$'으로부터 발생하는 s_i의 내생성을 어떻게 적절히 처리할 것인가 하는 문제로 귀결된다고 할 수 있다.

본 논문에서는 's_i의 내생성'을 통제하는 방법으로서 최근 실증경제학 연구에서 관측자료(observational data) 분석에 폭넓게 활용되는 도구변수법과 비모수 구간추정법을 사용한다.

1. 도수변수법

본 논문의 첫 번째 분석방법(방법 I)은 도구변수법(instrument variables methods)이다. 도구변수법에서는 식 (1)의 중요 내생변수(s_i)와는 관련성이

높고, 오차항(ϵ_i)과는 무관한 도구변수(instrumental variable: IV)를 사용하여 식 (1)을 2단계 추정법에 의하여 추정한다. 본 논문에서는 Black *et al.* (2005, p.695)의 제안에 따라, 학생의 출생순위(birth order)가 첫째인지의 여부(F_i)를 사교육비 지출(s_i)에 대한 도구변수로 활용한다. 한 개인의 출생순위는 의심할 여지없이 자연에 의하여 외생적으로 정해지기 때문에 '$Cov(F_i, \epsilon_i) = 0$'의 관계가 성립할 가능성이 있다. 그리고 한 시점에서 첫째 자녀(first-born child)인 학생에 대한 교육투자는 둘째 이상인 학생(later-born child)에 대한 교육투자보다 일반적으로 크게 나타나기 때문에, '$Cov(F_i, s_i) > 0$'가 성립할 가능성이 높다. 도구변수가 위의 두 가지 조건을 만족하는 경우 통계학의 2단계 최소자승 추정법(two-stage least squares methods: 2SLS)을 이용하면 β_2에 대한 일치추정량이 구해진다.

다음 장에서 확인되는 바와 같이, '$Cov(F_i, s_i) > 0$'의 조건은 대체로 성립한다. 그리고 가구 내에서 자녀 간 교육투자의 배분을 다룬 교육학, 사회학 및 경제학의 다양한 실증연구에서도 대체로 '$Cov(F_i, s_i) > 0$'의 관계가 관측된다. '$Cov(F_i, s_i) > 0$'이 조건이 성립하는 경우, β_1에 대한 바람직한 추정치를 얻기 위해 필요한 다른 조건은 '$Cov(F_i, \epsilon_i) = 0$'이다. 위의 조건은 관측되지 않는 오차항 ϵ_i을 동반하기 때문에 자료를 이용해서 직접적으로 이 조건을 검정하는 것은 불가능하다. 단지 이론적으로만 이 가정을 정당화할 수 있을 뿐이다.

출생순위와 지능 또는 교육성과를 다룬 교육학, 사회학 및 경제학의 많은 연구들이 양자 간에는 통계적으로 유의한 관계가 존재하지 않음을 보고하고 있다(Retherford and Sewell[1991]; Rodgers *et al.*[2000]). 이러한 연구 결과를 토대로 우리는 '$Cov(F_i, \epsilon_i) = 0$'이 성립한다고 가정할 수 있다. 그러나 본 논문에서는 '$Cov(F_i, \epsilon_i) = 0$'라는 강한 가정 대신에 '$Cov(F_i, \epsilon_i) \geq 0$'라는 다소 약화된 가정을 채택한다. 그 이유는 다음과 같다.

첫째, 출생순위와 지능 또는 교육성과 간에 유의한 상관관계가 존재하지 않는다는 연구 결과들과는 대조적으로 양자 간에 유의한 관계가 존재한다는 연구 결과들이 존재한다. 후자의 연구들은 대체로 첫째 자녀의 지능이나 교육성과가 둘째 이상의 자녀보다 평균적으로 높다고 보고하고 있다(Bjerkedal *et al.*[2007]; Black *et al.*[2007]; Zajonc[1976]; Zajonc and Mullahy[1997]). 결국 출생순위와 지능에 관한 연구성과들은 대체로 '$Cov(F_i, \epsilon_i) \geq 0$'의 관계가 성립하는 것으로 요약된다. 즉, 첫째 자녀의 지능이나 교육성과는 둘째 이상의 자녀보다 평균적으로 높거나 혹은 유의하게 다르지 않다는 것이 일반적인 견해이다. 반면에, 첫째 자녀의 지능이나 교육성과가 둘째 이상의 자녀보다 평균적으로 낮다고 보고하는, 즉 '$Cov(F_i, \epsilon_i) < 0$'를 지지하는 실증연구는 찾아보기가 대단히 어렵다.

둘째, 가족 내 교육자원의 배분을 다룬 실증연구들은 대체로 부모들이 둘째 이상의 자녀보다는 첫째 자녀에게 교육자원을 더 많이 투자한다고 보고한다(Behrman and Taubman[1986]; Black *et al.*[2005]). 본 논문의 분석에서도 첫째 자녀인 학생에 대한 평균 사교육비 지출액은 둘째 이상의 자녀인 학생에 대한 사교육비 지출액보다 크게 나타난다. 이와 같이 금전적인 측면에서 부모들이 둘째 이상의 자녀에 비하여 첫째 자녀에게 보다 많은 투자를 한다면, 교육적인 관심이나 정서적 측면과 같은 비금전적인 측면에서도 부모들은 첫째 자녀에게 상대적으로 더 많은 투자를 한다고 예상할 수 있다. 그리고 부모들의 이러한 선호는 결과적으로 '$Cov(F_i, \epsilon_i) > 0$'의 관계가 성립할 가능성을 보여준다.

이상에서 설명한 바와 같이, 도구변수 F_i에 대하여 '$Cov(F_i, s_i) > 0$'과 '$Cov(F_i, \epsilon_i) \geq 0$'의 조건이 성립한다면 2SLS 추정치 $\widehat{\beta_1}$은 β_1에 대한 일치추정량이거나 또는 과대추정치로 해석된다. 하지만 $\widehat{\beta_1}$이 β_1의 과소추정량이 될 가능성은 희박하다. 그리하여 만약 $\widehat{\beta_1}$가 0과 유의하게 다르지 않다는 검정 결과가 나타날 때, 우리는 사교육비 지출의 성적 향상효과가 그리 크지 않다고 결론 내릴 수 있다. 그리고 만약 $\widehat{\beta_1}$이 0과 유의하게 다른 양수라는 검정 결과가 나오면 우리는 사교육비의 진정한 효과는 2SLS의 추정치보다도 작을 것이라고 추측한다.

2. 비모수 구간추정법

본 논문의 두 번째 통계방법(방법 Ⅱ)은 비모수 구간추정법(nonparametric bounding method)이다. OLS나 도구변수법에서는 평균 처치효과(Average Treatment Effect: ATE)의 점추정치(point estimates)를 계산한 후 그것의 표준오차를 구하여 신뢰구간을 설정하고 특정 가설에 대하여 검정한다. 비모수 구간추정법에서는 평균 처치효과의 점추정치가 아니라 그것이 위치할 가능성이 있는 구간 범위(bounds)의 최솟값과 최댓값을 몇 가지 가정을 이용하여 추정한다. 그리고 이 구간에 대하여 신뢰구간을 설정하고 이로부터 특정 가설을 검정한다. 비모수적 구간추정법을 실제 자료에 대한 분석에 활용한 사례로는 Gonzalez(2005), Manski and Pepper(2000), 강창희・이삼호(2010) 등이 있다. 아래에서 기술되는 비모수 구간추정법은 위 논문들에 설명된 것을 참고하여 본 논문의 목적에 맞게 수정하였다.

먼저 처치수준과 성과수준을 연결시키는 반응함수를 다음과 같이 정의하자. $y_i(\cdot): T \rightarrow Y$. 실현된 성과(realized outcome) $y = y(z)$는 z라는 처치수준을 실제로 받은 학생의 성과수준을 나타낸다. 그리고 잠재성과(potential outcome) $y(t)\ (t \neq z)$는 동일한 학생이 t라는 가상적인 처치수준을 받았더라면 나타날 잠재적인 성과수준을 표시한다.

본 구간추정법에서는 이산적인 처치수준(discrete treatment levels)에 대한 인과효과를 추정하기 때문에 연속변수인 사교육비 지출(S_i)을 다음과 같이 세 가지 수준의 이산변수 T_i로 변형한다.

$$
\begin{aligned}
T_i &= 0 \qquad \text{if} \quad S_i = 0 \\
T_i &= 1 \qquad \text{if} \quad 0 < S_i \leq H_1 \\
T_i &= 2 \qquad \text{if} \quad H_1 < S_i
\end{aligned}
$$

아래의 실증분석에서 우리는 H_1을 국어 사교육비를 다루는 분석에서는 3만원, 수학과 영어의 사교육비를 다루는 분석에서는 9만원, 그리고 세 과목 전체의 사교육비를 다루는 분석에서는 20만원으로 설정한다.[2] 이 경우 개별

2 H_1에 대하여 다른 금액들을 사용할 수도 있다. 본 논문에서 명시적으로 보고하지 않았지만, H_1에 대해 위에서 제시된 금액 이외의 금액들(예를 들어 국어 사교육비 4만원, 수학 및 영어 사교육비 10만원, 그리고 세 과목 전체 사교육비 25만원)을 설정하여 동일한 분석을 실시하였다. 그러나 그 실증 결과들은 본 논문에 보고된 내용과 큰 차이를 보이지 않는다. H_1에 대해 다른 금액을 설정한 경우의 분석 결과는 독자의 요청이 있을 경우 제

학생에게는 0, 1, 2 중 하나의 처치수준(treatment level)이 적용된다.

평균 처치효과의 구간을 설정하기 위하여 먼저 $E[y(t)]$ $(t=0,1,2)$를 다음과 같이 분해한다.

$$E[y(t)] = E(y|z=t)\Pr(z=t) + E[y(t)|z \neq t]\Pr(z \neq t) \tag{3}$$

식 (3)의 $E(y|z=t)$, $\Pr(z=t)$ 및 $\Pr(z \neq t)$는 자료로부터 직접 계산이 가능하지만, 가상적 대응치(counterfactual)인 $E[y(t)|z \neq t]$는 자료로부터 계산이 불가능하다. 여기에서 y가 최소 K_0, 최대 K_1의 값을 취한다고 가정하자(즉. '$y \in [K_0, K_1]$'). 이 가정하에서 '$E[y(t)|z \neq t] \in [K_0, K_1]$'이 성립하기 때문에 우리는 $E[y(t)]$의 구간을 다음과 같이 구할 수 있다.

$$\begin{aligned}
&E(y|z=t)\Pr(z=t) + K_0 \cdot \Pr(z \neq t) \\
&\leq E[y(t)] \leq \\
&E(y|z=t)\Pr(z=t) + K_1 \cdot \Pr(z \neq t)
\end{aligned}$$

이 구간을 우리는 $E[y(t)]$의 최소가정 범위(Worst Case bounds: WC)라고 부른다.

$E[y(t)]$의 범위를 보다 좁히기 위하여 우리는 몇 가지의 가정들을 도입하고 이를 개별적으로 혹은 결합하여 활용한다. 첫 번째 가정은 '단조적 반응(monotone treatment response: MTR)의 가정'으로서, 수식으로는 다음과 같이 표현된다.

$$t_l < t_m \;\; \rightarrow \;\; y(t_l) \leq y(t_m)$$

즉, 처치수준이 높으면 성과는 불변이거나 또는 증가한다는 가정이다. 이 가정은 학생에 대하여 교육자원의 지출을 늘리면 그 학생의 학업성취도가 변하지 않거나 또는 향상되고, 최소한 줄어들지는 않을 것이라는 이론적 예측으로부터 도출된다. 교육투자의 성과를 다룬 대다수의 실증연구들로부터 이 가정의 유효성이 확인된다. 교육비 지출이 학생의 학업성과에 미치는 긍

공할 수 있다.

정적인 영향의 정확한 크기에 대해서 많은 논쟁이 있지만(Hanushek[1997, 2003]), 그럼에도 불구하고 교육비 지출이 학생의 학업성취도에 강한 부정적인 영향을 미친다는 실증연구는 대단히 드물기 때문이다.

구간추정법을 적용하기 위해 도입되는 두 번째 가정은 '단조적 선택(monotone treatment selection: MTS)의 가정'으로서 수식으로는 다음과 같이 표현된다.

$$t_l < t_m \rightarrow E[y(t)|z=t_l] \leq E[y(t)|z=t_m]$$

즉, 임의의 처치수준 t에 대한 평균 잠재성과 $E[y(t)]$는 사교육비를 적게 지출하는 부모를 둔 학생($z=t_l$)에서보다도 사교육비를 많이 지출하는 부모를 둔 학생($z=t_m$)에서 높게 나타난다는 가정이다. 예를 들어 소득이 높은 부모들은 소득이 낮은 부모들에 비하여 자녀에게 평균적으로 높은 사교육비를 지출할 것으로 예상할 수 있다. 이때 Haveman and Wolf(1995)의 연구결과에서와 같이, 유전적인 요인 또는 가정환경의 영향으로 소득이 높은 부모를 둔 학생들이 소득이 낮은 부모를 둔 학생들보다 평균적으로 지적인 능력이나 학업성취도가 높다면 위의 MTS 가정이 타당성을 가진다고 할 수 있다. 이 가정은 식 (1)을 OLS를 통해 추정할 때 문제가 되는 '$Cov(s_i, \epsilon_i)=0$'의 가정(즉, 외생성의 가정)을 '$Cov(s_i, \epsilon_i) \geq 0$'로 완화시키는 것으로 해석된다.

위에서 제시된 MTR과 MTS 가정을 결합하면 사교육비 평균효과의 범위가 상당히 축소된다. 두 가정을 결합함으로써 설정되는 $E[y(t)]$의 범위(MTR+MTS 범위)는 다음과 같이 주어진다.

$$\begin{aligned}&\sum_{h<t}\{E(y|z=h)\Pr(z=h)\}+E(y|z=t)\Pr(z\geq t)\\&\leq E[y(t)]\leq\\&\sum_{h>t}\{E(y|z=h)\Pr(z=h)\}+E(y|z=t)\Pr(z\leq t)\end{aligned}$$

아래 실증분석의 결과를 설명하는 과정에서 확인되는 바와 같이, MTR+MTS 범위는 개별적인 MTR 범위 및 MTS 범위에 비하여 상당히 좁은

범위로 줄어든다. 그런데 여기에서 '$E[y(t)|v=u_1]=E[y(t)|v=u_2]$ $(u_1 \neq u_2)$'로 표현되는 '평균 독립성(mean independence) 가정'을 만족하는 도구변수 v가 주어진다면 MTR+MTS 범위를 추가적으로 좁히는 것이 가능하다.

'평균 독립성 가정'하에서는 도구변수의 값이 '$v=u_1$'인 학생의 기대 학업성취도가 '$v=u_2$'인 학생의 것과 동일하다. 그러나 현실적으로 이러한 특성을 만족하는 외생적인 도구변수를 찾아내는 것은 상당히 어렵다고 알려져 있다. 이 어려움을 극복하는 방법으로 Manski and Pepper(2000)는 $E[y(t)|v=u_1] \leq \ E[y(t)|v=u_2]\,(u_1 < u_2)$'로 표현되는 '평균 단조성(mean monotonicity) 가정'을 만족하는 단조적인 도구변수(monotone IV: MIV) v를 활용하는 방법을 제안하였다. '평균 단조성 가정'하에서는 도구변수의 값이 '$v=u_1$'인 학생의 기대 학업성취도가 '$v=u_2$'인 학생의 것과 같거나 또는 작은 특성을 만족하는 것으로 충분하다. 본 논문에서는 '평균 단조성 가정'을 만족하는 단조적 도구변수로서 방법 I에서 활용된 바 있는 첫째 자녀 여부의 더미변수 F_i를 사용한다. '$Cov(F_i, \epsilon_{it}) \geq 0$'와 밀접하게 연관되는 가정으로서, '평균 단조성 가정'은 임의의 사교육비 수준(t)이 주어질 때 첫째 자녀인 학생($F_i=1$)의 평균성적은 둘째 자녀 이상인 학생($F_i=0$)의 평균성적과 동일하거나 높음을 의미한다.

MIV와 MTR+MTS 가정을 결합함으로써 우리는 $E[y(t)]$의 MIV+MTR+MTS 범위를 다음과 같이 구할 수 있다.

$$
\begin{aligned}
&\sum_{u \in F} \Pr(F=u) \times \\
&SUP_{u_1 \leq u}\left[\sum_{h<t} E(y|F=u_1, z=h)\Pr(z=h|F=u_1) + E(y|F=u_1, z=t)\Pr(z \geq t|F=u_1)\right] \\
&\leq E[y(t)] \leq \\
&\sum_{u \in F} \Pr(F=u) \times \\
&INF_{u_2 \geq u}\left[\sum_{h>t} E(y|F=u_2, z=h)\Pr(z=h|F=u_2) + E(y|F=u_2, z=t)\Pr(z \leq t|F=u_2)\right]
\end{aligned}
$$

여기에서 기댓값 $E(y| \cdot)$는 자료로부터 계산된 표본 평균값을 사용한다. 비모수적 구간추정법에서는 개별 가정들과 모든 조합의 결합 가정들에 대하여 $E[y(t)]$의 범위를 구할 수 있다. 그러나 공간 제약상 본 논문의 분석 결과는 MTR+MTS 결합 가정과 MIV+MTR+MTS 결합 가정을 사용한 경우에 대해서만 제시한다.

위에서 도입된 가정들을 이용하여 $E[y(t)]$의 범위를 구한 후, $E[y(t_m)-y(t_l)] \quad (t_m > t_l)$로 정의되는 ATE 범위는 $E[y(t_m)]$의 최솟값과 최댓값 및 $E[y(t_l)]$의 최솟값과 최댓값을 이용하여 계산한다. 즉, $E[y(t_m)-y(t_l)]$ 의 최솟값은 $E[y(t_m)]$의 최솟값에서 $E[y(t_l)]$의 최댓값을 차감함으로써, 그리고 $E[y(t_m)-y(t_l)]$ 의 최댓값은 $E[y(t_m)]$의 최댓값에서 $E[y(t_l)]$의 최솟값을 차감함으로써 구해진다. 그리고 50개의 bootstrap 샘플을 형성하여 $E[y(t_m)-y(t_l)]$ 의 최댓값에 대한 상위 5분위값과 $E[y(t_m)-y(t_l)]$ 의 최솟값에 대한 하위 5분위값을 구하여 $E[y(t_m)-y(t_l)]$의 범위에 대한 90% 신뢰구간을 설정하고 이를 가설검정에 이용한다.

제3절 분석자료: 학교교육 수준 및 실태 분석 연구

사교육비 지출의 효과를 분석하기 위하여 본고에서는 한국교육개발원이 2004년 중학교 3학년 재학생들을 대상으로 구축한 ‘학교교육 수준 및 실태 분석 연구: 중학교’의 원자료(이하 중학교 실태 연구자료)를 사용한다. ‘중학교 실태 연구자료’는 2004년 현재 우리나라 중학교 3학년에 재학 중인 14,372명의 학생들에 대하여 개인, 가정 및 학교의 배경을 조사한 횡단면 자료이다.[3] 본 표본자료는 전국 2,938개 중학교에 재학하는 총 1,933,543명의 중학생들을 대표할 수 있도록 지역별, 지역규모별, 학교 설립 유형별로 모집단에 비례하도록 무선 표집되었다. 좀 더 구체적으로는, 1단계로 학교 및 학생의 모집단 분포를 고려하여 전국적으로 200개의 표본 중학교가 선정된다.

---

3 ‘학교교육 수준 및 실태 분석 연구: 중학교’ 자료의 구성과 설문지의 구조는 김양분 외(2004)에 보다 자세히 기술되어 있다.

각 학교에서 무작위로 2개의 학급이 선정되고 그 학급에 재학하는 학생들 모두가 표본으로 추출된다. 이와 같이 선정된 학생들을 대상으로 개인 배경, 가족 배경 및 학교 배경을 조사한다. 학생의 학교생활 및 가정생활에 대한 보다 자세한 사항을 알기 위하여 학생이 재학하는 학교의 학교장, 교무부장, 교사 전체 그리고 부모 설문지가 따로 만들어져 조사된다.

학생에 대한 배경 질문과 더불어 '중학교 실태 분석 연구'에서는 서울시 교육청의 중학생 학력평가 문제지를 활용하여 개별 학생의 국어, 영어, 수학 성적을 0~100점 척도로 측정한다(학업성취도 평가일은 2004년 9월 14일). 이하의 분석에서는 결과 해석의 보편성을 위하여 각 과목의 원점수를 평균과 분산이 각각 0과 1이 되도록 정규화하여 사용한다. 그리고 각 과목의 점수와 더불어 세 과목의 평균점수(또는 두 과목의 점수만 있는 경우에는 두 과목의 평균점수)를 계산하고, 이 평균점수 또한 평균과 분산이 각각 0과 1이 되도록 정규화한다.

아래의 분석에 사용되는 중요 변수인 학생의 사교육 경험 및 사교육비 지출액과 학생의 출생순위 정보는 '중학교 실태 분석 연구'의 학부모 설문지로부터 구성된다. 이 설문지에서는 2004년 7월 현재 국어, 영어, 수학 각 개별 과목에 대하여 지출한 월평균 사교육비 액수가 사교육 유형별(예를 들어 학원, 개인과외, 학습지 등)로 조사된다. 이하의 분석에서 우리는 사교육비 지출액을 사교육 유형별로 구분하지 않고 모두 합하여 각 과목에 대한 월평균 사교육비 총지출액을 계산하여 사용한다. 그리고 세 과목의 평균점수를 종속변수로 이용하는 분석에서는 세 과목 전체에 대한 월평균 총사교육비를 계산하여 이용한다.

위에서 제시된 통계방법을 적용하기 위하여, 우리는 14,372명의 학생들에 대한 원자료를 다음의 과정을 통하여 축약하였다. 첫째, 학부모 설문에 대한 답변이 없는 2,564명의 학생들을 제외하였다. 둘째, 학부모 설문지(질문 3번)를 이용하여, 학생 가정에 아버지(또는 남성보호자)와 어머니가 동시에 부재한 총 456명의 학생을 분석에서 추가적으로 제외하였다. 왜냐하면 이러한 환경에 처한 학생의 가정환경은 정상적이라고 생각할 수 없기 때문이다. 그러나 아버지나 어머니 중 적어도 한 분과 함께 살고 있는 학생은 분석에

포함되었다. 셋째, 본고의 통계 분석에서는 외생적인 출생순위를 나타내는 변수로서 첫째아의 여부가 사교육비 지출액의 도구변수로 활용된다. 그런데 외자녀인 학생의 경우 정의상 첫째아 여부가 가정의 자녀 수가 1인 경우와 명확히 구분되지 않는다. 즉, 이들의 경우 첫째아 여부가 외생적인 출생순위와 동시에 자녀 수가 1인 경우를 동시에 표현한다. 일반적으로 자녀 수는 내생성이 있기 때문에, 첫째아 여부의 외생성을 명확하게 활용하기 위해서는 외자녀인 학생들을 분석에서 제외시킬 필요가 있다. 그리하여 총 812명의 외자녀 학생들이 분석에서 추가적으로 제외되었다. 넷째, 우리나라에서 국어, 영어 및 수학 등의 과목에 지출된 사교육비의 효과를 분석할 때 예체능 사교육을 받는 학생들을 분석에 포함시킬지의 여부가 논쟁거리이다. 예체능의 사교육을 받는 학생들은 일반적으로 고소득 가정 배경을 가진 학생들로서 다른 과목에 대한 사교육비 지출 또한 높다. 그러나 학업의 최종 목표가 예체능 계열로의 대학 진학일 가능성이 높아서 국어, 영어, 수학 등의 주요 과목 성적은 낮을 수 있다. 이러한 경향을 통계 분석에서 명시적으로 고려하지 않으면, 사교육비의 효과는 실제보다 과소추정될(understated) 가능성이 높다. 본고에서는 이러한 문제를 고려하여 예체능의 사교육을 받은 학생 모두를 분석에서 제외하였다. 총 1,354명의 학생들이 추가적으로 분석에서 제외되었다. 위의 네 가지 기준을 통하여 분석표본을 축소하고, 식 (1)의 변수에 대한 결측치를 제외시키면, 최종적으로 총 5,122명의 학생들에 대한 관측치가 남고 이 샘플이 아래의 통계 분석에서 사용된다. 사교육비의 정확한 액수에 대한 응답률이 낮기 때문에 마지막 단계에서 결측치가 많이 발생하였다.

〈표 1〉에는 이 최종 분석표본에 대한 각 변수들의 기술통계량이 제시되어 있다. (1)열에는 기술통계량 계산에 사용된 관측치의 수가, 그리고 (2)열에는 최종 분석표본에 나타난 각 변수들의 평균 및 표준편차가 보고되어 있다. 전체 분석표본에 포함된 학생들을 출생순위에 따라 나누어, (3)열에는 첫째아인 학생들에 대한 기술통계량, 그리고 (4)열에는 둘째아 이상인 학생들에 대한 기술통계량을 계산하였다. 마지막으로 (5)열에는 위의 두 하위표본 간 개별 변수의 평균 차이를 계산하고 그 차이가 통계적으로 유의한지를 검정

하는 T-검정통계량과 표준오차를 제시하였다. 최종 분석표본에 나타난 국어, 영어, 수학 과목 원점수의 평균(표준편차)은 각각 63.2(17.6), 65.1(24.0), 62.5 (24.0)이다. 세 과목의 평균 원점수에 대한 평균과 표준편차는 각각 63.6과 19.7이다.

먼저 학생의 출생순위에 따라 과목의 성적을 비교하면, 국어, 영어, 수학 과목 모두에서 첫째아의 평균성적이 둘째아 이상의 성적보다 통계적으로 유의하게 높다. 첫째아의 국어, 영어 및 수학 성적의 평균값은 각각 65.2, 68.6, 65.1인 반면에, 둘째아 이상의 각 과목에 대한 평균값은 각각 61.4, 62.1, 60.2이다. 세 과목의 평균성적 또한 첫째아의 평균성적(66.3)이 둘째아 이상의 평균성적(61.2)보다 유의하게 높다.

학생에 대한 사교육비 지출액을 살펴보면, 첫째아인 학생에 대한 월평균 사교육비 지출액이 둘째아 이상인 학생에 대한 사교육비 지출액보다 크다. 세 과목 전체의 월평균 사교육비는 평균 154,100원 정도이다. 그런데 첫째아에 대한 사교육비 평균은 178,300원으로서 둘째아 이상에 대한 사교육비 평균 133,000원을 약 34% 정도 상회한다. 각 과목별로 사교육비 지출액을 살펴보면, 국어보다는 영어와 수학에 대한 사교육비 지출이 높음을 알 수 있다. 국어에 대한 월평균 사교육비가 약 31,970원인 데 비하여, 영어 사교육비는 약 69,820원, 수학 사교육비는 약 65,430원으로서 국어 사교육비에 비하여 2배 이상 크다. 각 과목에 대해서든 세 과목 전체에 대해서든, 부모의 사교육비 지출은 둘째아 이상인 학생보다는 첫째아인 학생에게서 높게 나타난다.[4]

학업성적과 사교육비 이외의 변수들을 살펴보면, 첫째아인 학생들의 직전 학기 성적이 둘째아 이상의 학생들에 비하여 통계적으로 유의하게 높다. 1점(최하위)에서 9점(최상위) 사이에서 본인이 제시한 직전 학기 성적에 의하면, 첫째아 학생의 평균은 5.634이고 둘째아 이상 학생의 평균은 5.279이다.

4 사교육비 지출액의 평균값 계산에 사용된 관측치의 숫자가 각 변수별로 다른 것은 각 과목별로 사교육비 정보의 관측치 수가 다르기 때문이다. 그리하여 개별 과목의 평균 사교육비 계산에 사용된 관측치 수는 작고, 세 과목의 평균 사교육비 계산에 사용된 관측치의 수는 크다.

〈표 1〉 분석자료에 대한 기술통계량

변수명	관측치 (1)	분석표본 전체 평균(표준편차) (2)	첫째아 표본 평균(표준편차) (3)	둘째아 이상 표본 평균(표준편차) (4)	T-검정 차이(표준오차) (5)
세 과목 평균성적	5122	63.600(19.676)	66.299(19.302)	61.243(19.700)	5.056(0.547)**
국어성적	5122	63.176(17.645)	65.213(17.616)	61.397(17.480)	3.816(0.491)**
영어성적	5122	65.128(24.032)	68.551(23.264)	62.138(24.297)	6.414(0.667)**
수학성적	5122	62.497(24.041)	65.132(23.630)	60.195(24.164)	4.938(0.670)**
세 과목 사교육비 (W1,000)	5122	154.1(241.3)	178.3(269.3)	133.0(211.6)	45.29(6.729)**
국어 사교육비 (W1,000)	4662	31.97(76.88)	36.5(86.8)	28.1(66.9)	8.489(2.255)**
영어 사교육비 (W1,000)	4752	69.82(116.7)	82.1(128.9)	59.1(103.7)	22.966(3.378)**
수학 사교육비 (W1,000)	4716	65.43(105.5)	75.9(115.4)	56.5(95.3)	19.389(3.069)**
주평균 자기학습 시간	5122	5.441(7.030)	6.036(7.578)	4.921(6.471)	1.115(0.196)**
국어 자기학습시간	5122	1.411(2.259)	1.519(2.455)	1.316(2.067)	0.203(0.063)**
영어 자기학습시간	5121	2.002(2.959)	2.266(3.188)	1.771(2.724)	0.495(0.083)**
수학 자기학습시간	5121	2.029(2.844)	2.251(3.043)	1.836(2.645)	0.416(0.079)**
전기 성적	5122	5.445(2.002)	5.634(2.015)	5.279(1.976)	0.355(0.056)**
남성 여부	5122	0.473(0.499)	0.431(0.495)	0.510(0.500)	−0.078(0.014)**
만 나이	5122	14.845(0.327)	14.84(0.329)	14.85(0.325)	−0.014(0.009)
가구 내 자녀 수	5122	2.500(0.926)	2.284(0.548)	2.690(1.125)	−0.406(0.025)**
첫째아 여부	5122	0.466(0.499)	1.000(0.000)	0.000(0.000)	
부모의 결혼 여부	5122	0.955(0.206)	0.958(0.200)	0.953(0.211)	0.005(0.006)
부모 평균연령	5122	43.42(3.205)	42.04(2.589)	44.62(3.209)	−2.577(0.082)**
부친 교육연수	5122	12.85(2.752)	13.25(2.661)	12.50(2.782)	0.747(0.076)**
모친 교육연수	5122	12.06(2.336)	12.45(2.211)	11.71(2.387)	0.746(0.065)**
가구소득 결측	5122	0.144(0.351)	0.136(0.343)	0.151(0.359)	−0.015(0.010)
가구소득(W1,000)	4496	3300.4(4478.2)	3324.3(4284.9)	3279.2(4643.3)	45.05(133.84)

주: * $p < 0.05$; ** $p < 0.01$.

아울러 첫째아 학생들은 둘째아 학생들에 비하여 사교육시간을 제외한 주간 평균 자기학습시간도 국어, 영어, 수학 과목 모두에서 길게 나타난다.

위에서 설명된 변수들 이외의 다른 변수들을 살펴보더라도, 첫째아 학생들에게서는 둘째아 이상의 학생들에 비하여 높은 학업성적으로 연결되는 좋은 특성들이 발견된다. 예를 들어 첫째아 학생은 형제 수가 적어서 교육자원에 대한 경쟁이 덜하고, 첫째아 학생 부모들의 평균 교육수준도 둘째아 이상의 부모들에 비하여 높다. 첫째아 학생들과 둘째아 이상 학생들 간의 평균적인 특성의 차이에서 확인되는 바와 같이, '$Cov(F_i, \epsilon_{it}) = 0$'은 적절한 가정이라고 생각하기 어렵다. 그러나 〈표 1〉의 결과는 '$Cov(F_i, \epsilon_{it}) \geq 0$'의 가정과 크게 모순되지는 않는 것으로 생각된다. '$Cov(F_i, \epsilon_{it}) \geq 0$'이 내포하는 '평균 단조성(mean monotonicity) 가정'이 성립한다면, 우리는 도구변수법과 더불어 구간추정법을 통하여 사교육비의 효과에 관한 유용한 정보를 도출할 수 있을 것이다.

제4절 실증분석 결과

본 장에서는 앞에서 설명된 통계 분석방법들을 '중학교 실태 분석 연구' 자료에 적용하여 추정한 사교육비의 성적 향상효과가 제시된다. 먼저 사교육비 지출의 내생성이 통제되지 않는 OLS의 추정 결과를 살펴보고, 내생성을 명시적으로 통제하는 도구변수법과 구간추정법의 추정 결과를 설명한다.

1. OLS 및 도구변수법의 추정 결과

〈표 2〉에는 세 과목 평균의 Z-점수를 종속변수로 이용하는 식 (1)에 대한 OLS 및 도구변수법의 추정 결과가 보고되어 있다. 〈표 3〉의 패널 A, B 및 C에는 각각 국어, 영어, 수학 개별 과목의 Z-점수를 종속변수로 사용하는 경우의 추정 결과가 제시되어 있다. 〈표 2〉와 〈표 3〉에 제시된 사교육비 로그값의 추정계수는 $\widehat{\beta_1} = \dfrac{\Delta E(y_i)}{\Delta s_i}$로서 Δs_i가 1단위 증가할 때, 즉 사교육비가

100% 증가할 때 종속변수 Z-점수의 기댓값이 변화하는 양을 표현한다. 해석의 편의를 위하여 우리는 이 값을 사교육비가 10% 증가할 때 변화하는 원점수 기댓값의 퍼센트 변화량인 '탄력성 계수'로 변환하여 사교육비 로그값 추정계수의 아래에 보고하였다. 탄력성 계수는 분석자료에서 주어진 해당 원점수의 평균값에서 계산되었다.

〈표 2〉 OLS 및 도구변수법 추정 결과: 종속변수는 세 과목 평균 점수

통계방법:	OLS		도구변수법			
변수명	(1)		y_i 추정식 (2)		s_i 추정식 (3)	
사교육비 로그값	0.084	(0.006)**	0.241	(0.061)**		
[탄력성계수]	[0.259]		[0.744]			
첫째아 여부					0.345	(0.045)**
전기 성적	0.313	(0.005)**	0.296	(0.009)**	0.106	(0.011)**
자기공부시간	0.004	(0.001)*	0.002	(0.002)	0.007	(0.003)*
남성 여부	−0.133	(0.023)**	−0.170	(0.028)**	0.253	(0.051)**
만 나이	0.068	(0.031)*	0.081	(0.033)*	−0.077	(0.061)
가구 내 자녀 수	−0.041	(0.010)**	−0.030	(0.012)*	−0.048	(0.024)*
부모의 결혼 여부	0.091	(0.061)	0.018	(0.069)	0.475	(0.116)**
부모 평균연령	0.008	(0.003)**	0.008	(0.003)**	0.019	(0.007)**
부친 교육연수	0.026	(0.005)**	0.019	(0.006)**	0.042	(0.011)**
모친 교육연수	0.019	(0.006)**	0.013	(0.006)*	0.032	(0.012)**
가구소득 결측	0.364	(0.148)*	−0.287	(0.292)	4.245	(0.303)**
가구소득 로그값	0.046	(0.019)*	−0.042	(0.039)	0.572	(0.039)**
상수항	−4.093	(0.057)**	−3.977	(0.060)**	−1.901	(1.003)
학교특성변수들	통제됨		통제됨		통제됨	
F(excluded IV)					59.95	
R-square	0.597		–		0.235	
관측치 수	5,122		5,122		5,122	

주: 괄호 안에는 추정치의 표준오차가 표시됨. * p 〈 0.05; ** p 〈 0.01.

〈표 2〉의 (1)열에 제시된 바와 같이, OLS의 추정 결과에 의하면, 사교육비 지출과 성적 간의 연관관계는 통계적으로 0을 기각하지만 그리 크게 나타나지는 않는다. 즉, 10% 높은 사교육비 지출은 평균 0.008SD(표준편차) 정도 높은 점수로 연결된다. 이를 탄력성 계수로 변환하면, 10% 높은 사교육비 지출은 평균점수를 얻는 학생의 점수를 약 0.259% 정도 높이는 효과가 있다. 앞에서 강조된 바와 같이, OLS 방법은 s_i의 내생성을 적절히 고려하지 못하기 때문에, 추정 결과가 사교육비의 진정한 효과라고 확신하기 어렵다. $Cov(s_i, \epsilon_i)$의 방향에 따라 추정된 결과가 사교육비 지출의 진정한 효과를 과대추정할 수도 또는 과소추정할 수도 있다.

〈표 2〉의 (2)열에 제시된 도구변수법의 추정 결과는 OLS 결과에 비하여 강한 사교육비의 성적 향상효과를 보여준다. 즉, 10% 높은 사교육비 지출은 평균적인 학생의 세 과목 평균점수를 약 0.744% 정도 향상시키는 효과가 있다. 그리고 이 크기는 0과는 통계적으로 유의하게 다른 수준이다.

도구변수법의 추정 결과가 일치추정량을 보여주는 두 가지 조건은 '$Cov(F_i, s_i) \neq 0$'과 '$Cov(F_i, \epsilon_i) = 0$'이다. 〈표 2〉의 (3)열에는 s_i가 종속변수이고, F_i 및 X_i가 설명변수인 식에 대한 1단계(first-stage) 추정 결과가 보고되어 있다. 앞 절에서 예상된 바와 같이, 첫째아 학생에 대한 사교육비 지출액은 둘째아 이상의 학생에 비하여 약 34.5% 정도 높다. 이는 0보다 통계적으로 유의하게 높은 것이며, F-검정통계량은 59.95로서 F_i가 s_i의 적절한 도구변수일 첫 번째 조건을 만족한다.

그러나 앞 절에서 논의된 바에 따르면, F_i와 ϵ_i 사이에서는 '$Cov(F_i, \epsilon_i) = 0$'보다는 '$Cov(F_i, \epsilon_i) > 0$'의 관계가 성립할 가능성이 높다. 즉, 위에 제시된 사교육비의 효과에 대한 도구변수법의 추정 결과는 진정한 효과의 과대추정치일 가능성이 높다. 결국 10% 높은 사교육비 지출은 평균적인 학생의 점수를 0.744%보다는 작은 정도로 향상시키는 효과가 있다고 할 수 있다. 그러나 효과의 최소치가 어느 정도일지에 대해서 도구변수법의 결과를 통해서는 정확히 알 수 없다.[5]

---

5 〈표 2〉와 〈표 3〉에서 통제되는 설명변수에는 학생의 전기 성적, 자기공부시간, 성별, 나이, 부모의 연령 및 교육연수, 가구소득, 거주주택의 종류, 학교가 위치한 도시의 특성

사교육비가 세 과목의 평균성적에 미치는 효과가 그리 크지 않다는 사실은 사교육과 학업성적의 대상 과목을 국어, 영어, 수학으로 세분하는 경우에도 크게 다르지 않다. 〈표 3〉의 패널 A, B 및 C에는 각각 국어, 영어, 수학 개별 과목의 Z-점수를 종속변수로 사용하고 각 과목별 사교육비 지출액을 s_i로 설정한 모형의 추정 결과가 제시되어 있다.

먼저 OLS의 추정 결과에 의하면, 10% 높은 과목별 사교육비 지출은 평균적인 학생의 국어, 영어, 수학의 시험성적을 각각 0.08%, 0.33%, 0.40% 정도 향상시키는 효과가 있다. 그러나 s_i의 내생성으로 인하여 위의 결과들이 사교육비의 진정한 인과효과를 제시한다고 판단하기 어렵다. s_i의 내생성을 어느 정도 고려하는 도구변수법의 추정 결과에 의하면, 10% 높은 과목별 사교육비 지출은 평균적인 학생의 국어, 영어, 수학의 시험성적을 각각 1.24%, 1.28%, 0.75% 정도 향상시키는 효과가 있다. 이 정도 크기의 효과는 세 과목의 평균성적과 세 과목 전체의 사교육비 지출을 이용하는 모형에 비하여 약간 크다고 할 수 있다. 특히 국어와 영어에 대한 사교육비 지출의 효과는 세 과목 평균의 경우에 비하여 약 1.6배 정도 크게 나타난다. 그러나 도구변수법의 결과가 사교육비의 진정한 효과에 대한 과대추정치일 가능성이 있음을 감안하면 그 효과의 크기를 해석하는 데 주의할 필요가 있다. 참고로 변수 F_i는 세 과목 모두에 대하여 '$Cov(F_i, s_i) \neq 0$'의 조건을 만족하는 s_i의

(대도시, 중소도시), 사립 여부 및 남녀공학 여부 등 학교의 특성들이 포함된다. 본 논문에 사용되는 '첫째아 여부' 변수가 의문의 여지없이 외생적인 경우 사교육비의 인과효과를 추정하는 모형에서 X_i로 포괄되는 설명변수들이 추가로 통제될 필요는 없다. $\frac{E(y_i|F_i=1)-E(y_i|F_i=0)}{E(s_i|F_i=1)-E(s_i|F_i=0)}$로 표현되는 Wald 추정치에서는 추가적인 통제변수인 X_i가 전혀 고려되지 않는다. 그러나 이 경우는 도구변수가 무작위로 결정되는 지극히 예외적인 경우이다. 본 논문에서와 같이 '첫째아 여부' 변수에 어느 정도의 내생성이 있는 것으로 생각되는 경우에는 추가적인 X_i 변수를 통제하여 도구변수의 내생성을 어느 정도 줄이는 노력이 필요하다. 본 논문에서는 이러한 노력의 일환으로서 s_i와 더불어 추가적인 설명변수들 X_i를 통제하고 있다. 세 과목 평균성적을 사용하는 경우의 Wald 추정치는 0.550(se 0.061), 국어에 대한 Wald 추정치는 0.460(se 0.064), 영어에 대한 Wald 추정치는 0.568(se 0.068), 수학에 대한 Wald 추정치는 0.441(se 0.053)로서 〈표 2〉와 〈표 3〉의 2SLS 추정치들에 비하여 약간 크다. 이는 '첫째아 여부' 도구변수가 완전히 외생적이지는 않음을 시사한다. 그러나 Wald 추정치들과 〈표 2〉와 〈표 3〉에 제시된 2SLS 추정치들이 큰 차이를 보이지는 않기 때문에, 본 논문에서는 X_i를 통제하는 모형을 기초로 '첫째아 여부' 변수를 도구변수로 사용하는 분석방법을 채택한다.

〈표 3〉 OLS 및 도구변수법 추정 결과: 종속변수는 개별 과목 점수

통계방법:	OLS	도구변수법	
변수명		y_i 추정식	s_i 추정식
	(1)	(2)	(3)
		패널 A. 국어 점수	
사교육비 로그값	0.026 (0.011)*	0.403 (0.203)*	
[탄력성계수]	[0.079]	[1.244]	
첫째아 여부			0.137 (0.033)**
전기 성적	0.281 (0.006)**	0.266 (0.011)**	0.037 (0.008)**
자기공부시간	0.005 (0.005)	0.003 (0.006)	0.005 (0.007)
남성 여부	−0.235 (0.029)**	−0.337 (0.065)**	0.278 (0.038)**
만 나이	0.067 (0.036)	0.097 (0.043)*	−0.077 (0.046)
가구 내 자녀 수	−0.036 (0.013)**	−0.017 (0.018)	−0.038 (0.016)*
부모의 결혼 여부	0.151 (0.073)*	0.078 (0.088)	0.197 (0.073)**
부모 평균연령	0.006 (0.004)	0.009 (0.004)*	0.000 (0.005)
부친 교육연수	0.012 (0.006)*	0.012 (0.006)	−0.001 (0.008)
모친 교육연수	0.024 0.007)**	0.022 (0.008)**	0.004 (0.009)
가구소득 결측	0.269 (0.193)	−0.363 (0.396)	1.710 (0.208)**
가구소득 로그값	0.038 (0.025)	−0.048 (0.053)	0.231 (0.027)**
상수항	−3.473 (0.072)**	−4.432 (0.079)**	2.121 (0.083)**
F(excluded IV)			16.80
관측치 수	4,670	4,670	4,670
		패널 B. 영어 점수	
사교육비 로그값	0.107 (0.009)**	0.414 (0.092)**	
[탄력성계수]	[0.331]	[1.280]	
첫째아 여부			0.274 (0.037)**
전기 성적	0.272 (0.005)**	0.246 (0.010)**	0.080 (0.009)**
자기공부시간	0.016 (0.004)**	0.007 (0.005)	0.030 (0.007)**
남성 여부	−0.204 (0.026)**	−0.249 (0.032)**	0.159 (0.042)**
만 나이	0.051 (0.034)	0.065 (0.038)	−0.040 (0.050)
가구 내 자녀 수	−0.044 (0.011)**	−0.025 (0.015)	−0.043 (0.018)*
부모의 결혼 여부	0.041 (0.065)	−0.059 (0.076)	0.332 (0.090)**
부모 평균연령	0.009 (0.003)**	0.008 (0.003)*	0.020 (0.005)**
부친 교육연수	0.026 (0.005)**	0.015 (0.007)*	0.036 (0.009)**
모친 교육연수	0.023 (0.006)**	0.013 (0.007)	0.028 (0.010)**
가구소득 결측	0.479 (0.157)**	−0.537 (0.345)	3.384 (0.247)**
가구소득 로그값	0.056 (0.020)**	−0.079 (0.045)	0.450 (0.032)**
상수항	−3.873 (0.061)**	−3.620 (0.068)**	−1.741 (0.099)*
F(excluded IV)			54.28
관측치 수	4,755	4,755	4,755

〈표 3〉의 계속

통계방법:	OLS	도구변수법	
변수명		y_i 추정식	s_i 추정식
	(1)	(2)	(3)
패널 C. 수학 점수			
사교육비 로그값	0.130 (0.009)**	0.244 (0.097)*	
[탄력성계수]	[0.402]	[0.754]	
첫째아 여부			0.241 (0.037)**
전기 성적	0.290 (0.005)**	0.281 (0.010)**	0.071 (0.009)**
자기공부시간	0.017 (0.004)**	0.014 (0.005)**	0.031 (0.007)**
남성 여부	0.048 (0.026)	0.029 (0.031)	0.179 (0.042)**
만 나이	0.061 (0.034)	0.072 (0.036)	−0.097 (0.050)*
가구 내 자녀 수	−0.038 (0.012)**	−0.033 (0.012)**	−0.025 (0.019)
부모의 결혼 여부	0.050 (0.067)	0.019 (0.072)	0.270 (0.087)**
부모 평균연령	0.004 (0.003)	0.004 (0.003)	0.018 (0.005)**
부친 교육연수	0.025 (0.005)**	0.020 (0.007)**	0.043 (0.009)**
모친 교육연수	0.002 (0.006)	−0.001 (0.007)	0.026 (0.010)*
가구소득 결측	0.272 (0.166)	−0.105 (0.364)	3.379 (0.247)**
가구소득 로그값	0.040 (0.021)	−0.010 (0.048)	0.449 (0.032)**
상수항	−3.648 (0.060)**	−3.663 (0.061)**	−0.665 (0.089)
F(excluded IV)			43.26
관측치 수	4,721	4,721	4,721

주: 괄호 안에는 추정치의 표준오차가 표시됨. * p 〈 0.05; ** p 〈 0.01.
개별 과목에 대한 추정에서 학교의 특성들은 설명변수로서 통제됨.

강한 도구변수(strong IV)이다. F-검정통계량은 국어의 경우에는 16.8, 영어의 경우에는 54.3, 수학의 경우에는 43.3으로서, 세 과목 모두에서 통상적인 기준치 10을 넘어서고 있기 때문이다.

도구변수법의 추정 결과를 요약하면, 국어, 영어, 수학 개별 과목에 대해서든 혹은 세 과목의 평균에 대해서든 공통적으로, 사교육비 지출은 평균적인 학생의 성적을 약간 향상시키는 양의 효과가 존재하기는 하지만 그 효과의 정도는 그리 크지 않은 것으로 판단된다.

2. 비모수 구간추정법의 추정 결과

〈표 4〉에는 사교육비의 평균효과(Average Treatment Effect: ATE)에 대한 구간추정법의 추정 결과가 세 과목 평균점수 및 개별 과목별로 제시되어 있다. 각 패널별로 ATE에 대한 MTR+MTS 범위와 MIV+MTR+MTS 범위가 제시되어 있다. 이 추정치들에 대한 편리한 해석을 위하여 (5)열과 (6)열에는 각각 ATE 범위의 최댓값과 최댓값의 bootstrap 95분위값을 10%의 사교육비 지출 증가에 따른 성적 변화율로 변환한 탄력성 계수가 제시되어 있다.

앞에서 제시된 추정 결과들과 마찬가지로, 〈표 4〉의 패널 A에 나타난 세 과목 전체에 대한 구간추정법의 결과는 높은 사교육비 지출로 인하여 학생의 성적이 크게 향상되지는 않음을 암시한다. MTR+MTS 범위의 최댓값에 의하면, 10% 높은 사교육비 지출은 학생의 평균성적을 최대 1.86~2.36% 정도까지 향상시킨다. 그러나 MTR+MTS 범위의 최솟값은 사교육비 지출의 효과가 0일 가능성을 배제하지 못한다. 이는 사교육비 지출의 진정한 성적 향상효과가 그리 크지 않을 것임을 암시한다.

MTR+MTS의 결합 가정이 타당한지를 검증하기 위하여 Manski and Pepper (2000, p.1004)는 u와 u' $(u' \le u)$에 대하여 $E(y|z=u') \le E(y|z=u)$ 이 성립하는지의 여부를 조사하는 방법을 제안하였다. 왜냐하면 MTR+MTS 라는 결합 가정하에서는 다음이 성립하기 때문이다.

$$
\begin{aligned}
& u' \le u \ \Rightarrow E[y|z=u'] = \\
& E[y(u')|z=u'] \le_{MTR} \\
& E[y(u)|z=u'] \le_{MTS} \\
& = E[y(u)|z=u] = E[y|z=u]
\end{aligned}
$$

〈표 5〉에는 표본평균 $\hat{E}[y(0)]$, $\hat{E}[y(1)]$, $\hat{E}[y(2)]$이 각 과목별로 보고되어 있다. 표의 결과에 따르면, 세 과목의 평균점수를 사용하든 개별 과목의 점수를 사용하든 공통적으로 '$\hat{E}[y(0)] \le \hat{E}[y(1)] \le \hat{E}[y(2)]$'의 관계가 대체로 성립한다. 다만, 국어 과목에서는 $\hat{E}[y(1)] = 0.283$, $\hat{E}[y(2)] = 0.285$

〈표 4〉 사교육비 효과에 대한 구간추정법의 추정 결과

통계방법:	최솟값 (1)	최댓값 (2)	최솟값의 5 pctile (3)	최댓값의 95 pctile (4)	최댓값 (5)	최댓값의 95 pctile (6)
패널 A. 세 과목 평균 점수						
	MTR+MTS 범위				탄력성 계수	
$E[y(1)-y(0)]$	0.000	0.678	0.000	0.728	2.064	2.218
$E[y(2)-y(1)]$	0.000	0.454	0.000	0.507	1.866	2.083
$E[y(2)-y(0)]$	0.000	0.819	0.000	0.879	2.364	2.538
	MIV+MTR+MTS 범위				탄력성 계수	
$E[y(1)-y(0)]$	0.000	0.654	0.000	0.703	1.993	2.141
$E[y(2)-y(1)]$	0.000	0.444	0.000	0.494	1.823	2.027
$E[y(2)-y(0)]$	0.000	0.795	0.000	0.851	2.294	2.458
패널 B. 국어 점수						
	MTR+MTS 범위				탄력성 계수	
$E[y(1)-y(0)]$	0.000	0.296	0.000	0.347	0.875	1.025
$E[y(2)-y(1)]$	0.000	0.195	0.000	0.253	0.773	1.004
$E[y(2)-y(0)]$	0.000	0.295	0.000	0.354	0.872	1.046
	MIV+MTR+MTS 범위				탄력성 계수	
$E[y(1)-y(0)]$	0.000	0.271	0.000	0.323	0.801	0.955
$E[y(2)-y(1)]$	0.000	0.194	0.000	0.264	0.771	1.047
$E[y(2)-y(0)]$	0.000	0.285	0.000	0.356	0.843	1.052
패널 C. 영어 점수						
	MTR+MTS 범위				탄력성 계수	
$E[y(1)-y(0)]$	0.000	0.608	0.000	0.653	2.265	2.430
$E[y(2)-y(1)]$	0.000	0.522	0.000	0.575	2.359	2.598
$E[y(2)-y(0)]$	0.000	0.808	0.000	0.861	2.734	2.913
	MIV+MTR+MTS 범위				탄력성 계수	
$E[y(1)-y(0)]$	0.000	0.579	0.000	0.626	2.157	2.332
$E[y(2)-y(1)]$	0.000	0.512	0.000	0.556	2.314	2.510
$E[y(2)-y(0)]$	0.000	0.781	0.000	0.830	2.643	2.809
패널 D. 수학 점수						
	MTR+MTS 범위				탄력성 계수	
$E[y(1)-y(0)]$	0.000	0.739	0.000	0.784	2.708	2.874
$E[y(2)-y(1)]$	0.000	0.443	0.000	0.490	2.078	2.301
$E[y(2)-y(0)]$	0.000	0.822	0.000	0.875	2.898	3.085
	MIV+MTR+MTS 범위				탄력성 계수	
$E[y(1)-y(0)]$	0.000	0.726	0.000	0.770	2.658	2.822
$E[y(2)-y(1)]$	0.000	0.434	0.000	0.483	2.036	2.269
$E[y(2)-y(0)]$	0.000	0.805	0.000	0.856	2.837	3.018

〈표 5〉 처치수준별 성적의 평균, $\widehat{E}[y(t)]$

평균값	분석 표본			
	세 과목 평균	국어	영어	수학
$\widehat{E}[y(0)]$	−0.294	−0.012	−0.265	−0.327
$\widehat{E}[y(1)]$	0.336	0.283	0.260	0.382
$\widehat{E}[y(2)]$	0.525	0.285	0.543	0.495

로서 '$\hat{E}[y(1)] < \hat{E}[y(2)]$'의 관계가 통계적으로 유의하게 성립하지는 않는다. 그러나 '$\hat{E}[y(1)] > \hat{E}[y(2)]$'의 관계가 성립한다고 단정할 수도 없으므로, 본 자료에서 MTR와 MTS의 결합 가정에 문제가 있다는 명시적인 증거는 발견되지 않는다.

〈표 4〉에 제시된 세 과목 전체의 평균점수에 대한 MIV+MTR+MTS 범위의 추정 결과도 높은 사교육비 지출로 인하여 학생의 성적이 크게 향상되지는 않는다는 점을 보여준다. MIV+MTR+MTS 범위의 최댓값에 의하면, 10% 높은 사교육비 지출은 학생의 세 과목 평균성적을 최대 1.82~2.29% 정도까지 향상시킨다. 그러나 MIV+MTR+MTS 범위의 최솟값은 사교육비 지출의 효과가 0이라는 가설을 기각하지는 못한다. MIV+MTR+MTS 범위 또한 사교육비 지출의 진정한 성적 향상효과가 그리 크지 않음을 암시한다.

사교육비 지출의 그리 크지 않은 성적 향상효과는 사교육비와 성과변수를 개별 과목으로 한정하는 경우에도 비슷하게 나타난다. 국어 과목의 경우, MIV+MTR+MTS 범위의 탄력성 계수 최댓값은 약 0.77~0.84% 정도로서 사교육비 지출의 효과가 그리 크지 않음을 보여준다. 하지만 영어와 수학 과목에 대한 MIV+MTR+MTS 범위의 탄력성 계수 최댓값은 사교육비 지출이 그리 작지 않은 효과가 있을 가능성을 제기한다. 영어의 경우, 10% 높은 사교육비 지출이 영어의 평균성적을 최대 2.16~2.64% 정도까지 향상시킨다. 수학의 경우에는 최대 2.04~2.84% 정도까지 향상시킨다. 그러나 두 과목에서 공통적으로 MIV+MTR+ MTS 범위의 최솟값은 0과 다르지 않다. 결국 구간추정법의 결과를 보수적으로 해석하면, 국어, 영어, 수학 과목에서 모두 높은 사교육비 지출이 유의미한 정도의 성적 향상으로 연결된다는 확실한 증거를 발견하기는 어렵다.

사교육비 지출의 내생성을 명시적으로 고려하는 도구변수법 및 구간추정법의 추정 결과는 사교육비 지출이 학생의 학업성적에 다소 영향을 미치기는 하지만 그 영향이 그리 크지는 않은 것으로 요약된다. 도구변수법의 결과를 사교육비 지출의 평균효과에 대한 점추정치(point estimates)로 해석하면, 10% 높은 사교육비 지출은 국어성적을 약 1.24%, 영어성적을 약 1.28%, 수학성적을 약 0.75%, 그리고 세 과목 전체의 평균성적을 약 0.74% 정도 향상시킨다고 요약할 수 있다. 사용된 도구변수의 특성으로 인하여 이상의 추정치들은 실제 효과에 대한 과대추정치로 해석된다. 아울러 위의 점추정치들이 위치한 구간은 대체로 구간추정법을 통하여 추정된 ATE의 범위들 내에 위치함으로써 상이한 추정방법들은 사교육비의 평균효과에 대하여 대체로 일관된 결과를 보여준다. 이는 본 논문에서 추정한 사교육비의 평균효과의 크기가 상당히 신뢰할 만한 수치임을 암시한다.

사교육비 지출이 성적을 향상시키는 강력한 효과가 있을 것이라는 일부의 기대와는 다르게 본 논문의 추정 결과는 그 효과가 그리 크지 않음을 보여준다. 본 논문에 제시된 사교육비 효과의 추정치를 정확히 해석하기 위해서는 먼저 실제 지출된 사교육비의 분포를 자세히 살펴볼 필요가 있다. 〈표 1〉에 따르면, 분석자료에서 나타난 국어의 월평균 사교육비는 약 31,970원, 영어 사교육비는 약 69,820원, 수학 사교육비는 약 65,430원, 그리고 세 과목 전체에 대한 사교육비는 약 154,100원 정도이다. 그리고 각 과목에 대한 사교육비 분포의 95 퍼센타일을 구해 보면, 국어의 경우에는 약 132,200원, 영어의 경우에는 약 283,300원, 수학의 경우에는 약 283,300원, 그리고 세 과목 전체에 대한 사교육비는 약 566,600원 정도이다. 즉, 전체 학생의 약 95%는 세 과목 전체 사교육비 기준으로 월평균 566,600원보다 작은 금액을 국어, 영어, 수학 과목 전체의 사교육에 지출한다. 본 논문에서 구한 사교육비의 효과는 주로 월평균 사교육비로 이 정도의 금액을 지출하는 학생들에 대한 평균효과로 해석된다.

일반적으로 자료들이 풍부하게 관측되는 사교육비의 평균값 근처에서 계산된 탄력성 계수는 그 정확도가 상당히 높다. 그러나 사교육비의 평균값과 큰 차이가 나는 사교육비 증가의 한계효과는 그 정확도가 떨어질 가능성이

높다. 예를 들어 월평균 사교육비를 566,600원에서 1,133,200원으로 두 배 증가시키면 성적의 변화율이 어느 정도일지에 대한 한계효과는 그 정확도가 상당히 떨어진다고 할 수 있다.

한편, 월평균 사교육비로 95 퍼센타일 이하의 금액을 지출하는 학생들이 주로 받는 사교육은 성적 향상에 크게 도움이 되지 않는다 하더라도, 95 퍼센타일 이상의 월평균 사교육비에 해당하는 고액의 사교육은 학생의 성적을 크게 향상시키는 효과가 있을 수 있다. 이는 사교육비의 효과가 지출수준에 따라 비선형의 형태를 취할 수 있음을 의미한다. 월평균 60만원 이상 고액의 사교육비를 지출하는 학생들의 비중이 약 5% 이하에 불과한 자료상의 한계를 감안할 때, 고액의 사교육이 학업성적에 미치는 영향에 관하여 본 논문에서는 정확히 알 수 없다. 이를 규명하기 위해서는 고액의 사교육 부분에 특화된 새로운 자료가 일차적으로 요구된다.

제5절 결 론

본 논문에서 우리는 경제학의 실증분석방법들을 활용하여 사교육비의 성적 향상효과가 어느 정도인지를 추정하였다. 추정 결과에 의하면, 우리나라에서 사교육비 지출은 평균적인 학생의 학업성적에 다소간의 영향을 미치기는 하지만 그 영향이 그리 크지는 않은 것으로 요약된다. 도구변수법의 결과에 의하면, 평균 정도의 사교육비를 지출한 상태에서 지출을 10% 정도 높이면 국어성적은 약 1.24%, 영어성적은 약 1.28%, 수학성적은 약 0.75%, 그리고 세 과목 전체의 평균성적은 약 0.74% 정도 향상되는 것으로 추정된다.

본 논문에서 추정된 사교육비 지출의 탄력성 계수는 서구의 연구에서 보고된 바 있는 공립학교의 교육비 투자가 미치는 성적 향상효과와 크게 모순되지 않는다. 예를 들어 Guryan(2003)의 연구 결과에 의하면, 학생 일인당 교육비 지출을 10% 늘릴 때 학업성적은 약 0.77~1.15% 정도 향상된다. 그리고 Card and Krueger(1996)와 Grogger(1996)가 요약한 바에 따르면, 학생 일인당 교육비 지출이 10% 늘어날 때 노동시장에서의 임금으로 표시된 교육

성과는 약 0.7~1.1% 정도 향상된다. 한국에서 추정된 학생 일인당 사교육비의 성적 향상효과도 이 정도 수준에 대응된다고 요약할 수 있다.

본 논문에서 계산된 사교육비의 효과가 여타의 연구들과 크게 모순되지 않는다고 하더라도 본 논문의 결과는 사교육이 만연한 우리 사회의 현실과 조화되지 않는다는 비판이 있을 수 있다. 우리는 이에 대하여 아래에서 세 가지 정도의 이론적인 가능성을 검토해 보고자 한다.

첫째, 사교육비의 진정한 효과에 관하여 부모들이 정확히 알지 못할 가능성이 있다. 사교육비의 한계수익에 대한 정확한 정보가 없기 때문에 사교육비의 한계수익과 한계비용이 일치되는 최적의 수준에서 자녀의 사교육비 지출액이 결정되지 않을 수 있다. 그렇다면 사교육의 효과가 그리 크지 않더라도 부모는 자녀의 사교육비용을 과도하게 지출할 수 있다.

둘째, 사교육비의 진정한 효과의 크기를 알고 있다고 하더라도, 사교육비 지출 결정에 외부성이 존재하는 경우에 최적의 사교육비 지출이 일어나지 않을 수 있다. 본인의 자녀에 대한 사교육비 지출액의 결정에 타인 자녀의 사교육비 지출액(예를 들어 동료집단의 평균 사교육비 지출액)이 영향을 미친다면, 외부성으로 인하여 최적의 수준에서 본인 자녀에 대한 사교육비 지출액이 결정되지 않을 수 있다. 이 경우 사교육비의 낮은 한계수익과 높은 지출액이 공존할 수 있다.

셋째, 한 학생의 성적 결정에서 사교육과 자기학습(self-study) 간에 대체적인 관계가 존재하는 경우이다. 만약 학생이 사교육을 받는 양에 비례하여 자기학습의 강도나 시간을 줄인다면 사교육비 지출이 성적 향상에 미치는 영향이 그리 크지 않을 수 있다. 이러한 현상이 장기화될 경우 부모는 사교육과 자기공부 간의 대체관계를 인식하여 사교육비 지출을 최적화할 것이므로, 세 번째의 설명에는 첫 번째의 설명과 유사하게 사교육비의 진정한 효과에 관하여 부모들이 정확히 알지 못한다는 가정이 내포된다.

위에서 제시한 설명들은 현재로서는 풍부한 실증적인 증거로서 뒷받침되지는 못하는 이론적인 추측에 불과하다. 앞으로 보다 진전된 과학적인 연구들이 개발되어 사교육비의 효과와 사교육이 만연한 현실을 조화시킬 적절한 이론들이 제시되기를 기대한다.

참고문헌

강창희・이삼호, 『사교육에 대한 경제학적 분석 및 정책 제언』, 한국교육개발원 연구보고서, CR 2010-12, 2010.

김양분・류한구・김미숙・남궁지영, 『학교교육 수준 및 실태 분석 연구: 중학교』, 한국교육개발원 연구보고서, RR 2004-3, 2004.

김지하・김정은, 「Propensity Score Matching 방법을 사용한 사교육 유형별 효과 분석」, 『교육재정경제연구』, 제18권 제3호, 2009, pp.63~87.

김진영, 「고3 수험생들의 시간활용과 사교육의 효과」, 『한국교육』, 제34권 제4호, 2007, pp.57~78.

남기곤, 「사교육시간과 학업성적과의 관련성: PISA 자료를 이용한 국제비교 분석」, 『한국경제학보』, 제15권 제1호, 2008, pp.55~90.

박현정・상경아・강주연, 「사교육이 중학생의 학업성취에 미치는 효과」, 『교육평가연구』, 제21권 제4호, 2008, pp.107~127.

상경아・백순근, 「고등학생의 수학 과외가 학업성취도, 태도, 자기조절학습에 미치는 영향」, 『교육평가연구』, 제18권 제3호, 2005, pp.39~57.

오영수・윤정식, 「일반계 고등학생의 성적 결정요인 분석」, 『교육재정경제연구』, 제12권 제1호, 2003, pp.33~56.

이수정・임현정, 「중학생의 학업성취에 대한 사교육비 효과 분석」, 『교육재정경제연구』, 제18권 제1호, 2009, pp.141~166.

임천순・박소영・이광호, 「사교육이 학업성취에 미치는 영향」, 『교육재정경제연구』, 제13권 제1호, 2004, pp.331~356.

조혜영・이경상, 「사교육시간, 개인공부시간, 학교수업참여도의 실태 및 주관적 학업성적 향상효과」, 『한국교육』 제32권 제4호, 2005, pp.29~56.

최형재, 「사교육의 대학 진학에 대한 효과」, 『국제경제연구』, 제14권 제1호, 2008, pp.73~110.

최상근 외, 『사교육 실태 및 사교육비 규모 분석 연구』, 한국교육개발원 연구보고서, CR 2003-19, 2003.

한대동・성병창・길임주, 「고등학생 학업성취에 대한 학교효과와 과외효과의 비교

연구」, 『교육사회학 연구』, 제11권 제1호, 2001, pp.33~54.

Anderson, G. M., W. F. Shughart, and R. Tollison, "Educational Achievement and the Cost of Bureaucracy," *Journal of Economic Behavior & Organization* 15(1), 1991, pp.29~45.

Anderson, T. W. and H. Rubin, "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics* 20(1), 1949, pp.46~63.

Bedard, K., "School Quality and the Distribution of Male Earnings in Canada," *Economics of Education Review* 22, 2003, pp.395~407.

Behrman, J. R. and P. Taubman, "Birth Order, Schooling, and Earnings," *Journal of Labor Economics* 4(3), 1986, pp.S121~S145.

Betts, J. R., "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth," *Review of Economics and Statistics* 77(2), 1995, pp.231~250.

Betts, J. R., "Is There a Link between School Inputs and Earnings? Fresh Scrutiny of an Old Literature," G. Burtless (ed.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, Brookings, Washington, DC, 1996, pp.141~191.

Bjerkedal, T., P. Kristensen, G. A. Skjeret, and J. I. Brevik, "Intelligence Test Scores and Birth Order among Young Norwegian Men (Conscripts) Analyzed within and between Families," *Intelligence* 35(5), 2007, pp.503~514.

Black, S. E., P. J. Devereux, and K. G. Salvanes, "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education," *Quarterly Journal of Economics* 120(2), 2005, pp.669~700.

Black, S. E., P. J. Devereux, and K. G. Salvanes, "Older and Wiser? Birth Order and IQ of Young Men," NBER Working Paper 13237, 2007.

Blundell, R., A. Gosling, H. Ichimura, and C. Meghir, "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds," *Econometrica* 75(2), 2007, pp.323~363.

Briggs, D. C., “The Effect of Admissions Test Preparation: Evidence from NELS: 88,” *Chance* 14(1), 2001, pp.10~18.

Card, D. and A. B. Krueger, “Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States,” *Journal of Political Economy* 100(1), 1992, pp.1~40.

Card, D. and A. B. Krueger, “School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina,” *Journal of Economic Perspectives* 10(4), 1996, pp.31~50.

Dang, H., “The Determinants and Impact of Private Tutoring Classes in Vietnam,” *Economics of Education Review* 26(6), 2007, pp.683~698.

Dang, H. and F. H. Rogers, “How to Interpret the Growing Phenomenon of Private Tutoring: Human Capital Deepening, Inequality Increasing, or Waste of Resources?” Policy Research Working Paper Series of The World Bank, No. 4530, 2008.

Gerfin, M. and M. Schellhorn, “Nonparametric Bounds on the Effect of Deductibles in Health Care Insurance on Doctor Visits—Swiss Evidence,” *Health Economics* 15(9), 2006, pp.1011~1020.

Gonzalez, L., “Nonparametric Bounds on the Returns to Language Skills,” *Journal of Applied Econometrics* 20(6), 2005, pp.771~795.

Greenwald, R., L. V. Hedges, and R. D. Laine, “The Effect of School Resources on Student Achievement,” *Review of Educational Research* 66(3), 1996, pp.361~396.

Grogger, J., “School Expenditures and Post-schooling Earnings: Evidence from High School and Beyond,” *Review of Economics and Statistics* 78(4), 1996, pp.628~637.

Gurun, A. and D. L. Millimet, “Does Private Tutoring Payoff,” IZA Discussion Paper, No. 3637, The Institute for the Study of Labor (IZA), 2008.

Guryan, J., “Does Money Matter? Estimates from Education Finance Reform in Massachusetts,” mimeo, 2003.

Hanushek, E. A., “Assessing the Effects of School Resources on

Student Performance: An Update," *Educational Evaluation and Policy Analysis* 19(2), 1997, pp.141~164.

Hanushek, E. A., "The Failure of Input-based Schooling Policies," *Economic Journal* 113(485), F64-F98, 2003.

Harris, J. R., "Where Is the Child's Environment? A Group Socialization Theory of Development," *Psychological Review* 102(3), 1995, pp.458~489.

Harris, J. R., *The Nurture Assumption: Why Children Turn out the Way They Do*, New York, NY: Touchstone, 1998.

Haveman, R. and B. Wolfe, "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature* 33(4), 1995, pp.1829~1878.

Hedges, L. V., R. D. Laine, and R. Greenwald, "Does Money Matter? A Meta-analysis of Studies of the Effects of Differential School Inputs on Student Outcomes," *Educational Researcher* 23(3), 1994, pp.5~14.

Herrnstein, R. and C. Murray, *The Bell Curve*, The Free Press, New York, 1994.

Hoxby, C. M., "How Teachers' Unions Affect Education Production," *Quarterly Journal of Economics* 111(3), 1996, pp.671~718.

Kang, C., "Does Money Matter? The Effect of Private Educational Expenditures on Academic Performance," Departmental Working Papers from National University of Singapore, Department of Economics, 2007.

Kleibergen, F., "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression," *Econometrica* 70(5), 2002, pp.1781~1803.

Kreider, B. and J. V. Pepper, "Disability and Employment: Reevaluating the Evidence in Light of Reporting Errors," *Journal of the American Statistical Association* 102(478), 2007, pp.432~441.

Krueger, A. B., "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* 114(2), 1999, pp.497~532.

Krueger, A. B., "Economic Considerations and Class Size," *Economic Journal* 113(485), F34-F62, 2003.

Lechner, J., "Nonparametric Bounds on Employment and Income Effects of Continuous Vocational Training in East Germany," *Econometrics Journal* 2(1), 1999, pp.1~28.

Leuven, E., M. Lindahl, H. Oosterbeek, and D. Webbink, "The Effect of Extra Funding for Disadvantaged Pupils on Achievement," *Review of Economics and Statistics* 89(4), 2007, pp.721~736.

Manski, C. F., "Nonparametric Bounds on Treatment Effects," *American Economic Review* 80(2), 1990, pp.319~323.

Manski, C. F., "Monotone Treatment Response," *Econometrica* 65(6), 1997, pp.1311~1334.

Manski, C. F. and D. S. Nagin, "Bounding Disagreements about Treatment Effects: A Case Study of Sentencing and Recidivism," *Sociological Methodology* 28, 1998, pp.99~137.

Manski, C. F. and J. V. Pepper, "Monotone Instrumental Variables, with an Application to the Returns to Schooling," *Econometrica* 68(4), 2000, pp.997~1012.

Moreira, M. J., "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica* 71(4), 2003, pp.1027~1048.

Ono, H., "Does Examination Hell Pay Off? A Cost-benefit Analysis of 'Ronin' and College Education in Japan," *Economics of Education Review* 26(3), 2007, pp.271~284.

Pepper, J. V., "The Intergenerational Transmission of Welfare Receipt: A Nonparametric Bounds Analysis," *Review of Economics and Statistics* 82(3), 2000, pp.472~488.

Raudenbush, S. W. and A. S. Bryk, *Hierarchical Linear Models: Applications and Data Analysis Methods*, (2nd ed.), Sage:CA. 2002.

Retherford, R. D. and W. H. Sewell, "Birth Order and Intelligence: Further Tests of the Confluence Model," *American Sociological Review* 56(2), 1991, pp.141~158.

Rodgers, J. L., H. H. Cleveland, E. V. D. Oord, and D. C. Rowe, "Resolving the Debate Over Birth Order, Family Size, and Intelligence," *American Psychologist* 55(6), 2000, pp.599~612.

Stock, J. H., J. H. Wright, and M. Yogo, "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics* 20(4), 2002, pp.518~529.

Zajonc, R. B., "Family Configuration and Intelligence," *Science* 192(4236), 1976, pp.227~236.

Zajonc, R. B. and P. R. Mullally, "Birth Order Reconciling Conflicting Effects," *American Psychologist* 52(7), 1997, pp.685~699.

CHAPTER 3

Education and Globalization on Income Inequality: Cross-Country Evidence

by
*Kang H. Park**
(Missouri State University)

Abstract

This study examines the effects of education and globalization on income inequality using cross-section data. We confirm the existence of the Kuznets inverted-U curve for the relationship between income level and income inequality. However, the inverse U-shaped curve lacks robustness when additional variables are used in the model. The empirical results show that educational variables play an important role for better distribution of income. Our findings indicates that a higher level of educational attainment of the population has an equalizing effect on income distribution, while the larger the dispersion of educational attainment among the population, the greater the income inequality. The dispersion of schooling among the population has a much greater dis-equalizing effect on income inequality than previous studies have suggested.

It is also found that the higher the level of globalization, the more unequal income distribution, while freedom, whether economic freedom or political freedom, has marginal effects on income inequality.

* All correspondence to Professor Kang H. Park, Professor of Economics, Missouri State University, Cape Girardeau, MO 63701, Tel: 573-651-2942, e-mail: khpark@semo.edu

1. Introduction

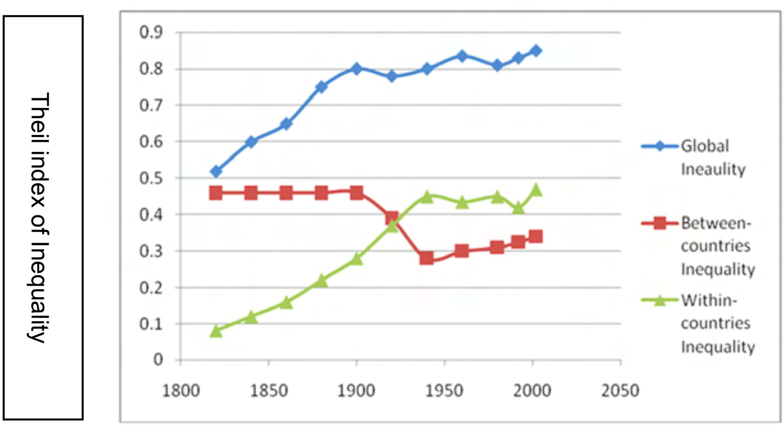
Income distribution and its economic and social effects has long been a topic of interest for many scholars. The recent resurgence of income inequality has raised concerns in regard to causes of rising income inequality. In his seminal 1955 article, Kuznets suggested that as economies develop, income inequality tends to rise first, reach its peak, and then decline. The causes of this trend have been attributed to dualism and structural changes that occur during the economic development process. Population growth and urbanization accompanying the early stage of economic development initially worsen income distribution, but later economic policies and political factors slow down income growth of the upper income group while improving the share of the lower income group. However, recent globalization and its resultant outsourcing and wage compression may have reversed trend in income inequality.

This well-known hypothesis has been the subject of much controversy in the economic literature and has inspired a great deal of research, particularly cross-country comparative empirical studies (Kravis, 1960; Kuznets, 1963; Adelman and Morris, 1973; Paukert, 1973; Ahluwalia, 1974; Chenery and Syrquin, 1975; Ahluwalia, 1976b; Robinson, 1976; Papanek, 1978; Stewart, 1978; Wright, 1978; Winegarden, 1979). More recently, with up-dated data on income distribution and other variables, there has been a surge of empirical research on Kuznets' inverted-U hypothesis with cross-country data (Ram, 1984; Papanek and Kyn, 1986; Tsakloglou, 1988; Ram, 1988; Bourguignon and Morrisson, 1990; Nielson and Alderson, 1995; Checchi, 2000; Barro, 2000; Alderson & Nielsen, 2002; De Gregorio and Lee, 2002; Ravallion, 2004; Wells, 2006).

Most of the cross-country empirical studies (Kuznets, 1963; Adelman and Morris, 1973; Ahluwalia, 1974; Papanek and Kyn, 1986; Tsakloglou, 1988; Ram, 1988; Bourguignon and Morrisson, 1990; Nielson and Alderson, 1995; De Gregorio and Lee, 2002 to name a few) have supported the inverted-U hypothesis with the exception of a few studies contesting this hypothesis (Saith, 1983; Anand and Kanbur, 1984; Ram, 1988, Ravallion, 2004). Recent research has suggested “the great U-turn" indicating that the relationship again reverses for the very high income countries (Galbraith *et al*., 2000; Alderson and Nielsen, 2002).

Because there are different types of income inequality, it is necessary to clarify the concept of income inequality to be used in this study. First, "world income inequality (or global income inequality)" considers all individuals in the world together and ranks them from the richest to the poorest, regardless of their country of origin. The citizen of the world rather than individual countries is the unit of analysis. The second concept is "international income inequality (or between countries income inequality)." This inequality refers to income inequality among countries due to differences in their per capita income or per capita GDP. Here the units of analysis are not individuals, but countries. The third and the most commonly used inequality is "national income inequality (or within countries income inequality)," referring to the dispersion of income distribution within a country. Yitzhaki (1994) showed that the index of world inequality can be decomposed as the sum of international income inequality and national income inequality plus residual. The trend of these three income inequalities from the period of 1820 to 1992 reported by Bourguignon and Morrisson (2002), based on the data of fifteen individual countries with abundant data and other eighteen

| Figure 1 | Three Income Inequalities by Bourguignon and Morrisson



country groups, is shown in Figure 1. In this study, our attention is on the national income inequality or within-countries income inequality.

Although the present study is another addition to numerous cross-sectional analyses done on Kuznets' inverted-U hypothesis, this study differs from the previous studies in the following respects. First, this study's focus is not on the Kuznets' inverted-U hypothesis itself, rather it emphasizes education variables as important explanatory variables for income inequality. Second, this study considers the impact of globalization and openness to trade in income inequality. Since the 1980s, many countries followed trade and financial liberalization and globalization is on the rise. Globalization affects income inequality directly and also indirectly through its effect on education.

In this study, we try to reestablish the relationship between education and income inequality in the context of an increasingly integrated and globalized world economy with more updated and expanded data. This paper is organized in the following manner. The next section briefly surveys the related literature on determinants of income inequality with focus on education and income inequality. This is followed by a section includes presentation of the models and description of data and the variables used in this study. Section 4 interprets the estimated results of the model. In the final section, summaries and policy suggestions will be made.

2. Review of Literature

Since Becker's human capital theory (1964), many studies have included education variables in explaining income distribution. The types of education variables used in explaining income distribution can be classified into four groups: (1) a flow variable of education such as enrollments at different levels of education, (2) a stock variable of education such as the mean or median years of schooling of the labor force, (3) the rate of return on education at different levels of education, and (4) the dispersion of educational attainment.

As a variable representing the level of educational attainment, many chose school enrollments at different levels of education (secondary and tertiary in Adelman and Morris, 1973; primary and secondary in Chenery and Syrquin, 1975, and Ahluwalia, 1976; female primary and secondary in Tsakloglou, 1988; secondary in Papanek and Kyn, 1986, and Bourguignon and Morrisson, 1990, both secondary and higher education in Barro, 2000, Alderson & Nielsen, 2002), while others chose average or median years of schooling (Marin and Psacharopoulos, 1976; Cromwell, 1977; Winegarden, 1979; Ram, 1984; Braun, 1988; Psacharopoulos, *et al*., 1995; Checchi, 2000; De Gregorio and Lee, 2002).

Psacharopoulos and Tilak (1991) used both mean years of schooling and enrollments at each level of education as explanatory variables. Tilak (1988) examined the effects of the rate of return to education on income distribution with cross-country data. There are a few studies that used the dispersion of educational attainment as an independent variable (Chiswick, 1971; Checchi, 2000). Some studies also considered both the level of educational attainment and the dispersion of educational attainment as independent variables. They include Tinbergen (1972 and 1975), Psacharopoulos (1977), Winegarden (1979), Ram (1984), and Bourguignon and Morrisson (1990) and De Gregorio and Lee (2002).

Because of the numerous cross-sectional studies done on Kuznets' inverted-U hypothesis and because of this study's emphasis is on the effect of educational inequality on income distribution, our review of the literature will be mainly limited to empirical studies that analyze the effects of the level of education and the inequality of education on income distribution. Chiswick (1971), with a sample of nine countries, found that the rate of growth of output and the relative dispersion of educational attainment in the labor force are positively related to income inequality while the level of economic development has ambiguous or insignificant effects on income distribution. Tinbergen (1972) also found considerable influence of both the level of education and the inequality of education on income distribution, using data for the U.S.A., Canada, and the Netherlands. An increase in years of schooling and a smaller dispersion of schooling contributed to a reduction in the degree of income inequality in his study.

Marin and Psacharopoulos (1976) found that one extra year of schooling is associated with a 10% decrease in income inequality that is measured by the variance of the log of earnings using data from the U.S., and noted the encouraging implication of educational expansion policy. However, Psacharopoulos (1977) could not come up with the same conclusion from the regressions fitted to the data for 49 countries. In his study, three factors are used as explanatory variables in determining income distribution: educational inequality, per capita income, and average level of schooling. The average level of schooling is measured by the weighted sum of enrollments at each level, and educational inequality is measured by the coefficient of variation of enrollments by school level. The educational inequality variable exhibits significant negative effects on income distribution consistently throughout various regressions. The average level of schooling has a worsening effect on income distribution when the variable enters the regression as an explanatory variable along with the educational inequality variable. It has an equalizing effect only when the educational inequality variable is omitted from the regression. The author attributes this to the high correlation between the average level of schooling and per capita income. However, we believe the cause of the conflicting coefficients for the average level of schooling is the collinearity between the average level of schooling and educational inequality. The collinearity problem arises because the variable chosen by the author as a proxy for educational inequality contains the average level of schooling in it.

Winegarden (1979) regressed the income share of the bottom 80% on the mean and variance of schooling along with many other explanatory variables, using data covering 32 countries, and concluded that higher average levels of schooling are an equalizer on income distribution, while educational inequality tends to generate income disparities to a considerable degree. However, Winegarden's method of calculating the mean and variance from the natural logarithm of the years of schooling instead of straight years was questioned by others.

On the other hand, Ram (1984), with a sample consisting of 28 countries, showed the impact of educational inequality on the income shares of the bottom 85% and 40%, contradicting previous studies. A higher level of schooling exerts mild equalizing effect as most studies have suggested, whereas a larger educational variance contributes to more equality in income distribution, contrary to previous findings. However, in Ram's study, the estimated coefficients of the educational inequality variable for both full sample and LDCs only are statistically insignificant.

Bourguignon and Morrisson (1990), using the rate of secondary education enrollments as a proxy for the schooling level with a sample of 35 developing countries, found a positive and significant effect of education on the income share of the bottom 40%. They tried to approximate educational inequality by the variance of the dichotomous variable indicating the level of individuals in the labor force, but had to omit the variable from the model because of a strong collinearity with the schooling variable.

Barro (2000) found different effects of education on income inequality, depending on the level of education used: a negative relationship between primary enrollment and income inequality, but a positive relationship between higher education enrollment and income inequality. Alderson & Nielsen (2002) found that income inequality is negatively affected by the average school years in developed countries.

Other potential determinants of income inequality have been studied. Li and Zou (2002) studied the relationship between economic freedom and income inequality while Li , *et al*. (1998) found no relationship between political freedom and income inequality. Barro (2000) see no clear relationship between democracy and income inequality. Many other variables were used in the previous studies: sector dualism and percentage of the labor force in agriculture in Nielson and Alderson (1995); female labor force participation and union density in Alderson and Nielson (2002); foreign investment dependence (Alderson & Nielsen, 1999); financial development in Dollar and Kraay (2002) and Beck, *et al*. (2004); and openness to trade and the level of tariffs in Dollar and Kraay (2002) and Milanovic and Squire (2005). Beck, *et al*. (2004) found that financial development reduces income inequality by disproportionately boosting the incomes of the poor. Milanovic and Squire found that more liberal policies were associated with increased inequality in poorer countries while decreased inequality in richer countries.

There have been some studies focusing on the relationship between globalization on income inequality. Alderson & Nielsen (2002) studied the effects of three aspects of globalization, that is, direct foreign investment, North-South trade, and migration. Heshmati (2003) found that globalization measured by the Kearney/Foreign Policy Magazine index explains only 7-11 percent of the variations in income inequality. Harjes (2007) suggested broad phenomena of globalization such as trade liberalization and technological changes may not be major drivers of inequality. Ruffin (2009) offered an explanation that globalization is likely to improve global income inequality because poor countries are likely to gain more from the trade gains due to lower living costs. Because of the inconclusive nature of much of the previous empirical studies, it is necessary to reestablish the relationship between education and income inequality in the context of an increasingly integrated and globalized world economy with more updated and expanded data.

3. Model, Data and Variables

Although there are various specifications to generate an inverse U-shaped curve, the following is a typical form used in many empirical studies:

$$YINEQ = a_0 + a_1 \ln Y + a_2 (\ln Y)^2 + u \qquad (1)$$

where YINEQ is a measure of income inequality, ln Y stands for the natural logarithm of per capita GDP (commonly used to indicate the level of economic development), and u is the error term. According to the Kuznets hypothesis, a positive sign for a_1 and a negative sign for a_2 are expected.

Many previous studies attempted to introduce additional explanatory variables along with these two income variables to explain income inequality within a cross-sectional framework. Since this study's focus is the effects of educational variables on income distribution, we specify the following equation.

$$YINEQ = b_0 + b_1 \ln Y + b_2 (\ln Y)^2 + b_3 ED + b_4 EDINEQ + u \qquad (2)$$

where ED stands for the average level of educational attainment and

EDINEQ is the relative dispersion of educational attainment. It is postulated that an increase in the level of educational attainment contributes to an improvement in income distribution. This postulation is based on a set of relationships: 1) the level of earnings is an increasing function of the level of education, which is an essence of human capital theory; and 2) The law of diminishing rates of return applies to investment in education. The diminishing returns to education have been evidenced by many studies for various countries. A comprehensive compilation on the returns to investment in education was done by Psacharopoulos (1985). We expect a direct relationship between educational inequality and income inequality in accordance with conventional human capital theory.

Other variables also affect income distribution especially in the increasingly integrated and globalized world. So, we add two important control variables to equation (2) as follows.

$$\text{YINEQ} = d_0 + d_1 \ln Y + d_2 (\ln Y)^2 + d_3 \text{ ED} + d_4 \text{ EDINEQ} + d_5 \text{ FREEDOM} + d_6 \text{ GLOBAL} + u \tag{3}$$

where FREEDOM indicates either the degree of economic freedom or economic freedom of a country and GLOBAL represents the globalization index of a country.

The proxy variable for income inequality, the dependent variable in this study is the Gini coefficient obtained from the World Income Inequality Database (WIID) by WIDER (2005). In addition, the income shares of the top 20%, and the bottom 40% of the population are used as an alternative variable. As a proxy variable for economic development, the natural log of GDP per capita from the World Bank's World Development Indicators (WDI) is used. Two educational variables, the average level of educational attainment and the relative dispersion of educational attainment, are obtained from the World Bank Education Gini index developed by Thomas, *et al*. (2003). They used educational attainment data over 15 years old population from 140 countries and set up quinquennial data from 1960 to 2000. They measured education Gini coefficients based on the years of schooling and the following equations.

$$E = [N/(N-1)] * 1/ \mu \Sigma_{i=1} \Sigma_{j=1} p_i \mid y_i - y_j \mid p_j = [N/(N-1)] * E_L$$

where E is the education Gini based on educational attainment distribution, μ is the average years of schooling for the concerned population, p_i and p_j stand for the proportion of population with certain levels of schooling, y_i and y_j are the years of schooling at different educational attainment levels, and N is the number of individuals in the concerned population. They divided the whole population of one country into mutually exclusive and collectively exhaustive seven groups – no schooling, partial primary, complete primary, partial secondary, complete secondary, partial tertiary and complete tertiary.

To measure the variable FREEDOM, two different types of freedom are used. Economic freedom is measured by the degree of freedom of individuals and businesses from government restraints on economic activities as well as legal and institutional frameworks to safeguard economic freedom. Each year since 1994, the Heritage Foundation and the Wall Street Journal publish an index of economic freedom in which they rate countries according to 50 independent variables organized into 10 broad factors of economic freedom. A key element of democracy is political freedom. Factors evaluating political freedom are how people are free in the areas of political rights and civil rights. Each year since 1978, Freedom House, a New York based nonprofit organization monitors political rights and civil liberties around the world and publishes Freedom in the World, listing countries ranked according to the degree to which these freedoms exist.

There are two indices developed to measure the degree of globalization of individual countries. One is the KOF globalization index and another is the A.T. Kearney/Foreign Policy Magazine globalization index. These indices are generally based on economic, cultural and political integration to the world and level of personal contact across national borders. Economic integration measures include movements of goods and services, convergence of domestic and international prices and inward and outward-directed foreign investment and portfolio capital flows while the measures for the level of personal contact across national borders include international travel, international phone calls, cross-border remittances, internet users and servers, and memberships in international organizations. In this study we use the KOF globalization index which is available on a yearly basis for 208 countries over the period 1970 - 2008.

4. Emprirical Results

The data for income inequality is from WIID. Even though the WIID data have been improved compared to the data available in the past, there are a number of countries that are not represented in the data and there are some variations in estimates for the same country. One way of analyzing the data is to use this unbalanced panel when there are missing observations for some variables. In our case, there are many missing observations for the Gini coefficient, the most important variable in this study. Therefore, a cross-sectional analysis is done in this study with 1990 and 2000 data instead of an unbalanced panel analysis. Missing observations of the Gini coefficient made the sample size to be drastically reduced.

Table 1 reports the regression estimation results of equation 1. As may be seen from Table 1, all the signs of the coefficients are as expected and most of them are statistically significant at the 5% level, regardless of the alternative measures used for the dependent variable.

Table 1 a. Regression of Income Inequality on Income (Year 1990)

	TOP 20%	BOTTOM 40%	GINI
Constant	-14.876	58.532	-34.451
	(33.702)	(42.672)	(39.735)
ln Y	25.004**	-9.228*	23.811**
	(8.239)	(5.114)	(8.276)
$(\ln Y)^2$	-2.056**	.788**	-2.901**
	(.634)	(.367)	(1.068)
N	64	64	64
Adj. R^2	.295	.212	.226

b. Regression of Income Inequality on Income (Year 2000)

	TOP 20%	BOTTOM 40%	GINI
Constant	-10.765 (27.362)	35.654 (24.781)	-31.651 (42.126)
ln Y	19.495** (8.160)	-13.145** (6.172)	25.558** (11.087)
(ln Y)2	-1.707** (.760)	1.091** (.413)	-2.101** (.879)
N	78	78	78
Adj. R2	.327	.234	.244

Note: The first entry for each predictor is the coefficient estimate, and the second in parentheses is the standard error of the coefficient estimate. * indicates significance at the 10% level and ** at the 5% level.

An inverse U-shaped curve relationship is observed for TOP20 and the Gini coefficient, while a U-shaped curve relationship is observed for BOTTOM40. This observation is held with both 1990 and 2000 data. The explanatory power of the model measured by the adjusted R^2 statistic is relatively small as in previous studies. This is partly due to the nature of cross-section data. This could also indicate the possibility of some important explanatory variables missing from the model specification.

Table 2 presents the estimation results of equation 2 which has two additional variables representing the level and the inequality of education as explanatory variables. We used the average years of schooling of the labor force (ED hereafter) as a proxy for the level of educational attainment. We used standard deviation of years of schooling as a proxy variable for the relative dispersion of educational attainment (EDINEQ hereafter). The addition of the two educational variables, ED and EDINEQ, changes the picture. First, there is an improvement in the explanatory power of the model as can be seen from the adjusted R^2 statistic. Second, the coefficients of the income variables, both ln Y and $(\ln Y)^2$, decreased considerably and lost their significance. On the other hand, both educational variables show expected and significant impacts on income distribution. ED has a negative effect on TOP20 and GINI and a positive effect on BOTTOM40, both of a considerable magnitude, while EDINEQ has a positive effect on TOP20

Table 2 a. Regression of Income Inequality on Income and Education Variables (Year 1990)

	TOP 20%	BOTTOM 40%	GINI
Constant	4.258 (27.702)	38.245 (22.461)	16.057 (31.475)
ln Y	12.775 (9.592)	-3.483 (4.241)	7.265 (11.217)
$(\ln Y)^2$	-1.201 (.786)	.368 (.529)	-.743 (1.408)
ED	-1.628** (.544)	.937** (.349)	-2.059** (.702)
EDINEQ	2.875** (1.238)	-2.443** (.757)	3.792** (1.364)
N	64	64	64
Adj. R^2	.412	.343	.372

b. Regression of Income Inequality on Income and Education Variables (Year 2000)

	TOP 20%	BOTTOM 40%	GINI
Constant	5.316 (22.372)	29.372 (31.223)	24.451 (42.735)
ln Y	14.504 (10.893)	-4.814 (3.667)	5.181 (9.726)
$(\ln Y)^2$	-1.067 (.756)	.529 (.432)	-1.293 (1.716)
ED	-1.426 (.614)	1.148 (.402)	-2.429 (.668)
EDINEQ	2.271 (.976)	-1.922 (.611)	2.864 (.821)
N	78	78	78
Adj. R^2	.407	.337	.389

Note: The first entry for each predictor is the coefficient estimate, and the second in parentheses is the standard error of the coefficient estimate. * indicates significance at the 10% level and ** at the 5% level.

and GINI and a negative effect on BOTTOM40, again both of a considerable magnitude. Both variables are statistically significant at the 1% level in most of the cases, and at least at the 5% level.

When we use the coefficient of variation of years of schooling in place of standard deviation of years of schooling to represent EDINEQ, we obtain disappointing results. Not only do the coefficients of the educational inequality variable turn out to be insignificant, but they also show the opposite sign, contrary to conventional human capital theory. One reason for poor performance of EDINEQ is a collinearity problem between the variables, ED and EDINEQ. As the coefficient of variation of years of schooling is calculated as standard deviation of years of

Table 3 a. Regression of Income Inequality on Income, Education and Globalization Variables (Year 1990)

	TOP 20%	BOTTOM 40%	GINI
Constant	-11.726 (22.113)	25.632 (31.264)	-27.164 (21.543)
ln Y	10.542 (7.257)	-4.894 (3.463)	8.388 (5.392)
$(\ln Y)^2$	-.918 (.633)	.527 (.634)	-1.332 (.945)
ED	-1.541** (.473)	1.036* (.515)	2.168** (.784)
EDINEQ	2.138** (.966)	-2.954** (1.332)	1.891** (.778)
ln ECONOMIC FREEDOM INDEX	1.945* (1.107)	-3.225 (2.864)	1.426 (1.549)
POLITICAL FREEDOM RATING	.124 (.097)	-.084 (.127)	-.215 (.152)
ln GLOBALIZATION INDEX	3.472** (1.560)	-2.992* (1.541)	4.661** (1.936)
N	57	57	57
Adj. R^2	.442	.391	.425

Note: The first entry for each predictor is the coefficient estimate, and the second in parentheses is the standard error of the coefficient estimate. * indicates significance at the 10% level and ** at the 5% level.

b. Regression of Income Inequality on Income, Education and Globalization Variables (Year 2000)

	TOP 20%	BOTTOM 40%	GINI
Constant	-13.168 (-20.839)	24.683* (13.237)	-35.883 (27.458)
ln Y	11.836 (8.229)	-4.179 (3.337)	12.908 (7.338)
(ln Y)2	-1.031 (.7019)	.615 (.507)	-1.695 (1.084)
ED	-2.174** (.664)	.887** (.402)	1.745* (.976)
EDINEQ	3.149** (1.096)	-1.973** (.786)	2.564** (1.027)
ln ECONOMIC FREEDOM INDEX	2.184 (1.437)	-3.105* (1.702)	1.883 (1.994)
POLITICAL FREEDOM RATING	.047 (.139)	-.081 (.116)	-.114 (.246)
ln GLOBALIZATION INDEX	5.337** (2.455)	-2.847* (1.592)	6.152** (2.368)
N	69	69	69
Adj. R2	.472	.419	.403

Note: The first entry for each predictor is the coefficient estimate, and the second in parentheses is the standard error of the coefficient estimate. * indicates significance at the 10% level and ** at the 5% level.

schooling divided by ED, there is a high negative correlation between ED and EDINEQ (r=-.87). So, the estimation results with the coefficient of variation of years of schooling as EDINEQ is not reported here. In regard to the effect of the level of education on income distribution, our results are similar to most of the previous studies, but contrary to Psacharopoulos (1977). Ram (1984) had a limited sample size and did not get strong results to dispute the dis-equalizing effect of the schooling inequality on income distribution.

Table 3 presents the estimation results of equation 3 which has two more additional control variables representing the degree of freedom and the degree of globalization of individual countries. We used both economic freedom index by the Heritage Foundation and political freedom index by the Freedom House. To measure the effect of globalization, we used the KOF globalization index. There is a moderate improvement in the explanatory power of the model. The significance of the two educational variables has unchanged while the two income variables exhibit insignificant coefficients, though they show expected signs.

Economic freedom is positively associated with income inequality, though not significant. No meaningful relationship between political freedom and income inequality can be obtained. The KOF globalization index is positively associated with income inequality and explains the variations in income inequality significantly, supporting the great U-turn suggested by Alderson and Nielson (2002). The longitudinal trend toward increasing inequality can be partly explained by the globalization variable. The impact of globalization may come through a widening of wage differentials in favor of skilled and highly educated individuals. Furthermore the technical changes from globalization tend to be biased against unskilled workers. However, the globalization index is composed of many different factors of globalization. To find out which factors are major determinants, breakdown of globalization components would be needed.

5. Conclusions

In this paper, we have tried to analyze the effects of education on income inequality using cross-section data in the context of globalization. First, we tested the Kuznets inverse U-hypothesis. We confirmed the existence of the inverted-U curve when only the income variables, that is, $\ln Y$ and $(\ln Y)^2$, are used as explanatory variables. However, the coefficients of the two income variables, though having the expected signs, become smaller and statistically insignificant as the two additional variables for the level of education and the dispersion of educational attainment are added. Thus, the inverse U-shaped curve lacks robustness.

Second, we have tried to capture the effects of education on income distribution. The econometric results from a cross-section analysis support our postulation that a higher level of educational attainment of the labor force has an equalizing effect on income distribution. Our estimates also suggest that the larger the dispersion of schooling among the labor force, the greater the income inequality. This study shows that the dispersion of educational attainment has a much stronger dis-equalizing effect on income equality than previous studies have suggested.

Third, we also examined the effects of globalization and freedom on income distribution. The results show that the higher the level of globalization is, the more unequal is income distribution. On the other hand, freedom, whether economic freedom or political freedom, has marginal effects on income inequality.

With intrinsic limitations of this study imposed by the availability and reliability of the data, the findings in this study are tenuous and may need to be scrutinized with further research. Nonetheless, we obtained very clear and strong empirical results from the largest sample of cross-country data that we are aware of. Some development policy implications can be drawn from this study. If a developing country is committed to achieving an egalitarian society, it should emphasize educational policies that aim for educational expansion along with equal access to education in addition to its various direct redistributive policies. Emphasis on equity may not necessarily conflict with the objective of economic growth. While redistributive policies tend to be detrimental to economic growth, educational expansion and equal access to education have been identified as major factors contributing to economic growth by many studies. Growth with equity has been demonstrated by the experiences of Japan, Taiwan, and South Korea just to name a few. All three countries have emphasized educational expansion and equal access to education in their economic development process.

References

Adelman, I. and C. Morris*, Economic Growth and Social Equity in Developing Countries* (Stanford: Stanford University Press, 1973).

Ahluwalia, M., "Income Inequality, Some Dimensions of the Problem," in: H. Chenery *et al*. (eds.), *Redistribution with Growth* (New York: Oxford University Press, 1974), 3-37.

Ahluwalia, M., "Income Distribution and Development: Some Stylized Facts," *American Economic Review* 66 (1976a), 128-135.

Ahluwalia, M., "Inequality, Poverty and Development," *Journal of Development Economics* 3 (1976b), 307-342.

Alderson, A. and F. Nielsen, "Income Inequality, Development, and Dependence: A Reconsideration," *American Sociological Review* 64 (1999), 606-631.

Alderson, A.and F. Nielsen, "Globalization and the Great U-turn: Income Inequality Trends in 16 OECD Countries," *American Journal of Sociology*, 107 (2002). 1244-1299.

Anand, S. and S. M. R. Kanbur, "Inequality and Development: A Reconsideration," in: H. P. Nissen (ed.), *Towards Income Distribution Policies*, Book Series 3 (Tilbury: EADI, 1984), 131-167.

Barro, R., "Inequality and Growth in a Panel of Countries," *Journal of Economic Growth* 5 (2000), 5-32.

Beck, T., Demirgue-Kunt, A, and R. Levine, Finance, *Inequality and Poverty*: Cross-country Evidence. World Bank Discussion Paper (Washington D.C.: World Bank,2004)

Becker, G., *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education* (New York: Columbia Univ. Press, 1964).

Bourguignon, F. and C. Morris, "Income Distribution, Development and Foreign Trade: A Cross-sectional Analysis," *European Economic Review* 34 (1990), 1113-1132.

Bourguignon, F. and C. Morris, The Size Distribution of Income Among World Citizens," American Economic Review (2002)

Braun, D., "Multiple Measurements of U.S. Income Inequality," *Review of Economics* *and Statistics* 60 (1988), 398-405.

Checchi, D., *Does Educational Achievement help to explain Income Inequality?* Working Paper No. 208 (Helsinki: United Nations University, World Institute for Development Economics Research, 2000).

Chenery, H. and M. Syrquin, *Patterns of Development 1950-1970* (Oxford: Oxford University Press, 1975).

Chiswick, B. R., "Earnings Inequality and Economic Development," *Quarterly Journal of Economics* 85 (1971), 21-39.

Cline, W., "Distribution and Development: A Survey," *Journal of Development Economics* 1 (1975), 359-400.

Cromwell, J., "The Size Distribution of Income: An International Comparison," *Review of Income and Wealth* 23 (1977), 291-308.

De Gregorio, J. and J. Lee, “Education and Income Inequality: New Evidence from Cross Country Data,” *The Review of Income and Wealth* 48 (2002), 395-416.

Dollar, D. and A. Kraay, “Growth is Good for the Poor. *Journal of Economic Growth* 7 (2002), 195-225.

Fields, G., "A Welfare Economic Analysis of Growth and Distribution in the Dual Economy, *Quarterly Journal of Economics* 93 (1979), 325-353.

Galbraith, J. K., Conceicao, P., and H. Kum, *Inequality and Growth Reconsidered Once Again: Some New Evidence from Old Data.* UTIP Working Paper No. 17 (Austin, TX: the University of Texas at Austin, 2000).

Heshmati, A., *The Relationship between Income Inequality and Globalization*. WIDER Working Paper (Helsinki: The United Nations University, 2003)

Harjes, T., *Globalization and Income Inequality: A European Perspective*. IMF Working Paper (Washington, D.C.: IMF, 2007)

Kravis, I.G., "International Differences in the Distribution of Income," *Review of Economics and Statistics* 42 (1960), 408-416.

Kuznets, S., "Economic Growth and Income Inequality," *American Economic Review* 45 (1955), 1-28.

Kuznets, S., "Quantitative Aspects of the Economic Growth of Nations VIII: Distribution of Income by Size," *Economic Development and Cultural Change* 11 (1963), 1-80.

Li, H., Squire, L., and H. Zou, “Explaining International and Intertemporal Variations in Income Inequality,” *Economic Journal*, 108 (1998), 26-43.

Li, H., and H. Zou, “Inflation, Growth, and Income Distribution: A Cross-country Study,” *Annals of Economics and Finance* 3 (2002) , 85-101.

Loehr, W., "Economic Growth, Distribution and Incomes of the Poor," *Journal of Economic Studies* 7 (1980), 127-139.

Marin, A. and G. Psacharopoulos, "Schooling and Income Distribution," *Review of Economics and Statistics* 58 (1976), 332-338.

Milanovic, B., and L. Squire, *Does Tariff Liberalization Increase Wage Inequality? Some Empirical Evidence.* World Bank Policy Research Working Paper No. 3571 (Washington DC: World Bank, 2005).

Nielsen, F., and A. Alderson, “Income Inequality, Development, and Dualism: Results from an Unbalanced Cross-national Panel. *American Sociological Review* 60 (1995), 674-701.

Papanek, G., "Economic Growth, Income Distribution and the Political Process in LDCs," in: Z. Griliches, et. al. (eds.), *Income Distribution and Economic Inequality* (Frankfurt: Campus Verlag-Halstead Press, 1978), 259-273.

Papanek, G. and O. Kyn, "The Effect on Income Distribution of Development, the Growth Rate and Economic Strategy," *Journal of Development Economics* 23 (1986), 55-65.

Paukert, F., "Income Distribution at Different Levels of Development," *International Labor Review* 108 (1973), 97-125.

Psacharopoulos, G., "Unequal Access to Education and Income Distribution," *De Economist* 125 (1977), 383-392.

Psacharopoulos, G., "Returns to Education: A Further International Update and Implications," *Journal of Human Resources* 20 (1985), 581-604.

Psacharopoulos, G. and A. Arrigada, "The Educational Attainment of the Labor Force: An International Comparison," *World Bank Discussion Paper (*Washington, D.C., 1986).

Psacharopoulos, G. and J.B.G. Tilak, "Schooling and Equity," in: G. Psacharopoulos (ed.), *Essays on Poverty, Equity and Growth* (Oxford: Pergamon Press, 1991), 53-78.

Ram, R., "Population Increase, Economic Growth, Educational Inequality and Income Distribution," *Journal of Development Economics* 14 (1984), 419-428.

Ram, R., "Economic Development and Income Inequality: Further Evidence on the U-curve Hypothesis," *World Development* 16 (1988), 1371-1376.

Ram, R., "Educational Expansion and Schooling Inequality: International Evidence and Some Implications," *Review of Economics and Statistics* 62 (1990), 266-274.

Ram, R., "Kuznets's Inverted-U Hypothesis: Evidence from a Highly Developed Country," *Southern Economic Journal* 57 (1991), 1112-1123.

Ravallion, M., *Competing Concepts of Inequality in the Globalization Debate*. World Bank Policy Research Working Paper No. 3038 (Washington DC: World Bank, 2004).

Robinson, S., "A Note on the U-hypothesis Relating Income Inequality and Economic Development," *American Economic Review* 66 (1976), 437-440.

Rodgers, G., "Population Growth, Inequality and Poverty," *International Labor Review* 122 (1983), 443-460.

Ruffin, R., “Globalization and Income Inequality,” *Trade and Development Review* 2 (2009), 56-69.

Saith, A., "Development and Distribution: A Critique of the Cross-country U-hypothesis, *Journal of Development Economics* 13 (1983), 367-382.

Stewart, F., "Inequality, Technology and Payment System," *World Development 6* (1978), 275-293.

Thomas, V., Wang, Y., and X. Fan, "Measuring Education Inequality: Gini Coefficients of Education for 140 countries, 1960-2000", *Journal of Education Planning and Administration* 17 (2003), 5-33.

Tilak, J.B.G, "Rates of Return to Education and Income Distribution," *De Economist* 137 (1989), 454-465.

Tinbergen, Jan, "The Impact of Education on Income Distribution," *Review of Income and Wealth* 19 (1972), 255-265.

Tinbergen, Jan, *Income Distribution: Analysis and Policies* (Amsterdam: North-Holland, 1975).

Tsakloglou, P., "Development and Inequality Revisited," *Applied Economics* 20 (1988), 509-531.

Wells, R., “Education's Effect on Income Inequality: an Economic Globalisation Perspective,” *Globalisation, Societies and Education* 4 (2006), 371 – 391.

Winegarden, C. R., "Schooling and Income Distribution: Evidence from International Data," *Economica* 46 (1979), 83-87.

World Institute for Development Economics Research (WIDER), *World Income Inequality Database* (United Nations University, 2005).

Wright, C. L., "Income Inequality and Economic Growth: Examining the Evidence, *Journal of Developing Areas* 13 (1978), 49-66.

Yitzhiki, S. “Economic Distance and Overlapping of Distributions,” *Journal of Econometrics* 61 (1994), 147-159.

CHAPTER 4

Should I Marry Again?

by
*Jinyoung Kim**
(Korea University)

Abstract

In the last several decades, the prevalence of remarriages, coupled with fertility in remarriage, has resulted in increasing complexities in family structure. Studies have considered child rearing, division of labor, sharing consumption among others as gains from marriage, regardless of first or subsequent marriages. This paper recognizes additional benefit of marriage, unique in subsequent marriages. If (1) the quality of children is uncertain when parents invest in children's human capital, and (2) the quality distributions of two step siblings of the same biological mother (or father) are less correlated than two children from an intact family, due to genetic inheritance, a parent has an incentive to hedge risk in terms of children's qualities, and diversify the parent's portfolio of children by having children in remarriage. We show that a child born after parents' remarriage in a reconstituted or blended family receives less human capital investment than a child in an intact family as long as parents have a motive for precautionary saving. The paper

* I thank Gary Becker, Marco Caliendo, Michael Greenstone, David Jaeger, Kamhon Kan, Douglas Krupka, Derek Neal, Solomon Polachek, David Ribar, Mark Rosenzweig, Bruce Weinberg, and Dennis Yang for helpful comments for the earlier version of this paper. All errors are exclusively the responsibility of the author.
Comments welcome at Jinyoung Kim, Department of Economics, Korea University, Anam-dong, Seongbuk-gu, Seoul, Korea 136-701, jinykim@korea.ac.kr.

theoretically identifies and empirically tests socio-economic variables that increase this benefit of remarriage and describes who will be more likely to remarry.

1. Introduction

The second half of the twentieth century has witnessed marked changes in living arrangements and family composition, characterized by falling marriage rates and rising divorce rates. As divorce rates have climbed up during the period, remarriages, especially those following divorces, have also become prevalent.[1]

According to data from 1995 cycle of the National Survey of Family Growth (NSFG), 54 percent of divorced women remarried within 5 years and 75 percent of divorced women remarried within 10 years (Bramlett and Mosher, 2002). In 2001, 17.7 percent of all American wives were in their second marriage and 3.6 percent of wives were in their third or later marriage (Kreider, 2005). Younger cohorts show higher tendency of remarriage: 18.7 percent of women born between 1935 and 1939 were married two times or more by age 45 while the corresponding percent for women born between 1950 and 1954 is 26.6 percent (SIPP, U.S. Census Bureau, 2001). Along with increasing remarriages, a noticeable change is that more children are born to parents in remarriage. Wineberg (1990) reports that more than 50 percent of remarried women gave birth after their second marriage.

The prevalence of remarriages following divorces, coupled with fertility in remarriages, has resulted in increasing complexities in family structure. The NSFG data indicate that 38.2 percent of American women who were 15-44 years of age in 1995 and who had more than one child had children by at least two different fathers. This percentage was reported as 56.9 percent for blacks and 31.1 percent for whites. Among

---

1 Until recently, the majority of remarriages followed the death of a spouse rather than a divorce. From colonial times to as late as the 1920s, more remarriages followed widowhood than divorce. By 1987 only 9% of men and women remarried due to the death of a spouse (Cherlin, 1992).

those women in 38.2 percent, 50.6 percent of them were always married at the time of childbirths with different partners, 44.3 percent were always unmarried, and the rest 5.1 percent were married for some childbirths with different fathers and unmarried for the others childbirths. The data also show that among those children who have at least one sibling on the maternal side, only in 54.6 percent do all siblings have the same father.

Previous theoretical studies on remarriage do not typically make any distinction between first and subsequent marriages in terms of gains from marriage. The economics literature views marriage as a partnership for joint production and joint consumption. The production and rearing of children is the most commonly recognized role of marriage, but other functions of marriage include division of labor, sharing collective goods, and risk pooling (Weiss, 1993). Marriage will be resolved when the productivity in joint production or consumption of a spouse is materialized to be less desirable, and partner search in the remarriage market can ensue following divorce. In this paper we recognize additional benefit of marriage (or partnership) unique in subsequent marriages: diversification in the quality (or outcome of human capital investment) of children.

This benefit is predicated on two premises. First, the quality of children as an outcome of human capital investment in them is uncertain when parents make the investment decision. Traditional models, static or dynamic, in the literature since the seminal paper of Becker (1960) have assumed for simplicity's sake that parents know *a priori* what will be the outcome of human capital investment in children, i.e. the quality of children, when decisions regarding fertility and human capital investment are made.

However, most dimensions of child quality are unknown at the point at which parents make a decision regarding childbirth or human capital investment, and are subject to a variety of determinants which cannot be predicted perfectly when the investment decision is made. For instance, parents spend resources to ensure that young children acquire skills needed afterwards in the labor market, but children's labor market performances are not solely influenced by the parents' investment, but other determinants, such as unobservable abilities and the market price of the acquired skills when the children enter the labor market, also play a role. More often than not, parents can observe how successful their children are in the labor market only a long time after they have committed to the decision to invest in the human capital of their children.

One immediate implication of this premise is that parents maximize the *expected* utilities that are determined by children's quality, which will be realized in the future. In our model, we assume that parents do not know *a priori* their children's quality, but have knowledge of the distribution of a child's possible quality in the future, which pertains to the concept of risk in the Knightian sense (Knight, 1921).

The second premise of the paper is that children of a parent will not be homogenous in terms of their quality, but their quality distributions should correlate with each other to varying degrees. Children differ in quality as the result of the different amounts of parental resources invested, but also as the result of differences in the abilities that are genetically inherited from either of their parents, which implies that each child should have a different distribution of expected quality. The quality distributions of two children, however, may not be independent. If two children share the same biological mother and father, their distributions may be highly correlated, because they have inherited genes from the same pool. Conversely, the correlation between two step siblings of the same mother or father may be lower.

If parents are risk-averse, and their utilities are influenced positively by the qualities of the children born (due to either altruism or financial support from children), a straightforward consequence of these two premises is that a parent has an incentive to hedge risk in terms of children's qualities, and diversify the parent's portfolio of children. In particular, a parent reduces the risk in child quality by having children in remarriage with a different partner.[2] Demand for gender balance in children may also be due to the same motive of quality diversification, rather than due to a difference in preference or cost between sons and daughters.[3]

---

2 If children are considered as assets that yield parents' future utilities, this argument is closely related to the notion of asset diversification in financial portfolio theory.

3 For evidence concerning diversification behaviors in the context of reproduction in nature and in human society, see Appendix A.

This idea brings forth an interesting implication with respect to human capital investment in children. If there is a precautionary motive for saving, a child born in a reconstituted or blended family after parents' remarriage will receive less human capital investment than a child in an intact family because children in the latter type of family will have higher variance in the outcomes of human capital investment due to higher correlation in quality among siblings. Parents in an intact family thus invest more in children as a precautionary measure. Earlier theoretical and empirical studies in the literature have been mostly concerned with welfare or performances of stepchildren after their mothers remarry stepfathers because their biological fathers may not be willing to transfer money to the custodial mothers (Chiappori and Weiss, 2007).[4] In contrast to this *Cinderella effect*, our model implies that a child born in remarriage will receive less investment even by his/her biological parent and thus show lower performance.

We formalize our idea along the lines of Chiappori and Weiss (2006, 2007). These authors construct a model of marriage in a general equilibrium setup where resource allocation to each member of a family is determined by strategic behavior of partners. To highlight our main idea, our model simplifies their framework considerably in several aspects but extends it in others. We do away with the general equilibrium effect through marriage market interactions in that the pool of suitable matches in the remarriage market will be smaller if more marriages are intact. We also ignore post-divorce transfer to children that is determined by bargaining between ex-partners. On the other hand, we explicitly model human capital investment in children where the outcome of this investment is uncertain. Moreover, all adults are assumed to be identical ex ante with respect to income in each period that is a random variable and drawn from a non-negative distribution different for each period. In contrast to the model in Chiappori and Weiss, our model allows childbearing in remarriage.

4 There are a large number of empirical studies on performances of children in stepfamilies after their parents' divorces. See, for example, Hetherington and Stanley-Hagen (1999), Lamb et al. (1999), Painter and Levine (2000), Lang and Zagorsky (2001), Gruber (2004), Page and Stevens (2004), Francesconi et al. (2010).

This paper is organized as follows. Section 2 lays out a formal model of marriage and human capital investment in an environment where the quality of children is uncertain and heterogeneous across siblings. We identify in this section socio-economic conditions that increase this benefit of remarriage and thus describes who will be more likely to remarry. Section 3 describes the data and explains our strategies for empirical investigation. We report our empirical findings in Section 4. Finally, Section 5 presents the conclusions of this paper.

2. The Model

Each adult lives through three periods in our model and can be married in the first two periods. At the beginning of the first period each adult is matched with one potential partner of opposite sex who is randomly selected in the pool of eligible mates. The quality of match between the two matched partners is a random variable and drawn from a time-invariant distribution that is known a priori. After observing the match quality, the adult decides whether to marry or not. If married in the first period, he/she makes a two-step decision in the second period: first, whether to stay married with the same partner or divorce, and then, if divorced, whether to marry a new partner matched in the remarriage market or not. To keep things simple, we assume that the distribution of the match quality in the *remarriage* market is also time-invariant and is not affected by how many divorcees enter this market. As in the first period, the decision of remarriage is made after the quality of match is revealed. A single in the first period repeats the marriage decision in the second period.

To simplify the analysis substantially, we assume that fertility is exogenous and a married couple has one child per period.[5] Parents jointly decide on the amount of human capital investment for each child and the investment cost is borne only in a single period. This implies that in the next period after birth there is no additional cost involved

5 Effectively, marriage means in this model any type of partnership with propagation of children, including legal union and cohabitation.

with a child regardless of whether parents stay married or divorce. In the third period, all adults are assumed to be single and bear no child. Therefore, the maximum number of children a parent can have is two.

All individuals, males and females, are assumed to be homogenous ex ante in every dimension. Income in period t, y_t (t = 1, 2, 3), is independently drawn from a known distribution at the beginning of each period. We leave the analysis of resource allocation between partners in marriage or after divorce due to systematic differences by gender in income and custody right to earlier studies such as Aiyagari et al. (2000) and Chiappori and Weiss (2006, 2007).

In our model, the gains from marriage come from two sources. First, consumption in the household is a public good. A couple can pool their incomes and enjoy a greater level of consumption, achieving the scale economies within a household. Next, marriage provides non-pecuniary benefits from love or companionship which are uncertain at the time of search in the marriage market and revealed only after two potential partners are matched. Each married couple derives utility from companionship which is common to both partners and denoted by θ. We now describe the optimization problem for a representative adult by backward induction, starting with the period-2 problem and then going back to period 1. Since all adults are assumed to be identical ex ante, we illustrate our model hereafter from the perspective of a representative woman because only women are surveyed in most fertility data bases such as ours, which allows us only to empirically investigate *women*'s decisions on remarriage and human capital investment.[6]

2. 1. Married in Period 1

A woman who was married in period 1 makes decisions in two steps in period 2. First, she decides whether to stay married with the same spouse or divorce. Next, if she divorces, she decides whether to marry a

6 In fact, our model is squarely applicable to men, too. Unlike women, uncertainty in child quality can arise from one more source for men. A man is never 100 percent certain that a newborn child is a biological relative. This fact should lead to an added incentive for men to diversify. (See Cox, 2003, for discussion of paternity uncertainty.)

new partner matched in the remarriage market or become single.

Before the decision in the first step is made, the quality of match with the present spouse is revealed to be deteriorated or not. Deterioration of the match quality is stochastic with a constant probability q and the quality is lowered to $\theta_1 - \delta$ where θ_1 was drawn in the previous period. With the probability of $(1 - q)$, the match quality with her spouse remains θ_1. Before the second-step decision of remarriage, she observes the quality of match with a new mate.

We consider first the second-step decision on whether to remarry or not. The expected lifetime utility from remarriage is

$$\underset{h_{2s}}{Max} \quad u(2y_2 - v_2 h_{2s}) + \theta_2 - \lambda + E[W(y_3, H_f + H_s)], \tag{2.1}$$

where u and W denote the intertemporally-additive utilities in period 2 and 3, respectively, and E is an expectation operator. Note we are ignoring discounting for simplicity. The control variable, h_{2s}, represents human capital investment in a child where subscript 2 denotes time period and subscript s implies her second child. v_2 represents the net cost of one unit of human capital investment borne in the second period. The quality of match in period 2, θ_2, is an independent draw from a time-invariant distribution with the support in $(-\infty,\infty)$. We assume that divorce involves only non-pecuniary cost, λ. H_f and H_s are human capital levels ("quality") of the first and the second child, respectively, materialized in period 3. For simplicity's sake, we assume that the parent's utility in the last period depends on the sum of the human capital levels of all children as well as own consumption.

Note that consumption in period 2 for the married woman equals to the pooled income $2y_2$, net of the cost of human capital investment, because all adult members of the same household share consumption. In period 3, the woman as a single obtains utility from private consumption (which is equal to y_3) and the human capital levels of all children. The utility function W is further assumed to take the following form:

$$W(y_3, H_f + H_s) = w(y_3 + H_f/\alpha + H_s/\alpha), \tag{2.2}$$

in which α is a positive constant. If the source of utility for parents is financial support from children and the wage rate of human capital is normalized to 1, $(1/\alpha)$ in this form can be interpreted as a constant fraction of the earnings of a child that each parent receives. The functions u and w are assumed to satisfy the usual conditions: $u'(c)>0$, $u''(c)<0$, $w'(c)>0$, and $w''(c)<0$.

The human capital (or earning capacity in the labor market) of a child is accumulated through a production technology linking the parents' investment in the child's education with the child's human capital:

$$H_i = f(h_{ti}) + \varepsilon_i, \text{ where } i = f \text{ (first child) or } s \text{ (second child), and } t = 1 \text{ or } 2. \quad (2.3)$$

In equation (2.3), h_{ti} denotes parents' investment in child i's human capital in period t. The human capital production function, f, is a continuously increasing and concave function of h: $f'(h) > 0$ and $f''(h) < 0$.

Equation (2.3) demonstrates that human capital is not determined exclusively by the parents' investment, but is also influenced by the stochastic component, ε, which is revealed only in the old-age period after the human capital investment had already been committed. This component reflects uncertainty in returns to the parents' investment in the children's human capital or earning capacity in the labor market, which may possibly be the consequence of several factors. First, the spouse's influences on children's earning capacity in adulthood through genetic inheritance cannot be known with any degree of certainty. Second, a parent is not able to thoroughly observe the spouse's effort in the formation of children's human capital. Finally, earning capacity in the labor market can be influenced by fluctuations in production possibilities and the prices of goods and factors of production that are generally revealed only after children have received their education and much of their other training, and have entered the labor market (Becker, 1991). This "risk" in returns to human capital investment in children is reflected by the stochastic term, ε, the distribution of which is known a priori to the parents. Similar to Becker (1991), we assume that this stochastic term is additive.[7] We expect this stochastic term should vary across children.

If a divorced woman does not remarry in period 2, her expected utility is $u(y_2) - \lambda + E[w(y_3 + H_f/\alpha)]$. Note that she still receives utility from her only child from the marriage in period 1. She decides to remarry in period 2 if her θ_2 is greater than a threshold value, θ_2^*, where θ_2^* satisfies:

$$\underset{h_{2s}}{Max}\ u(2y_2 - v_2 h_{2s}) + \theta_2^* + E[w(y_3 + H_f/\alpha + H_s/\alpha)] = \qquad (2.4)$$
$$u(y_2) + E[w(y_3 + H_f/\alpha)].$$

Now consider the first-step decision on divorce. In the following analysis we make an assumption for convenience that the probability of deterioration in match quality is one so that all couples experience the quality of match deteriorated at the beginning of period 2. Relaxing this assumption does not alter the model's qualitative implications. The expected lifetime utility when she stays married with the same spouse is

$$\underset{h_{2s}}{Max}\quad u(2y_2 - v_2 h_{2s}) + \theta_1 - \delta + E[w(y_3 + H_f/\alpha + H_s/\alpha)], \qquad (2.5)$$

where h_{2s} denotes human capital investment in the second child born in period 2.

The expected lifetime utility when she divorces is

$$\Pr(\theta_2 \geq \theta_2^*)\ \{\underset{h_{2s}}{Max}\ u(2y_2 - v_2 h_{2s}) + E[\theta_2|\theta_2 \geq \theta_2^*] - \lambda \qquad (2.6)$$
$$+ E[w(y_3 + H_f/\alpha + H_s/\alpha)]\} + \Pr(\theta_2 < \theta_2^*)\ \{u(y_2) - \lambda +$$
$$E[w(y_3 + H_f/\alpha)]\}.$$

The first term with braces is the expected utility in remarriage before

7 Becker (1991) considers two additive stochastic terms in determination of children's income which are called "endowed luck" and "market luck". The stochastic component ε in our model does not distinguish these two terms.

she enters the remarriage market, and the corresponding probability is $Pr(\theta_2 \geq \theta_2^*)$. The second term with braces is the expected utility when she becomes single. We denote by $\theta_1^\dagger$ the threshold value of θ_1 that equalizes the two expected utilities in equations (2.5) and (2.6). If she draws θ_1 in period 1 which is greater than $\theta_1^\dagger$, she will stay married with the same spouse in period 2 since the match quality with the spouse is high enough even after its deterioration.

2. 2. Singles in Period 1

A single woman in period 1 is matched with a new partner in period 2 and decides to marry or not after observing the quality of match. The expected lifetime utility from marriage in this period is

$$\underset{h_{2f}}{Max} \quad u(2y_2 - v_2\, h_{2f}) + \theta_2 + E[w(y_3 + H_f/\alpha)], \tag{2.7}$$

She seeks an optimal value of human capital investment in her first child (h_{2f}) to maximize this utility. The net cost of human capital investment in this case is v_2 incurred in period 2.

If a single woman remains single in period 2, her expected utility is $u(y_2) + E[w(y_3)]$. Therefore, she decides to get married in period 2 if her θ_2 is greater than a threshold value, θ_2^{**}, where θ_2^{**} satisfies:

$$\underset{h_{2f}}{Max}\, u(2y_2 - v_2\, h_{2f}) + \theta_2^{**} + E[w(y_3 + H_f/\alpha)] \tag{2.8}$$
$$= u(y_2) + E[w(y_3)].$$

2. 3. Decision in Period 1

After the quality of match with a partner paired in the marriage market is known, a woman decides in period 1 whether to marry the partner or not. The expected lifetime utility when married in period 1 is

$$\underset{h_{1f}}{Max}\; u(2y_1 - v_1\, h_{1f}) + \theta_1 \tag{2.9}$$

$$+ \{\underset{h_{2s}}{Max}\ u(2y_2 - v_2 h_{2s}) + \theta_1 - \delta + E[w(y_3 + H_f/\alpha + H_s/\alpha)]\},$$

$$\text{if } \theta_1 \geq \theta_1^{\dagger},$$

$$\text{or, } \underset{h_{1f}}{Max}\ u(2y_1 - v_1 h_{1f}) + \theta_1$$

$$+ \Pr(\theta_2 \geq \theta_2^{*})\{\underset{h_{2s}}{Max}\ u(2y_2 - v_2 h_{2s}) + E[\theta_2|\theta_2 \geq \theta_2^{*}] - \lambda$$

$$+ E[w(y_3 + H_f/\alpha + H_s/\alpha)]\} + \Pr(\theta_2 < \theta_2^{*})\ \{u(y_2) - \lambda$$

$$+ E[w(y_3 + H_f/\alpha)]\}, \quad \text{otherwise.}$$

The first of the two equations in (2.9) pertains to the case when the woman stays married with the same partner in period 2 since the match quality turns to be high enough. The second refers to the case of divorce. If the woman does not get married in period 1, she gets

$$u(y_1) + \Pr(\theta_2 \geq \theta_2^{**})\{\underset{h_{2f}}{Max}\ u(2y_2 - v_2 h_{2f}) + E[\theta_2|\theta_2 \geq \theta_2^{**}] + E[w(y_3 + H_f/\alpha)]\} + \Pr(\theta_2 < \theta_2^{**})\{u(y_2) + E[w(y_3)]\} \quad (2.10)$$

The marriage decision in period 1 boils down to whether the value of the match quality (θ_1) is greater than some threshold value. If this threshold value is bigger than $\theta_1^{\dagger}$, all married adults will stay married with the same partner in the next period, and no divorce and thus no remarriage will take place in period 2. For analytical convenience we assume that deterioration of the match quality (δ) is large enough and/or the cost of divorce (λ) is small enough that $\theta_1^{\dagger}$ is greater than the threshold value for marriage in period 1, denoted by θ_1^{*}.

2. 4. Uncertainty in the outcome of human capital investment in children

In general, differences in the characteristics of children as well as parents can result in variations in ε_i. Focusing on the issue of childbearing with different partners under uncertainty, we make simplifying assumptions with regard to this variable. We assume that all the variations in this variable are due exclusively to the heterogeneity among partners. Specifically, our assumptions are as follows:

(i) For all i (= f or s), ε_i takes a binomial distribution where ε_i is either ε_H (with probability p) or ε_L (with probability (1–p)), $\varepsilon_H > \varepsilon_L$, and $p \in [0,1]$.
(ii) The value of ε_i is identical for two children from the same spouse. On the other hand, ε_f and ε_s are independent when the second child is born in remarriage.

The rationale for the identical distribution in quality among children regardless of their biological parents (assumption i) is that we ignore the effect of quality heterogeneity on the model's implications, and focus on the impact of uncertainty in child quality.

Our analysis does not require for its main implications an assumption as strong as (ii). In fact, we need only to assume that the correlation of qualities of two children from the same biological parents is higher than that of two children from different parents. That is, two children from the same parents are more alike than two children from different parents. [8] In order to more simply explain our results, we make assumption (ii), namely that the former correlation is 1 and the latter is zero. We note that the sole method in our model by which the children's quality can be diversified under uncertainty is to remarry.

What is the effect of uncertainty in child quality on human capital investment in children? In the following proposition, we compare the human capital investment in the second child in an intact family with that in remarriage.

Proposition 1. If the third derivative of w is positive such that there is a precautionary motive for saving, the second child in an intact family receives more investment than the second child in remarriage.

8 The genetic incompatibility hypothesis by Trivers (1972) and Zeh and Zeh (1996) supports this idea.

Proof. See Appendix B. ■

The reason for this result is straightforward. As a parent who stays married in an intact family is more likely to observe the worst outcome (in which both children draw ε_L such that the rate of returns to human capital investment is the highest) than a remarried parent, more is invested in a child as a precautionary measure in case of an intact family.

It is plausible that the investment cost for the second child in remarriage is higher than that for the second child in an intact family because there can be benefits of repeated cooperation in educating children between spouses in an intact family. If this is true, higher cost of human capital investment for the second child in remarriage generates two familiar effects: income and substitution effects. In both effects, the investment in the second child in remarriage will be less than the second child in an intact family, which reinforces the result in this proposition. We assume, hereafter, that there is a precautionary motive for saving such that $w'''(c)>0$.

In what follows we investigate the effects of changes in various parameters in our model on decisions regarding divorce and remarriage.

Proposition 2. An increase in the cost of human capital investment v_2 will reduce the probability of remarriage for divorcees. The effect of v_2 on the probability of divorce is ambiguous with two opposing effects.

Proof. See Appendix C. ■

Rising cost of human capital investment in children will lower the utility from remarriage while a single's utility is intact. This will increase θ_2^* and divorcees will be less likely to remarry. When v_2 rises, choices such as staying married with the same spouse and remarrying that involve childbearing become less desirable. However, we can show that the adverse effect on the former choice is greater because parents invest more in an intact family. Divorces are thus more likely. On the other hand, since the expected utility from remarriage always exceeds that from singlehood, the option value of divorce falls as the probability of remarriage falls with rising v_2. This will render divorce less likely.

Therefore, the net effect of v_2 on divorce will be ambiguous.

Proposition 3. If an adult's consumption in singlehood is greater than that in remarriage due to human capital investment in children, higher income in period 2 (y_2) will raise the probability of remarriage for divorcees. Rising y_2 has two opposing effects on divorce and can thus increase or decrease the probability of divorce.

Proof. See Appendix D. ■

Lower consumption in remarriage makes the marginal utility of income higher in remarriage, and rising income will therefore result in an increase in the probability of remarriage. As for the effect of y_2 on divorce, since an adult's consumption is lower in an intact family than that in remarriage due to more investment in children in an intact family, and the latter consumption is lower than that in singlehood, the marginal utility of income in an intact family is higher than that in remarriage or in singlehood. Rising y_2 thus encourages spouses to stay married. However, an increase in y_2 can encourage divorce via the other channel. Since the expected utility from remarriage always exceeds that from singlehood, the option value of divorce rises if the probability of remarriage rises with increasing income, which makes divorce more likely. The net effect of y_2 on divorce will thus be ambiguous.

Proposition 4. An increase in income in period 3 (y_3) will reduce the incentive for divorcees to remarry. The effect of an increase in y_3 on divorce is ambiguous.

Proof. See Appendix E. ■

Since an adult's consumption in period 3 when remarried in the previous period is higher and hence the marginal utility of income is lower than when not married in period 2, an increase in y_3 will result in a decrease in the probability of remarriage. The effect of rising y_3 on divorce is ambiguous. On one hand, the divorce probability is likely to rise since the marginal expected utility of income in period 3 is higher when she is single in the previous period than when she stays married to the same spouse. On the other hand, divorce is less likely because, first, the option value of divorce falls with falling remarriage probability since the expected utility from remarriage always exceeds that from singlehood and, second, the marginal expected utility of income in period 3 is lower when she remarries in the previous period than when she stays married to the same spouse (see proposition 1 for proof).

2. 5. Income Uncertainty

How will uncertainty in future income affect decisions regarding divorce and remarriage? Suppose that adults are uncertain about income in period 3 and know *a priori* its distribution, which is independent of the risk in child quality:

$$
\begin{aligned}
y_3 &= y_{3H} \text{ with the probability } \pi \\
&= y_{3L} \text{ with the probability } (1-\pi),
\end{aligned}
$$

where $y_{3H} > y_{3L}$ and $\pi \in (0,1)$. In the following proposition, we consider the effect of rising income uncertainty on the quality of children.

Proposition 5. Consider an increase in income uncertainty in the sense that the distribution of y_3 becomes wider with its mean intact. This increase in future income uncertainty will raise human capital investment in children in period 2.

Proof. See Appendix F. ■

The intuition here is simple. Human capital investment in children will increase with rising income uncertainty because the increase in income uncertainty will raise the expected rate of returns on children's human capital investment, so long as the parent's utility function is concave and the parents have a motive for precautionary saving.

The next proposition addresses the effect of income uncertainty on decisions regarding remarriage.

Proposition 6. An increase in future income uncertainty will unambiguously raise the benefit of remarriage and thus increase the probability of remarriage for divorcees. (EFFECT ON DIVORCE NOT DISCUSSED YET)

Proof. See Appendix G. ■

We know the expected utility is lower when one is single in period 2 than when one remarries and bears an additional child. The reason for proposition 6 is essentially that rising income risk reduces the expected utility more when the level of the expected utility is lower. Rising income uncertainty therefore exerts a more adverse effect for singles, which makes remarriage more likely.

3. Empirical Implementation

We test the model's propositions against individual longitudinal data from the National Survey of Family Growth (Cycle 5, 1995), which cover 10,847 women from 15-44 years of age and contain the interviewee's detailed history of childbearing, marital and cohabitation status, and employment in the labor market, in addition to information regarding socio-economic variables such as education level, age, residence, race, religion, and occupation.

In Table 2, we test our theoretical implications on human capital investment in Propositions 1 and 5. The dependent variable is the number of weeks during which a woman breastfed a child (BRSTFEED). We consider this variable to be a proxy for human capital investment in children, because the short-term and long-term health benefits of breastfeeding to children are well documented in the medical science, and because breastfeeding is time-intensive, and thus costly, for women. Studies have determined that breast milk is associated with lower rates of a variety of diseases, including urinary-tract and respiratory-tract infections, diarrhea, allergies, bacterial meningitis, and botulism (Lawrence, 2000; Haider *et al*., 2003). In addition to these physiological health benefits, breast milk also improves children's cognitive and academic abilities (Horwood and Fergusson, 1998). The empirical results reported by Blau *et al.* (1996) demonstrate that breastfeeding until the age of six months exerts large and positive effects on a child's weight and height.

Our basic specification in Table 2 is the Tobit model, with woman-specific random effects:

$$\begin{aligned} &BRSTFEED_{it} = \beta' X_{it} + u_i + \varepsilon_{it} && \text{if } y_{it}^* = \beta' X_{it} + u_i + \varepsilon_{it} > 0, \\ &BRSTFEED_{it} = 0 && \text{if } y_{it}^* = \beta' X_{it} + u_i + \varepsilon_{it} \leq 0. \end{aligned} \qquad (3.1)$$

where u_i is the random-effects term for woman i, normally distributed i.i.d. with mean zero, and the error term ε_{it} is normally distributed i.i.d. with mean zero. We favor the Tobit model because our dependent variable is censored and it is zero for a significant fraction of the observations. X_{it} is a vector of woman i's characteristics at the time of the t-th childbirth, which includes a measure of woman's education level, age, earnings, future income uncertainty, binary variables for urban residency, religion and race, the number of children already born, and the birth weight of a child.

We also include an indicator for a child born in remarriage as a regressor in order to determine whether a child in remarriage gets less investment (see proposition 1). In practice, we use a binary variable (NEWDAD) which assigns a value of one if a woman gives birth to a child whose father is different from the father of the last child born, and assigns zero if she gives birth to a child with the same father of the last child. We use this variable instead of one related with remarriage in the legal sense because remarriage in our theory effectively means any type of partnership with propagation of children with a new spouse, including cohabitation. Our dataset contains three pieces of information that can be utilized to construct this variable: (1) time elapsed since conception in the last childbirth until the conception in the current childbirth, (2) father's age at the time of conception in the last childbirth, and (3) father's age at the time of conception in the current childbirth. We identify as the same father both the father in the last childbirth and that in the current childbirth, if the father's age in (3) equals the sum of

(1) and (2).[9]

We include a woman's schooling years (EDUC) as an explanatory variable to control for the cost of educating children. Higher wages for more educated women imply a higher opportunity cost for time-intensive childrearing (Becker, 1991). A woman's age at the time of conception (AGE) is also included as an explanatory variable to control for the cost of children, as it has been well documented in the literature that the age-earnings profile assumes an inverted-U shape (Murphy and Welch, 1990). At the same time, AGE may control for the marriage market condition, as young women may not have many eligible partners in the remarriage market.

Earnings in the year of conception (INC) are estimated from the results of the Mincerian regression of the reported average monthly earnings of women over the survey year. The average monthly earnings from the last jobs for those women who worked in that survey year are regressed on such explanatory variables as education level, labor market experience, and binary variables for occupations, industries, residential areas, and races. Then, the average monthly earnings in the year of conception are projected on the basis of the regression coefficients and the values of the explanatory variables in that year. This variable thus measures the level of income during childbearing periods. We proxy income uncertainty by the number of changes in labor market status (either job finding or separation) after the t-th childbirth as a ratio to the number of months elapsed since the t-th childbirth until the survey month (INC_RISK). This variable approximates the probability of the woman's future status change in the labor market over one month. The construction of this variable is predicated on the assumption that women have rational expectations of future changes in labor market status.

Our explanatory variables include a couple of binary variables for residential areas: namely, whether living in the central city of a Standard Metropolitan Statistical Area (CITY), or in the suburb of a SMSA

---

9 We have verified the validity of this method independently by checking whether at the time of conception in the current childbirth the woman was living together with (either cohabitating with or married to) the same man with whom she had lived together at the time of conception in the last childbirth.

(SUBURB). The cost of raising children may be higher in urban areas, as the housing price in urban areas is higher and children in rural areas can be helpful in agriculture. Moreover, the cost of partner searching may be lower in urban areas.

We also include as regressors binary variables for Protestantism (PROTST), for Roman Catholicism (ROMAN), for blacks (BLACK), and for whites (WHITE). Religion may influence fertility and marriage decisions because, for example, the psychological cost of divorce may be higher for Roman Catholics, owing to the Catholic Church's teachings against divorce. The binary variables for races are introduced in order to account for differences in unobserved variables, including different marriage market situations by race. In order to control for the starting state, we also include the number of children already born to a woman (PARITY) as an explanatory variable. We prefer a log transformation for all the regressors in X_{it} with the exception of binary variables, variables with possible zero values, and the age variable (AGE). Variables of AGE in a linear form and in a quadratic form are included as regressors in order to account for possible non-monotonic influences on the dependent variable, as demonstrated in the age-earnings profile. For variable construction and sample statistics, see Table 1.

An alternative dependent variable used in Table 2 is a woman's schooling years (EDUC). To obtain the years of schooling completed, we utilize a subsample of women age 25 and over. The specification in this table is a linear model only with cross-sectional data in the survey year. This model includes as regressors the education levels of a woman's parents, her age, binary variables for her residency and race, binary variables for the religions in which she was raised, and the number of children born to her mother (including herself). The income of a woman's parents is not available in our data, and is thus excluded from this specification. We include as an added regressor a binary variable representing whether the woman's mother has remarried at least once (M_REMARRY) to determine whether children of remarried mothers have less human capital, as was predicted in Proposition 1.

In Table 3, we test our theoretical implications in Propositions 2-4 and 6 with regard to the remarriage probability, using a binary choice

| Table 1 | Variable Descriptions and Summary Statistics

Variable	Description	Mean	Std. ev.
BRSTFEED	Number of weeks for breastfeeding	13.487	24.94
NEWDAD	= 1 if a woman gives birth to a child whose father is different from the father of the last child born, or = 0 if she gives birth to a child with the same father of the last child.	0.3036	0.460
EDUC	Years of schooling	11.918	2.957
AGE	Age at the time of conception	24.958	4.895
INC	Woman's average monthly earnings in the year of conception	1370.6	435.2
INC_RISK	Number of changes in labor market status after a childbirth as a ratio to the number of months elapsed since the childbirth until the survey month	0.0181	0.024
CITY	= 1 if living in the central city of a SMSA	0.3582	0.479
SUBURB	= 1 if living in the suburb of a SMSA	0.4231	0.494
PROTST	= 1 for Protestants	0.5768	0.494
ROMAN	= 1 for Roman Catholics	0.2952	0.456
BLACK	= 1 for black women	0.2758	0.447
WHITE	= 1 for white women	0.6621	0.473
PARITY	Number of children already born	1.5975	0.960
BABYWT	Birth weight of a child in ounces	119.21	20.73
MARRIED	= 1 if currently married	0.7271	0.445
FSTCHILD	Mother's age at the time of the first childbirth	19.871	4.104
F_EDU	Schooling years of the father of a woman	11.112	4.200
M_EDU	Schooling years of the mother of a woman	11.117	3.640
P_PROT	= 1 if a woman was raised in Protestantism	0.5603	0.496
P_ROMAN	= 1 if a woman was raised in Roman Catholicism	0.3567	0.479
P_JEW	= 1 if a woman was raised in Judaism	0.0130	0.113
HISPANIC	= 1 if a woman is a Hispanic	0.1332	0.340
SIBLING	Number of children who were born to the mother of a woman	4.6153	2.800
M_REMARRY	= 1 if a woman's mother had remarried at least once	0.1355	0.342

model. The dependent variable is a binary variable NEWDAD, and our basic specification in Table 3 is a probit model, with woman-specific random effects:

$$\begin{aligned} NEWDAD_{it} &= 1 \quad \text{if } \lambda' X_{it} + u_i + \varepsilon_{it} > 0, \\ NEWDAD_{it} &= 0 \quad \text{if } \lambda' X_{it} + u_i + \varepsilon_{it} \le 0, \end{aligned} \tag{3.2}$$

where u_i is the random-effects term for woman i, normally distributed i.i.d. with mean zero, the error term ε_{it} is normally distributed i.i.d. with mean zero, and X_{it} is a vector of woman i's characteristics at the time of the t-th childbirth, which includes all the regressors in equation (3.1) except for the birth weight of a child.

4. Empirical Findings

4. 1. Human Capital Investment in Children: Tobit Model

The estimation results of the specification in equation (3.1) are reported in models (columns) 1 and 2 of Table 2. Model 2 of this table includes as an additional regressor a binary variable NEWDAD, in addition to the regressors included in model 1.

The estimated effects of remarriage and income uncertainty on children's human capital investment as reported in Models 1 and 2 are consistent with the predictions in Propositions 1 and 5. The effect of NEWDAD is demonstrated to be significantly negative, which implies that investment in a child fathered by a new mate is lower, in accordance with the predictions of Proposition 1. Table 2 also demonstrates that INC_RISK exerts a significantly positive impact, which verifies the prediction in proposition 5: higher income uncertainty in the future will increase human capital investment in children.

Mother's education level (EDUC) is shown in Table 2 to exert a positive impact on the duration of breastfeeding. This may be attributable to a couple of reasons: (1) women with more education have a better knowledge of the benefits of breastfeeding, and (2) the health benefits of breastfeeding have become widely recognized only in the last few decades, and later cohorts of women tend to be more educated. This variable is shown in the study of Roe *et al*. (1999), which utilizes the U.S. data collected by the Food and Drug Administration, to exert the same qualitative effects on breastfeeding. Note that the time cost of breastfeeding is higher for more educated women, which can produce an adverse effect of education on breastfeeding. In spite of this effect, EDUC exhibits a positive impact on breastfeeding. Income level (INC) is also shown in this table to have a positive effect on breastfeeding, albeit insignificant.

Table 2 demonstrates that women residing in cities are likely to breastfeed their children for a longer period of time. This may be due to the fact that information about the benefits of breastfeeding was more readily available in urban areas. Protestant or Roman Catholic women are shown in this table to breastfeed their children for a shorter period of time. Our results indicate that black women have a short duration of breastfeeding, and this was also reported by Roe *et al*. (1999). Children with more siblings are shown to be breastfed for a shorter period of time, possibly because a mother with more children has to spend less time per child. Babies who were heavier at birth are likely to have a longer duration of breastfeeding.

Although the American Academy of Pediatrics (AAP) has officially recommended breastfeeding since 1948, the benefits of breastfeeding have become well recognized since 1980's. We find that the estimated coefficients are not changed qualitatively when we restrict the sample to be childbirths in 1980 or later (results not reported to save space).

In models 3 and 4 of Table 2, the dependent variable is a woman's schooling years, and this variable functions as a proxy for her human capital level. We find that a binary variable representing whether the woman's mother had remarried at least once (M_REMARRY) exerts a significantly adverse effect on the woman's education. The finding that children of remarried mothers have less human capital supports our prediction in Proposition 1. These models demonstrate that women with more educated parents, women who reside in urban areas, women who were raised under Roman Catholicism or Judaism, and women with fewer siblings all tend to have higher educational levels. Among ethnic

Table 2 Human Capital Investment

Dep. variable	BRSTFEED (1)	BRSTFEED (2)		EDUC (3)	EDUC (4)
EDUC	1.7321	1.6274	F_EDU	0.1739	0.1769
	3.09	2.90		18.63	19.07
AGE	1.2800	1.3058	M_EDU	0.2128	0.2114
	1.30	1.32		17.79	17.74
AGE2	-0.0014	-0.0016	AGE	-0.0456	-0.0536
	-0.08	-0.09		-0.69	-0.81
LnINC	7.4357	7.7740	AGE^2	0.0012	0.0013
	1.19	1.25		1.25	1.33
INC_RISK	65.3387	67.3438	CITY	0.3696	0.3734
	2.35	2.42		4.51	4.59
CITY	4.5789	4.5255	SUBURB	0.2774	0.2846
	2.08	2.06		3.82	3.95
SUBURB	1.6426	1.4735	P_PROT	-0.0294	-0.0361
	0.77	0.69		-0.25	-0.31
PROTST	-5.4791	-5.5644	P_ROMAN	0.2742	0.2326
	-2.60	-2.64		2.26	1.94
ROMAN	-6.5184	-6.6244	P_JEW	1.6689	1.5814
	-2.87	-2.92		6.66	6.39
BLACK	-36.3285	-35.9289	BLACK	-0.7353	-0.6893
	-10.92	-10.80		-3.84	-3.61
WHITE	-6.5128	-6.6223	WHITE	-0.7964	-0.7326
	-2.22	-2.26		-4.39	-4.05
PARITY	-1.8562	-1.9530	HISPANIC	-1.1521	-1.0952
	-2.87	-3.01		-5.59	-5.34
LnBABYWT	16.9790	17.1234	SIBLING	-0.1037	-0.1102
	5.84	5.89		-8.85	-9.41
NEWDAD		-2.2834	M_REMARRY		-0.7457
		-1.99			-10.02
Log likelihood	-17152.5	-17150.5	R-squared	0.3046	0.3127
Wald $\chi2$ (d.f.)	661.1 (13)	665.5 (14)	Observations	7,216	7,216
Observations	6,887	6,887			

Note: Rows show the estimated coefficient and the ratio of coefficient to standard error for each independent variable.

groups, Hispanic women were the least educated according to our data.[10]

4. 2. Remarriage for child diversification: Dynamic Binary Choice Model

Table 3 shows the results of our estimation of the specification in equation (3.2). Models 1 and 2 present the estimates from a probit and a logit model, respectively, both with woman-specific random effects. In model 3, a binary variable for marital status (MARRIED) and a mother's age when her first child was born (FSTCHILD) are added to the specification as regressors. A currently married woman should incur a higher childbearing cost with a new partner than a single woman, possibly as the consequence of a loss of marriage-specific human capital. Those who had the first child early may have given birth to a child unintentionally, not fathered by a most desired partner, which increases her likelihood of remarriage afterwards.

Both EDUC and INC are shown in Table 3 to exert significant effects on remarriage for child diversification. Moreover, the estimated coefficients associated with the two variables have opposite signs: a woman with less education or more income is more likely to use a different partner to father her next child. If EDUC and INC proxy the net cost of one unit of human capital investment (v_2) and income in adulthood (y_2), respectively, this result is consistent with Propositions 2 and 3.

The results in models 1 and 2 indicate a U-shaped association between a woman's age (AGE) and remarriage, with NEWDAD reaching the bottom at age 25, which indicates that 57% of the observations lie below this age level. However, the relationship is shown to be monotonically positive, albeit not significant, in model 3, when the binary variable for marital status (MARRIED) is included. Not surprisingly, the coefficient associated with MARRIED is negative: those who are married tend to have a lower probability of remarriage

10 The estimation results are qualitatively similar to those reported in models 3-4 of Table 2 when we take logarithmic values of EDUC as the dependent variable.

❙ Table 3 ❙ Childbearing with a New Partner: Binary Choice Model

Dependent variable: NEWDA Random Effects Model

	(1) Probit	(2) Logit	(3) Probit
EDUC	-0.1475	-0.2512	-0.0746
	-8.96	-8.89	-4.70
AGE	-0.0725	-0.1230	0.1249
	-2.18	-2.18	3.60
AGE2	0.0015	0.0025	-0.0004
	2.38	2.39	-0.64
LnINC	0.5537	0.9437	0.4470
	3.09	3.08	2.56
INC_RISK	2.7162	4.5969	1.4586
	3.27	3.26	1.81
CITY	-0.1320	-0.2274	-0.1829
	-2.13	-2.16	-3.03
SUBURB	-0.2697	-0.4624	-0.2188
	-4.33	-4.34	-3.63
PROTST	-0.1126	-0.1852	-0.0733
	-1.81	-1.74	-1.21
ROMAN	-0.1192	-0.2021	-0.0145
	-1.75	-1.73	-0.22
BLACK	0.5395	0.9092	0.0905
	5.77	5.72	0.99
WHITE	-0.1596	-0.2731	-0.1475
	-1.85	-1.86	-1.77
PARITY	-0.1016	-0.1754	-0.3739
	-4.51	-4.56	-13.89
MARRIED			-0.7944
			-16.51
FSTCHILD			-0.1718
			-18.99
Log likelihood	-4121.79	-4121.98	-3717.33
Wald χ^2 (d.f.)	383.70 (12)	361.61 (12)	837.18 (14)
Observations	7,223	7,223	7,223

Note: Rows show the estimated coefficient and the ratio of coefficient to standard error for each independent variable. All models are based on the Probit specification except model 2 (logit model).

due to higher costs. As the probability of being married rises with age in our data, this channel of the age effect will evidence a negative relationship with age and remarriage. On the other hand, older women tend to have more eligible partners in the remarriage market, which can result in a positive relationship. When marital status is controlled for as in model 3, the age variable takes up only the latter effect, and demonstrates a positive relationship with remarriage.

The effect of income uncertainty (INC_RISK) on remarriage for child diversification is shown in Table 3 to be statistically significant and positive in all models. According to proposition 6, this result may be attributable to the incentive to reduce one type of risk in child quality when there is an increase in risk of another type (risk in income). The result in Table 3 bolsters this notion.[11]

Table 3 shows that remarriage for child diversification is less likely in urban areas. The coefficients associated with the binary variables, CITY and SUBURB, are significant and negative in all models. This may be due to a higher childrearing cost in urban areas. We find that Roman Catholics and Protestants have a lower likelihood of remarriage, although these effects are only marginally significant. This may be because Christians, and particularly Catholics, have a higher cost for divorce or remarriage than do non-Christians (including people with no religion) as the consequence of the Christian Church's teachings against family dissolution.

The results in Table 3 indicate that blacks have a significantly higher probability of remarriage for child diversification, and that whites are less likely to remarry than the control group (mostly Hispanics). It has been extensively documented that black and Hispanic females are in excess supply in the marriage market, as higher incarceration rates and unemployment rate among black and Hispanic males significantly reduces the number of eligible participants in the marriage market.[12]

---

11 As a sensitivity test, we introduced the number of changes in labor-market status in the past as a regressor, instead of the number of status changes in the future (INC_RISK).The former variable was shown to have no significant effect on diversification.

12 The number of jail inmates per 100,000 people in 2005 is 166 for whites, 800 for blacks, and 268 for Hispanics (Bureau of Justice Statistics, **Prison and Jail Inmates**

This results in a higher cost of mating for black and Hispanic women, and thus to a higher tendency of child diversification for them. Black and Hispanic women have less education and higher income uncertainty, which also contributes to a higher probability of remarriage for child diversification.

The cost of childbearing and mate search for the next child can vary with the number of children already born to a woman. Table 3 shows the significant and adverse effect of the parity of children on child diversification. This result indicates that the cost of a new mate search rises with parity and reduces the likelihood of remarriage.

In addition to marital status, model 3 includes a mother's age when her first child was born (FSTCHILD) as an additional regressor. Women who experienced an unplanned childbirth at young ages are more likely to change partners afterwards and the estimated coefficient for FSTCHILD corroborates this expectation. Note that the estimated coefficient associated with the binary variable for blacks is no longer significant in this model, and those explanatory variables in model 3 can thus account for a significantly higher level in our dependent variable, NEWDAD, among blacks.

5. Concluding Remarks

This paper develops a dynamic model of decisions regarding remarriage and the quality of children with uncertainty in child quality. So long as the quality distributions of two children are correlated less profoundly when they are fathered by two different spouses than when they are fathered by the same spouse, a woman naturally seeks to hedge her risks regarding child quality by remarrying to have children with different partners, or by diversifying her portfolio of children.

Our model shows that a parent invests less in children born in remarriage and invests more when future income is less certain.

at Midyear 2005). The unemployment rate is 4.8% for whites, 10.4% for blacks, and 7.0% for Hispanics in 2004 (US Census Bureau, Statistical Abstract of the United States, 2006).

Remarriage for child diversification is more likely to take place for divorcees with lower cost of educating children, higher income for parenting, and higher income uncertainty in the future. Our empirical analysis generally corroborates these predictions.

Our theory on remarriage for child diversification can shed some light on out-of-wedlock childbirths. Since the 1960's, the fraction of out-of-wedlock births has grown steadily in Western Europe and in the U.S. (Hotz, Klerman, Willis, 1997; Willis, 1999; Neal, 2004). In the U.S., the fraction of out-of-wedlock births reached a rate of 36.9% in 2005, when this fraction was 25.3 percent for whites, 69.9 percent for blacks, and 48.0 percent for Hispanics. The fraction of out-of-wedlock births had also increased during this period throughout Western Europe, but at different rates in different countries. The fraction was 2% in Italy in 1970, and reached a level of 7% in 1992, while in Sweden, it had increased from 18% in 1970 to over 50% in 1992. Willis (1999) has argued that the prevalence of out-of-wedlock births among blacks in the U.S. is due principally to the sex ratio imbalance in the marriage market. This may be an important explanation for out-of-wedlock births among blacks or Hispanics. However, this explanation alone clearly cannot apply to childbirths in unmarried white women in the U.S. or women in Western European countries. In fact, the out-of-wedlock birth rate among white women in the U.S. has been rising even more rapidly than that of blacks over the past few decades. This paper offers a different perspective on this issue. Women exercising child diversification may end up with out-of-wedlock births in cases in which the cost of legal remarriage is prohibitively steep. This explanation is fairly consistent with the rising fraction of out-of-wedlock births among white women in the U.S. and in Western Europe who have not experienced sex ratio imbalances in the marriage market, unlike black women in the U.S. Our data from the National Survey of Family Growth show that 55% of the cases where a woman gives birth to a child whose father is different from the father of the last child born (that is, NEWDAD equals 1) were out-of-wedlock childbirths while only 18% of the cases with the same father were out of wedlock.

Our narrative also helps to refute racial stereotypes about promiscuity. Black women, like black men, are subject to a sexual stereotype that portrays them as sexually “loose” (Hernton, 1971; Staples, 1972). Harrison *et al*. (1969) have shown higher levels of permissiveness among blacks, and Singh *et al*. (1976) have found blacks more accepting of extramarital sex. Zelnik & Kantner (1977) and Houston (1981) have reported black women to have a greater number of premarital partners. Those black women, typically with less education and higher levels of uncertainty regarding future income than the average in the population, may look to be promiscuous as they undertake child diversification strategies. This is, then, not a consequence of biological differences in promiscuity, but rather, a consequence of rational behaviors.

References

Baker, R. Robin, and Mark A. Bellis. 1990. “Do females promote sperm competition? Data for humans.” *Animal Behavior* 40: 997-999.

Bellis, Mark A., Karen Hughes, Sara Hughes, John R Ashton. 2005. “Measuring paternal discrepancy and its public health consequences.” *Journal of Epidemiology and Community Health* 59: 749-754.

Becker, Gary S. 1960. “An Economic Analysis of Fertility.” *Demographic and Economic Change in Developed Countries*.

Becker, Gary S. 1991. *A Treatise on the Family, Enlarged Edition*, Chapter 6, Harvard University Press, Cambridge, MA.

Birkhead, Tim. 2000. *Promiscuity: An Evolutionary History of Sperm Competition*, Harvard University Press, Cambridge, MA.

Bramlett M. D. and W. D. Mosher. 2002. "Cohabitation, Marriage, Divorce, and Remarriage in the United States," National Center for Health Statistics. Vital Health Stat 23(22).

Cartwright, John. 2000. *Evolution and Human Behavior*, MIT Press, Cambridge, MA.

Cherlin, A. J. 1992. *Marriage, Divorce, Remarriage: Revised and Enlarged Edition*. Harvard University Press: Cambridge, MA.

Cox, Donald. 2003. “Private Transfers within the Family: Mothers, Fathers, Sons, and Daughters.” *Death and Dollars: The Role of Gifts and Bequests in America*, edited by A. H. Munnell and A. Sunden, Brookings Institution Press, Washington, D.C.

Francesconi, Marco, Stephen Jenkins, and Thomas Siedler. 2010. "Childhood family structure and schooling outcomes: evidence for Germany," *Journal of Population Economics*, vol. 23(3), pages 1073-1103.

Gruber, Jonathan. 2004. "Is making divorce easier bad for children? The long-run implications of unilateral divorce," *Journal of Labor Economics* 22, no. 4:799–833.

Haider, Steven J., Alison Jacknowitz, Robert F. Schoeni. 2003. “Welfare Work Requirements and Child Well-Being: Evidence from the Effects on Breast-Feeding.” *Demography* 40 (3): 479-497.

Harrison, D. E., W. H. Bennett, G. Gobinetti. 1969. “Attitudes of rural youth towards premarital sexual permissiveness.” *Journal of Marriage and the Family* 31: 783-787.

Hetherington, Mavis, and Margaret Stanley-Hagan. 1999. "Stepfamilies," In *Parenting and child development in non-traditional families,* ed. M. Lamb. Mahwah, NJ: Erlbaum.

Hotz, V. Joseph, Jacob A. Klerman, Robert J. Willis. 1997. “The Economics of Fertility in Developed Countries”, *Handbook of Population and Family Economics*, edited by M.R. Rosenzweig and O. Stark, Elsevier.

Horwood, L.J. and D.M. Fergusson. 1998. “Breastfeeding and Later Cognitive and Academic Outcomes.” *Pediatrics* 101(1): e9.

Houston, L. N. 1981. “Romanticism and eroticism among black and white college students.” *Adolescence* 16: 263-272.

Kleiman, D.G. 1977. “Monogamy in mammals.” *Quarterly Review of Biology* 52: 39-69.

Knight, Frank H. 1921. *Risk, Uncertainty, and Profit*. Houghton Mifflin Company, Boston, MA.

Kreider, R. M. 2005. "Number, Timing and Duration of Marriages and Divorces: 2001," Current Population Reports, P70-97.

Lamb, Michael, Kathleen Sternberg, and Ross Thompson. 1999. "The effects of divorce and custody arrangements on children’s behavior, development, and adjustment," In *Parenting and child development in nontraditional families,* ed. Michael Lamb. Mahwah, NJ: Erlbaum.

Lang, Kevin, and Jay L. Zagorsky. 2001. "Does Growing up with a Parent Absent Really Hurt?" *Journal of Human Resources*, University of Wisconsin Press, vol. 36(2), pages 253-273.

Lawrence, R.A. 2000. “Breastfeeding: Benefits, Risks, and Alternatives.” *Current Opinion in Obstetrics and Gynecology* 12: 519-24.

Murphy, Kevin, Finis Welch. 1990. “Empirical Age-Earnings Profiles.” *Journal of Labor Economics* 8 (2): 202-229.

Neal, Derek. 2004. “The Relationship between Marriage Market Prospects and Never-Married Motherhood.” *Journal of Human Resources* 39 (4), 938-957.

Page, Marianne E., and Ann Huff Stevens. 2004. "The Economic Consequences of Absent Parents," *Journal of Human Resources*, University of Wisconsin Press, vol. 39(1).

Painter, Gary, and David I. Levine. 2000. "Family Structure and Youths' Outcomes: Which Correlations are Causal?" *Journal of Human Resources*, University of Wisconsin Press, vol. 35(3), pages 524-549.

Ridley, Matt. 1995. *The Red Queen: Sex and the Evolution of Human Nature*. Penguin Books.

Roe, Brian, Leslie A. Whittington, Sara Beck Fein, Mario F. Teisl. 1999. "Is There Competition between Breast-Feeding and Maternal Employment?" *Demography* 36 (2), 157-171.

Singh, K. B., B. L. Walton, J. S. Williams. 1976. "Extramarital sexual permissiveness: Conditions and contingencies." *Journal of Marriage and the Family* 38: 701-712.

Staples, R. 1972. "The sexuality of black women." *Sexual Behavior* 2, 4-15.

Trivers, R. L. 1972. "Parental investment and sexual selection," *Sexual Selection and the Descent of Man, 1871-1971*, edited by B. Campbell, Aldine-Atherton, Chicago, IL: 136-79.

U.S. Census Bureau. 2001. Survey of Income and Program Participation (SIPP), 2001 Panel, Wave 2 Topical Module.

Weiss, Yoram, 1993. "The formation and dissolution of families: Why marry? Who marries whom? And what happens upon divorce," in M. R. Rosenzweig & O. Stark (ed.), *Handbook of Population and Family Economics*, vol. 1, ch. 3, 81-123, Elsevier.

Willis, Robert. 1999. "A Theory of Out-of-Wedlock Childbearing." *Journal of Political Economy* 107: S33-64.

Wineberg, H. 1990. "Childbearing after remarriage," *Journal of Marriage and the Family*, 52, 31-38.

Zeh, J.A., and D. W. Zeh. 1996. "The evolution of polyandry I: intragenomic conflict and genetic incompatibility," *Proceedings of the Royal Society of London*: 1711-17.

Zelnik, M., and J. F. Kantner. 1977. "Sexual and contraceptive experience of young married women in the United States, 1976 and 1971." *Family Planning Perspectives* 9: 55-71.

Appendix

A. Diversification behavior in nature and in humans

Rather than reproducing via a process of self-duplication, many species on this planet, including mammals, rely on sexual reproduction, which is itself all about genetic mixing in order to diversify genes (Ridley, 1995)—this process of swapping paternal and maternal genes is referred to as recombination. One of the major theories thus far advanced to explain the persistence of sexual reproduction is the lottery principle as elucidated by Williams (1975): sex introduces genetic variety in order to enable genes to survive in changing or novel environments, since organisms have a better chance of producing offspring that survive when they produce a range of types, rather than more of the same. Another prominent theory is the Red Queen hypothesis initially put forth by Van Valen (1973). After discovering that the probability of a family of marine organisms becoming extinct at any point of time is completely unrelated to how long it has already survived, Van Valen argued that variability through sex enables organisms to remain competitive in a world in which other organisms are constantly poised to take advantage of any weakness—thus, organisms should "run fast" like the Red Queen in the story "Alice in Wonderland", in order to "stay still" in the constantly-shifting genetic arms race (Cartwright, 2000). We should note that all these theories share in common the notion of sex as a mechanism for the maintenance of genetic variability or diversification.

Throughout the animal kingdom, species that practice sexual monogamy are very rare. Paternity studies have confirmed this for many invertebrate groups and every major vertebrate group: fish, amphibians, reptiles, birds, marsupials and mammals (Birkhead, 2000). Birds were once considered to be models of monogamy because most bird species are observed to breed in pairs, or to be *socially* monogamous. However, detailed paternity studies began to reveal that the females of most socially monogamous birds are not *sexually* monogamous at all (Kleiman, 1977). In the majority of mammals, males are polygynous and very few species are socially, and thus sexually monogamous. While males are renowned for their promiscuity, it was not until the 1980's when researchers started to look at sperm competition in animals, and in birds in particular, that it was realized that females of almost all animal species routinely copulate with several different male partners (Birkhead, 2000). Trivers (1972) and Zeh and Zeh (1996) suggest that females copulate with different males to minimize their chances of being fertilized by an incompatible male. They point out that there is abundant evidence for genetic incompatibility: inbreeding depression, the existence of genetic conflicts within and between the nucleus and cytoplasm. In some instances, females may be able to discriminate between males if males advertise their types as in mice and humans which signal their MHC (major histocompatibility complex) type via their odor.[14]

In the case of our closest relatives among the primates, the great apes, including chimpanzees and bonobos, have been shown to utilize a promiscuous mating system with multi-male and multi-female groups, whereas gorillas and orangutans practice polygyny (Ridley, 1995). Many primate species, including humans, exhibit the characteristic of concealed ovulation, which is believed to afford some adaptive advantage in evolution. One prominent theory that has been proposed to account for this feature is referred to as the "nice daddy" theory. Hrdy (1979) has suggested that concealed ovulation functions to confuse paternity and enables a female to select desirable males for mating without the risk of infanticide from suspicious males. Benshoof and Thornhill (1979) have argued that a female can mate with another male without alerting her first mate by concealing her ovulation from her (ostensibly monogamous) partner. According to this theory, concealed

14 The MHC consists of a set of genes in vertebrates that are associated with disease resistance and the immune system (Birkhead, 2000).

ovulation becomes a strategy that is employed by females to allow them to mate with many males, and yet remain secure in the knowledge that the confusion over paternity would extract maximum paternal care. By investigating data regarding the mating system and signs of ovulation in 68 extant primate species, Sillen-Tullberg and Moller (1993) concluded that the concealed ovulation strategy emerged in the process of evolution more often in a non-monogamous context, which bolsters the "nice daddy" hypothesis.

In the case of our species in modern times, many surveys have previously reported that the rate of "paternal discrepancy" or "misattributed paternity", both of which terms refer to cases in which a woman's husband is not her child's biological father, is sizable. Baker and Bellis (1990) estimated from the England survey data that the rate was between 6.9 and 13 percent. Using studies conducted in the U.K., U.S., Europe, Russia, Canada, South Africa, South America, New Zealand, and Mexico from the 1950's through 2002 that mentioned paternal discrepancy, Bellis *et al*. (2005) determined that paternal-discrepancy estimates vary wildly, from less than 1% to more than 30%, and that the average paternal discrepancy is 3.7%. They reported that the rates were higher for disadvantaged people, for those with more than one sex partner at a time, and for younger women.

B. Proof for Proposition 1

The values of human capital investment in the second child in remarriage and in an intact family are denoted by h_{2s}^{R} and h_{2s}^{I}, respectively, and they are found from the following first-order optimality conditions if we have internal solutions:

$$\text{(A.1)}\quad -v_2\, u'(2y_2 - v_2\, h_{2s}^{R}) + E[w'(y_3 + (f(h_{1f}) + f(h_{2s}^{R}) + \varepsilon + \varepsilon')/\alpha)\, f'(h_{2s}^{R})/\alpha] = 0,$$

$$\text{(A.2)}\quad -v_2\, u'(2y_2 - v_2\, h_{2s}^{I}) + E[w'(y_3 + (f(h_{1f}) + f(h_{2s}^{I}) + \varepsilon + \varepsilon)/\alpha)\, f'(h_{2s}^{I})/\alpha] = 0,$$

Note that the stochastic terms for the first and the second child are the same in equation (A.2) while they are different and independent in equation (A.1) since the two children are fathered by different partners.

Suppose $h_{2s}^{R} > h_{2s}^{I}$. Comparing equations (A.1) and (A.2), we have

$$\text{(A.3)}\ E[w'(y_3 + (f(h_{1f}) + f(h_{2s}^{I}) + \varepsilon + \varepsilon)/\alpha)] / E[w'(y_3 + (f(h_{1f}) + f(h_{2s}^{R}) + \varepsilon + \varepsilon')/\alpha)]$$

$$= [u'(2y_2 - v_2\ h_{2s}^{I}) / u'(2y_2 - v_2\ h_{2s}^{R})]\ [f'(h_{2s}^{R}) / f'(h_{2s}^{I})],$$

The right hand side of equation (A.3) is less than 1. To compare the numerator and the denominator of the left hand side, we take

$$E[w'(y_3 + (f(h_{1f}) + f(h_{2s}^{R}) + \varepsilon + \varepsilon')/\alpha)] - E[w'(y_3 + (f(h_{1f}) + f(h_{2s}^{I}) + \varepsilon + \varepsilon)/\alpha)]$$

$$= p^2\ [w'(\ y_3 + (f(h_{1f}) + f(h_{2s}^{R}) + \varepsilon_H + \varepsilon_H)/\alpha) - w'(\ y_3 + (f(h_{1f}) + f(h_{2s}^{I}) + \varepsilon_H + \varepsilon_H)/\alpha)]$$

$$+ p(1\text{-}p)\ [w'(\ y_3 + (f(h_{1f}) + f(h_{2s}^{R}) + \varepsilon_H + \varepsilon_L)/\alpha) - w'(\ y_3 + (f(h_{1f}) + f(h_{2s}^{I}) + \varepsilon_H + \varepsilon_H)/\alpha)]$$

$$+ p(1\text{-}p)\ [w'(\ y_3 + (f(h_{1f}) + f(h_{2s}^{R}) + \varepsilon_L + \varepsilon_H)/\alpha) - w'(\ y_3 + (f(h_{1f}) + f(h_{2s}^{I}) + \varepsilon_L + \varepsilon_L)/\alpha)]$$

$$+ (1\text{-}p)^2\ [w'(\ y_3 + (f(h_{1f}) + f(h_{2s}^{R}) + \varepsilon_L + \varepsilon_L)/\alpha) - w'(\ y_3 + (f(h_{1f}) + f(h_{2s}^{I}) + \varepsilon_L + \varepsilon_L)/\alpha)].$$

The first and fourth terms with brackets are negative since $h_{2s}^{R} > h_{2s}^{I}$. If $w''' > 0$, we can show the second and third terms are negative. This means the left hand side of equation (A.3) is bigger than 1, which contradicts that the right hand side should be less than 1. Therefore, $h_{2s}^{R} \leq h_{2s}^{I}$.

C. Proof for Proposition 2

Suppose we have an unanticipated change in v_2. Total differentiating equation (2.4) and using the envelope theorem, we have

$-\ h^R_{2s}\ u'(2y_2 - v_2\ h^R_{2s})\ dv_2 + d\theta_2^* = 0$, and

$d\theta_2^*/dv_2 = h^R_{2s}\ u'(2y_2 - v_2\ h^R_{2s}) > 0$. Remarriage is thus less likely as v_2 rises.

In equation (2.6), we denote M_1 and M_2 by

$$M_1 = \underset{h_{2s}}{Max}\ u(2y_2 - v_2\ h_{2s}) + E[\theta_2|\theta_2 \geq \theta_2^*] + E[w(y_3 + H_f/\alpha + H_s/\alpha)],$$

and

$$M_2 = u(y_2) + E[w(y_3 + H_f/\alpha).$$

Differentiating the equation for $\theta_1^\dagger$ implied by equations (2.5) and (2.6), we get

$$-\ h^I_{2s}\ u'(2y_2 - v_2\ h^I_{2s})\ dv_2 + d\theta_1^\dagger = \Pr(\theta_2 \geq \theta_2^*)\ h^R_{2s}\ u'(2y_2 - v_2\ h^R_{2s})$$
$$\{-1 + \frac{\partial}{\partial\theta_2^*}\ E[\theta_2|\theta_2 \geq \theta_2^*]\}\ dv_2$$
$$+ \frac{\partial\theta_2^*}{\partial v_2}\ \frac{\partial \Pr(\theta_2 \geq \theta_2^*)}{\partial\theta_2^*}\ (M_1 - M_2)\ dv_2\ \text{, and}$$

$$d\theta_1^\dagger/dv_2 = \{h^I_{2s}\ u'(2y_2 - v_2\ h^I_{2s}) - \Pr(\theta_2 \geq \theta_2^*)\ h^R_{2s}\ u'(2y_2 - v_2\ h^R_{2s})$$
$$\{1 - \frac{\partial}{\partial\theta_2^*}\ E[\theta_2|\theta_2 \geq \theta_2^*]\}\}$$
$$+ \{\ \frac{\partial\theta_2^*}{\partial v_2}\ \frac{\partial \Pr(\theta_2 \geq \theta_2^*)}{\partial\theta_2^*}\ (M_1 - M_2)\ \}$$
$$\equiv \{N_1\} + \{N_2\}.$$

In term N_1, $h^I_{2s}\ u'(2y_2 - v_2\ h^I_{2s}) \geq h^R_{2s}\ u'(2y_2 - v_2\ h^R_{2s})$ since $h^I_{2s} \geq h^R_{2s}$, and $\{1 - \frac{\partial}{\partial\theta_2^*}\ E[\theta_2|\theta_2 \geq \theta_2^*]\} < 1$ since $\frac{\partial}{\partial\theta_2^*}\ E[\theta_2|\theta_2 \geq \theta_2^*] > 0$. Therefore, term N_1 is positive. We can show that

$M_1 - M_2 = E[\theta_2|\theta_2 \geq \theta_2^*] - \theta_2^* > 0.$

Since $\frac{\partial \theta_2^*}{\partial v_2} > 0$ and $\frac{\partial \Pr(\theta_2 \geq \theta_2^*)}{\partial \theta_2^*} \leq 0$, term N_2 is negative.

Any choice like staying married in an intact family or remarrying that involves childbearing become less desirable when v_2 rises. We find from term N_1 that the adverse effect on the former choice is greater because parents invest more in an intact family, which makes divorces more likely. On the other hand, since the expected utility from remarriage always exceeds that from singlehood (i.e., M_1 is greater than M_2), the option value of divorce falls as the probability of remarriage falls with rising v_2, and this makes divorce less likely (as illustrated by negative N_2). Therefore, the net effect of v_2 on divorce is ambiguous.

D. Proof for Proposition 3

We can show from equation (2.4) that

$d\theta_2^*/dy_2 = u'(y_2) - u'(2y_2 - v_2\ h_{2s}^R) < 0,$

if an adult's consumption in remarriage $(2y_2 - v_2\ h_{2s}^R)$ is smaller than that in singlehood (y_2) due to human capital investment in children. Remarriage is thus more likely when y_2 rises. Hereafter we consider the case when $y_2 \geq (2y_2 - v_2\ h_{2s}^R)$.

For the effect on the probability of divorce, we get the following from the equation for $\theta_1^\dagger$ implied by equations (2.5) and (2.6):

$d\theta_1^\dagger/dy_2 = \Pr(\theta_2 \geq \theta_2^*)\ \{u'(2y_2 - v_2\ h_{2s}^R) - u'(2y_2 - v_2\ h_{2s}^I) + \frac{\partial \theta_2^*}{\partial y_2} \frac{\partial}{\partial \theta_2^*}$

$E[\theta_2|\theta_2 \geq \theta_2^*]\}$

$+ \Pr(\theta_2 < \theta_2^*)\ \{u'(y_2) - u'(2y_2 - v_2\ h_{2s}^R)\}$

$+ \{\ \frac{\partial \theta_2^*}{\partial y_2} \frac{\partial \Pr(\theta_2 \geq \theta_2^*)}{\partial \theta_2^*}\ (M_1 - M_2)\ \}$

$\equiv \Pr(\theta_2 \geq \theta_2^*)\ \{N_3\} + \Pr(\theta_2 < \theta_2^*)\ \{N_4\} + \{N_5\}$

Since $h^R_{2s} \le h^I_{2s}$, h^I_{2s} $u'(2y_2 - v_2\, h^R_{2s}) - u'(2y_2 - v_2\, h^I_{2s}) \le 0$. Also, we have $\frac{\partial \theta_2^*}{\partial y_2} < 0$ and $\frac{\partial}{\partial \theta_2^*}$ $E[\theta_2|\theta_2 \ge \theta_2^*] > 0$, and therefore term N_3 is negative. So is term N_4 as long as $y_2 \ge (2y_2 - v_2\, h^R_{2s})$. We can show in the same way as in proposition 2 that term N_5 is positive.

Since the marginal utility of income in an intact family is higher than that in remarriage or in singlehood, rising y_2 encourages spouses to stay married (as illustrated in terms N_3 and N_4). On the other hand, the option value of divorce rises with y_2 since the expected utility from remarriage always exceeds that from singlehood, which makes divorce more likely (as illustrated in term N_5). The net effect of y_2 on divorce is thus ambiguous.

E. Proof for Proposition 4

We can show from equation (2.4) that

$d\theta_2^*/dy_3 = E[w'(y_3 + H_f/\alpha)] - E[w'(y_3 + H_f/\alpha + H_s/\alpha)] > 0$, and

remarriage is thus less likely when y_3 rises.

For the effect on the probability of divorce, we get the following from the equation for $\theta_1^\dagger$ implied by equations (2.5) and (2.6):

$$d\theta_1^\dagger/dy_3 = \Pr(\theta_2 \ge \theta_2^*)\ \{E[w'(y_3 + (f(h_{1f})+f(h^R_{2s})+\varepsilon+\varepsilon')/\alpha)] - E[w'(y_3 + (f(h_{1f})+f(h^I_{2s})+\varepsilon+\varepsilon)/\alpha)]\}$$

$$+ \Pr(\theta_2 \ge \theta_2^*)\ \frac{\partial \theta_2^*}{\partial y_3} \frac{\partial}{\partial \theta_2^*}\ E[\theta_2|\theta_2 \ge \theta_2^*]$$

$$+ \Pr(\theta_2 < \theta_2^*)\ \{E[w'(y_3 + (f(h_{1f})+\varepsilon)/\alpha)] - E[w'(y_3 + (f(h_{1f})+f(h^I_{2s})+\varepsilon+\varepsilon)/\alpha)]\}$$

$$+ \left\{ \frac{\partial \theta_2^*}{\partial y_3} \frac{\partial \Pr(\theta_2 \ge \theta_2^*)}{\partial \theta_2^*}\ (M_1 - M_2) \right\}$$

$$\equiv \Pr(\theta_2 \ge \theta_2^*)\ \{N_6\} + \Pr(\theta_2 \ge \theta_2^*)\ \frac{\partial \theta_2^*}{\partial y_3} \frac{\partial}{\partial \theta_2^*}\ E[\theta_2|\theta_2 \ge \theta_2^*] + \Pr(\theta_2 < \theta_2^*)\ \{N_7\} + \{N_8\}$$

Term N_6 is shown in proposition 1 to be negative. We can show that

$\frac{\partial\theta_2^*}{\partial y_3}\frac{\partial}{\partial\theta_2^*}E[\theta_2|\theta_2 \geq \theta_2^*] > 0$ and term N_7 is also positive. Term N_8 is negative since $\frac{\partial\theta_2^*}{\partial y_3} > 0$, $\frac{\partial Pr(\theta_2 \geq \theta_2^*)}{\partial\theta_2^*} \leq 0$, and $M_1 > M_2$.

The probability of divorce is likely to rise as y_3 increases since the marginal utility of income in period 3 is higher when she is single in period 2 than when she stays married in an intact family (term N_7). On the other hand, divorce is less likely because (i) the option value of divorce falls with falling remarriage probability since the expected utility from remarriage always exceeds that from singlehood (term N_8), and (ii) the marginal expected utility of income in period 3 is lower when she remarries in the previous period than when she stays married to the same spouse (term N_6).

F. Proof for Proposition 5

Taking the total differentiation on the first-order optimality condition for h_{2s}^R in equation (A.1), we get

$$
\begin{aligned}
0=&\{v_2^2u''(2y_2-v_2\, h_{2s}^R) + \pi[f'(h_{2s}^R)/\alpha]^2\, E[w''(y_{3H}+H_f/\alpha+H_s/\alpha)] + \pi f''(h_{2s}^R)/\alpha\, E[w'(y_{3H}+H_f/\alpha+H_s/\alpha)] \\
&+ (1-\pi)\, [f'(h_{2s}^R)/\alpha]^2\, E[w''(y_{3L}+H_f/\alpha+H_s/\alpha)] + (1-\pi)\, f''(h_{2s}^R)/\alpha\, E[w'(y_{3L}+H_f/\alpha+H_s/\alpha)]\}\, dh_{2s}^R \\
&+ \pi\, f'(h_{2s}^R)/\alpha\, E[w''(y_{3H}+H_f/\alpha+H_s/\alpha)]\, dy_{3H} \\
&+ (1-\pi)\, f'(h_{2s}^R)/\alpha\, E[w''(y_{3L}+H_f/\alpha+H_s/\alpha)]\, dy_{3L} \\
\equiv\, &\Omega_1\, dh_{2s}^R + \pi\, \Omega_2\, dy_{3H} + (1-\pi)\, \Omega_3\, dy_{3L},
\end{aligned}
$$

where Ω_1 is negative to satisfy the second-order optimality condition. We can then show

$$
[\partial h_{2s}^R/\partial y_{3H} - \pi/(1-\pi)\partial h_{2s}^R/\partial y_{3L}] = -\pi f'(h_{2s}^R)/\alpha\, \{E[w''(y_{3H}+H_f/\alpha+H_s/\alpha)] - E[w''(y_{3L}+H_f/\alpha+H_s/\alpha)]\}/\Omega_1,
$$

where the term in braces is positive since $w''' > 0$. Therefore, $[\partial h_{2s}^R/\partial y_{3H} - \pi/(1-\pi)\partial h_{2s}^R/\partial y_{3L}] > 0$. In the same manner, we can show that $[\partial h^{I}_{2s}/\partial y_{3H} - \pi/(1-\pi)\partial h^{I}_{2s}/\partial y_{3L}] > 0$.

G. Proof for Proposition 6

Differentiating θ_2^* with respect to y_{3H} and y_{3L} with the condition $\pi dy_{3H} + (1-\pi)\, dy_{3L} = 0$,

$$
\begin{aligned}
&[\partial\theta_2^*/\partial y_{3H} - \pi/(1-\pi)\, \partial\theta_2^*/\partial y_{3L}]/\pi \\
= \quad & E[w'(y_{3H} + H_f/\alpha)] - E[w'(y_{3H} + (f(h_{1f})+f(h^{R}_{2s})+\varepsilon+\varepsilon')/\alpha)] \\
& + E[w'(y_{3L} + (f(h_{1f})+f(h^{R}_{2s})+\varepsilon+\varepsilon')/\alpha)] - E[w'(y_{3L} + H_f/\alpha)] \\
= \quad & p^2 \; \{w'(y_{3L} + (f(h_{1f})+f(h^{R}_{2s})+\varepsilon_H+\varepsilon_H)/\alpha) - w'(y_{3H} + (f(h_{1f})+f(h^{R}_{2s})+\varepsilon_H+\varepsilon_H)/\alpha)\} \\
& + 2p(1-p)\; \{w'(y_{3L} + (f(h_{1f})+f(h^{R}_{2s})+\varepsilon_H+\varepsilon_L)/\alpha) - w'(y_{3H} + (f(h_{1f})+f(h^{R}_{2s})+\varepsilon_H+\varepsilon_L)/\alpha)\} \\
& + (1-p)^2\; \{w'(y_{3L} + (f(h_{1f})+f(h^{R}_{2s})+\varepsilon_L+\varepsilon_L)/\alpha) - w'(y_{3H} + (f(h_{1f})+f(h^{R}_{2s})+\varepsilon_L+\varepsilon_L)/\alpha)\} \\
& - p\; \{w'(y_{3L} + (f(h_{1f})+\varepsilon_H)/\alpha) - w'(y_{3H} + (f(h_{1f})+\varepsilon_H)/\alpha)\} \\
& - (1-p)\; \{w'(y_{3L} + (f(h_{1f})+\varepsilon_L)/\alpha) - w'(y_{3H} + (f(h_{1f})+\varepsilon_L)/\alpha)\} \\
\equiv \quad & p^2\, \{\Omega_4\} + 2p(1-p)\, \{\Omega_5\} + (1-p)^2\, \{\Omega_6\} - p\, \{\Omega_7\} - (1-p)\, \{\Omega_8\} \\
= \quad & p^2\, \{\Omega_4 - \Omega_7\} + p(1-p)\, \{\Omega_5 - \Omega_7\} + p(1-p)\, \{\Omega_5 - \Omega_8\} + (1-p)^2\, \{\Omega_6 - \Omega_8\}.
\end{aligned}
$$

We can show that $(\Omega_4 - \Omega_7)$, $(\Omega_5 - \Omega_7)$, $(\Omega_5 - \Omega_8)$, and $(\Omega_6 - \Omega_8)$ are all negative as long as $w''' > 0$. Therefore, $[\partial\theta_2^*/\partial y_{3H} - \pi/(1-\pi)\, \partial\theta_2^*/\partial y_{3L}]/\pi < 0$. This implies that the probability of remarriage for divorcees rises as income uncertainty is increased.

CHAPTER 5

Stereotypes and Inequality: A Structure of Identity Choice

by
Young Chul Kim.*
(Korea Development Institute)
*Glenn C. Loury***
(Brown University)

Abstract

When identity is exogenous and if the ability distributions within groups are the same, then inequality of group reputations in equilibrium can only arise if there is a positive feedback between group reputation and individual human capital investment activities (Arrow, 1973; Coate and Loury, 1993). When group membership is endogenous, the logic of individuals' identity choices leads there to be a positive selection of higher ability individuals into the favored group. As a result, ability distributions within distinct groups can endogenously diverge, reinforcing incentive-feedbacks. We develop the theoretical framework that can examine the positive selection and the endogenous group formation, and examine the existence and stability of stereotyping equilibria. We show inequality deriving from stereotyping of endogenously constructed social groups is at least as great as the inequality that can emerge between exogenously given groups. Also, the

* Korea Development Institute, Dongdaemoon-gu, Seoul 130-868, South Korea. Email: yckim@kdi.re.kr.
** Department of Economics, Brown University, Providence, RI 02912, USA. Email: Glenn Loury@brown.edu.

model implies that the equal state is not sustainable when the society has enough fraction of members whose identity choice cost is sufficiently low.

1. Introduction

We develop an identity choice model based on a stereotyping-cum-signaling framework. If a worker's productivity is not perfectly observable, employers have an incentive to use the collective reputations of the identity groups to which job applicants belong in the screening process. A key feature of our model is that individuals belonging to a group with a good collective reputation have a greater incentive to invest in skills than do those who belong to a group with a poor reputation. And yet, given its greater rate of investment in skills, the former group will tend to maintain a better collective reputation than the latter. That is, there can be more than one self-confirming equilibrium in group reputations (Arrow, 1973; Coate and Loury, 1993). Previous work on such statistical discrimination has taken group identity as immutable, so that each individual is affected by the collective reputation of his own group only. We handle the dynamics between the collective reputation and the identity choice problem by relaxing this immutability assumption.

To illustrate the relevance of our model, consider one concrete example -the phenomenon commonly known as "passing." This widely observed behavior is evidently an instance of identity choice. Talented young members in the group with a worse collective reputation may consider "passing" into the group with a better reputation when the return for such "passing" (e.g., better treatment in the labor market) outweighs its cost (e.g., loss of ties to one's own kind.) Thus, it has been noted that a significant number of the blacks in the U.S. population consistently passes for White or some other race (Sweet, 2005).[1]

1 The National Longitudinal Survey conducted by the Department of Labor of the US shows that 1.87 percent of those who had originally answered "Black" to the interviewer's race question in 1979 switched to either "White," "I don't know, " or

Moreover, many ethnic Koreans in Japan (most of whom descend from forced laborers in mines and factories who were brought to Japan from the Korean peninsula during the period of Japanese imperialism) are “passing” for native Japanese -by changing both their surnames and their given names when seeking formal employment or marriage. They have done so in order to escape negative stereotypes and prejudices against ethnic Koreans in Japan.[2]

When passing for a member of the advantaged group with high reputation is not possible due to immutability, the most talented of the stereotyped group are more likely to seek styles of self-presentation that aim to communicate “I’m not one of THEM; I’m one of YOU!” because they are the ones who gain most by separating themselves from the mass (Loury, 2002). Taking the example of the Black population in the US, methods that are known to be used for “partial passing” are: affections of speech, dressing up rather than wearing casual clothes, spending more on conspicuous consumption, and migration to affluent residential areas(Charles et al., 2009; Goffman, 1959). There is evidence that the more educated (or talented) blacks tend to speak standard American English rather than African American English (Grogger, 2011). That is, the most talented of the stereotyped group “pass for” the slightly better-off subgroup that maintains a higher reputation than the stereotyped population as a whole by adopting the cultural traits of the subgroup.

This selective out-migration to the better-off subgroup may undermine solidarity in the disadvantaged population and cause conflict among them, such as the accusation of “Acting White” against the ones who practice the “partial passing” methods (Fryer and Torelli, 2010). However, there might be a social gain through this practice: at least some cultural subgroups of the stereotyped population might be able to recover their reputation when the talented young members gather around certain cultural traits. The usage of the observable cultural traits in the screening process can to some extent cure the social inefficiency

---

“other,” by 1998.

2 Every year about 10,000 Koreans living in Japan, out of around 600,000 Korean descendants holding Korean nationality, choose to be naturalized, giving up their names and original nationality (Fukuoka et al., 1998).

caused by imperfect information about the true characteristics of workers.

It is not only backward or disadvantaged groups whose behavior is captured by our model. The emergence of an elite social group out of a population can also be explained through identity choice behavior. Fang (2001) discussed the use of cultural instruments that are intrinsically irrelevant for productivity to form an elite group. He notes that this may help to an account for the complexity of elite etiquette in European (or Confucian) societies, as well as for the greater respect afforded to those with an "Oxford Accent." Skilled and unskilled workers have different incentives to join a group with unique cultural traits that are *expensive* to obtain. Thus, the cultural group is treated preferentially by employers due to the higher fraction of skilled workers, even though the cultural traits of the group are not relevant for productivity. We may see an autonomously growing elite subgroup with differentiated cultural traits whose members are considered as distinguished from their peers.

The identity choice model in this paper starts with a standard statistical discrimination framework (Coate and Loury, 1993). We identify multiple self-confirming prior beliefs, which we call Phenotypic Stereotyping Equilibria (PSE). This multiplicity of phenotypic equilibrium explains inequality of collective reputations between exogenous and equally endowed identity groups as being due to positive feedback between a group's reputation and its members' investment incentives. But it entails no selection into or out of the groups. However, when membership is endogenous and if the groups' reputations were to differ in equilibrium, then not only will members of a favored group face greater incentives to invest in human capital. It is also the case that the better-regarded group will, in equilibrium, come to consist disproportionately of high ability/low human capital investment cost types. These are the types who gain most from joining a favored group. The result is that human capital cost distributions between groups endogenously diverge, which reinforces incentive-feedbacks.

We call such a group-disparate equilibrium with positive selection a (non-trivial) Affective Stereotyping Equilibrium (ASE).

For the development of our theoretical model, we introduce two affects (i.e., possible group identities), A and B. We assume that the cost to choose affect A rather than B varies across the population. Agents choose affect A if and only if the anticipated return exceeds the agent's cost of choosing affect A. We further assume that the cost distribution among agents for this affect choice is statistically independent of the cost distribution of human capital investment. (That is, a person's identity orientation cannot be used to predict their economic abilities.) Nevertheless, we show that in the model's equilibrium the more able individuals -those with lower human capital investment costs -will tend to choose affect A whenever the collective reputation of the affect A group is better than that of the affect B group. But, of course, their behaving in this way is what causes group A to have a better reputation in the first place!

The main result which we demonstrate with this model is the following: group inequality that derives from the ascriptive stereotyping of *endogenously constructed* social groups is at least as great as the group inequality that can emerge from the phenotypic stereotyping of *exogenously given* groups. Again, this is due to the fact that when groups are endogenous then low human capital cost types are disproportionately drawn to the group with a better reputation, causing a skill disparity between groups to endogenously diverge, thereby justifying the disparity of group reputations.

The model also implies that there exist multiple (non-trivial) Affective Stereotyping Equilibria whenever multiple Phenotypic Stereotyping Equilibria can be constructed in the labor market structure. Even more, in the overlapping generation framework, those non-trivial ASE are the only stable equilibria when the society has a critical fraction of newborns whose identity choice cost is sufficiently low. That is, the skill composition of the society converges to a non-trivial ASE in the long run. In addition, we show that non-trivial ASE can exist even under the unique PSE. Even when phenotypic discrimination cannot generate inequality between any identity groups, which could happen due to the uniqueness of the PSE, Affective discrimination may bring about inequality between Affective groups that are endogenously being constructed in a society.[3]

3 The example of Fang (2001) is a special case in which non-trivial ASE exists given

The paper is organized into the following sections. Section 2 develops the model with the identify choice and skill investments. Section 3 defines Phenotypic and Affective Stereotyping Equilibria. Section 4 and Section 5 each identify Affective Stereotyping Equilibria with multiple PSE and with unique PSE. Section 6 presents further discussions and Section 7 concludes.

2. Model with the Identity Choice

In this section, we display the general framework of the model that includes agents' decision making and the payoff structure. The workers make an investment decision on skill acquisition and choose the identity type before they enter the labor market. Employers set the wage for each worker to be proportional to the worker's expected productivity using both the productivity-related noisy signal and the identity type. The workers' decision makings, the employers' wage setting and the consequent expected payoff are discussed below.

Workers′ Affective/Expressive Behavior: Agents choose affect $i \in \{A, B\}$. The cost to choose the affect A is $k \in$ R. k can be negative: the affect A can generate benefits for some agents. CDF of the Affective behavior cost is denoted by $H(k)$. We assume the Affective symmetry: $H(k)$=1- $H(-k)$. Agents choose the affect A if and only if the anticipated return exceeds the agent's cost k. Otherwise, they choose the affect B. WLOG, it is natural to assume that PDF of the cost k, $h(k)$ has one peak at $k = 0 : h'(k) > 0$ for any $k \in (-\infty, 0)$ and $h'(k) < 0$ for any $k \in (0, -\infty)$.[4]

Workers′ Skill Acquisition Behavior: Agents choose whether to be skilled or not: $e \in \{0, 1\}$. The cost to be skilled is c, which is non-negative. CDF of the skill acquisition cost is $G(c)$, in which $G(0) \geq 0$ and

the uniqueness of the PSE that is zero.

4 This is not a critical assumption in the model. We will have the identical conclusions with the uniform distribution of k.

$G(\infty) = 1$.[5] WLOG, it is natural to assume that PDF of the cost *c*, $g(c)$, has one peak at. $\hat{c} : g'(c) > 0$ for any $c \in (0,\ \hat{c})$ and $g'(c) < 0$ for any $c \in (c',\ \infty)$. An agent chooses (e=1) if the return from doing so exceeds that agent′s cost for the skill acquisition (c). We impose that *c* and *k* are independently distributed.

Employers′ Wage-setting Behavior: Skill *e* is not fully identified. Employers observe group identity and noisy signal $t \in R^{+}$ distributed conditional on *e*. PDF of the signal conditional on *e* is $f_e(t)$ and its CDF is $F_e(t)$. Let us define the function $f(\pi,\ t)$ as $f(\pi,\ t) \equiv \pi f_1(t) + (1-\pi) f_0(t)$, which indicates the distribution of signal *t* of agents belonging to a group of which the skill level(the fraction of the skilled workers) is believed to be *π*. WLOG, we assume that $f_1(t)/f_0(t)$ increases with respect to *t*, which is denoted by MLRP: Monotonic Likelihood Ratio Property. The employers' belief that an agent with signal *t* is skilled is $\rho(\pi,\ t)(\equiv \Pr[e=1|\pi,\ t]) = \dfrac{\pi f_1(t)}{f(\pi, t)}$.

Under MLRP, $\rho(\pi,\ t)$ increases with both *π* and *t*. The productivity of a skilled worker is *w* and that of an unskilled worker is zero. We assume that the wage is proportional to the expected skill level:

$$\begin{aligned} W(\pi,\ t) &= w \cdot \rho(\pi,\ t), \quad \text{for some} \quad w > 0 \\ &= w \cdot \frac{\pi f_1(t)}{\pi f_1(t) + (1-\pi) f_0(t)} \end{aligned} \tag{1}$$

Workers' Payoffs: The expected wage from acquiring skill level e is denoted $V_e(\pi)$:

$$V_e(\pi) \int_0^1 f_e\,(t) W(\pi,\ t) \tag{2}$$

in which $V_e'(\pi)$ is positive for any $e \in \{0,\ 1\}$. Workers' expected return acquiring human. capital $(R(\pi))$ is define d as

5 With G(*c*) $\geq$ 0, we allow that a fraction of workers always invest for skills.

$$R(\pi) \equiv V_1(\pi) - V_0(\pi) \tag{3}$$

$R(\pi)$ is expressed as

$$R(\pi) = \int_0^1 f_1(t) - f_0(t)W(\pi,\ t)dt$$
$$= w\pi \int_0^1 \frac{f_1(t) - f_0(t)f_1(t)}{f(\pi,\ t)} dt. \tag{4}$$

The followings can be easily seen

$$R'(\pi) = w\int_0^1 \frac{f_1(t) - f_0(t)f_1(t)f_0(t)}{f(\pi,\ t)^2} dt, \tag{5}$$
$$R''(\pi) = -2w\int_0^1 \frac{(f_1(t) - f_0(\mathrm{t}))^2 f_1(t)f_0(t)}{f(\pi,\ t)^3} dt(< 0) \tag{6}$$

Thus, $R(\pi)$ is concave and $R(0) = R(1) = 0$, which implies that $\lim_{\pi\to 0} R'(\pi) > 0$ and $\lim_{\pi\to 1} R'(\pi) > 0$. Let us denote *argmax* $\{R(\pi)\}$ by $\bar{\pi}$: $R'(\bar{\pi}) = 0$

The first derivatives of $V_0(\pi)$ and $V_1(\pi)$ are

$$V_0'(\pi) = \int_0^1 wf_1(t)f_0(t)^2 f(\pi,\ t)^{-2} dt, \tag{7}$$

$$V_1'(\pi) = \int_0^1 wf_1(t)^2 f_0(t)f(\pi,\ t)^{-2} dt. \tag{8}$$

Note that $V_0'(\pi) = w$ and $V_1'(\pi) = w$. Since we know $R'(0) > 0$ and $R'(1) < 0$, we have $V_0'(1) = w$ and $V_1'(0) = w$. It is more likely that $V_0'(\pi)$ tends to increase as π increases and $V_1'(\pi)$ tends to decrease as π increases. WLOG, we impose that relative marginal benefits $(V_1'(\pi)/V_0'(\pi))$ declines over π. Let us call the property the Marginal Benefits Ratio Property (MBRP):

$$\frac{V_1'(\pi)}{V_0'(\pi)} > \frac{V_1'(\pi+\delta)}{V_0'(\pi+\delta)} \quad \text{for any } \delta > 0.^{6}$$

Thus, a worker with cost c, in a group believed to be investing at rate π has the payoff:

$$U(\pi,\ c) = \max\{V_1(\pi) - c;\ V_0(\pi) \qquad (9)$$

in which the function $U(\pi,\ c)$ is increasing in π ($\because\ V_e'(\pi) > \forall e \in \{0,\ 1\}$) and non-increasing in c.

3. Stereotyping Equilibrium

In this section, we define the Phenotypic and Affective Stereotyping Equilibria and search for the properties of the equilibria.

3.1. Phenotypic Stereotyping Equilibria

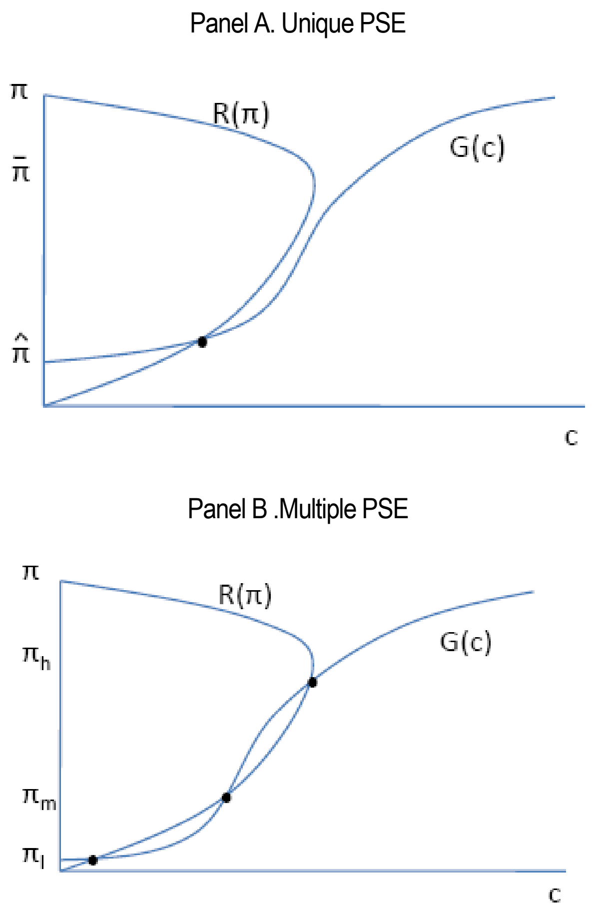
Given the employers' prior belief (π) about human capital investment rate in a population, the fraction of workers who choose (e = 1) is $G(R(\pi))$. Let us denote an equilibrium belief/investment rate by $\hat{\pi} \in [0,\ 1] : \hat{\pi} = G(R(\hat{\pi}))$. The set of all such equilibria is denoted by Ω_{CL} (Coate and Loury 1993). Let us call them Phenotypic Stereotyping Equilibria (PSE). Inequality of collective reputation between exogenous groups in equilibrium is due to feedback between group reputation and individual investment activities. The individuals in a group with a better

6 Consider a simple example that $f_1(t) = 1 - P_1$ for any $t \in (0,\ 1)$ and $f_1(t) = 1 - P_1$ for any $t \in (1,\ 2)$ together with $f_0(t) = 1 - P_0$ for any $t \in (0,\ 1)$ and $f_1(t) = P_0$ for any $t \in (1,\ 2)$. Define $\Pr(e = 1 | 0 < t < 1,\ \pi) = W_N(\pi)$ and $\Pr(e = 1 | 1 < t < 2,\ \pi) = W_P(\pi)$. It is easily seen that $W_N'' > 0$ and $W_P'' < 0$. We have $V_1(\pi) = (1 - P_1)\ W_N(\pi) + P_1 W_P(\pi)$ and $V_0(\pi) = (1 - P_0)\ W_N(\pi) + P_0 W_P(\pi)$. Using these results, we can confirm that the following MBRP property is true for this example: $\frac{\partial[V_1'(\pi) / V_0'(\pi)]}{\partial \pi} < 0$.

collective reputation have a greater incentive to invest in skills, and with their greater skill investment rate, the group maintains a better collective reputation, (and Vice Versa).

It is most likely that there exists either one or three equilibria in the economy, because $G(c)$ is $S-$shaped as displayed in Figure 1. Multiple equilibria $\hat{\pi} \in \Omega_{CL}$ create possibility of Phenotypic Stereotyping (PS) wherein groups are exogenously and visibly distinct, though equally well endowed. Nevertheless, they fare unequally in the equilibrium (Panel 2 of Figure 1).

| Figure 1 | Phenotypic Stereotyping Equilibria

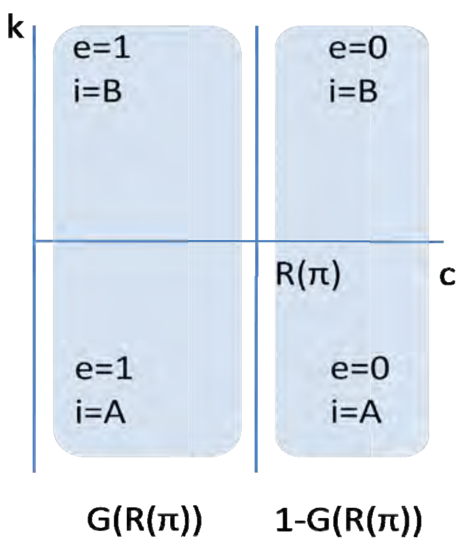


It is clear that the socially optimal level of human capital investment is $G(w)$. However, human capital investment is socially inadequate in any PSE $\hat{\pi}: G(R(\hat{\pi}) < G(w)$ any $\hat{\pi} \in \Omega_{CL}$, because $R(\pi) > w$ as $w\int_0^1 f_1(t)dt - R(\pi) = \quad w\int_0^1 f_1(t)f_0(t)f(\pi,\ t)^{-1}dt > 0$. This implies the social optimality can be further improved through the relaxation of the given constraints.

3.2. Affective Stereotyping Equilibria

Absent Affective discrimination, workers choose 'affect' based on their "natural" orientation: $i = B$ if $k > 0$ and $i = A$ if $k < 0$. This implies that the human capital cost distribution, namely $G(c)$, is the same for both Affective groups. Refer to Figure 2 for the case with Affective discrimination absent. However, when Affective discrimination prevails in the labor market and group membership is endogenous, this is not true any more. Given that the initial reputations of the two Affective groups differ, the favored group not only faces great human capital

| Figure 2 | Human Capital Investment with Affective Discrimination Absent



investment incentives, but it also consists disproportionately of low human capital investment cost types, who gain more from joining a favored group. Thereby, it will cause human capital cost distributions between groups to endogenously diverge, reinforcing incentive-feedbacks.

The economic analysis of the story can be developed in the following way. Let π_i be employer belief about human capital investment rate in Affective group *i*. Consider two affective groups *A* and *B*. Let us define a function $\Delta U(\pi_A,\ \pi_B;\ c)$ as the payoff difference between a A-type worker and a B-type worker given their skill acquisition cost level c: $\Delta U(\pi_A,\ \pi_B;\ c) \equiv \Delta U(\pi_A,\ c) - U(\pi_B,\ c)$ Given $\pi_A > \pi_B$, $\Delta U(\pi_A, \pi_B;\ c)$ is positive because $\partial U(\pi,\ c) \partial \pi > 0$ Note that $\Delta U(\pi_A,\ \pi_B;\ c) \equiv -\Delta U(\pi_B,\ \pi_A;\ c)$ and $\Delta U(\pi,\ \pi;\ c) = 0$.

An agent with the cost set (*c, k*) chooses Affective behavior *i* = *A* if and only if $\Delta U(\pi_A,\ \pi_B;\ c) \geq k$. Otherwise, he chooses Affective behavior *i* = *B*. Given that *c* and *k* are independent, the fraction of agents choosing (*i* = *A*) is given by

$$\Sigma^A = \int_0^\infty H(\Delta\, U(\pi_A,\, \pi_B\,;\ c)\mathrm{d}G(c) \tag{10}$$

The fraction of workers choosing (*i* = *A*) and (*e* = 1) is given by

$$\sigma^A = \int_0^{R(\pi_A)} H(\Delta\, U(\pi_A,\, \pi_B\,;\ c)\mathrm{d}G(c) \tag{11}$$

Then, the fraction of agents choosing (*i* = *B*) is obtained using $\sigma^A = 1 - \Sigma^A$ and $\Delta U(\pi_A,\ \pi_B;\ c) = \Delta U(\pi_A,\ \pi_B;\ c)$

$$\Sigma^B = \int_0^\infty H(\Delta\, U(\pi_A,\, \pi_B\,;\ c)\mathrm{d}G(c) \tag{12}$$

Consequently, the fraction of workers choosing (*i* = *B*) and (*e* = 1) is given by

$$\sigma^B = \int_0^{R(\pi_B)} H(\Delta U(\pi_A, \pi_B; c)\mathrm{d}G(c) \tag{13}$$

Given the employer belief about human capital investment rates (π_A, π_B), the actual investment rates for the Affective groups denoted by $\phi(\pi_A, \pi_B)$ and $\phi(\pi_B, \pi_A)$ are

$$\begin{cases} \Pr\{e = 1 | e = A, \pi_A, \pi_B\}(\equiv \phi(\pi_A, \pi_B)) = \sigma^A / \Sigma^A, \\ \Pr\{e = 1 | e = B, \pi_B, \pi_A\}(\equiv \phi(\pi_B, \pi_A)) = \sigma^B / \Sigma^B. \end{cases} \tag{14}$$

It is noteworthy that when employers' belief is the same for both Affective groups $(\pi_A = \pi_B)$, $\Delta U(\pi_B, \pi_A; c)$ is zero and we have $R(\pi_A) = R(\pi_B)$ This implies that the Affective behavior does not affect the human capital investment activities: $\phi(\pi_A, \pi_B) = \phi(\pi_B, \pi_A)$ $(= G(R(\pi_A)))$.

An equilibrium with Affective stereotyping (ASE) is define d as a pair of investment rates for the Affective groups $(\pi_A^*, \pi_B^*) \in [0, 1]^2$ such that $\pi_A^* = \phi(\pi_A^*, \pi_B^*)$ and $\pi_B^* = \phi(\pi_B^*, \pi_A^*)$. The set of all such equilibria is denoted by Ω_{KL}. Note that every PSE corresponds to trivial ASE where differences in affect are uninformative:

$(\hat{x}, \hat{x}) \in \Omega_{KL}$ If $\hat{x} \in \Omega_{CL}$ because $\phi(\hat{x}, \hat{x}) = G(R(\hat{\pi})) = \hat{x}$. Affective stereotyping discrimination occurs if and only if $\pi_A^* \neq \pi_B^*$.

For notation simplicity, we use *a* and *b* instead of π_A and π_B $\Delta U(a, b; c)$ can be expressed by

$$\Delta U(a, b; c) = \max\{R(a) - c; 0\} + V_0 - \max\{R(b) - c; 0\} - V_0(b). \tag{15}$$

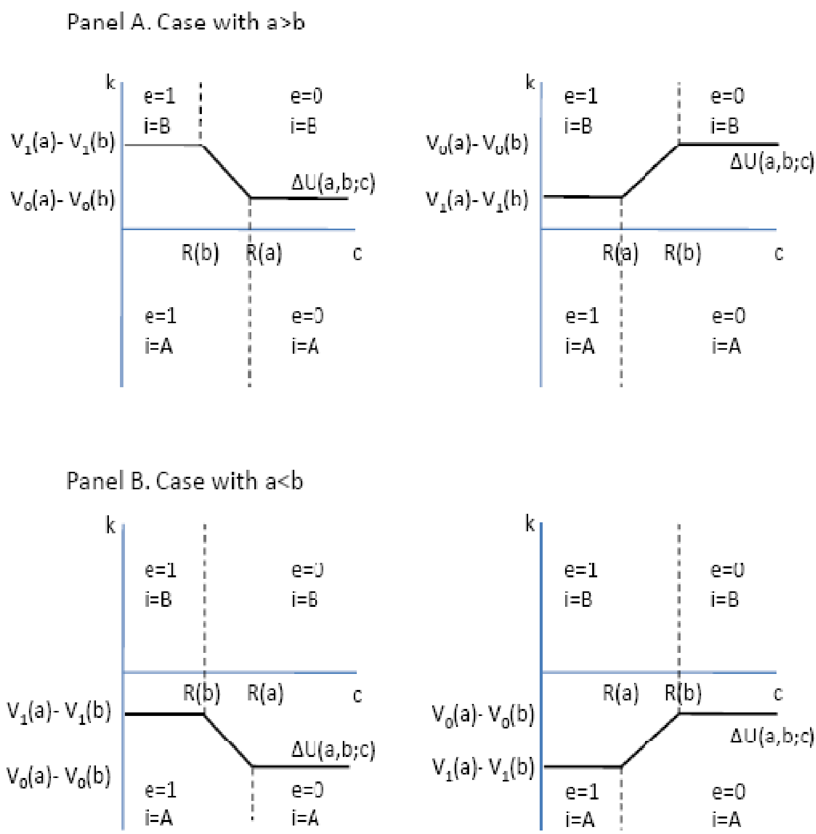
Using $R(a)$ and $R(b)$, we have the following lemma concerning $\Delta U(a, b; c)$:

Lemma 1. *For any c* ≤ $\min\{R(a), R(b)\}$, *ΔU(a, b; c)*= $V_1(a) - V_1(b)$. *For any c* ≥ $\max\{R(a), R(b)\}$, *ΔU(a, b; c)*= $V_0(a) - V_0(b)$. *For any c such that* $\min\{R(a), R(b)\} < c < \max\{R(a), R(b)\}$, *we have*

$$\Delta U(a,\ b;\ c)\begin{cases} V_1(a)-V_0(b)-c & \text{if } R(a)\geq R(b), \\ V_0(a)-V_1(b)+c & \text{if } R(a)<R(b). \end{cases} \tag{16}$$

The above lemma is summarized in Figure 3. Panel 1 of the figure displays the case with *a*>*b* and panel 2 does the case with *a*<*b*. It is easily seen that $\Delta U(a,\ b;\ c)>0$ for any c if and only if *a*>*b*. Therefore, we have the following result.

| Figure 3 | Human Capital Investment and Affective Behavior



Proposition 1. *When employers have different beliefs about two Affective groups* $(\pi_A \neq \pi_B)$, *the number of workers who adopt the 'affect' corresponding to the favored employers' belief is greater than that of workers who adopt the 'affect' with the less favored employers' belief:* $\Sigma^i > \Sigma^j$ *if* $\pi_i > \pi_j$ *for any* $i, j \in \{A, B\}$

That is, in the current setting with symmetric cost distribution, more than half workers adopt the 'affect' that corresponds to the more favorable employers' belief: Σ^i .5 and $\Sigma^j < .5$ if $\pi_i > \pi_j$. The Lemma 1 implies that $\Delta U(a, b; c)$ is non-increasing with respect to c whenever *R*(*a*)>*R*(*b*), and non-decreasing whenever *R*(*b*)>*R*(*a*). It leads to the following useful lemma.

Lemma 2. *Whenever R*(*a*)= *R*(*b*), *the following holds:* $\phi(a, b) > \phi(a, a)$ *and* $\phi(b, a) < \phi(b, b)$. *In a symmetric way, whenever R*(*a*)< *R*(*b*), *the following holds:* $\phi(a, b) < \phi(a, a)$ *and* $\phi(b, a) > \phi(b, b)$. *When R*(*a*)= *R*(*b*) *and* $a \neq b$, *the following holds:* $\phi(a, b) = \phi(b, a) = \phi(a, a) = \phi(b, b)$

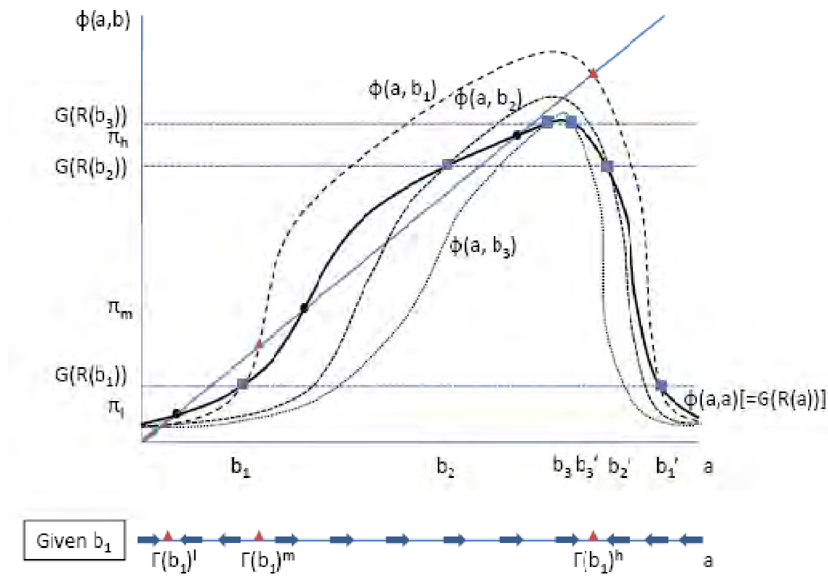
The above lemma implies the following proposition.

Proposition 2. *The disproportionately more talented workers, whose human capital investment costs (c) are relatively lower, choose the 'affect' that corresponds to the greater return to human capital investment: given R*(*i*) >*R*(*j*), ϕ (*i, j*) >*G*(*R*(*i*)) *and* ϕ (*j, i*) <*G*(*R*(*j*)) *for any i, j* $\in \{a, b\}$.

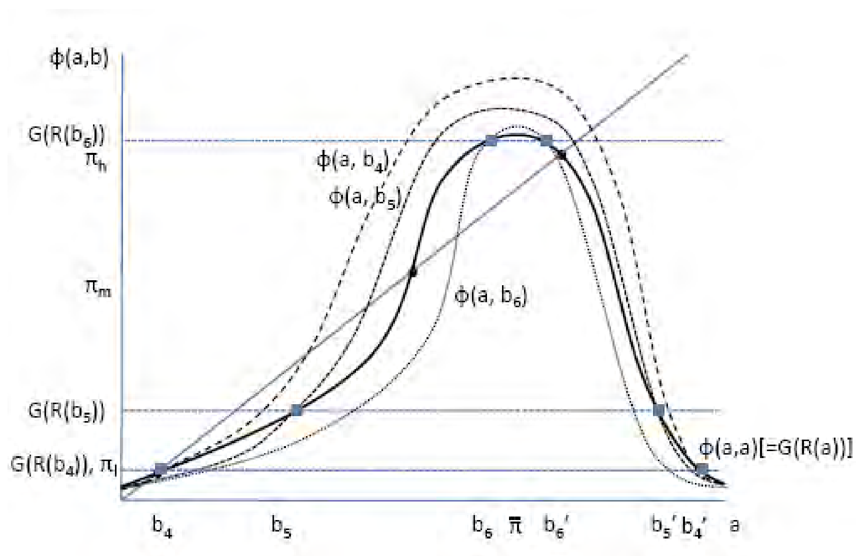
For any *b* except for $\bar{\pi}$, we can find b'. such that $R(b) = R(b')$. The following should hold for the combination $(b, b') : \phi(b', b) = G(R(b))$. The overall shape of $\phi(a, b)$ is displayed in Panel A of Figure 4 for three different levels of *b* below $\bar{\pi}$, $b_1 < b_2 < b_3 < \bar{\pi}$, together with the the shape of $\phi(a, a)(= G(R(b)))$, in which $\bar{\pi} > \pi_h$. Also, Panel B of the figure displays the shape of $\phi(a, a)$ for the case with $\bar{\pi} > \pi_h$ and the overall shape of $\phi(a, b)$ for three different levels of b below $\bar{\pi}$, $b_4 < b_5 < b_6 < \bar{\pi}$. Note that the $\phi(a, b)$ curve

| Figure 4 | Human Capital Investment Rate

Panel A. Case with $\bar{\pi} > \pi_h$



PanelBA. Case with $\bar{\pi} > \pi_h$



intercepts the $\phi(a, a) = G(R(a))$ curve at $a = b$ and $a = b'$. We have the following lemma for the relative positions of $\phi(a, b)$ s.

Lemma 3. *For any* b_1 *and* b_2 *such* $b_1 < b_2$, $\phi(a, b_1)$ *is placed above* $\phi(a, b_2)$: $\phi(b_2, \phi(a, b_1) > \phi(a, b_2), \forall a \in (0, 1)$. *Also, for any* b_1 *and* b_2 *such that* $\bar{\pi} < b_1 < b_2$, $\phi(a, b_{12})$ *is placed above* $\phi(a, b_1)$: $\phi(a, b_2) > \phi(a, b_1), \quad \forall a \in (0, 1)$.

Proof. See the proof in the appendix. ■

The following lemma helps us understand how the $\phi(a, b)$ curve cross over the $\phi(a, a)$ curve:

Lemma 4. *The slope of the* $\phi(a, b)$ *curve at the point where it crosses over the* $\phi(a, a)$ *curve is*

$$\frac{\partial \phi(a, b)}{\partial a}\Big|_{a=b} \approx g(R(b))R'(b) + 2H'(0)R'(b)G(R(b))(1 - G(R(b)). \quad (17)$$

Proof. See the proof in the appendix. ■

The above lemma implies that the slope of $\phi(a, b)$ at the crossing point is positive (negative) whenever $R'(b)$ is positive (negative). Also, the slope of $\phi(a, b)$ at the crossing point is greater (smaller) than the slope of $\phi(a, a)(= g(R(b))R'(b)$ whenever $R'(b)$ is positive (negative).

4. Affective Stereotyping Equilibria with Multiple PSE

Let us define a correspondence $\Gamma(y)$

$$\Gamma(y) = \{x : x = \phi(x, y)\}. \quad (18)$$

Note that any $\hat{\pi} \in \Omega_{CL}$ satisfies $\hat{\pi} \in \Gamma(\hat{\pi})$ and any $\hat{\pi} \in \Gamma(\hat{\pi})$ satisfies $\hat{\pi} \in \Omega_{CL}$. Thus, the set of PSE is represented as follows using the correspondence: $\Omega_{CL} = \{x : x \in \Gamma(x)\}$. The set of Affective stereotyping equilibria can be expressed as $\Omega_{KL} = \{x : y) : x \in \Gamma(y)$ and $y \in \Gamma(x)\}$.

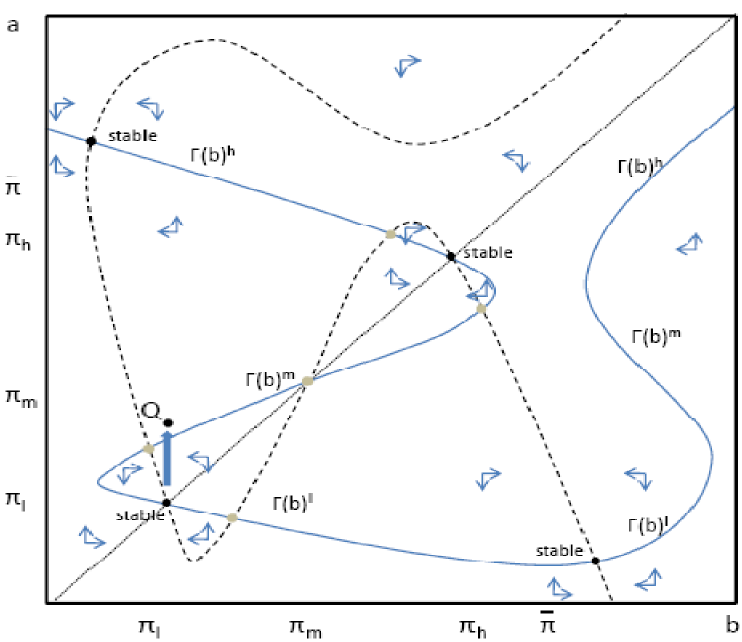
Consider the case with multiple PSE. WLOG, we assume that there are three: π_h, π_m and π_l. We will examine the case with a unique PSE in the next section.

4.1. Existence of Affective Stereotyping Equilibria

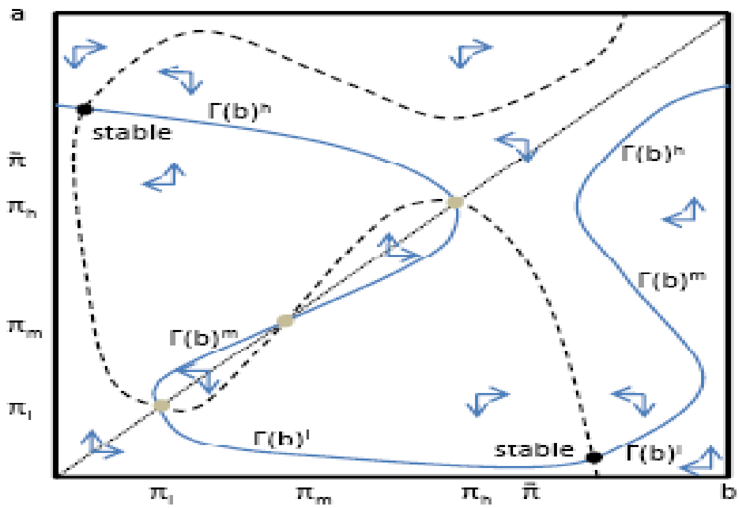
It is most likely that there exist either one or three values in $\Gamma(y)$. Let us denote the three values by $\Gamma(y)^h$, $\Gamma(y)^m$ and $\Gamma(y)^l$ as displayed in Figures 5 and 6. Panels A, B and C of Figure 5 describe the case with $\bar{\pi} > \pi_h$ and Panel A of Figure 6 the case with $\bar{\pi} < \pi_h$. If there exists a unique value for some range of y, $\Gamma(y)$ with its unique value is denoted by $\Gamma(y)^i$ as it is connected to nearby $\Gamma(y)^i$ for $i \in \{h, m, l\}$, which is an element of $\Gamma(y)$ with multiple values. We can infer the following result using Lemma 3.

| Figure 5 | ASE given Multiple PSE: Case with $\bar{\pi} > \pi_h$

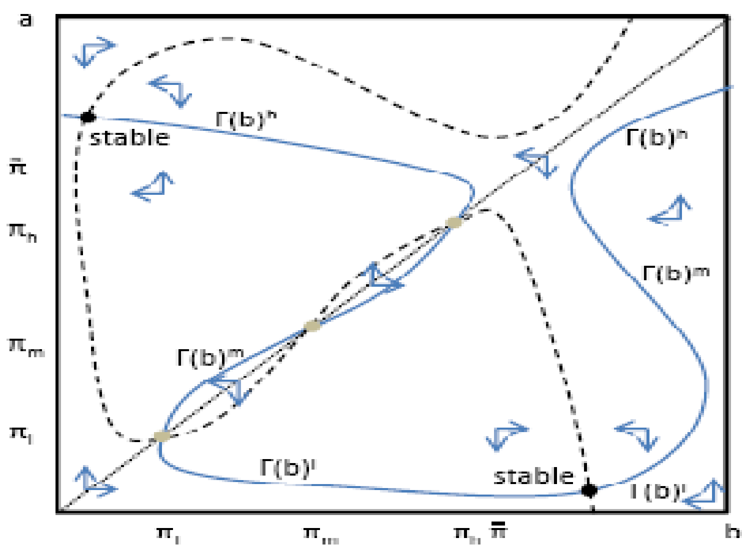
Panel A. Given both $-1 < \Gamma'(\pi_h) < 0$ and $-1 < \Gamma'(\pi_l) < 0$



Panel B. Given both $\Gamma'(\pi_h) < 1$ and $\Gamma'(\pi_l) < -1$



Panel C. Given both $\Gamma'(\pi_h) > 1$ and $\Gamma'(\pi_l) > 1$



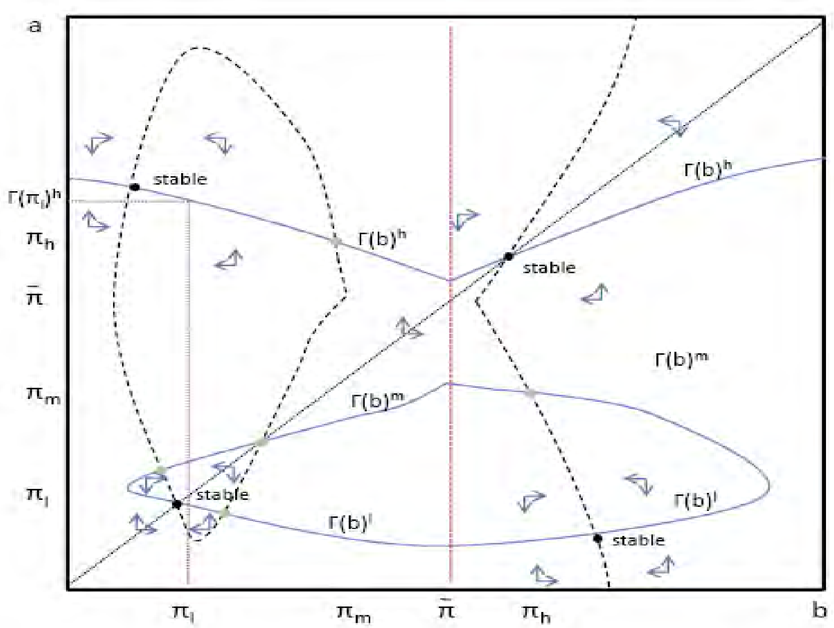
Lemma 5. *For any y below* $\Gamma(y)^h$ *and* $\Gamma(y)^l$ *decrease in y and* $\Gamma(y)^m$ *increases in y, while* $\Gamma(y)^h$ *and* $\Gamma(y)^l$ *increase in y and* $\Gamma(y)^m$ *decreases in y for any y above π. Also, we have* $\pi_h < \Gamma(0)^h < 1$ *and* $\pi_h < \Gamma(1)^h < 1$.

This lemma also implies that min $\Gamma(y)^l = \Gamma(\pi)^l$ and arg min $\Gamma(y)^l = \pi$. $\Gamma(a)$ and $\Gamma(b)$ are overlapped in Figure 5. Using the local linearization described in Appendix Figure 2, we can calculate the slope of correspondence curve at each trivial ASE, $\Gamma'(\pi)$.

Lemma 6. The slope of correspondence curve at a trivial ASE $(\hat{x},\ \hat{x})$, which is denoted by $\Gamma'(\hat{x})$, is approximated by

$$\Gamma'(\hat{x}) \approx \frac{2H'(0)R'(\hat{x})\hat{x}(1-\hat{x})}{g(R(\hat{x}))R'(\hat{x}) - 1 + 2H'(0)R'(\hat{x})\hat{x}(1-\hat{x})} \tag{19}$$

| Figure 6 | ASE given Multiple PSE: Case with $\bar{\pi} < \pi_h$
PanelA. Given $0 < \Gamma'(\pi_h) < 1$ and $-1 < \Gamma'(\pi_l) < 0$



Proof. See the proof in the appendix. ■

Using the above lemma, we can describe the correspondence curves more accurately. First, note that the slope of $\Gamma(\pi)$ at trivial ASE (π_m, π_m) always satisfies $0<\Gamma'(\pi_m)<1$, because the slope of the $\phi(a,\ a)$ curve at $a=\pi_m$ is greater than one: $g(R(\pi_m))R'(\pi_m)>1$. Secondly, only when $\overline{\pi}<\pi_h$ as shown in Panel B of Figure 4, we have $R'(\pi_h)<0$ Then, we have $0<\Gamma'(\pi_m)<1$ of Lemma 6, as displayed in Panel A of Figure 6. (However, note that even when $\overline{\pi}<\pi_h$, we have $R'(\pi_l)<0$.) Thirdly, as far $\overline{\pi}<\pi_h$, any PSE $\hat{\pi}$ satisfies $R'(\hat{\pi})>0$. We have the following summary for all of the above cases.

Lemma 7. *The slope of correspondence at trivial ASE (π_m, π_m) always satisfies* $0<\Gamma'(\pi_m)<1$ *Given* $R'(\pi_h)<0$*, the slope of correspondence at trivial ASE (π_h, π_h) is* $0<\Gamma'(\pi_h)<1$. *Given* $R'(\hat{x})>0$ *for* $\hat{x}\in\{\pi_h,\ \pi_l\}$, *the slope of correspondence at a trivial ASE* $(\hat{x},\ \hat{x})$ *depends on the the density of identity cost k around zero,* $H'(0)$:

$$\begin{cases} -1<\Gamma'(\hat{x})<0 \quad \text{if} \quad H'(0)<\dfrac{1-g(R(\hat{x}))R'(\hat{x})}{4R'(\hat{x})\hat{x}(1-\hat{x})} \\[2ex] \Gamma'(\hat{x})<-1 \quad \text{if} \quad \dfrac{1-g(R(\hat{x}))R'(\hat{x})}{4R'(\hat{x})\hat{x}(1-\hat{x})}<H'(0)<\dfrac{1-g(R(\hat{x}))R'(\hat{x})}{2R'(\hat{x})\hat{x}(1-\hat{x})}, \quad \forall\hat{x}\in\{\pi_h,\ \pi_l\}. \\[2ex] \Gamma'(\hat{x})>1 \quad \text{if} \quad H'(0)>\dfrac{1-g(R(\hat{x}))R'(\hat{x})}{2R'(\hat{x})\hat{x}(1-\hat{x})} \end{cases} \tag{20}$$

Proof. Given $R'(\hat{x})>0$ for $\hat{x}\in\{\pi_h,\ \pi_l\}$, we have $0<g(R(\hat{x})R'(\hat{x})<1$. Under this condition, Lemma 6 derives the given result. ■

The lemma implies that given $R'(\hat{x})>0$, when the sensitivity of identity choice represented by $H'(0)$ is above a certain level, $\dfrac{1-g(R(\hat{x}))R'(\hat{x})}{4R'(\hat{x})\hat{x}(1-\hat{x})}$, the absolute value of the slope of correspondence Curve $|\Gamma'(\hat{x})|$ at the trivial ASE $(\hat{x},\ \hat{x})$ is greater than one.

Theorem 1. *Given multiple PSE, there always exist at least two non-trivial ASE.*

Proof. See the proof in the appendix. ■

At least two non-trivial ASE exist as far as multiple PSE exist. Whether there are more than two ASE or not depends on the curvature of $\Gamma(a)$ and $\Gamma(b)$ around trivial ASE $(\hat{x},\ \hat{x})$. The slope of correspondence at a trivial ASE is important to examine the exact number of non-trival ASE. WLOG, the condition $\left|\Gamma'(\hat{x})\right| < 1$ for $\hat{x} \in \{\pi_h,\ \pi_l\}$ generates two more nontrivial ASE near to a trivial ASE $(\hat{x},\ \hat{x})$, while the condition $\left|\Gamma'(\hat{x})\right| < 1$ for $\hat{x} \in \{\pi_h,\ \pi_l\}$ does not generate such non-trivial ASE around a trivial ASE $(\hat{x},\ \hat{x})$. Refer to Panel A of Figure 5 and Panel A of Figure 6 for the case with the condition $\left|\Gamma'(\hat{x})\right| < 1$ for $\hat{x} \in \{\pi_h,\ \pi_l\}$, and Panels B and C of Figure 5 for the case with the condition $\left|\Gamma'(\hat{x})\right| > 1$ for $\hat{x} \in \{\pi_h,\ \pi_l\}$

Proposition 3. *WLOG, it is most likely that the number of non-trivial ASE is six when both* $\left|\Gamma'(\pi_h)\right| < 1$ *and* $\left|\Gamma'(\pi_l)\right| < 1$ *and it is only two when both* $\left|\Gamma'(\pi_h)\right| > 1$ *and* $\left|\Gamma'(\pi_l)\right| > 1$.

Panel A of Figure 5 and Panel A of Figure 6 display six non-trivial ASE given $\left|\Gamma'(\pi_h)\right| < 1$ and $\left|\Gamma'(\pi_l)\right| < 1$, and Panels B and C of Figure 5 display two non-trivial ASE given $\left|\Gamma'(\pi_h)\right| > 1$ and $\left|\Gamma'(\pi_l)\right| > 1$. Let us call the two non-trivial ASE that exist regardless of the curvatures of the correspondences $\Gamma(a)$ and $\Gamma(b)$ "Persistent ASE," and denote them $(\pi_H^{**},\ \pi_L^{**})$ and $(\pi_L^{**},\ \pi_H^{**})$.

Proposition 4. *The two "Persistent ASE",* $(\pi_H^{**},\ \pi_L^{**})$ *and* $(\pi_L^{**},\ \pi_H^{**})$, *that consistently exist given multiple PSE (regardless of* $\left|\Gamma'(\pi_h)\right|$ *and* $\left|\Gamma'(\pi_l)\right|$ *satisfy*

$$\pi_L^{**} < \min\{\Omega_{CL}\} << \max\{\Omega_{CL}\}\ \pi_H^{**} \tag{21}$$

Proof. See the proof in the appendix. ■

The proposition implies that inequality between endogenous groups in some non-trivial ASE can be greater than inequality between exogenous groups in any PSE.

4.2. Stability of Affective Stereotyping Equilibria

Consider an intergenerational population structure. Every period, the randomly chosen α fraction of the workers die and the same number of agents are newly born. The newborn agents incur the cost c of skill achievement and the cost k to choose the affect *A: k* can be negative. Each newborn agent with his cost set (*c, k*) decides whether to invest for skills or not and which 'affect' to choose among *A* and *B* in the early days of his life. After those days of education and affect adaption, newborns join the labor market and receive wage set by employers. We assume that employers set the newborns' lifetime wage $W(\pi, t)$ proportional to the estimated skill level $\rho(\pi, t)$: $W(\pi_j, t) = w \cdot \rho(\pi_j, t)$ for the entering newborns with group identity $j \in \{A, B\}$ and the noisy signal *t*, given $\rho(\pi_j, t) = \pi_j f_1(t) / f(\pi_j, t)$. Employers use the skill composition of the current workers belonging to identity group *j* to estimate π_j. Therefore, we have the following dynamics:

$$\dot{\pi}_A > (<) \Leftrightarrow \phi(\pi_A, \pi_B) > (<) \pi_A, \tag{22}$$
$$\dot{\pi}_B > (<) \Leftrightarrow \phi(\pi_B, \pi_A) > (<) \pi_B, \tag{23}$$

The direction arrows in Panel A of Figure 4 describe the law of motions of π_A given π_B fixed as $\dot{\pi}_A > 0$ for any b_1: $\pi_A \in (0, \Gamma(b_1)^l)$ for any $\pi_A \in (\Gamma(b_1)^m, \Gamma(b_1)^h)$ and $\dot{\pi}_A < 0$ for any $\pi_A \in (\Gamma(b_1)^l, \Gamma(b_1)^m)$, and any $\pi_A \in (\Gamma(b_1)^h, 1)$. Therefore, direction arrows of $\dot{a}$ are upward between $\Gamma(b)^h$ and $\Gamma(b)^m$ and below $\Gamma(b)^l$ in the (*b, a*) plain, and downward between $\Gamma(b)^m$ and $\Gamma(b)^l$ and above $\Gamma(b)^h$. The direction arrows of b˙are rightward between $\Gamma(a)^h$ and $\Gamma(a)^m$ and at the lefthand side of $\Gamma(a)^l$ in the (*b, a*) plain, and leftward between $\Gamma(a)^m$ and $\Gamma(a)^l$ and at the righthand side of $\Gamma(a)^h$, as displayed in Figures 5 and 6. From the described direction arrows, we can infer the following theorem.

Theorem 2. *Given Multiple PSE, two "Persistent ASE",* $(\pi_H^{**},\ \pi_L^{**})$ *and* $(\pi_L^{**},\ \pi_H^{**})$, *are stable and all other non-trivial ASE are unstable.*

The theorem together with Proposition 4 implies that inequality between endogenous groups in non-trivial ASE should be greater than inequality between exogenous groups in any PSE in the long run, because stable non-trivial ASE must be "Persistent ASE."

Proposition 5. *The middle trivial ASE* $(\pi_m,\ \pi_m)$ *is always unstable. Other trivial ASEs,* $(\pi_h,\ \pi_h)$ *and* $(\pi_l,\ \pi_l)$, *are stable if* $\left|\Gamma'(\hat{x})\right| \leq 1$ *and unstable if* $\left|\Gamma'(\hat{x})\right| > 1$

Using the direction arrows, we can easily confirm the above proposition as well. Therefore, given $\overline{\pi} < \pi_h$, the trivial ASE $(\pi_h,\ \pi_h)$, is stable because of $0 < \Gamma'(\pi_h) < 1$ (Lemma 7). Using Lemma 7 and the above proposition, we have the following result.

Theorem 3. *Given* $R'(\hat{\pi}) > 0,\ \forall \hat{x} \in \Omega_{CL}(=\{\pi_l,\ \pi_m,\ \pi_h\})$, *the trivial ASE* $(\pi_m,\ \pi_m)$ *is unstable and other trivial ASE, ,* $(\pi_h,\ \pi_h)$ *or* $(\pi_l,$ $\pi_l)$, *is stable if and only if* $H'(0) \leq \dfrac{1-g(R(\hat{x}))R'(\hat{x})}{4R'(\hat{x})\hat{x}(1-\hat{x})}$, *for* $\hat{x} \in \{\pi_h,\ \pi_l\}$.

The theorem implies the following interesting result:

Corollary 1. *Given* $R'(\hat{\pi}) > 0,\ \forall \hat{x} \in \Omega_{CL}(=\{\pi_l,\ \pi_m,\ \pi_h\})$, *the stable ASE are "Persis tent ASE",* $(\pi_H^{**},\ \pi_L^{**})$ *and* $(\pi_L^{**},\ \pi_H^{**})$, *and all other ASE are unstable if and only if* $H'(0) \leq \dfrac{1-g(R(\hat{x}))R'(\hat{x})}{4R'(\hat{x})\hat{x}(1-\hat{x})},\ \forall \hat{x} \in \{\pi_h,\ \pi_l\}$

Therefore, when the society has enough fraction of newborns whose identity choice cost k is very low (i.e. $H'(0)$ is sufficiently big), balanced skill rates between two identity groups, $(\pi_h,\ \pi_h)$ *or* $(\pi_l,\ \pi_l)$, are not sustainable due to the incentives for the talented members to choose the "affect" associated with the slightly better collective reputation. The skill composition of the society converges to a non-trivial ASE in the long run, in which inequality between endogenous identity groups is greater than that of exogenous groups in any PSE:

$$\left|\pi_H^{**},\ \pi_L^{**}\right| > \left|\pi_i - \pi_j\right|,\ \forall i \in \{l,\ m,\ h\}$$

Now imagine that the society is trapped by the low skill investment rates: the society is placed in a stable ASE (π_l, π_l). As far as two identity groups are feasible and the identity choice is available for a fraction of workers, the social coordinator such as a government can mobilize the society to move out of the low investment trap by treating one of the identity groups favorably. The favorable treatment will lead more talented newborns to join the selected identity group. The skill level of the group can improve quickly with the higher skill investment activities of the newborns and by joining disproportionately more talented newborns to the group. However, the skill level of the other group which is not supported by the social coordinator may continue to be left behind in the low skill investment trap. For example, as shown in Panel A of Figure 5, the governmental intervention to relocate the skill composition from (π_l, π_l) to the point Q in the basin of attraction to $(\pi_L^{**},\ \pi_H^{**})$ can mobilize the society to carry the much enhanced skill investment activities and, consequently, to arrive at a "Persistent ASE" $(\pi_L^{**},\ \pi_H^{**})$ in which overall skill rate of the economy is much greater than the original skill rate π_l .

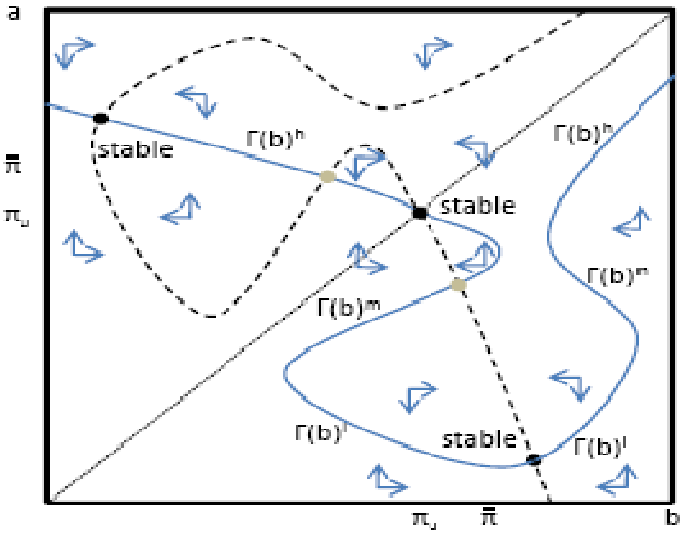
Proposition 6. *When the society is in low skill investment trap (*π_l*, *π_l*), the Affective stereotyping may improve the social e.ciency as the skill composition of the society can move to a "Persistent ASE" with a little push for an identity group to advance.*

5. Affective Stereotyping Equilibria with Unique PSE

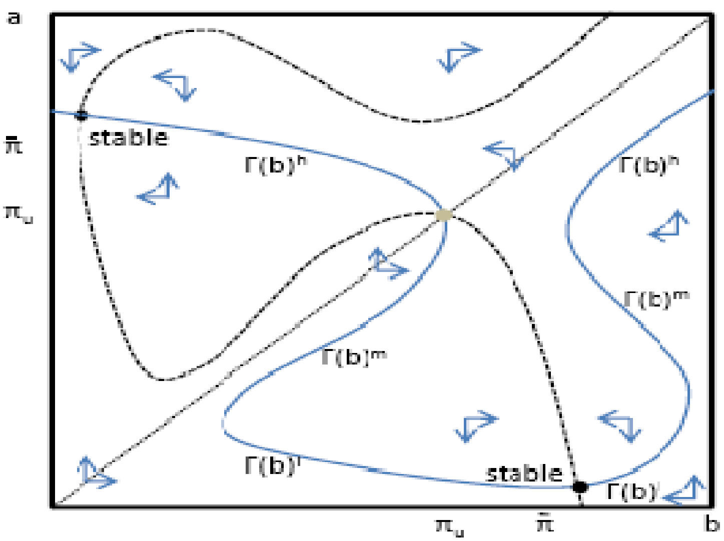
In this section, we consider the case with unique PSE Let us denote it by π_u : $G(R(\pi_u) = \pi_u$. We show that non-trivial ASE can exist even under the unique PSE. It is surprising that even when phenotypic

| Figure 7 | ASE given Unique PSE: Case with $\bar{\pi} < \pi_u$

Pane lA. Given $-1 < \Gamma'(\pi_u) < 0$



Panel B. Given $\Gamma'(\pi_u) < -1$



discrimination cannot generate the inequality between any groups, Affective discrimination may bring about the inequality between Affective groups forming endogenously in a society.

5.1. Existence of Affective Stereotyping Equilibria

Every PSE corresponds to trivial ASE: a trivial ASE $(\pi_u,\ \pi_u)$ exists which satisfies $\phi(\pi_u,\ \pi_u) = \pi_u$. $\Gamma(b)$ passes through the points $(\pi_u,\ \pi_u)$ and a-intercept $(b,\ a) = (0,\ \Gamma(b)^h)$, in which $\pi_u < \Gamma(0)^h < 1$. $\Gamma(a)$ passes through $(\pi_u,\ \pi_u)$ and b-intercept $(a,\ b) = (1,\ \Gamma(1)^h)$, in which $\pi_u < \Gamma(1)^h < 1$. Therefore, as far as $|\Gamma'(\pi_u)| > 1$, there should be at least one non-trivial ASE which satisfies $\pi_A^*,\ \pi_B^*$ and at least one non-trivial ASE which satisfies $(\pi_B^*,\ \pi_A^*)$.(An example is described in Panel B of Figure 7 given $\Gamma'(u_u) < -1$.) Because $\Gamma(b)^h$ is decreasing when $b < \overline{\pi}$, WLOG, there are two non-trivial ASE given $|\Gamma'(\pi_u)| > 1$.

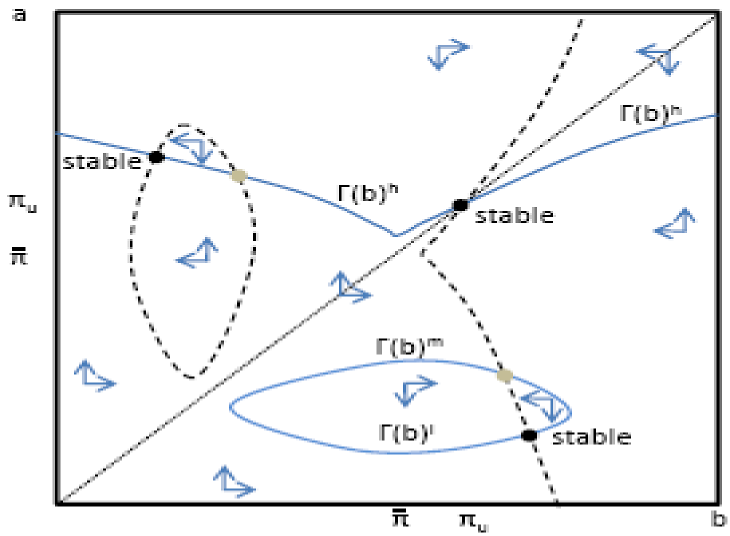
Proposition 7. *Given unique PSE (π_u) and $|\Gamma'(\pi_u)| > 1$, WLOG, there exist two non-trivial ASE.*

However, the existence of non-trivial ASE is not guaranteed when $|\Gamma'(\pi_u)| > 1$. Panel A of Figure 7 and Panel A of Figure 8 show cases with existing non-trivial ASE while Panel B of Figure 8 shows a case without existing non-trivial ASE. Given $|\Gamma'(\pi_u)| < 1$, the curvature of $\Gamma(\pi)$ is critical for the determination of non-trivial ASE's existence: the closer the $\phi(x,\ y)$ curve is to the 45 degree line, the more likely that non-trivial ASE exist. If any non-trivial ASE exists, WLOG, it is most likely that there are four non-trival ASE given $|\Gamma'(\pi_u)| > 1$.

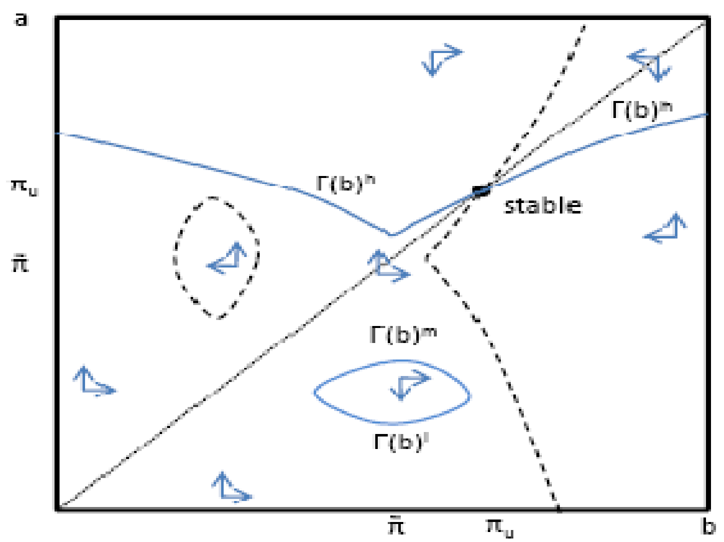
Corollary 2. *Given unique PSE (π_u) and $|\Gamma'(\pi_u)| > 1$, the existence of non-trivial ASE depends on the curvature of $\Gamma(\pi)$. Once they exist, WLOG, there are four non-trivial ASE.*

| Figure 8 | ASE given Unique PSE: Case with $\overline{\pi} < \pi_u$

Panel A. Multiple ASE given $0 < \Gamma'(\pi_u) < 1$



Panel B. Unique ASE given $0 < \Gamma'(\pi_u) < 1$



With the careful examination of the relative position of $\Gamma(a)$ and $\Gamma(b)$, we can confirm the following result:

Proposition 8. *Given unique PSE* ($\pi_u \in (0, 1)$)*, any pair of non-trivial ASE,* (π_H^{**}, π_L^{**}) *and* (π_L^{**}, π_H^{**}) *, satisfies the following condition:*

$$\pi_l^* < \pi_h^* < \pi_u \ . \tag{24}$$

At any non-trivial ASE, the collective reputation of an Affective group is better than the PSE level π_u and that of the other Affective group is worse than the level π_u.

5.2. Stability of Affective Stereotyping Equilibria

Using the direction arrows in phase diagrams in Figures 7 and 8, we can confirm the following results:

Proposition 9. *When two non-trivial ASE exist, both of them are stable. When four nontrivial ASE exist, two of them closer to the 45 degree line are unstable and the other two near the corners are stable*.

Lemma 8. *The trivial ASE* (π_u, π_u) *is stable if* $|\Gamma'(\pi_u)| \leq 1$ *and unstable if* $|\Gamma'(\pi_u)| > 1$

Note that, given $R'(\pi_u) > 0$, we have $|\Gamma'(\pi_u)| > 1$ if and only if

$$H'(0) \leq \frac{1 - g(R(\pi_u))R'(\pi_u)}{4R'(\pi_u)\pi_u(1-\pi_u)}$$

(Lemma 7). Therefore, we achieve the following interesting result:

Proposition 10. *Given* $R'(\pi_u) < 0$*, the trivial ASE* (π_u, π_u) *is stable because* $0 < \Gamma'(\pi_u) < 1$*. Given* $R'(\pi_u) > 0$*, it is stable if and only if*

$$H'(0) \leq \frac{1 - g(R(\pi_u))R'(\pi_u)}{4R'(\pi_u)\pi_u(1-\pi_u)}$$

Theorem 4. Given $R'(\pi_u) > 0 \quad H'(0) \leq \dfrac{1 - g(R(\pi_u))R'(\pi_u)}{4R'(\pi_u)\pi_u(1-\pi_u)}$

the only stable ASE are nontrivial ones while the trivial ASE $(\pi_u,\ \pi_u)$ *is unstable.*

Therefore, when the society has enough newborns whose ‘affect’ choice cost k is low, the equal society cannot be stable due to the emerging Affective stereotyping. The society must converge to a non-trivial ASE, in which one group’s skill level is greater than πu and another group’s skill level is less than π_u.

6. Further Discussions

To see how our model can be used to shed light on labor market phenomena other than ethnic/racial group inequality, let us consider the problem of “re-branding ex-cons.” The main feature of the labor market for ex-convicts is that employers wish to avoid associating with those who end up returning to criminal activity, but employers cannot be certain from information which among the convicts will and which will not do so. However, the convicts themselves are presumed to know their own intentions. Under this informational asymmetry, employers choose not to hire any ex-cons leaving all of them unemployed.

Let us imagine a *re-branding program* run by the government along the following line: There is to be a certifiableand costly activity such that, before going into the labor market, ex-cons can elect to join this program or not. Those who will go straight are more willing to join the program than those who will return to crime because those going straight have greater gain from having a job in the market. The collective reputation of the ex-con subgroup with the program certi.cate improves with the greater fraction of ex-cons going straight joining the program, so that those with the program certi.cate will be hired and those without it will not be hired by employers in the labor market. That is, notwithstanding the informational asymmetry and the adverse selection issue in the market, a government can nevertheless design a costly program by means of which some ex-cons can credibly convey their good intentions to employers. The social e.ciency can be achieved by the introduction of the costly program.

Also, Fang(2001) discusses the emergence of an elite social group out of a population. Eventually, the example examined by Fang(2001) is a special case of the given model that there exists a unique PSE which is zero: $\pi_u = 0$. Fang(2001)'s Proposition 2 proves that there exists at least one non-trivial ASE if and only if $\phi(a,\ 0) > a$ a for some $a \in (0, 1)$. (Refer to Panel A of Appendix Figure 3.)

Using Lemma 6, we know $\Gamma'(0) = 0$. We also confirm $\Gamma(b)^l = 0$, $\forall b \in [0,\ 1]$ from the $\phi(a,\ b)$ curves in Panel A of the figure. Given $\phi(a,\ 0) > a$ for some $a \in (0,\ 1)$, we have both $\Gamma(0)^h > 0$ and $\Gamma(0)^m > 0$. Existence of non-trivial ASE is easily confirmed from the $\Gamma(a)$ and $\Gamma(b)$ curves in Panel B of the figure: $\Gamma(0)^j \in \Gamma(0)$ and $0 \in \Gamma(\Gamma(0)^j)$, $\forall_J \in \{m,\ h\}$. Corollary 2 shows that, WLOG, there are four non-trivial ASE once any non-trivial ASE exists. Those four non-trivial ASE are denoted in Panel B of the figure. According to Proposition 9, two of them closer to the 45 degree line are unstable and the other two near the corners are stable, as displayed in the panel.

7. Conclusion

In this paper, we develop an identity choice model that can explain social activities such as passing and selective out-migration from a stereotyped group as well as the emergence of elite cultural group, loosening the assumption of group identity immutability in standard statistical discrimination models. More talented members with low human capital investment cost have a greater incentive to identify themselves with a group that has a better collective reputation. The positive selection into a favored group plays a critical role in causing human capital cost distributions between groups to endogenously diverge. This model can be applied to many other social settings such as code switching (Go.man, 1959) and generating certi.cates to fight negative stereotypes (e.g. re-branding ex-cons).

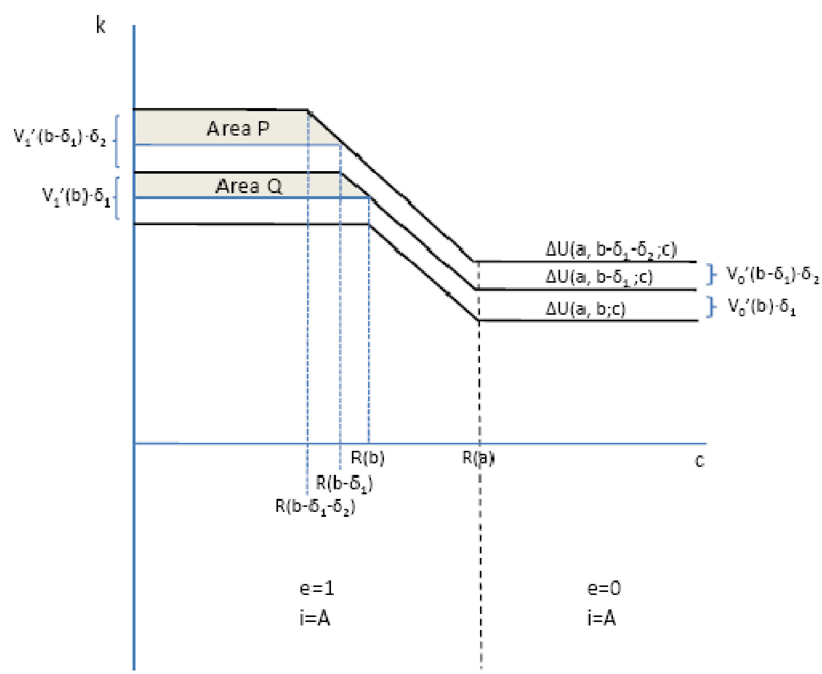
8. Appendix: Proofs

8.1. Proof of Lemma 3

Let us prove the first part. First, consider an arbitrary level of b such that $b < a < \bar{\pi}$. For very small δ_1 and δ_2, the following approximation holds: $h(\Delta U(a,\ b;\ c)) \approx h(\Delta U(a,\ b-\delta_1; c)) \approx h(\Delta U(a,\ b-\delta_1-\delta_2; c))$, which is denoted by $\tilde{h}(a,\ b,\ c)$. The small incremental decrease of b as much δ leads $V_0'(b)\cdot\delta$ increase of ΔU for any $0 \in (R(a),\ \infty)$ and $V_1'(b)\cdot\delta$ increase of ΔU for any $c \in (0,\ R(b-\delta))$. Therefore, the incremental decrease as much as δ_1 and subsequent decrease as much as δ_2 generate the different levels of ΔU as shown in Appendix Figure 1.

❙ Appendix Figure 1 ❙ Proof of Lemma 3

Panel A. Case with b $b,\ a,\ \bar{\pi}$



Let us impose that $V_0'(b-\delta_1)\cdot\delta_2 = V_0'(b)\cdot\delta_1$. Then, the incremental impact of decreased b on the overall human capital investment rate depends on the relative size of the skilled population of area P and that of area Q. As far as the skilled population of area P is greater than the skilled population of area Q, it is assured that the incremental decrease of b leads to the increase of $\phi(a,\ b)$: $\partial\phi(a,\ b)/\partial b < 0$ for any $b < a < \overline{\pi}$.

Let the skilled population in area P and area Q be denoted by $\sigma^A[P]$ and $\sigma^B[Q]$:

$$\begin{aligned}\sigma^A[P] &\approx [V_1'(b-\delta_1) - V_0'(b-\delta_1]\delta_2 \cdot \widetilde{h}(a,\ b,\ 0)\cdot G(R(b-\delta_1)) \\ &\quad -[V_1'(b-\delta_1) - V_0'(b-\delta_1]^2\delta_2^2 \cdot \widetilde{h}(a,\ b,\ 0)\cdot g(R(b-\delta_1)) \\ &\approx \left[\frac{V_1'(b-\delta_1)}{V_0'(b-\delta_1)} - 1\right]\ \delta_1 V_0'(b)\cdot h(a,\ b,\ 0)\cdot G(R(b))\end{aligned} \tag{25}$$

$$\begin{aligned}\sigma^B[Q] &\approx [V_1'(b) - V_0'(b)]\delta_1 \cdot \widetilde{h}(a,\ b,\ 0)\cdot G(R(b)) \\ &\quad -[V_1'(b) - V_0'(b)]^2\delta_2^2 \cdot \widetilde{h}(a,\ b,\ 0)\cdot g(R(b)) \\ &\approx \left[\frac{V_1'(b)}{V_0'(b)} - 1\right]\ \delta_1 V_0'(b)\cdot h(a,\ b,\ 0)\cdot G(R(b))\end{aligned} \tag{26}$$

Using the declining marginal benefits ratio property (MBRP), $\dfrac{V_1'(\pi)}{V_0'(\pi)} > \dfrac{V_1'(\pi+\delta)}{V_0'(\pi+\delta)}$ for any $\delta > 0$, we confirm that $\sigma^A[P] > \sigma^B[Q]$. Therefore, given $b_1 < b_2 < a < \overline{\pi}$, $\phi(a,\ b_1)$ is placed above $\phi(a,\ b_2)$: $\phi(a,\ b_1) > \phi(a,\ b_2)$. In the identical way, we can show the same results for other levels of a given $b_1 < b_2 < \overline{\pi}$. Also, we can prove the second part of the lemma (concerning the cases under $\overline{\pi} < b_1 < b_2$) using the similar methodology. (The proof needs to be improved further.) Q.E.D.

8.2. Proof of Lemma 4

Consider a very small $\delta > 0$ such that $a = b + \delta$. Define $\Delta(\delta)$ as $\Delta(\delta) \equiv R(b+\delta) - R(b)$: $\Delta' = \dot{R}(b+\delta)$. We have $H'(k) \approx H'(0)$ for small enough k. Using Lemma 1 and Panel A of Figure 3, we can calculate $\sigma^A(\delta)$ and $\Sigma^A(\delta)$, and consequently $\sigma^{A'}[\delta]$ and $\Sigma^{A'}[\delta]$

$$\sigma^{A'}[\delta] \approx G(R(b)+\Delta)\cdot[.5+H'(0)(V_0(b+\delta)-V_0(b+\Delta)]-.5H'(0)g(R(b))\Delta^2, \tag{27}$$

$$\sigma^{A'}[\delta] \approx G(R(b)+\Delta)R'(b+\delta)[.5+H'(0)(V_0(b+\delta)-V_0(b)+\Delta)]$$
$$+G(R(b)+\Delta)H'(0)(V_0'(b+\delta)-R'(b+\delta)-H'(0)g(R(b))\Delta R'(b+\delta). \tag{28}$$

(Note that the last terms, $-.5+H'(0)R'(0)(b)\Delta^2$ and $-H'(0)g(R(0)(b))\Delta R'(b+\delta))$, are added only when $R'(b)>0$).

$$\Sigma^A[\delta] \approx .5+H'(0)(V_0(b+\delta)-V_0(b)+G(R(b)+\Delta)H'(0)$$
$$\Delta-.5H(0)g(R(b))\Delta^2, \tag{29}$$

$$\Sigma^A[\delta] \approx H'(0)V_0'(b+\delta)+G(R(b)+\Delta)H'(0)R'(b+\delta)+$$
$$g(R(b)+\Delta)R'(b+\delta)H'(0)\Delta-H'(0)g(R(b))\Delta(b+\delta) \tag{30}$$

The slope of the $\phi(a, b)$ curve given a=b can be expressed as follows:

$$\frac{\partial\phi(a,\ b)}{\partial a}\Big|_{a=b} = \lim_{\delta\to 0}\frac{\phi(b+\delta,\ b)-\phi(a,b)}{\delta}$$
$$= \lim_{\delta\to 0}\frac{\sigma^A(\delta)/\Sigma^A(\delta)-\sigma^A(0)/\Sigma^A(0)}{\delta}$$
$$= \lim_{\delta\to 0}\left[\frac{[\sigma^A(\delta)-\sigma^A(0)]\Sigma^A(0)}{\delta}-\frac{[\Sigma^A(\delta)-\Sigma^A(0)]\sigma^A(0)}{\delta}\right].$$
$$\frac{1}{\Sigma^A(\delta)\Sigma^A(0)}$$
$$\approx \frac{\sigma^{A'}(0)/\Sigma^A(0)-\sigma^A(0)/\Sigma^{A'}(0)}{\Sigma^A(0)^2} \tag{31}$$

We can achieve the following results:

$$\begin{cases} \sigma^{A}(0) \approx .5G(R(b) \\ \sigma^{A'}(0) \approx .5g(R(b)R'(b) + G(R(b))H'(0)(V_0'(b) + R'(b)) \\ \Sigma^{A}{}'(0) \approx .5 \\ \Sigma^{A}{}'(0) \approx H'(0)(V_0'(0) + G(R(b))H'(0)R'(0) \end{cases} \tag{32}$$

Consequently, we ave $\frac{\partial\phi(a,\ b)}{\partial a}\Big|_{a=b} \approx g(R'(b))R(b) + 2H'(0)R'(b)$ $[1 - G(R(b))]$. Q.E.D.

8.3. Proof of Lemma 6

Given the slope of $\phi(x,\ y)$ at $(\hat{x},\ \hat{x})$ denoted by $\frac{\partial\phi(x,\ y)}{\partial x}\Big|_{x=y=\hat{x}}$ and the slope of $\phi(x,\ x)$ at the same point, $G(R(\hat{x}))R'(\hat{x}))$, we can find a correspondence value x' such that $x' = (x',\ \hat{x} + \Delta)$ using the following equation:

$$x' - [\hat{x} + g(R(\hat{x}))R'(\hat{x})\Delta] = \frac{\partial\phi(x,\ y)}{\partial x}\Big|_{x=y=\hat{x}} \cdot [x' - (\hat{x} + \Delta)]. \tag{33}$$

Therefore, we have $\Gamma'(\hat{\pi})$, which is approximately equal to $\frac{x' - \hat{x}}{\Delta}$

$$\Gamma'(\hat{x}) \approx \left[g(R(\hat{x}))R'(\hat{x})) - \frac{\partial\phi(x,\ y)}{\partial x}\Big|_{x=y=\hat{x}} \right] \Big/ \left[1 - \frac{\partial\phi(x,\ y)}{\partial x}\Big|_{x=y=\hat{x}} \right]. \tag{34}$$

From Lemma 4 and $G(R(\hat{x}))R' = (\hat{x})$, we have $\frac{\partial\phi(x,\ y)}{\partial x}\Big|_{x=y=\hat{x}}$ $= g(R(\hat{x}))R'(\hat{x}) + 2H'(0)R'(\hat{x})\hat{x}(1 - \hat{x})$. Then, we have the given result for $\Gamma'(\hat{\pi})$. Q.E.D.

8.4. Proof of Theorem 1

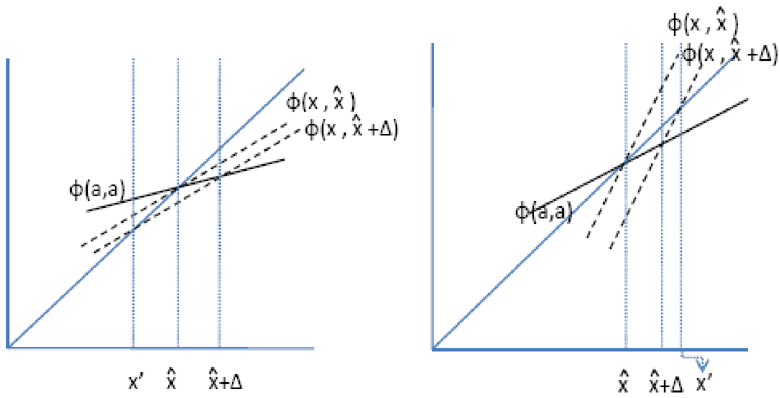
First, consider the case with $\overline{\pi} > \pi_{h.}$ $\Gamma(b)$ passes through the points $(\pi_h,\ \pi_h)$ and a-intercept $(a,\ b)$ $(0,\ \Gamma(0)^h)$, in which $\pi_h < \Gamma(0)^h < 1$. $\Gamma(a)$ passes through $(\pi_l,\ \pi_l)$ and b-intercept $(b,\ a)$ $=(1,\ \Gamma(1)^h)$, in which $\pi_h < \Gamma(1)^h < 1$. Thus, there should be at least one ASE which satisfies $(\pi_A^* > \pi_B^*)$. In the same way, we can findat least one ASE which satisfies $(\pi_B^* > \pi_A^*)$ Secondly, consider the case with $\overline{\pi} > \pi_{h.}$. Using Lemma 3, we can find that $\pi_h < \Gamma(\pi_l)^h < b_4'$, in which as $\phi(b_4',\ \pi_l) = \pi_l$ shown in Panel A of Figure 6. We can draw the shape of $\phi(a,\ \Gamma(\pi_l)^h)$, which pass through the $\phi(a,\ a)$ curve both at some $a > \pi_l$ and at some $a < \pi_l$, as well as the point $(\Gamma(\pi_l)^h,\ \phi(\Gamma(\pi_l)^h,\ \Gamma(\pi_l)^h))$ on the curve. This implies that $\Gamma(\Gamma(\pi_l)^h)^l < \pi_l$ and $\Gamma(\Gamma(\pi_l)^h)^m < \pi_l$. From this, we can infer that there exist at least four non-trivial ASE. (Note that even when the number of PSE is two instead of three (for example, $\phi(a,\ a)$ is tangent to the 45 degree line at $a = \pi_l$), the proof goes in the same way.) Q.E.D.

8.5. Proof of Proposition 4

Given multiple PSE, using Lemma 3, we can findthat $\pi_h < \Gamma(\pi_l)^h < \widetilde{b}$, in which $\phi(\widetilde{b},\ \pi_l) = \pi_l$ as shown in Figure 4. We can draw the shape of $\phi(a,\ \Gamma(\pi_l)^h)$, which pass through the $\phi(a,\ a)$ curve both at some $a > \pi_l$ and at some $a < \pi_l$. This implies that $\Gamma(\Gamma(\pi_l)^l < \pi_l$. Since $\Gamma(b)^h$ decreases over $b \in (0,\ \overline{\pi})$, there must be an intercept of $\Gamma(b)$ and $\Gamma(a)$, $(\pi_L^{**},\ \pi_H^{**})$ which satisfies $\pi_L^{**},\ \pi_l$ and and $\pi_H^{**},\ \pi_h$. Out of the symmetricity, there must be another ASE $(\pi_H^{**},\ \pi_L^{**})$ Q.E.D.

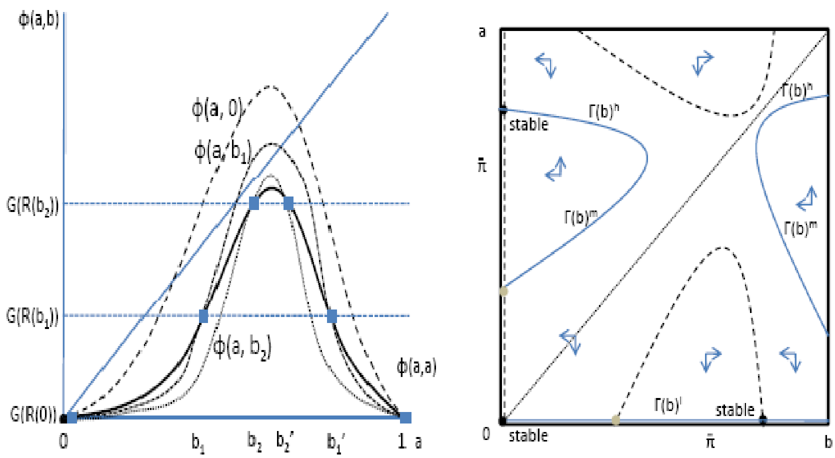
❙ Appendix Figure 2 ❙ Slope of Correspondence at trivial ASE

Panel A. Example for $\Gamma'(\hat{x}) < 1$ Panel B. Example for $\Gamma'(\hat{x}) > 1$



❙ Appendix Figure 3 ❙ An example of Fang (2001)

Panel A. Given Unique PSE: π_u=0 Panel B. Multiple ASE (When φ(a, 0)>a for some a)



References

Arrow, K.J. (1973). \The Theory of Discrimination," in Ashenfelter and Rees (eds.), *Discrimination in Labor Markets*.

Charles, Kerwin K., Erik Hurst and Nikolai Roussanov (2009). "Conspicuous Consumption and Race," *The Quarterly Journal of Economics* 124: 425-467.

Coate, Stephen and Glenn Loury (1993). "Will A.rmative-Action Policies Eliminate Negative Stereotype?" *American Economic Review* 83: 1220-1240.

Fang, Haming (2001). "Social Culture and Economic Performance," *American Economic Review* September: 924-937.

Fukuoka, Yasunori, Yukiko Tsujiyama and John Russell (1998). "MINTOHREN: Young Koreans Against Ethnic Discrimination in Japan," *The Bulletin of Chiba College of Health Science* 10.

Fryer, Roland and Paul Torelli (2010). "An Empirical Analysis of 'Acting White'," *Journal of Public Economics* 94: 380-396.

Go.man, Erving (1959). *The Presentation of Self in Everyday Life*, New York: Doubleday.

Grogger, Je.rey (2011). "Speech Patterns and Racial Wage Inequality," *The Journal of Human Resources* 46: 1-25.

CHAPTER 6

The Dynamics of Income Inequality in Korea

by
*Jongsung Kim**
(Department of Economics Bryant University)

Abstract

This paper reviews the trends of income inequality in Korea and examines the sources of income inequality from 1998 to 2008. The findings confirm the importance of earned income as a major component in total income. It was found that an increase in transfer income reduces the income inequality most significantly among all income sources. The findings have important policy implications. In order to raise the overall economic status, policy efforts should be exerted to create more stable employment opportunities with better compensation. Also called for is more active policy intervention to increase the transfer payment. The nature of tax policy is also important. Despite the government's claim that the recent tax cuts are not for the rich, to the extent that the tax cuts are implemented differently across income distribution in favor of the rich, the income inequality will persist.

It is encouraging that in 2010, the Gini coefficients, quintile ratios, and relative poverty rates all improved for the first time since 2006. However, it is too early to tell whether or not this improvement was a temporary result from the active government intervention with transfers and tax cuts. To continue this improving momentum, the pressing

* Address: Smithfield, RI 02917, USA. E-mail: jkim@bryant.edu.

policy challenge now is more than implementing stronger economic policies for growth and redistribution. What is urgently needed is to reduce the prevalent uncertainty and to create social and public consensus to share the economic pain and gain.

Given the importance of transfer income to reduce the inequality, some form of tax raise may be inevitable. The social consensus to share "the pain and gain" will lighten the potential tax resistance from the rich and induces the poor to be more patient. Conscientious political leadership based on reality and feasibility is also needed since the public policies based on populism only retard individual incentives and further exacerbate the income inequality. For policies to be implemented successfully, they need to be supported by sufficient resources and adequate coordination with other relevant policy aspects. If and only if these prerequisites are met, economic and public policies to reduce labor market inequality and polarization will work, and the lower and the poor class will feel the "trickle-down" effect.

1. Introduction

The economic growth of the Republic of Korea (hereafter, Korea) is often hailed as a miracle. Korea has ascended from the destruction and shambles of Korean War in the 1950s to the 15th largest economy in 2009 in GDP terms. The data from the Statistics Korea (formerly the Korean National Statistical Office) show that from 1982 to 2009, Korea's nominal (real) GDP has increased almost eighteen folds (five folds) and nominal GDP per capita jumped from mere \$1,927 to \$17,175 after reaching \$21,695 in 2007. This remarkable economic growth contributed to the substantial reduction in Korea's labor income inequality[1].

Keeping up with this trend, in 2010 Korea attained an impressive 6.3% rise in GDP, its fastest growth since 2002. With this strong rebound and speedy recovery from the 2008 economic crisis, attributed

1 Loosely speaking, income inequality represents interpersonal income differences within a given population (Chakravarty 2009: 1).

mostly due to robust exports, the Korean economy experienced improved consumer spending and increased facilities investment. In 2010, export increased 29 percent to $467 billion.[2]

It is reported that many of Korea's companies affiliated with *Chaebol* (loosely translated as Korean conglomerates of many companies strongly clustered around one parent company) have performed much better than they did after the financial crisis in 1997 when the IMF had to step in. Many *Chaebol*-affiliated firms have taken advantage of weak Korean currency (KRW) to attain the competitive edge over the competing Japanese manufacturers.

However, Korea's outwardly robust economy did not ameliorate the financial woes of the ordinary people and small-to-midsize companies who once represented Korea's thriving middle class. Lying beneath the economic success in the nation were dormant cries of distress among indebted households and financially-strapped small and medium enterprises (SME). Korean households are now facing record level of debts. The Bank of Korea recently reported that in March 2011, the total debt of Korean households reached 801.4 trillion KRW, surpassing the threshold of 800 trillion KRW for the first time and increased 17.3% in two years.[3]

During the last decade following the Asian crisis, many Koreans in their 30s and 40s, encouraged by government policies and generous credits, have heavily borrowed to buy their homes.[4] Many of them are reported to have been struggling to make their mortgage payments, falling into the category of "House Poor". Swamped in heavy mortgage debts, those in their 30s and 40s are the most vulnerable group signaling that increase in their debt burden would trigger an onslaught of bankruptcy much like the housing crisis in the U.S. after the bubble burst. As the economic inequality in Korea has grown significantly over the past decade, the growing disparity is observed in major aspects of

---

2 Financial Times, "South Korea: An economy divided" (May 29, 2011).

3 Chosun English Daily, "Household Debt Spirals to W800 Trillion" (May 26, 2011).

4 According to OECD statistics, the household savings rate (% of disposable household income) in Korea has plummeted from a world-beating 23.1% in 1993 to 3.5% in 2011 – one of the lowest among OECD countries.

social life such as consumption pattern and educational opportunities (Koo 2007: 1).

Against the backdrop of this major change in Korean labor market, this paper aims to review the trends of income inequality in Korea and investigate the dynamics of the composition of income from different sources. This paper adds to the existing literature by providing new evidence of the dynamics of income inequality in Korea with the most recent data set from Korean Labor Income and Panel Study (KLIPS). Section 2 reviews the background and previous research evidence on inequality in Korea. Section 3 lists statistics to explain the trends in income inequality. Section 4 provides new evidence of income inequality and its decomposition. Section 5 concludes and provides policy implications and recommendations.

2. Income Inequality in Korea

After the 1997 financial crisis, the Korean economy experienced an increase in inequality among workers and households, and deepening labor market polarization[5] gave birth to a distinct two-tier labor market. The primary labor market is characterized by regular and secure employment, relatively higher compensation, and highly educated skilled workers. In contrast, the secondary labor market is represented by temporary or contingent jobs, lower compensation, and unskilled workers. The faults between these two markets are increasingly evident and the upward mobility from the secondary to the primary labor market proves to be extremely difficult.

The 1997 financial crisis and the economic restructuring program mandated by the International Monetary Fund (IMF) inflicted major consequences on Korean economy and society. Bank credit squeeze

---

5 Broadly speaking, polarization is concerned with appearance (or disappearance) of groups in a distribution (Chakravarty 2009: 105). One notion of income polarization, often referred to as bipolarization, is concerned with the decline of the middle class. For more details on the measurement of inequality and polarization, refer to Chapter 4 in Chakravarty (2009).

and contracting financial policies led a sharp increase of business bankruptcies. In addition, the IMF regulations spurted a progeny of economic policies of neo-liberalistic nature that emphasize the market efficiency and labor market flexibility. As a result, company practices of lifetime job security in exchange of employee loyalty and fidelity began to disappear, expanding the share of the secondary labor market.

The sharp decline in job stability experienced by Korean workers during the 1997 financial crisis never bounced back to the previous level. The recovery process was much slower for irregular, short-tenured or less educated workers, further reinforcing the polarization of job stability (Cho and Keum 2009). A direct consequence of this is an increase in income inequality and the polarization of the income distribution. Such disparity in income created from sudden shifts in company practices is a foreseeable consequence which resulted from restructuring the nation's economy after financial crisis when the government is forced to implement new policies for a quick recovery.

The importance of economic growth as a tool to improve the economic status of the members in a society and as a weapon against poverty has been widely discussed among policymakers and academicians. The idea that an improving economy benefits all members in that economy is most famously and laconically summarized in the former U.S. President John F. Kennedy's remark "A rising tide lifts all boats". What is left out in this aphorism, however, is the reality that the benefits of economic growth may not be evenly distributed across people and some members fare better than others. Moreover, in some cases, it is also possible that certain members' economic position even deteriorates when an economy expands and they are left worse off than before. This, unequivocally, also leads to an increase in income inequality.

Focusing on the relationship between economic growth and income inequality, Kuznets (1955) predicted that as economies develop, income inequality will first rise and reach the peak and fall after a certain critical threshold development stage and income level. Kuznets documented this argument using both cross-country and time series data. This inverted U-shaped pattern of income inequality (often measured by the Gini coefficient, a scale on which zero is perfect equality and one is perfect inequality) is now famously known as the Kuznets curve, becoming one of the major stylized facts about long-run processes of economic development.

Income inequality in Korea has risen rapidly since the early 1960s when the government began to implement a series of five-year economic development plans. Consistent with the Kuznets curve, income inequality declined during the 1980s and until 1990s after reaching its critical peak. According to Fields and Yoo (2000: 139), the Gini coefficients based on Korea's labor income declined by 11 Gini points (or 27%) between 1976 and 1993. However, recent research findings consistently show that income inequality has rebounded sharply around late 1990s when Korean economy fell victim to the Asian financial crisis. Also present were the patterns of Gini coefficients, based on longitudinal data, which is U-shaped for approximately two decades from early 1980s to late 1990s (Sung 2010 and references therein).

Much research has investigated various aspects of income inequality in Korea. The main factors of recent increase in income inequality are believed to be mostly from economic reasons such as rapid industrial changes/developments, restructuring, business cycles, asymmetric changes in wage discrepancies, and skill-biased development. Growing labor market inequality is sometimes blamed on increased trade competition from China. In a slightly different vein, Yun (2009) also argues that labor market inequality is not simply driven by such structural changes but by the nature of the ways in which new labor market regulations were created and the resulting regulatory contradictions. For example, despite the presence of active labor market policies in Korea designed to help the labor market entrance of marginalized workers and further stabilize those workers' income, these policies, such as the Employment Stabilization Program (ESP) and the Job Skill Development Program (JSDP), were poorly implemented and failed to provide adequate financial resources and incentives to marginalized workers.

Changing demographics such as age structure also have affected income inequality. Korea, one of the most rapidly aging countries, has already seen its share of population over 65 years of age increase from 5% to more than 9% over the last 15 years. With an increase to 35% projected for 2050, median age in Korea will have increased 20 years to 55. By mid 21st century, Korea will even replace Italy as the world's second-oldest country (Hayutin, 2007). Sung (2010) looks into the effects of population aging on income distribution and statistical relationship between income inequalities and estimate their longitudinal change by focusing on the effects of quarterly income mobility on annual income inequality. The population aging in Korea was found to account for approximately 7.7% or 39.7% of total change in the Squared Coefficient of Variation (SCV) ratio between 1994 and 2009, depending upon the base-year income distribution condition. This finding implies that the annual income inequality could have been reduced by those amounts if the population aging had not occurred.

While the number of jobs has increased, those newly-created jobs are disproportionately concentrated in the small-scale establishments. From 2003 to 2008, over 80% increase in jobs were found in the establishments that hire fewer than 5 people. The average workers in the sector where the size of the firm is less than 5 workers earned only 46.7% of the workers in the sector where the firms hire 300 or more workers. The ratio improved from 43.6% in 2006, but the workers are still making less than half. This trend of new job creation in small businesses has also contributed to the increase in income inequality.[6]

In response to the public efforts to reduce income inequality, research also look into the public policy implications on the magnitude of income inequality. Sung and Park (2011), for example, examined the redistributive effects of Korea's fiscal policies, including consumption taxes and in-kind benefits. Using the 2007 Household Income and Expenditure Survey, they found that taxes and transfers reduce income inequality in Korea by 13.8%. Sung and Park (2009) also found that contrary to the popular belief that implementation of direct taxes is the

---

6 2011 Population Survey for Economic Activities (Additional Surveys), Statistics Korea.

key in effective redistribution, in-kind benefits, direct taxes, and social security contributions all decrease the Gini coefficient by 6.7, 4.7 and 2.9 percentage points, respectively.

Before further ado, the distinction between income inequality and polarization is in order. Although these two concepts are sometimes used synonymously, Wolfson (1994: 354) showed that polarization and inequality are demonstrably different and also pointed out the potential problem of using the conventional scalar measures of inequality to assess the extent and trend in polarization. In Korean context, however, empirical evidence has shown mixed results. Shin and Shin (2007: 111) find that the polarization poses more serious problem than inequality on the income distribution in Korea. According to Shin and Shin (2007: 81), from 1997 to 2003 the polarization index based on household total income increased by 67%~310% depending on the value of polarization sensitivity of the polarization index introduced in Esteban and Ray (1994). The corresponding Gini index increased only 7% during the same period. On the other hand, despite their conceptual differences, Yoo (2007: 47) found no significant statistical difference between relative income inequality index and polarization index.[7]

3. Trends in Income Inequality in Korea

The data collected by the Statistics Korea show that the decile ratio (P90/P10) increased slightly from 2009 to 2010 for both market income and disposable income when only urban and non single-person households are considered. But when all household, including agricultural and single-person households, are considered, the decile ratio slightly declined from 2009 to 2010 (Figure 2).[8] During the same period, the P50/P10 ratio also increased from 2.5 to 2.53, indicating an

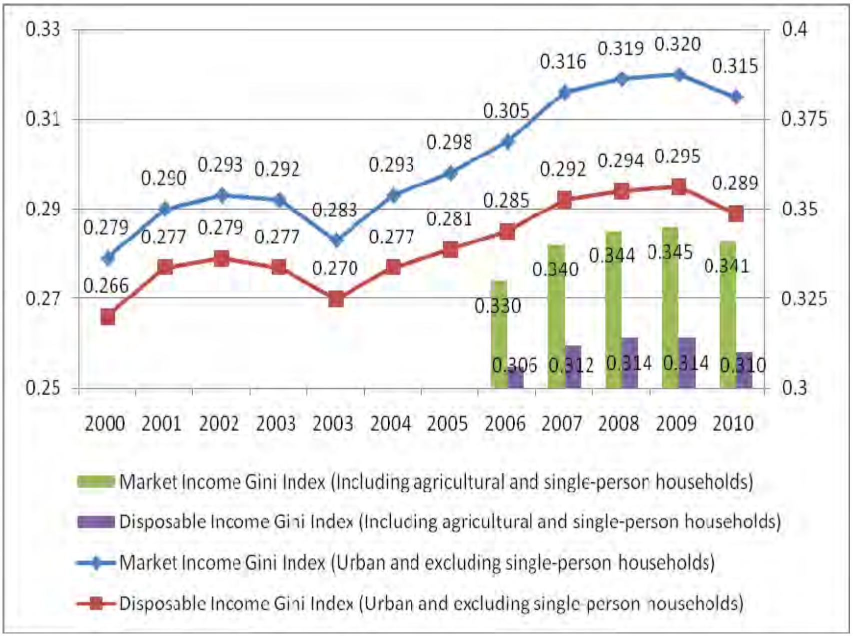
7 Yoo (2007: 35) pointed out that rounding errors in KLIPS may have deepened the clustering of income distribution, leading to overestimate the polarization index while the Gini coefficients can be underestimated.

8 In 2010, in individual equivalized disposable income, P10 amounts to KRW 620,000 and P90 amounts to KRW 2,979,000.

increase in income inequality between the middle class and the lower class. On the other hand, the P90/P50 ratio declined from 1.92 to 1.9. This pattern reflects the relative improvement of the middle class income relative to those of the rich and the poor.

The improvement is tracked by P10 increase of 5.08%, while P50 and P90 increased by 6.37% and 5.30% respectively. OECD defines that the upper class (or the rich) as households with income greater than 150% of the median income and the low income class (or the poor) is defined as the household with income less than 50% of the median income. Households with income between 50% and 150% of the median income are defined as the middle class. The data from the Statistics Korea shows the diminishing share of the middle class in the early 1990s and the acceleration of this pattern after the late 1990s when the Korean economy experienced the Asian currency crisis.

❙ Figure 1 ❙ Gini Index

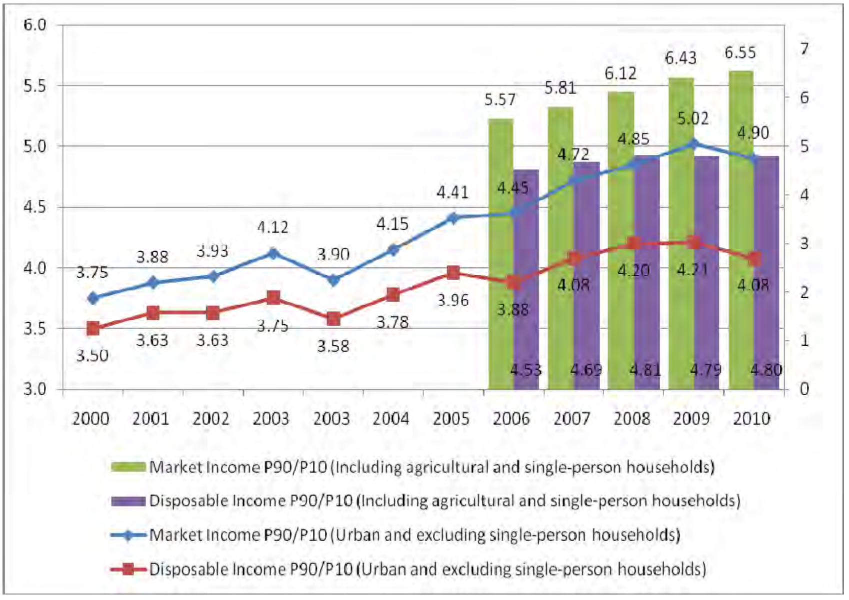


Source: Statistics Korea (obtained from Korea Statistical Information Service)

But for the past two years, the share of middle class income to the total household increase has seen a marginal rebound. Statistics Korea recently announced that the share of disposable income of the middle class to the total household income was 66.7% in 2010. This is a marginal increase from the 66.2% in 2008 but still lower than the share of 70.1% in 2003.[9]

Figure 1 shows the pattern of the Gini coefficient for four categories. In all categories, the Gini coefficients reached the maximum in 2009 and declined slightly in 2010. For example, Gini coefficient for market income (urban and excluding single-person households) stood at 0.315 after reaching 0.320 in 2008. The pattern shows the increase in income inequality after the Asian financial crisis, although the inequality slightly declined in 2010.

❙ Figure 2 ❙ Decile Ratio (P90/P10)

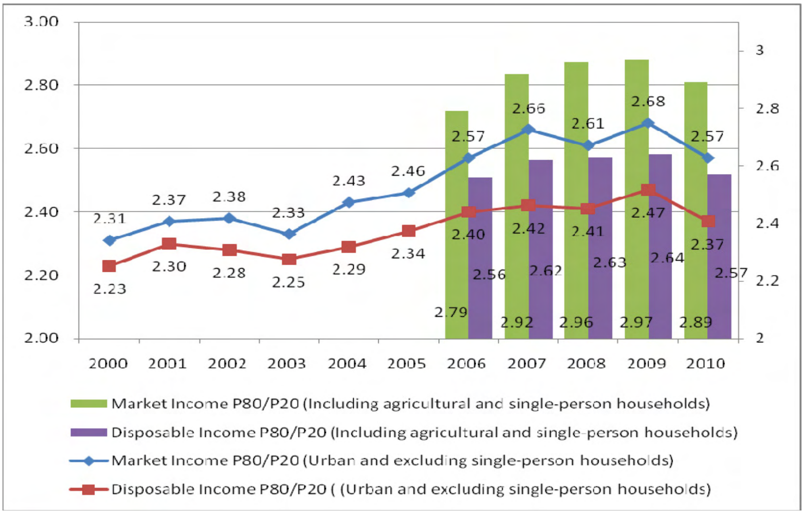


Source: Statistics Korea (obtained from Korea Statistical Information Service).

9 Korea Herald, "Shrinking of Middle Class" (July 23, 2010).

The decile ratio (P90/P10) portrays another side of income inequality and income polarization. Figure 2 shows that the decile ratio has consistently increased after 2006 in both market income and disposable income. Although the income inequality measured by the Gini coefficient declined slightly in 2009 and 2010 as shown in Figure 1, the income (including agricultural and single-person households) gap between the rich (P90) and the poor (P10) has still widened. When agricultural and single-person households are included, the rich earned 4.8 times more than the poor in disposable income.[10] In market income terms, the decile ratio is even greater at 6.55 in 2010. However, the decile ratio declined slightly from 2009 to 2010 for market income and disposable income when only urban and non-single-person households were considered.

❙ Figure 3 ❙ Quintile Ratio (P80/P20)



Source: Statistics Korea (obtained from Korea Statistical Information Service).

Figure 3 presents four quintile ratios for household market income and household disposable income with different status of urban residency and single-person households. All four trends confirm the

10 The OECD average ratio is 4.2.

decline in income inequality in the most recent year. In 2010, the quintile ratios for all four categories declined. The quintile ratios and two of decile ratios when combined with the Gini coefficients tell us that the income inequality has slightly declined in 2009 and 2010 (with lower Gini coefficients and the quintile ratios).

Table 1 Changes of income in the 2009 first quarter from the previous year (%)

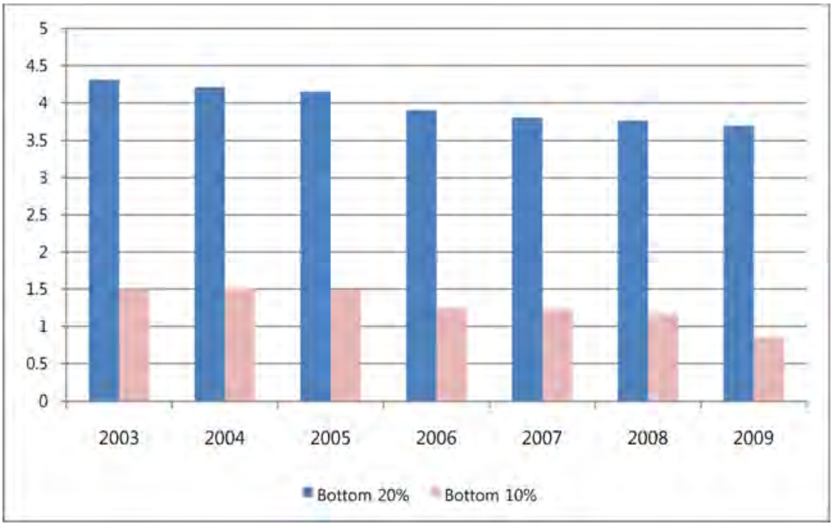
10th decile (P90 and above)	3.4
9th decile	-1.1
8th decile	0.5
7th decile	0
6th decile	-1.1
5th decile	-0.8
4th decile	-1.2
3th decile	-1.6
2nd decile	-2.8
1st decile (P10 and below)	-9.7

Source: Statistics Korea

Table 1 lists that in the first quarter of 2009, the income of the top 10% (P90 and above) increased 3.4% from 2008. Income of almost all other deciles declined. The upper middle class (8th and 7th deciles) reported almost no increase in income. The income for all the lower deciles declined and the extent of decrease is more conspicuous for the 1st and 2nd deciles at 9.7% and 2.8% respectively. This pattern shows that the 2008 financial crisis actually benefited the rich class and punished almost all others. This pattern is indicative of more than an income inequality, extending to the level of income polarization.[11]

11 The polarization of income distribution is one of the components of neo-liberalism along with financial deregulation (Stockhammer 2010: 3). The rapid increase of P90 income was also seen in the United States.

| Figure 4 | Proportion of household income of the lower-income households

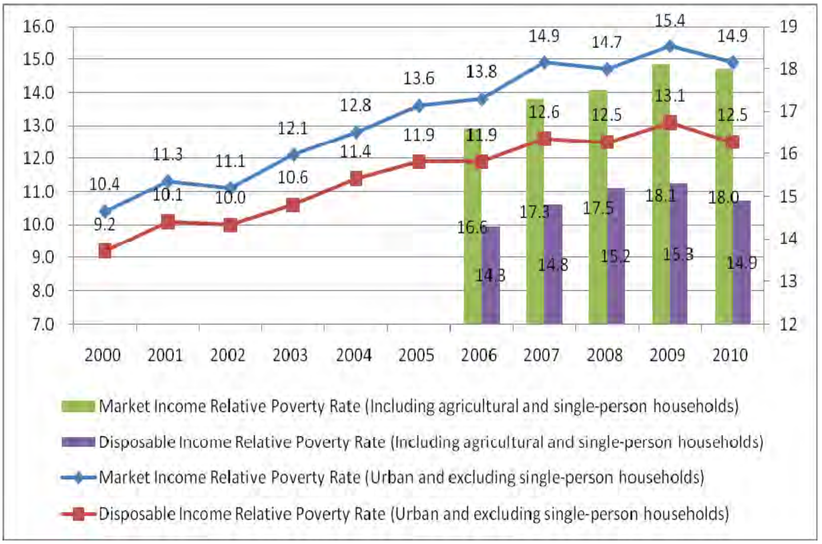


Source: Statistics Korea, Korea Institute for Health and Social Affairs (KIHASA)
2003 – 2005: Agricultural and fishery households, single-person households excluded
2006 - : Agricultural and fishery households excluded

Another problem worthy of more attention is that employment per se does not necessarily improve the economic status. More than half of the households below the poverty line remain in that situation despite the presence of an employed family member in the household, falling into the working poor category. Figure 4 presents the pattern of the consistent decline in income share by the bottom 10% and 20% households, another indication of deepening income inequality. The vulnerable economic position of the poor class is also shown in Figure 5 which displays the increasing pattern of relative poverty rate since 2000. The relative poverty is defined as the share of households whose income is less than 50% of median income. The decline in the relative poverty rate in 2010 is encouraging but it is premature to predict whether this improving momentum will continue.[12]

12 As Gyeongjoon Yoo of KDI was quoted to have stated, "since 2000, the relative poverty rate has increased faster than the Gini index that measures the inequality of the whole income distribution," and "in the 2000s, economic growth dropped and more fruits of growth went to those who are not poor" (Various sources).

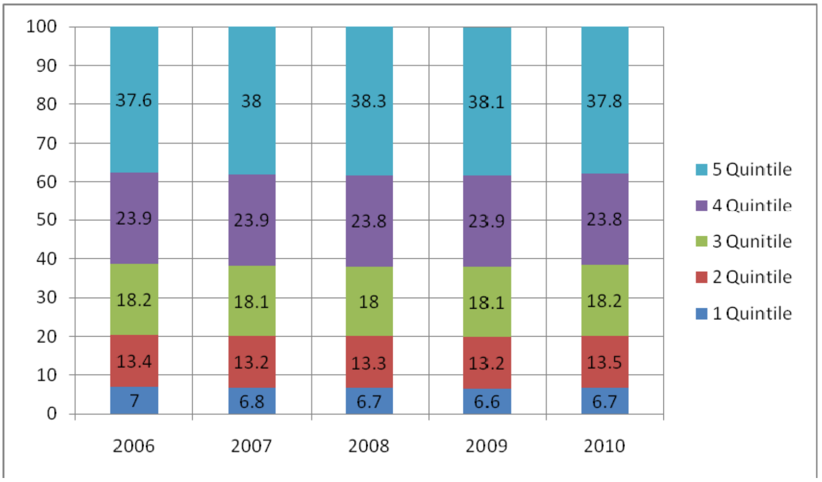
❙ Figure 5 ❙ Pattern of Relative Poverty Rate



Source: Statistics Korea (obtained from Korea Statistical Information Service).
Relative poverty rate = * Equivalized individual disposable income.

❙ Figure 6 ❙ Proportion of Household Income of Each Quintile Group*

(Unit: percent)



Note: * Equivalized individual disposable income (individual, and agricultural households included).
Source: Statistics Korea.

The increase of absolute poverty rate is also problematic. The absolute poverty rate is the share of households whose inflation-adjusted income is lower than a certain income threshold or the number of households unable to afford certain basic goods and services. The absolute poverty rate of the urban household reached as high as 16.4% in 1999, right after the financial crisis and declined to 10.5% in 2000 and since then, it has hovered around 9% until 2003 when the absolute poverty rate shot up to 11.2% and reached at 14.4% in 2009. This pattern also reflects the aggravating income inequality.

Figure 6 shows the proportion of household income of each quintile group. Although the gap between the top quintile and bottom quintile slightly widened from 2006 to 2010, the gap narrowed from 2009 to 2010 by approximately similar magnitudes without any major structural change. This reflects that although the momentum of aggravating income inequality was somewhat curved in 2010, the pattern since 2006 still shows an underlying force of increasing income inequality and the shrinking middle class.[13] According to the KDI findings, in terms of disposable income, the proportion of middle class declined from 58% in 2007 to 56.4% in 2008. On the other hand, the poor class increased from 18.3% to 19%, and the rich class increased from 23.7% to 24.6%. In comparison with 1996 pattern, the share of middle class declined by 12.1 percentage points of which 7.7 percentage points fell down into the poor class while 4.4 percentage points moved up to the rich class.

One of the emerging concerns from the recent pattern of income inequality is that the contraction of the middle class and the widening income gap will shrink both domestic consumption and retard investment incentives, consequently weakening the growth potential of the Korean economy. To provide new evidence and information about income inequality in Korea, this paper investigates the dynamics of household income inequality in Korea from 1998 to 2008 using the

---

13 The pattern of the average disposable income for each quintile groups of urban households (agricultural and fishery households, single-person households excluded) from 2000 to 2010 (including window from 2009 to 2010) shows that the gap between the top quintile and the bottom quintile has widened (from comments by discussant).

KLIPS data collected by the Korea Labor Institute. Also discussed will be the inequality of the various income components that constitute the total household income.

4. New Evidence of Income Inequality and its Decomposition

Lerman and Yitzhaki (1985: 152) show that

$$(1) \quad G = \sum_{k=1}^{K} R_k G_k S_k$$

where the Gini coefficient is the product of three components: R_k, G_k and S_k. Stark, Taylor and Yitzhaki (1986) noted that the relation among these three terms has the following interpretations. The influence of any income source on total income inequality depends on three components: R_k is the Gini correlation between income source *k* and total income, G_k is the relative Gini of source *k,* and S_k is source *k*'s share of total income.[14]

If an income source represents a large share of total income, it may potentially have a large impact on inequality. However, if income is equally distributed ($G_k = 0$), it cannot influence inequality even if its share is large. On the other hand, if this income source is large and unequally distributed (i.e. S_k and G_k are large), it may either increase or decrease inequality. This depends on which households, at which points in the income distribution, earn it. If the income source is unequally distributed and flows disproportionately toward those at the top of the income distribution (R_k is positive and large), its contribution to inequality will be positive. However, if it is unequally distributed but targets poor households, the income source may have an equalizing effect on the income distribution (López-Feldman 2006; Azam and Shariff 2011).

14 When this paper was completed, the author learned that Jang and Lee (2010) also used the Lerman and Yitzhaki's model for 1997-2006 KLIPS data.

A key rationale for studying and examining decompositions by source is to learn how changes in particular income sources will affect overall income inequality. Considering a change in income from source k equal to eYk where e is close to 1, Lerman and Yitzhaki (1985) derived an expression for the partial derivative of the overall Gini with respect to a percent change (e) in source k as follows.

$$(2) \quad \frac{\partial G}{\partial e_k} = S_k (R_k G_k - G)$$

where G is the Gini coefficient of total income inequality before the income change. Dividing (2) by G yields the source's marginal effect relative to the overall Gini (G), which can be further written as the source's inequality contribution as a percentage of the overall Gini minus the source's share of total income:

$$(3) \quad \frac{\partial G / \partial e_k}{G} = \frac{S_k R_k G_k}{G} - S_k$$

In the decomposition exercise, the Stata code **descogini** written by López-Feldman (2006) was used.

Data

The Korea Labor Institute began to collect detailed data for households and individuals starting in 1998. KLIPS is a longitudinal survey of the labor market and income activities of households and individuals residing in urban areas. This data collection is modeled after and is similar to the Panel Study of Income Dynamics (PSID) in the United States. This paper uses the 11 waves from 1998 to 2008. The 11^{th} wave data for 2008 is the most recent data available. All income variables in the KLIPS are after-tax incomes and include the amount of previous year's income.[15]

15 The target population includes all persons living in urban areas, except Jeju Island

For household income, KLIPS includes the following six categories: earned income, financial income, real estate income, social insurance income, transfer income, and income from other sources. Each income category includes various relevant items. For example, the earned income includes wages and compensations received from the workplaces and the self-employment income. Details for other categories are available in the KLIPS User's guide and Questionnaires.

Results

Table 2 lists the income decomposition for the total household income and income sources from 1998 and 2008.[16] Since the Gini coefficients (Gk) in table 2 are based on the limited sample of approximately 5,000 observations collected by the KLI, these coefficients are not to be directly compared to the Gini coefficients released by the Statistics Korea.[17]

From 1998 to 2008, income inequality in total income actually declined from 0.487 to 0.429. This is not surprising since the income inequality reached its highest level after the 1997 Asian financial crisis. Of the income sources, Gini coefficients for earned income and financial income increased from 1998 to 2008 from 0.443 to 0.463. Gini coefficients for other income components (financial, real estate income, social insurance income, transfer income, and other incomes) declined during the sample period. The Gini coefficients for social insurance income and transfer income declined 5.6 and 15 percentage

---

and those institutionalized. Individuals selected for the survey are interviewed once per year. For the first panel of 1998, the KLIPS collected retrospective data on past work experience. Compared with other longitudinal surveys such as NLSY and PSID, the attrition rates of the KLIPS are known to be reasonable, and sample attrition problem is less likely to bias the empirical results of the studies using the KLIPS (Cho and Keum 2009).

16 To save space, only the results from 5 waves are reported. The results for all waves are available from the author by request.

17 There is a possibility that observation fluctuations in the data (4,123 in 2001 vs. 5,116 in 2008) may have hurt the consistency of the results of the Gini coefficient decomposition (from comments by discussant).

❙ Table 2 ❙ Income Decomposition in selected waves

Wave	Year	Source of Income	Sk	Gk	Rk	Share	% change	Change of Gk from 1998 to 2008
1	1998	Total		0.487				
# of observation = 4884		Earned	0.833	0.443	0.948	0.718	-0.115	
		Financial	0.024	0.967	0.481	0.023	-0.001	
		Real estate	0.064	0.992	0.859	0.111	0.048	
		Social insurance	0.056	0.998	0.963	0.110	0.054	
		Transfer	0.015	0.990	0.813	0.024	0.010	
		Others	0.009	0.993	0.716	0.014	0.004	
3	2000	Total		0.421				
# of observation = 4242		Earned	0.862	0.433	0.938	0.831	-0.031	
		Financial	0.032	0.980	0.705	0.052	0.020	
		Real estate	0.026	0.976	0.575	0.035	0.009	
		Social insurance	0.014	0.980	0.379	0.012	-0.002	
		Transfer	0.026	0.928	-0.214	-0.012	-0.038	
		Others	0.041	0.986	0.865	0.082	0.042	
6	2003	Total		0.441				
# of observation = 4574		Earned	0.861	0.450	0.945	0.831	-0.030	
		Financial	0.012	0.975	0.524	0.014	0.002	
		Real estate	0.049	0.976	0.771	0.083	0.035	
		Social insurance	0.017	0.972	0.199	0.007	-0.009	
		Transfer	0.032	0.911	0.086	0.006	-0.026	
		Others	0.030	0.990	0.874	0.059	0.029	
9	2006	Total		0.437				
# of observation = 4981		Earned	0.812	0.465	0.923	0.797	-0.015	
		Financial	0.014	0.970	0.598	0.018	0.005	
		Real estate	0.052	0.971	0.737	0.085	0.033	
		Social insurance	0.022	0.949	0.178	0.009	-0.014	
		Transfer	0.048	0.832	-0.054	-0.005	-0.052	
		Others	0.052	0.977	0.823	0.096	0.044	
11	2008	Total		0.429				0.429-0.487= -0.059
# of observation = 5116		Earned	0.843	0.463	0.939	0.854	0.011	0.463-0.443= 0.019
		Financial	0.013	0.965	0.486	0.014	0.001	0.965-0.967= -0.002
		Real estate	0.044	0.974	0.740	0.073	0.030	0.974-0.992= -0.018
		Social insurance	0.025	0.942	0.123	0.007	-0.018	0.942-0.998= -0.056
		Transfer	0.051	0.841	0.083	0.008	-0.043	0.841-0.990= -0.150
		Others	0.025	0.982	0.772	0.044	0.019	0.982-0.993= -0.011

Source: KLIPS.

points respectively. The decline is also not surprising since the number of recipients of these benefits sharply increased immediately after the 1997 crisis but has gradually declined since. Because the most recent wave in the KLIPS data is the 11th wave for 2008, this paper unfortunately is unable to investigate the impact of 2008 financial crisis on the income inequality in Korea.

Table 2 also reports the impact on the Gini coefficients from a change in income sources for five waves. For all six income sources (earned, financial, real estate, social insurance, transfer, and others), table 2 lists 1) the source k's share of total income (S_k), 2) the relative Gini of component k (G_k) , 3) the Gini correlation between income component k and total income (R_k), 4) the share of each income source in total inequality (Share) , and 5) % change on Gini from a 1% increase in income source k (% Change).

For all years, the earned income accounts for the largest shares in total income in the range between 81% and 86 %. The share of transfer income increased from 1.5% in 1998 to 5.1% in 2008, signaling the increasing importance of transfer income.[18] All other income sources, except real estate income, report relatively small and almost negligible share in the total inequality. Earned income became more unequal from 1998 (0.443 for G_k) to 2008 (0.463 for G_k). Its share in total income also increased from 0.833 to 0.843. Although the Gini correlation between earned income and total income declined (from 0.948 to 0.939), "0.939" is still shows a strong correlation. The strong correlation of earned income's R_k reflects that the earned income is unequally distributed and flows disproportionately toward those at the top of the income distribution.[19] At the risk of stating the obvious, the earned

18 Transfer income includes 1) auxiliary income from relatives and acquaintances for living and education expenses; 2) unilateral monetary assistance from government and social organization, including in-kind assistance. Traffic fare discount for senior citizens and lunch delivery from local government on a constant basis are also included in transfer income.

19 The Gini correlation of income component *k* with total income R_k is defined as the ratio of the covariance between the income component and the cumulative distribution of total income to the covariance between the income component and the cumulative distribution of that income source: $R_k = \mathrm{cov}[y_k, F(y)] / \mathrm{cov}[y_k, F(y_k)]$ (Stark *et al.* 1986: 725).

income favors the rich. These all imply the importance of earned income in understanding the overall household income inequality.

The last column of Table 2 lists the % change on Gini from a 1% income in income sources. In 2008, the impact of a 1% increase of the four types of incomes (earned income, financial income, real estate income and other incomes) would have increased overall income inequality. Holding other income components constant, a 1% change in real estate income will positively affect the overall income inequality by 3%. As expected, the increase in social insurance income and transfer income will reduce the overall income inequality. This is because the lower income class is more likely to receive transfer income and social insurance income, and these extra incomes would noticeably boost their total income. A 1% increase in transfer income will reduce the overall income inequality by 4.3%, holding other income components constant. This result confirms the importance of redistribution in general and transfer payment in particular to reduce the overall income inequality.

5. Discussion and Conclusions

This paper reviews the trends of income inequality in Korea and examines the sources of income inequality from 1998 to 2008. The findings confirm the importance of earned income as a major component in total income. It was found that an increase in transfer income reduces the income inequality most significantly among all income sources. The findings have important policy implications. In order to raise the overall economic status, policy efforts should be exerted to create more stable employment opportunities with better compensation. However, just because an individual is hired at a workplace with better pay does not necessarily mean that income inequality decreases. Nevertheless, despite the possibility of increasing inequality when other people earn more, the benefit of overall increase

Therefore, a large positive value of R_k of an income component shows that the income component favors those whose total income is high.

in economic status from more employment opportunities with better pay will outweigh the potentially increasing income inequality.

In order to reduce overall income inequality, more active policy intervention to increase the transfer payment is called for.[20] The increase of transfer payment will inevitably require an increase in tax. Although this paper does not address the distributive aspects of specific form of tax, there is much research addressing this issue. For example, Hyun and Lim (2005) find that Korea's income tax system can have more redistributive effect by increasing the level of horizontal equity, leading to the equal tax treatment of equal income group. This requires the abolishment of such tax incentives as allowance, deduction and exemption. The nature of tax policy is also important. Despite the government's claim that the recent tax cuts are not for the rich, to the extent that the tax cuts are implemented differently across income distribution in favor of the rich, the income inequality will persist.

It is encouraging that in 2010, the Gini coefficients, quintile ratios, and relative poverty rates for both market and disposable income improved for the first time since 2006 (Figures 1, 3 and 5) However, it is still premature to predict whether or not this improvement was a temporary result from the active government intervention with transfers and tax cuts. To continue this improving momentum, the pressing policy challenge is more than merely implementing stronger economic policies for growth and redistribution. What is urgently called for is to reduce the prevalent uncertainty and to create social and public consensus to share the economic pain and gain. For example, while the government is calling for "sharing growth" with SMEs, business leaders in general are dismissive of the notion of profit share. One prominent business leader was even quoted to have protested that the government's proposal of profit sharing reeks of socialism. This is a telling example that shows how hard it is to change the mindset of "those who have" and lead them to share the outcomes.

---

20 There is an intense debate over the relationship between welfare spending and growth. One view is that welfare spending reduces income inequality and thus leads to growth. The opposing view is that unproductive welfare spending results in tax increases and thus hurts economic growth (Sung and Park 2011).

As Korea has become more global in almost all aspects, the existing inequality structure “becomes both complicated and hardened through the intricate interconnections of domestic and global factors in favor of the reproduction of class privilege” (Koo 2007: 31). For example, the acquisition of often-coveted skills such as English speaking proficiency or exposure to the environment of advanced countries that are usually attained through study-abroad program at the college level require a sizeable amount of financial support. To the extent that this type of skill and investment are closely related to the future labor market success, the chances for the younger members from the low income households to move up the income ladder and socio-economic hierarchy would remain slim.

OECD (2008) finds that social mobility is generally higher in countries with lower income inequality, and vice versa. As Korea is forging its way ahead toward economically developed, socially mature and cohesive country, the findings of this paper shed a light for future policy directions. The current income inequality, if left unchecked, would slow down the social mobility and potentially create class conflicts which will hamper the achievement of the cohesive society. In the long run, this will possibly perpetuate and even exacerbate income inequality and may shake social stability.

Given the importance of transfer income to reduce inequality, some form of tax raise appears to be inevitable. The social consensus to share “the pain and gain” will lighten the potential tax resistance from the rich and induces the poor to be more patient. Conscientious political leadership based on reality and feasibility rather than crowd-pleasing rhetoric is also urgently needed since the public policies based on populism and myopic special interests only retard individual incentives and further exacerbate the income inequality and income polarization. It is also important to note that policies should be supported by sufficient resources and adequate coordination with other relevant policy aspects. If and only if these prerequisites are met, economic and public policies to reduce labor market inequality and polarization will work, and the lower and the poor class will feel the “trickle-down” effect.

As Stockhammer (2010: 21) argues, three major building blocks of neo-liberalism – globalization, financialization and rising inequality – are closely intertwined to create the imbalances that caused the most recent global economic crisis. As Korea has become heavily linked with world economy, Korea is not insusceptible to economic turbulence in global setting. Although it remains uncertain how these turbulences affect inequality in Korea, the existing evidence suggests that the workers and households in the lower tier of income distribution will most likely to bear brunt of the burden. To counter these ramifications of factors and events working against income inequality, active intervention toward better-targeted income support and distribution, employment security and training program, more transparent tax codes are strongly called for.

Acknowledgement

I thank the participants of the 2011 KDI Journal of Economic Policy Conference and Dr. Young-Chul Kim for his helpful comments on a previous version of this paper.

References

Azam, Mehtabul and Abusaleh Shariff, 2011, "Income Inequality in Rural India: Decomposing the Gini by Income Sources", Economics Bulletin 31(1), 739-748.

Chakravarty, Satya R. 2009. Inequality, Polarization and Poverty: Advances in Distributional Analysis. Springer, New York, NY.

Cho, Joonmo and Jaeho Keum. 2009. “Dualism in Job Satisfaction of the Korean Labour Market: The Impact of the 1997 Financial Crisis,” Pacific Economic Review 14(1): 155-175.

Esteban, Joan-María and Debraj Ray. 1994. “On the Measurement of Polarization.” Econometrics 62(4): 819-51.

Fields, Gary S. and Gyeongjoon Yoo, 2000, “Falling Labor Income Inequality in Korea's Economic Growth: Patterns and Underlying Causes,” Review of Income and Wealth 46(2), 139-159.

Hayutin, Adele M. 2007. Global Demographic Shifts. PREA quarterly (Fall 2007), 46-53.

Hyun, Jin Kwon and Byung-In Lim. 2005. “The Financial Crisis and Income Distribution in Korea: The Role of Income Tax Policy,” The Journal of the Korean Economy 6(1), 51-65.

Jang, Yeon-Joo and Min-Hwan Lee. 2010. “An Analysis on the Korean Household Income Inequality by Using the Gini Decomposition.” Presented at 2010 Asia-Pacific Economic Association.

Lerman, Robert and Shlomo Yitzhaki. 1985. “Income Inequality Effects by Income Sources: A New Approach and Applications to the United States.” The Review of Economics and Statistics 67(1), 151-156.

López-Feldman, Alejandro. 2006. “Decomposing Inequality and Obtaining Marginal Effects.” The Stata Journal 6(1), 106-111.

Koo, Hagen. 2007. “The Changing Faces of Inequality in South Korea in the Age of Globalization.” Korean Studies 31, 1-18.

Kuznets, Simon. 1955. Economic Growth and Income Inequality. American Economic Review 45: 1-28.

OECD. 2011. "A Framework for Growth and Social Cohesion in Korea."

OECD. 2008. Growing Unequal? Income Distribution and Poverty in OECD Countries.

Shin, Kwanho and Donggyun Shin. 2007. "Bi-Polarization of the Income Distribution and Its Socio-Economic Consequences in Korea," [in Korean] The Analysis of Korean Economy 13(1): 63-123.

Stark, Oded, Edward Taylor and Shlomo Yitzhaki. 1986. "Remittances and Inequality." The Economic Journal 96(383), 722-740.

Stockhammer, Engelbert. 2010. "Neo-liberalism, Income Distribution and the Causes of the Crisis," Research on Money and Finance Discussion Paper no 10. Available at http://www.researchonmoneyandfinance.org/media/papers/RMF-19-Stockhammer.pdf

Sung, Myung Jae. 2010. "Population Aging, Mobility of Quarterly Incomes, and Annual Income Inequality: Theoretical Discussion and Empirical Findings," Presented at the 66th Annual Congress of IIPF.

Sung, Myung Jae and Ki-baeg Park. 2011. "Effects of Taxes and Benefits on Income Distribution in Korea." Review of Income and Wealth 57(2), 345-363.

Sung, Myung Jae and Ki-baeg Park. 2009, "Effects of Demographic Changes on Income Inequality in Korea." Kyong Je Hak Yon Gu (Korean Economic Association) 57(4), 5-37.

Wolfson, Michael. 1994. "When inequalities diverge," American Economic Review, Papers and Proceedings, 84 (2): 353-8.

Yoo, Gyeongjoon. 2007. "The Measure of Income Inequality and Polarization," [in Korean] Korea Development Institute Policy Research Series 2007-01.

Yun, Ji-Whan. 2009. "Labor Market Polarization in South Korea: The Role of Policy Failures in Growing Inequality," Asian Survey 49(2): 268-290.

CHAPTER 7

The Impact of Mortgage Securitization on the Housing Bubble and Subprime Mortgage Crisis: A Self-organization Perspective

by
*Hong Y. Park**
(College of Business and Management Saginaw Valley State University)
Hyejung Chang
(School of Management Kyung Hee University)
Kaustav Misra
(College of Business and Management Saginaw Valley State University)

Abstract

The subprime mortgage crisis has been analyzed from many different perspectives. The securitization of subprime mortgages has emerged as the leading cause of the subprime mortgage crisis. This securitization is a complex process that involves a number of different players (Ashcraft and Schuerman, 2008).

Securitization of subprime mortgages, which are a part of mortgage-backed securities (MBSs) also led to further complexity by the introduction of collateralized debt obligations (CDOs) and credit default swaps (CDSs). MBSs, CDOs and CDSs became sources of adverse selection and moral hazard which have contributed significantly to the current subprime mortgage crisis. Securitization of mortgages also made the mortgage market global, which provided opportunities for

* The first author would like to express his appreciation for the financial support of Saginaw Valley State University.

homebuyers in the U.S. to draw funds from all over the world. Our study investigates the impact of securitization of mortgages on mortgage rates, the housing bubble and the subprime mortgage crisis.

The study found that securitization of mortgages has an inverse relationship with mortgage rates and that securitization of subprime mortgages triggered the housing bubble in 1995. The housing bubble contributed to the construction boom and economic growth while it was expanding, but it caused catastrophic adverse effects on the U.S and global economies when it popped.

We found that an application of the self-organization principle in biology and thermodynamics to the analysis of the current housing bubble provides insight in the current subprime mortgage crisis and other bubbles in general. This study may offer a fresh new perspective for policy makers.

1. Introduction

The subprime mortgage crisis caused a major slowdown in the U.S. economy and economies around the world, imposing severe adverse effects on millions of people (Wu and Yang 2008). Numerous economists have analyzed the nature of the subprime mortgage crisis and offered solutions to the crisis. Shiller (2008) argues that the housing bubble, caused by irrational exuberance, led to the dangerous over-expansion of credit, which resulted in a global credit crunch. In a similar vein, Morris (2008) points out that the expansion of credit generated by home equity credit fueled high consumption and strong economic performances. During the housing bubble, credit was extended for "ninja" loans—no income, no job, no assets—and ninja loans became a part of subprime mortgage loans.

Some writers argue that the subprime mortgage crisis was caused by the housing bubble; others seek the origin of the crisis in the securitization of mortgages. It is likely that the housing bubble is related to the securitization of mortgages; thus treating the housing bubble and securitization of mortgages independently may not give a full picture of the current subprime mortgage crisis. Securitization of mortgages increases the liquidity of mortgages by creating secondary markets for them. As the improved liquidity of mortgages becomes a more desirable asset, demand for the asset increases. As demand increases, prices of mortgages increase and mortgage rates decline. The resulting housing prices and positive feedback of higher housing prices reinforce the housing bubble, demonstrating the relationship between the housing bubble and the securitization of mortgages.

The securitization of mortgages transformed the “originate and hold” mortgage model to the “originate and distribute” model (Mizen, 2008, p. 538). This model has, however, contributed to mispricing of risk (Mizen, 2008), moral hazard and adverse selection (Ashcraft and Schuermann, 2008). The collapse of the housing bubble and mispricing of risk appear to be key culprits of the subprime mortgage crisis. This paper, therefore, examines the process of the transformation of the originate and hold model to the originate and distribute model, and the relationship between the housing bubble and the securitization of mortgage credit.

We can draw an analogy between the formation of the housing bubble and the self-organization principle in biology and thermodynamics. Self-organization in biological systems refers to a broad range of pattern-formation processes in both physical and biological systems (Camazine et al., 2003). The formation of the bubble pattern in the housing industry can be analyzed from this self-organization perspective, which provides a useful analytical framework to explain the housing bubble pattern and offer a solution to the subprime mortgage crisis.

The paper examines the origin and evolution of mortgage securitization and the impact of securitization of mortgages on the housing bubble and subprime mortgage crisis. We also discuss solutions to the subprime mortgage crisis from a self-organization perspective and an alternative to self-organization.

2. Securitization of Mortgages: Its Complexity and Impact

2.1. Transformation to the originate and distribute model

Securitization of subprime mortgages is a result of the transformation of mortgage financing from the originate and hold model to the originate and distribute model. The complexity of this model grew over the years as financial institutions attempted to solve problems stemming from mortgage securitization.

This history may be traced to 1968, when the Government National Mortgage Association (Ginnie Mae) securitized Federal Housing Administration and Veterans Administration (FHA/VA) mortgages backed by the "full faith and credit" of the U.S. government for resale in the secondary market (Mizen, 2008, p. 536). Mizen (2008) indicates that government sponsored enterprises (GSE) such as the Federal National Mortgage Association (Fannie Mae) and Freddie Mac then began to securitize prime mortgages in 1980. Ginnie Mae is a government agent and Fannie Mae and Freddie Mac are government sponsored enterprises (GSEs). Securitization by a government agent and the GSEs meant that prime mortgage products and securitized prime mortgage products were subject to almost zero default risk. In contrast, private sector financial institutions' involvement in securitization, including high quality (prime) loans, subprime loans and Alt-A loans and their MBSs, are subject to significant default risk (Mizen, 2008). In 1984 securitization of prime mortgage loans by private sector financial institutions emerged and in 1995 private financial institutions began to securitize subprime and Alt A mortgages (see Appendix A and B).

By 2006 according to Rosen (2007), Ginnie Mae's guaranteed mortgages accounted for 4% of all mortgage backed securities (MBSs) issued. The GSEs involved 40% of MBSs; the remaining 56% were repackaged by private sector financial institutions. The mortgage backed securities (MBSs) share of total mortgage debt outstanding was about 56%, the non-MBSs share about 44%. MBSs, a kind of asset-backed securities (ABS), became more complex as private sector financial institutions issued more complicated new products, and pools of MBS were collected and securitized. Various special purpose vehicles (SPVs) were established to create new asset-backed securities from complex mixtures of residential MBSs, credit card and other debt receivables; To avoid banks' capital requirements, special purpose vehicles (SPVs) were treated as off-balance sheet items. Most MBSs included securities backed by prime loans, subprime loans or Alt-A loans, which are "issued to borrowers that appear to have good credit, but these loans do not meet the definition of prime or conforming" (Rosen, 2007, p. 2).

In this process, private financial institutions categorized asset-backed securities (ABS) into three tranches: senior, mezzanine, and equity levels based on the priority of the claim that holders of these financial instruments can make in case of bankruptcy. Bonds that are themselves backed by pools of bonds are referred to as collateralized debt obligations (Rosen, 2007); a number of the collateralized debt obligations (CDOs) purchased MBSs and securities of other CDOs. Banks hold asset-backed securities in warehouses before intermediating credit to end investors. Securitization gets complex as financial institutions develop a market for collateralized debt obligations, and Mizen's (2008) statement on the complexity of securitization illustrates the process of complexity in mortgage securitization:

> Some tranches of CDOs were then pooled and resold as CDOs of CDOs (the so-called CDOs-squared); CDOs-squared were even repackaged into CDOs-cubed (Mizen, 2008, p. 538).

These CDOs were distributed to final investors and various entities including primary lenders, mortgage brokers, bond insurers and credit rating agencies (OECD, 2008), all of whom participated in this process at various stages from origination to final distribution. Unlike the originate and hold model, which does not involve these stages and thus generates less credit risk, the originate and distribute model involves credit risk at each stage, and this mispricing of credit risk has been a key element in the current subprime mortgage crisis.

2.2. Problems in the originate and distribute model

What are the sources of credit risk? Economists contribute the mispricing of risk to information asymmetry and frictions stemming from it. According to Ashcraft and Scheurmann (2008), the securitization process is subject to seven key frictions:

(1) Frictions between the mortgagor and the originator: predatory lending because subprime borrowers can be financially unsophisticated.
(2) Friction between the originator and the arranger: predatory borrowing and lending; the originator has an information advantage over the arranger with regard to the quality of the borrower.
(3) Frictions between the arranger and third-parties: adverse selection; the arranger has more information about the quality of the mortgage loans, which creates an adverse selection problem: the arranger can securitize bad loans (the lemon) and keep the good ones.
(4) Frictions between the servicer and the mortgagor: moral hazard.
(5) Frictions between the servicer and third-parties: moral hazard.
(6) Frictions between the asset manager and investor: principal-agent.
(7) Frictions between the investor and credit rating agencies: model error (Ashcraft and Scheurman, 2008, pp. i and ii).

Because of these frictions, the originate and distribute model of mortgage lending creates opportunities for multiple problems: predatory lending and borrowing, moral hazard, adverse selection, principal-agent problem, and model error in credit rating. These problems led to mispricing of risk, which together with the housing bubble, caused the subprime mortgage crisis (Ashcraft and Scheurmann, 2008; Mizen, 2008; Morris, 2008; Rosen, 2007; Schiller, 2008). Since a significant portion of banks' revenues are generated by fees on originating mortgages, servicing mortgages and issuing MBSs, these fee-generating activities have changed the nature of banking. Banks began to engage in more fee-generating activities as they serviced mortgages and issued MBS. Therefore, securitization of mortgages likely fostered changes in banking practice to fee-generating banking. Banks continue to involve in various fee generating activities such as ATMs and automatic loans.

Furthermore, the originate and distribute model creates a larger number of steps and opacity of the financial system between the originator and the final holder of mortgages. Mizen cites comments from Alexander Lamfalussy and William Buiter, the former general manager of the Bank for International Settlements and former chief economist of the European Bank for Reconstruction and Development, respectively, who note that "banks have replaced the 'originate and hold' model of lending long and borrowing short, with an 'originate and distribute' model in which they lend and then sell the claims to someone else" (Mizen, 550). The originate and distribute model was designed to solve problems stemming from lending long and borrowing short in the originate and hold model. However, this model has accompanied problems which were not solved during its development. Mizen argues that a larger number of steps between the originator and holder added greater opacity to the process and contributed to the mispricing of risk that was not properly appraised. He further points out that the extension of originate and distribute banking to subprime mortgage securities created an asset class with an opaque ownership structure, and is ultimately responsible for the subprime mortgage crisis, as banks created an asset class of special purpose vehicles (SPVs) and put them on the off-balance sheet to avoid their capital requirements. According to Mishkin (2010), avoiding regulation is a typical behavior of financial institutions, but the consequences are often overlooked.

2.3. Problems in managing mortgage credit risk

A further complexity is added to the originate and distribute model (Mizen, 2008) in management of mortgage credit risk: a credit default swap (CDS) is utilized for management of credit risk. The CDS is a credit derivative contract between two counterparties: the buyer makes periodic payments (premiums) to the seller, and in return receives a payoff (protection) if any underlying financial instrument defaults during the term of the CDS contract. CDSs can be bought by most institutional investors, but it is not necessary for the buyer to own any CDO. The cost of insurance to cover default risk using CDSs had become much more expensive as subprime mortgage default increased and the ABX declined during 2008. The ABX index launched in January 2007, and is used as an indicator of default risk. The ABX index serves as a benchmark of the market for securities backed by home loans issued to borrowers with weak credit. The ABX index tracks the performance of a basket of credit default swaps (CDSs) based on U.S. subprime loans, and traders and investors are allowed to take positions without actually holding CDSs (Wong, 2008). Sellers of CDS were not anticipating the collapse of the housing bubble and there were also misleading or fraudulent opportunistic positions taken by financial institutions. Financial institutions on the wrong sides of positions and sellers of CDS became insolvent as the housing bubble popped. These financial institutions were bailed out by the government. According to the Bank for International settlements report, there was an estimated $62.2 trillion worth of CDSs contracts outstanding worldwide in 2008 (Morgensen, 2008-02-17, New York Times, ISDA market survey).

The originate and distribute model was designed to solve the problems in the originate and hold model of high interest rate risk and low liquidity in lending long and borrowing short. However, the illiquidity and high interest rate risk of the originate and hold model led to high mortgage rates, which reduced housing demands. The problems of interest rate risk of lending long and borrowing short became more severe in the early 1980s because of wide swings in short-term interest rates. However, the new solution came with new problems, such as increased opacity and mispricing of risk in the financial system. Problems generated by this new model have still not been adequately addressed. Popper's (1982) fundamental evolutionary sequence of events illustrates that solutions eliminate errors, but also generate new problems that need to be solved. Therefore, "all organisms are constantly, day and night, engaged in problem-solving" (Popper, p. 110). Policy makers might be advised to adopt Popper's constant problem-solving framework throughout their entire evolution of any new lending model or policy.

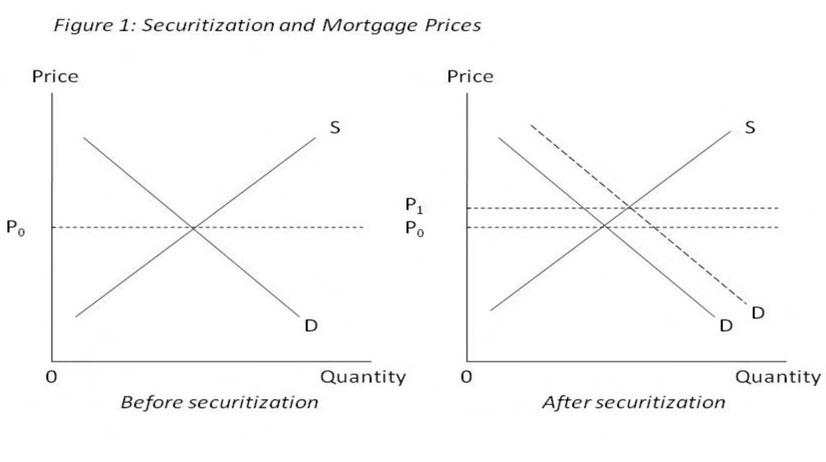
3. Securitization of Mortgages, Mortgage Rates and Housing Bubble

3.1. Securitization of mortgages and mortgage rates

Models such as Figure 1 demonstrate how securitization influenced mortgage rates and the housing bubble. Prices of an asset are determined by the demand and supply of the asset. Determining factors of the demand for mortgages are relative expected returns, taxes, liquidity and wealth (Mishkin, 2010). Securitization of mortgage credit increases the liquidity of mortgages and shifts the demand for mortgage to the right. As the demand for mortgages increases, the price will go up and mortgage rates will decline, due to an inverse relationship between prices of bonds and interest rates (Mishkin, 2010). The prices of mortgages will rise from P_0 to P_1 in Figure 1 due to mortgage securitization, and mortgage rates will decline as a result of an increase in the price of mortgages. Based on the effect of liquidity on mortgage rates, we can formulate the following hypothesis:

Hypothesis 1: Securitization of mortgage credit lowers mortgage rates.

Figure 1 Securitization and Mortage Prices



3.2. Securitization of Subprime Mortgages and the Housing Bubble

The housing bubble can be further explained by the spontaneous responses of actors in the mortgage market. What was the triggering or tuning mechanism of the housing bubble? Based on the model used above, securitization of subprime mortgages was the likely source. The questions are: What comes after the triggering or tipping point in a bubble and How does a bubble form? Shiller's irrational exuberance (2008) and Keynes' animal spirit (1936) attempt to answer the question. We argue that the self-organization theory in biology and thermodynamics (Kauffman, 1993; Prigogine and Stenger, 1984) help explain the formation of the bubble.

Shiller's home price index shows that the current nominal housing bubble started in 1995 and ended in 2006 (Shiller, 2008). He argues that the housing bubble was caused by irrational exuberance and the social contagion of boom thinking. He stresses the feedback effect in producing speculative bubbles:

> Psychological, epidemiological, and economic theory all point to an environment in which feedback of enthusiasm for speculative assets, or feedback of price increases into further price increases, can be expected to produce speculative bubbles from time to time. They make clear that these bubbles can have complicated—sometimes random and unpredictable—dynamics. (Shiller, 2008, p. 47)

Akerlof and Shiller (2009) argue a similar point in their newly published book, Animal Spirits. They point out that a key to address the current problem is to recover Keynes's (1936) insight about "animal spirits"—the spontaneous attitudes and ideas that guide economic action:

> Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many years to come, can only be taken as a result of animal spirits of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of

quantitative benefits multiplied by quantitative probabilities. (Keynes, 1936, p. 161)

The guiding principles of animal spirits are relatively simple and spontaneous. The self-organization theory is similar to animal spirits in terms of responding to information only in the area immediately around you. We argue that the self-organization perspective (Kauffman, 1993; Prigogine and Stenger, 1984) can explain the formation and expansion of the housing bubble well. Self-organization refers to a broad range of pattern formation processes in both physical and biological systems (Camazine et al., 2003). Camazine et al. define self-organization:

> Self-organization is a process in which a pattern at the global level of a system emerges safely from numerous interactions among the lower-level components of the system. Moreover, the rules specifying interactions among the system's components are executed with only local information, without reference to the global pattern. (Camazine et al., 2003, p. 8)

Furthermore, "The multiplicity of interactions that characterizes self-organization systems emphasizes that such systems are dynamic and require continual interactions among lower-level components to produce and maintain structure" (Camazine et al., 2003 p. 8). Home buyers, banks and regulators react to information on increasing housing prices; home buyers buy houses anticipating the appreciation of housing prices; and banks make mortgage loans with the expectation that mortgage loans are secured because home values will continue to exceed mortgage values. These spontaneous interactions and cascades of events form a housing bubble, which expands until the bubble reaches its maximum.

The concept of feedback in biological and physical systems (Kauffman, 1993; Prigogine and Stenger, 1984) can add more insight to analysis of the housing bubble pattern. Camazine et al. (2003) explain the relationship between feedback and change.

> Feedback can have two basic values: positive or negative. Feedback is positive if the recurrent influence reinforces or amplifies the initial change.

> The snowballing effect of positive feedback takes an initial change in a system and reinforces that change in the same direction as the initial deviation. Self-enhancement, amplification, facilitation, and autocatalysis are all terms used in positive feedback. (Camazine et al., 2003, p.17)

> The amplifying nature of positive feedback means that it has the potential to produce destructive explosions or implosions in any process where it plays a role. How can such snowballing be kept under control? This is where negative feedback plays a critical role, providing inhibition to offset the amplification and helping to shape it into a particular pattern. (Camazine et al., 2003, p. 19)

Camazine et al. (2003) illustrate self-organization phenomena using a school of fish and a herd of reindeer. A school of fish maneuvers gracefully, with all its members moving in parallel in the same direction. The reason that fish do not run into each other is due to a negative feedback, which helps maintain equilibrium for a school of fish. Similarly, a housing bubble moves upward, with all home buyers moving in the same direction. In other words, the housing bubble can be explained by persistent housing price increases due to an extended duration of positive feedback.

This raises a question: What triggered the initial change in the housing bubble? We believe that the initiation of private financial institutions' involvement in subprime mortgage credit securitization was a tipping point in the current housing bubble. As stated before, the subprime MBSs by private financial institutions grew rapidly after their inception in 1995. The increasing MBSs by private financial institutions made credit available to subprime mortgagors and created high demands for housing, thus increasing housing prices. Home buyers then responded to the appreciation of housing prices, and financial institutions responded to opportunities for generating fees in mortgage origination, servicing, issuing MBSs and CDOs, and underwriting mortgages and MBSs. Both home buyers and financial institutions reacted based on their local information, and the positive feedback of the initial change formed and amplified the housing price bubble. Thus we can observe that the housing bubble pattern emerged from numerous interactions among the lower level of components (e.g., home buyers), as seen in the self-organization of biology and physics (Camazine et al., 2003; Haken, 1977; Kauffman, 1993; Prigogine and Stengers, 1984). Interactions among home buyers and financial institutions are made based on local information and they are components of housing and financial markets. The pattern of the current housing bubble has emerged from interactions among home buyers and financial institutions.

What then made the housing bubble collapse and reverse itself? There are two factors: the housing bubble exhausted credit availability and household debts reached capacity limits (Morris, 2008). To maintain equilibrium negative feedback is required (Kauffman, 1993). Positive feedback accompanies negative feedback and all bubbles, therefore, eventually collapse (Schiller, 2008). Negative feedback offers an opportunity for error elimination (Plotkin, 1982; Popper, 1982). Errors in the securitization and housing bubble were mispricing of risk and lack of due diligence in all actors in the originate and distribute model, including government regulatory agents. As housing prices decreased, speculative downward demand reinforced a decline in housing prices, creating a gap between the value of the mortgage and the value of a house.

Based on the impact of securitization and self-organization on the housing bubble, we can draw the following hypothesis:

Hypothesis 2: The current housing bubble is triggered by securitization of subprime mortgages.

Hypothesis 3: Positive feedback in securitization of subprime mortgages amplifies the Housing bubble and negative feedback reverses the housing bubble: positive and negative feedback make the pattern of the housing bubble Sigmoidal.

4. Model and Empirical Results

4.1. Model

We can test these hypotheses with models. In the mortgage rate model (equation 1), we include the mortgage securitization ratio to test hypothesis 1; the inflation rate variable is included based on the Fisher effect on nominal interest rates. The GDP growth rate variable is used to capture the relationship between the demand for an asset and income. We expect that there is an inverse relationship between the mortgage rate and securitization ratio, and the inflation rate and GDP growth rate variables are expected to show a positive relationship with the mortgage rate variable.

The housing price model (equation 2) is specified to test hypothesis 2 and 3, as well as four theories that have been frequently cited as explanations for the current housing bubble. The housing price model includes the mortgage yield rate, mortgage securitization, real GDP growth rate, building cost index, and a dummy variable for subprime mortgage securitization. Except for the building cost index variable, these variables are determining factors of the demand for housing.

Models include the following elements to reflect theories of the mortgage rate and housing price:

1. Securitization of subprime and Alt A mortgage loans: This securitization started in 1995, the current housing bubble started at the same time. Subprime mortgage securitization increased housing demand by offering mortgages to less qualified home buyers.
2. The irrational exuberance hypothesis (Robert Shiller, 2008)
3. The low mortgage rate (interest rate) hypothesis: The Fed Chairman, Greenspan, maintained lower interest rates, which led to lower mortgage rates during the housing bubble. There is an inverse relationship between housing prices and mortgage rates.
4. Securitization of mortgages: Securitization of mortgages provides more funds for home buyers; mortgage-backed securities (MBSs) were sold all over the world and drew savings from the U.S. as

well as other countries. Securitization may also lower mortgage rates.

We specified our models to reflect these theories and models as follows:

MRT = f(INF, MSR,RGDPG) (1)
HPI = f(MSR, RGDOG, BCI, SUBP) (2)

Where, MRT: the 10 year mortgage yield rate
INF: the inflation rate
MSR: the mortgage securitization ratio (the ratio of securitized mortgages to total mortgage outstanding)
RGDPG: the real GDP growth rate
HPI: the housing price index
BCI: the building construction cost index
SUBP: the dummy variable for securitization of subprime mortgages
(SUBP=0 before 1995 and SUBP=1 after 1995)

4.2. Data and empirical results

Data used in estimation of models are from Shiller (2008), the 2008 Presidential Report and the 2008 mortgage market statistical annual (Inside Mortgage Finance Publication). Detailed sources of each variable are as follows:

HPI, BCI: Robert Shiller (2008)
MRT, INF, RGDPG: the 2008 Presidential report
MSR: estimated by the author based on the 2008 mortgage market statistical annual (Inside Mortgage Finance Publication)

| Table 1 | Estimation Results

MRT = 16.20 - 0.157 MSR – 0.1972 RGDPG – 0.0547 INF

(0.08) (0.00) (0.05) (0.58)

R^2 = .90 R^2 (adj) = .89 F (3, 24) = 76.16

HPI = - 503.29 + 275.40 MSR – 3.7625 RGDPG + 5.832 BCI + 28.389 SUB

(0.06) (0.02) (0.26) (0.04) (0.16)

R^2 = .63 R^2 (adj) = .62 F (4, 23) = 15.19

Structural Change Test for Subprime Mortgage

Test	Break Point	Num DF	Den DF	F Value	P Value
Chow	15	5	15	14.16	0.0001

Note: The numbers in parentheses are p-values.

These data are time series data from 1980 to 2007. We employed both a multiple regression estimation method and the SAS statistical package to estimate regression coefficients and Chow test statistics. Regression coefficients and p-values are presented in Table 1. The mortgage rate variable is deleted from the housing price model because of a strong correlation between the mortgage rate and the securitization of mortgage. The mortgage securitization variable reveals a statistically significant inverse association with the mortgage rate and a statistically significant positive association with the housing price variable. These results support the main thrust of our investigation on the subprime mortgage crisis. Detailed discussions are presented in the following section.

5. Empirical Analysis and Discussion

5.1. Securitization and the Mortgage Rate

We specified the mortgage rate model to find the impact of mortgage securitization on the mortgage yield rate. Independent variables included in the model are mortgage securitization ratio, the real GDP growth rate, and the inflation rate. Securitization reveals a statistically significant inverse association with the mortgage rate, which conforms to the expectation and supports hypothesis 1 that mortgage securitization has an inverse association with the mortgage rate. However, the real GDP variable has a statistically significant negative association with the mortgage rate, and the inflation variable does not show a statistically significant association with the mortgage rate variable. Signs of these two variables do not conform to the expectations of the theoretical model.

5.2. Subprime Mortgage Securitization, the Housing Bubble and Self-organization

Model (2) tests whether the subprime mortgage securitization was a triggering factor in the U.S. housing bubble; regression coefficients are estimated based on data from 1980 to 2008. We included additional variables such as the building construction cost index and real GDP growth rate variables in this model. The real GDP growth rate variable is not statistically significant and does not reveal the expected positive relationship with the housing price. The building cost index variable is statistically significant with the expected positive sign.

The mortgage securitization variable shows an expected sign and is statistically significant. The dummy variable representing securitization of subprime mortgages reveals a positive sign, although it is not statistically significant at the conventional α level. To find out whether subprime mortgage securitization led to a structural change in the housing price, we conducted a Chow test. The Chow statistics show that the break point is the 15^{th} observation (1995), which is the year that subprime mortgage securitization started. The Chow test statistics are statistically significant and the test results suggest that subprime mortgage securitization was a trigger point for the current U.S. housing bubble. Figure 2 confirms the breaking point for the structural change.

When we examined data from 1980 to 2007, we found that the current housing bubble started in 1995. The mortgage securitization ratio started to rise significantly in 1990 due to a sharp increase in both government sponsored enterprises (GSEs) and non-agency mortgage securities. Securitization of non-agency mortgage securities (private sector financial institutions) had started in 1984 with a securitization ratio of 19.91%.; by 2007 the securitization ratio rose to 59.59%. The non-government agency mortgage securities were only 4.2% in 1984, but rose to 31.98% in 2007, which can be calculated from columns (2) and (3) in Appendix B. Securitization of mortgages increases the liquidity of mortgages and reduces the interest rate risk of lending long and borrowing short. The problem of lending long and borrowing short became more apparent in the early '80s because of the wilder swing of the short-term interest rates. The securitization rate of mortgages, a parameter which changes between 0 and 1, is comparable to a tunable parameter in biology (Camazine et al., 2003). There is no securitization of mortgage when the parameter is 0 and 100 percent securitization when it is 1. Changes in the securitization parameters result in changes in housing prices. Securitization is a strategic or policy variable in management or economics.

What did trigger the current housing bubble? One answer may be securitization of subprime mortgage credits. Securitization of mortgage credit by Ginnie Mae started in 1968; other GSEs became involved in securitization in 1980. Securitization of subprime mortgages by private financial institutions started in the mid-1990s and increased rapidly. We argue that this is a tipping point. When a person raises the temperature, water does not boil until the temperature reaches the threshold point. Although securitization began earlier, the housing bubble did not start until the sharp increase in securitization of subprime mortgages in 1995. In thermodynamics and biology, the tipping point is referred to as the threshold point of bifurcation (Heylighen, 2008; Kauffman, 1993; Prigogine and Stengers, 1984). The snowballing effect of positive feedback takes an initial change in a system and reinforces that change in the same direction as the initial deviation (Camazine, et al., 2003, p. 17). In the same way, the initial housing bubble triggered by initial external subprime mortgage securitization in 1995 amplified the housing bubble and the pattern of housing prices changed in 1995.

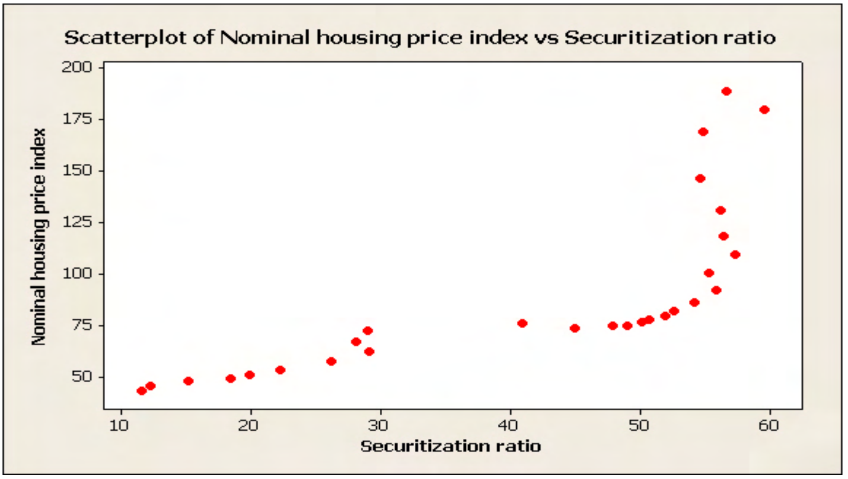
To test the impact of subprime mortgage securitization on the housing bubble, we used a dummy variable for the year that the securitization of subprime mortgages was introduced. Empirical results show that this variable had a positive impact on the housing bubble. The downward snowballing effect of negative feedback can provide inhibition to offset the amplification. The onset of the decline in housing prices triggered by the subprime crisis (unsustainable debt) can be amplified as time passes; home buyers act on information about increasing housing prices, they interact with each other, and they also interact with financial institutions. Consequently, the pattern of the housing bubble becomes nonlinear (see Figure 2).

The pattern of this housing bubble forms a sigmoidal curve. Kauffman's (1993) explanation of a sigmoidal curve in self-organization can be applicable to the sigmoidal curve of the housing bubble pattern:

> The sigmoidal function is initially below the proportional response. Here a given output levels to an output that is less than the input. Were that reduced output fed back as the next input, then the subsequent response would be even less. Over iterations, the response would dwindle to zero. The sigmoidal response becomes steep in its midrange, however, and crosses above the proportional response. An input above this critical crossing point leads to an output that is greater than the proportional-response output. In turn, were that output fed back as a next input, the output would be still greater than input. Over iterations the response would climb to a maximum. (Kauffman, 1993, p. 184)

As discussed earlier, the pattern of the housing bubble was formed by numerous interactions of home buyers, financial institutions and credit rating agencies. As long as a housing bubble continues to expand, MBSs are secured debts because they are asset-backed securities (ABSs). Securitization of subprime mortgages offers more mortgage loans to home buyers, and housing prices increase as a result of increased demand. As housing prices climbed to a maximum in 2006, the housing bubble reversed itself due to constraint of household debt capacity. These behaviors of housing prices make the pattern of the housing bubble sigmoidal, as we can see in Figure 2. Camazine et al. (2003) have shown how a small change in a system parameter can result in a large change in the overall behavior of the system (p. 35). Figure 2 shows that a change in securitization caused by subprime mortgage

| **Figure 2** | A Scatter Plot of Nominal Housing Prices and Mortgage Securitization



securitization resulted in large changes in housing prices. This also confirms the butterfly effect, the analogy that a butterfly flapping its wings could cause hurricanes in another part of the world (Lorenz, 1963; Rosser, 2000). Miller and Page (2007) also illustrate how systems of interacting agents can lead to emergent phenomena. Their linking individually based micro processes to macrosocial outcomes is a useful analogy to the formation of housing bubble phenomena. This result shows a nonlinear relationship between inputs and outputs. Thermodynamics and self-organization offer possibilities those small inputs can generate large outcomes and large inputs may result in small outputs; they provide a different perspective from classical Newtonian physics. In the Newtonian universe outputs are proportional to inputs and the perturbation eventually gravitates toward equilibrium. However, in thermodynamics and self-organization the perturbation reaches to a new attractor (equilibrium) by going through rugged landscapes. The originate and distribute model perturbed the originate and hold model and went through rugged landscapes. The new attractor (equilibrium) will have characteristics that are different from the originate and hold model. Policy makers and managers need to become aware of characteristics of the new attractor as they address problems stemming from the new order.

Securitization of home mortgages has continued to increase by adding more institutions, such as the government agency (Ginnie Mae), government sponsored enterprises (Fannie Mae and Freddie Mac) and private sector financial institutions. Securitization of private financial institutions includes prime mortgages, subprime mortgages and Alt A mortgages. The pattern of the housing bubble changes as securitization of mortgages increases. Figure 2 also illustrates the nature of the relationship between self-organization and housing prices. This relationship is supporting evidence for hypothesis 3.

The maximum point is another bifurcation point where negative feedback reinforces the downward change or amplifies the initial change. In the case of a school of fish, positive feedback and negative feedback work with little time lag, so that each fish maintains its distance with neighboring fish and equilibrium (Camazine et al., 2003). However, the housing bubble pattern shows that positive feedback amplifies a housing bubble and controlling or negative feedback operates with a significant time lag. The housing bubble started in 1995 and lasted until 2006. Although operating rules for home buyers and financial institutions are simple, the global pattern of spontaneous interactions of home buyers and financial institutions became complex. This complex global pattern of spontaneous interactions led to the crisis. Data for Figure 2 are from Shiller (2008) and the authors' estimation based on the Inside Mortgage Finance Report (2008).

5.3. Securitization of Subprime Mortgages and Subprime Mortgage Crisis

Column (5) in Appendix B shows the share of the subprime and Alt A mortgages out of total non-agency issuance (private financial institutions). The share of the subprime and Alt A mortgages has been growing since 1995 and peaked in 2006, with 78.8% of the total MBS issuance amount by private financial institutions.

When securitization of subprime and Alt A loans by private sector financial institutions started in 1995, mortgage backed securities became a significant part of the collateralized debt obligations (CDOs). Resecuritization of CDOs and a derivative of CDOs, credit default swaps (CDSs), further complicated the subprime mortgage problem. As stated before, there was an estimated $62.2 trillion worth of CDSs contracts outstanding worldwide in 2006 (Morgensen, 2008). Private sector financial institutions' involvement in subprime mortgage securitization caused subprime loans and Alt A loans to increase due to the transferability of risk in the originate and distribute model. Thus, securitization of subprime loans and Alt A loans became a main source of the current subprime mortgage crisis. MBS involves three risks: interest rate risk, prepayment risk and default risk (Rosen 2007). Securitization of subprime mortgages increases the default risk. A recourse clause (in the representations and warranties) that obligates originators (lenders) to buy back loans that are later discovered not to have originated with proper due diligence became ineffective at reducing the risk because so many originators became insolvent.

Securitization of mortgage was able to draw funds from the global financial market and fueled the housing bubble for an extended period. The globalization of mortgages made the mortgage market more opaque and the people with superior information took advantages of the less informed people.

Therefore, those who had superior information increased their income at the cost of less informed people.

6. Solutions to the Subprime Mortgage Crisis

6.1. A Self-organization Perspective

One self-organization solution to the current financial crisis is to envisage the solution from the perspective of an individual financial institution, home buyer, regulation agency and investors in the financial system. A collapse of the housing bubble can be seen as the operation of negative feedback (Camazine, 2003), which offers an opportunity for error elimination (Plotkin, 1982; Popper, 1982). Individual financial institutions will react to the financial crisis based on their local information regarding the financial crisis. They are likely to scrutinize mortgage loans and to require a higher percentage of down payments. They will also interact with other actors in the system such as loan applicants, savers, other financial institutions and regulatory agencies, based on their local information. Financial institutions will exercise more due diligence in dealing with loan applicants and with mortgage-backed security issuers such as MBS and CDO issuers. They will pay special attention to CDSs. Investors will also scrutinize these financial instruments before they invest in them.

A general characteristic of a self-organizing system is robustness or resiliency (Camazine et al., 2003; Heylighen, 2008; Kauffman, 1993; Prigogine and Stengers, 1984). Self-organization in biology states that living organisms react to changes in environment, and their reactions lead to a global order (Camazine et al., 2003). The current perturbation will push financial systems into a better state or equilibrium, a state where individual financial institutions, consumers and regulators mutually adapt. Foster (1992, 2000) argues that the firm, as a complex adoptive system with self-organizational qualities, can develop a range of forward looking contractual arrangements in the context of transaction cost economics. The same can be said about the current changes in the financial environment. Financial firms will develop a range of new financial arrangements in dealing with the mispricing of risk and the opacity created by the originate and distribute mortgage financing model. Financial systems will establish a new global order stemming from the current perturbations.

As Kauffman (1993) argues, "evolution is a complex combinatorial optimization process in each of the coevolving species in a linked ecosystem, where the landscape of each actor deforms as the other actors move" (p. 644). Financial institutions securitized subprime mortgages and made mortgage loans available to low income home buyers, which made home buyers buy more homes. The result deformed the final holders of CDOs and led to the current crisis. The crisis, a stimulus to the financial system and a complex combinatorial optimization process in each of the coevolving home buyers, financial institutions, final investors (holders) and regulatory agencies, will help create new systems. The mutual interactions of these actors help generate more healthy and efficient financial intermediaries. Kauffman (1993) also points out that evolution is an emergent order honored and honed by selection. Emergence of a new financial system likewise will have the order honed by selection, and the new order will show how order emerges from the chaos. Moreover, the situation can become worse before it gets better, and the U.S. financial systems will go through the rugged landscape until it reaches an equilibrium or attractor. The experience on the 1997 Korean economic crisis provides an evidence of the self-organization perspective solution. However, alternative solutions may be required to reduce pain in the short-run and to trigger economic recovery.

6.2. Alternative Solutions to Self-organization

Alternative solutions to self-organization perspectives include central authorities, blueprints, templates and leadership. The U.S. government and the Federal Reserve are central authorities; they are working to solve problems created by the housing bubble formed by positive feedback of self-organization. The collapse of the housing bubble has created enormous adverse effects on the U.S and world economies. This negative feedback is a step in the process of reaching a new equilibrium from the collapse of the housing bubble. Camazine et al. (2003) point out that the "individual acquires and processes information that elicits a negative feedback response: A small perturbation applied to the system triggers an opposing response that counteracts the perturbation" (p. 16). In a biological system, this negative feedback prevents an implosion. An ideal solution to problems in the housing bubble might have been due diligence exercised by the regulatory agencies while the housing bubble was forming and expanding. Can regulatory agencies play this role? This may be a challenging task, because non-linear dynamic self-organization models are capable of generating catastrophic discontinuities, chaotic dynamics and a variety of other complex dynamics, as noted by Rosser (2000). Home buyers and financial institutions were interacting with each other based on local information. Home buyers responded to increasing housing prices, and financial institutions responded to opportunities for fee generation from mortgage origination, servicing mortgages, issuing MBSs and CDOs and underwriting mortgages and MBSs. They did not see problems stemming from the global pattern, group characteristics and complexity. Regulatory agencies thought that actors in markets are smart and control themselves to maintain balance. However, markets failed to self-regulate and negative feedback did not operate until the housing bubble reached a maximum. Therefore, a new regulatory regime may be required to prevent problems of bubbles from recurring.

Today almost all large financial institutions in the U.S. and the world are involved in securitization as MBS and CDO issuers and underwriters. They actively participate in CDSs to manage their risk on CDOs. Consequently, problems faced by financial institutions, the U.S. economy and other economies in the world are severe and widespread. The crisis involved trillions of dollars worldwide, thus the problem requires massive coordinated efforts by governments in the world. To address this problem the U.S. Congress passed the Troubled Assets Recovery Program (TARP) and has designated $750 billion to purchase assets and equity from financial institutions. TARP, designed to strengthen the financial sector, allows the U.S. Department of Treasury to buy illiquid, toxic assets from banks and other financial institutions and to provide relief to homeowners who are facing mortgage bankruptcies. TARP also encourages banks to resume lending both to each other and to consumers and businesses. Interbank lending will restore financial market stability and make bank loans available to consumers and businesses, which will help increase consumer spending on durable goods such as automobiles, housing and furniture. However, toxic assets are likely to rise unless policy makers take measures to stop and reverse the trend. The U.S. government is taking steps to stem the foreclosure trend; it is planning to spend $75 billion for home owners who are facing home mortgage foreclosures and to offer subsidies to new home buyers to boost new home construction.

However, the costs of stabilizing financial systems are expected to rise significantly higher than the currently appropriated dollar amounts. U.S. policy makers may draw lessons from the Japanese real estate crisis in the 1990s. According to the New York Times (February 13, 2009: Hiroko Tabuchi, B 1), the Japanese economy endured a "lost decade" of economic stagnation as Japanese banks and policy makers were slow to recognize the magnitude of their banking problems and wasted trillions of Yen on half-measures. U.S. policy makers and banks need to confront the issues directly and should take effective steps fast. When the Korean government, banks and business firms took bold measures as they faced the 1997 financial crisis, they recovered relatively quickly (Park, 2008). Therefore, the timing and implementation of appropriate policy measures are crucially important in recovery of the U.S. economy from the current crisis. Current policy debates and institutional interventions on problems of self-organization have the same familiar tone of Keynesian and classical policy debates (Rosser, 1999).

However, hindsight on the current financial crisis offers some clues for government interventions in the future. Since design of the originate and distribute model, the development of off-balance sheet special purpose vehicles (SPVs) and the securitization of subprime mortgages have contributed to the current financial crisis, these developments should have been properly monitored as they were developing. Are government agencies capable of doing the job? Scholars continue to debate this question.

The self-organizational policy prescription would further suggest that policy makers study the evolving nature of patterns (Colander, 2000) and address problems as they arise.

Rather than bounding after the unknowable, and try to deduce analytically models that hold for all times, economics has reduced its search to what it believes is knowable. New Millennium economists search for patterns in data, try to find temporary models that fit the patterns, and study the changing nature of those patterns as institutions change. (Colander, 2000, p. 131)

Economists could not have known all potential problems when they designed and implemented the originate and distribute mortgage model. Problems emerged in the process of evolution in the new model. The securitization of subprime mortgage accompanied the cascade of events and a complex system has emerged as agents in financial institutions managed their credit risks and consumers acted on their local information of subprime mortgage securitization. The complexity of subprime mortgages led to catastrophic adverse effects on the global economy because the securitized subprime mortgages were sold to savers all over the world. Asymmetry of information between the originators of subprime mortgages and its final holders in other countries was more pronounced than other securities because of the complexity of subprime mortgage securitization. Therefore, government interventions require better understanding of the global patterns of bubbles (Rosser, 1999) and mechanisms for selecting a solution among tentative solutions (Popper, 1982) to reduce policy errors in mortgage financing changes.

Policy makers need to develop a process to mobilize knowledge from all knowledgeable people because people working in the field have concrete experiences and knowledge that the new system is experiencing and its future direction. Actions generate intended and unintended consequences (Giddens, 1984). Intended consequences of mortgage securitization are low mortgage rates, higher liquidity of mortgages and global market access to credit. Unintended consequences are special investment vehicles (SIVs) and complex CDOs and CDSs. Unintended consequences necessitate further actions which can stem the adverse effects of actions. The U.S. financial market and government have failed to take steps to stem the adverse effects of the unintended consequences. Reasons for the failure of corrective actions or inactions are numerous, but the condition that prevailed at the time in the U.S. was the idea that the market knows the best. The current debt of the U.S. federal government limits fiscal policy options to address problems of the subprime mortgage crisis at this critical juncture. This illustrates that corrective action alternatives are conditioned on the situation of the time and space (now and in the U.S.). Making a correct policy choice requires consciousness and knowledgeability of actors (policy makers) and actors tend to have knowledge in the field. There were conscious voices among actors who were working in the field and became aware of the problems of subprime mortgages, but we ignored their voices. Therefore, it is necessary to establish channel knowledge of conscious actors to address problems of adverse consequences. We tend to pay more attention to the intended consequences of new policy, but often fail to have due diligence on the unintended consequences of new policy.

When the Federal Reserve and the new administration began to take more aggressive steps, the Fed made over one trillion dollars of loans to financial intuitions in 2008, the Fed also purchased subprime mortgages and it is likely to maintain an easy monetary policy until early 2011. The new administration and Congress in 2009 worked out massive fiscal policy stimulus measures to boost the U.S. economy. The Bush administration had implemented tax cuts, but the amount of the cuts might not have been sufficiently large enough to trigger economic stimulation. The Obama administration had taken more aggressive stimulus fiscal policy measures. They proposed over $800 billion of recovery and reinvestment programs, and the U.S. Congress actually adopted the $787 billion stimulus and recovery package. Some economists argued that this amount was inadequate to address the problem of shortfalls in private sector spending. They argue that the spending gap was 1.3 trillion dollars.

Regulatory agencies are also scrutinizing financial institutions and their transactions more closely. Financial regulatory agencies are criticized for their lack of regulations and due diligence. The increasing tendency of deregulation and strong belief in a free market system in the past two decades created an environment for less stringent regulatory implementation. However, these trends are reversing now and some economists are proposing a financial product consumer protection act, equivalent to the current consumer protection act. Opaqueness stemming from information asymmetry among actors in securitization of mortgage credit created moral hazard, adverse selection and conflict of interest. These problems may have contributed to the current subprime mortgage crisis. New or existing regulations need to reduce or eliminate the opacity of the system and increase transparency of the securitization process. The system should encourage the originator to take more risk rather than transferring it. Securitization of subprime mortgages and new issuances of CDOs and CDSs need to be monitored with due diligence. Shiller (2008) has proposed that rather than more regulatory measures, we should have more democratization of financial systems to provide independent advice and information to financial consumers, so that consumers can make informed financial decisions. He also points out that the collapse of the housing bubble makes housing more affordable to home buyers. Ultimately, it is important to design an efficient regulatory system, and the current subprime mortgage crisis presents an opportunity for decision makers to create efficient new orders for the U.S. and the global financial systems. The U.S. Congress actually passed a financial consumer protection act to protect financial consumers and to prevent recurring future financial problems. The new regulatory institution for monitoring financial market is the consumer finance protection bureau (CFPB).

7. Summary and Conclusion

The originate and distribute model was originated to increase the liquidity of mortgages (lower mortgage rates) and to address interest rate risk stemming from lending long and borrowing short in mortgage financing. The new model facilitated the securitization of mortgages. However, it has created new problems of opacity, mispricing of risk and complexity in mortgage financing, as Popper (1982) predicted. Regulatory agencies need to understand these sources of opacity, mispricing of risk and complexity which developed underneath the housing bubble.

The pattern of the housing bubble or any bubble for that matter can be analyzed from a self-organization perspective. The current housing bubble was triggered by the securitization of subprime and Alt A mortgage loans by private financial institutions in 1995. The subprime mortgage crisis arose as the housing bubble popped, and the current crisis provides a good lesson why due diligence is required in new mechanism design (Maskin, 2008). Since no one would know all potential problems stemming from the emerging complexity of a new originate and distribute model at the time of its adoption, monitoring in a new model needs to be flexible and ongoing throughout the development of complexity in the model. Characteristics of complexity in self-organization are not universal and therefore controlling any bubble needs to be case-specific (Colander, 2000; Rosser, 1999).

Although economic bubbles are recurring frequently and impose enormous adverse effects on millions of people as an economic bubble collapses, not enough studies have been conducted on economic bubbles. There is a need for more research to address problems effectively, and the self-organization perspective provides an appropriate framework for both the analysis and policy prescriptions of economic bubbles.

References

Akerlof, George and Robert Shiller, 2009, Animal Spirits: How Human psychology Drives the Economy, and Why It Matters for Global Capitalism, Princeton: Princeton University Press.

Ashcraft, Adam B. and T. Scheurman, 2008, "Understanding the securitization of subprime mortgage credit," Federal Reserve Bank of New York Staff Reports no. 318, 1-76.

Camazine, Scott, J. Deneubourg, N. Franks, J. Sneyd, G. Gheraulaz and E. Bonabeau, 2003, Self-organization in Biological Systems, Princeton and Oxford: Princeton University Press.

Colander, David, 2000, "New millennium economics: How did it get this way, and what way is it?" The Journal of Economic Perspectives, Vol. 14(1), 121-133.

Foster, J., 1992, "Economics and the self-organization approach: Alfred Marshall revisited?" The Economic Journal, 103, 975-991.

Foster, J., 2000, "Is there a role for transaction cost economics if we view firms as complex adaptive system?" Contemporary Economic Policy, 18(4), 369-385.

Giddens, Anthony, 1984, The constitution of society, Berkeley and Los Angeles, University of California Press

Haken, Herman, 1977, Synergestics, Berlin: Springer-Verlag Publishing Co.

Heylighen, Francis, 2008, "The Science of self-organization and adaptivity," Free University of Brussels, Belgium, Working Paper, 1-26.

Inside Mortgage Finance, 2008, The 2008 Mortgage Finance Market Statistical Annual, Vol. I and II.

Kauffman, Stuart A., 1993, The Origins of Order: Self-organization and selection in evolution, New York and Oxford: Oxford University Press.

Keynes, John M., 1936, The General Theory of Employment, Interest, and Money, New York: Harcourt, Brace and Company.

Lorenz, E.N. 1963, "Deterministic non-periodic flow," Journal of Atmospheric Science Vol. 20 (March), 130-41.

Maskin, E.S., 2008, "Mechanism design: How to implement social goals," American

Economic Review, 98(3), 567-576.

Miller, John, H. and Scott E. Page, 2007, Complex adaptive systems, Princeton and Oxford: Princeton University Press.

Milne, Alistair and Geoffrey Wood, 2008, "Banking crisis solutions old and new," Federal Reserve Bank of St. Louis Review, 90(5), 517-530.

Mishkin, Frederic S. , 2010, The Economics of Money, Banking and Financial Markets, 9th edition, Boston: Pearson. Addison Wesley Publisher.

Mizen, Paul, 2008, "The credit crunch of 2007-2008: A discussion of the background, market reactions, and policy responses," Federal Reserve Bank of St. Louis Review, 90(5), 531-567.

Morgensen, 2008-02-17, New York Times, ISDA market survey.

Morris, Charles R., 2008, The Trillion Dollar Meltdown, New York: Public Affairs Publishing Co.

Organization for Economic Cooperation and Development (DECD), 2008, Financial Market Trends, 94(1), 13.

Park, Hong Y., 2008, "Economic crisis, learning and innovation in Korean firms," 2008 KDI-KAEA Conference Proceedings, Seoul, Korea: Korea Development Institute, 1-42.

Plotkin, H.C., 1982, "Evolutionary epistemology and evolutionary theory," in Learning, Development, and Culture (ed) H.C Plotkin, Chichester, U.K.: John Wiley & Sons, 3-16.

Popper, Karl, R., 1982, "Of clouds and clocks," in Learning, Development, and Culture, (ed) H.C. Plotkin, Chichester, U.K.: John Wiley & Sons, 109-119.

2008 Presidential Report, U.S Governmental Printing Office.

Prigogine, I. and I. Stengers, 1984, Order Out of Chaos, New York: Bantam Books

Rosen, Richard J., 2007, "The role of securitization in mortgage lending," Chicago Fed Letter No. 244, 1-4.

Rosser Jr., J. Barkley, 1999, "On the complexities of complex economic dynamics," Journal of Economic Perspectives, Vol. 13(4), 169-192

Rosser, Jr., J. Barkley, 2000, "Aspects of dialectics and non-linear dynamics," Cambridge Journal of Economics, Vol. 24(3), 311-324.

Shiller, Robert J., 2008, The Subprime Solution, Princeton and Oxford: Princeton University Press.

Tabuchi, Hiroko, 2009, New York Times, February 13, B 1.

Wong, Grace, 2007, Behind Wall Street's subprime fear index, CNNmoney.com.

Wu, Deming and Jiawen Yang, 2008, "Securitization, liquidity, risk and banks' risk exposure," Working paper, 1-46.

Appendix

Appendix A Outstanding Mortgage Securities and Securitization Rate

	(1) GSE	(2) Non-agency	(3) Total Security	(4) Total Mortgage	(5) Securitization
1980	111,086		$111,086	$957,900	11.59%
1981	126,186		126,187	1,030,200	12.25
1982	162,829		162,892	1,070,200	15.22
1983	219,201		219,201	1,186,100	18.48
1984	252,007	11,000	263,007	1,132,100	19.91
1985	314,554	24,000	338,554	1,518,600	22.29
1986	434,884	16,600	451,484	1,722,000	26.22
1987	531,867	27,800	559,667	1,920,500	29.14
1988	570,733	34,900	605,633	2,154,100	28.12
1989	646,759	43,300	690,057	2,378,900	29.01
1990	1,013,920	55,000	1,068,920	2,614,700	40.88
1991	1,152,453	96,700	1,249,153	2,781,700	44.91
1992	1,268,777	142,300	1,411,077	2,947,300	47.88
1993	1,354,638	167,900	1,522,538	3,106,200	49.02
1994	1,461,376	183,000	1,644,376	3,283,200	50.08
1995	1,554,901	193,800	1,748,701	3,451,200	50.67
1996	1,691,357	215,400	1,906,757	3,674,700	51.88
1997	1,801,319	253,500	2,054,819	3,900,600	52.56
1998	1,985,873	321,500	2,307,373	4,258,500	54.18
1999	2,255,520	353,200	2,608,720	4,674,200	55.81
2000	2,449,956	377,500	2,827,456	5,110,300	55.33
2001	2,791,287	463,200	3,254,487	5,678,000	57.32
2002	3,087,430	544,100	3,631,530	6,437,400	56.41
2003	3,394,021	664,000	4,058,021	7,227,800	56.14
2004	3,467,047	1,049,800	4,516,847	8,270,500	54.61
2005	3,607,558	1,536,600	5,144,158	9,374,300	54.87
2006	3,904,911	1,991,500	5,896,411	10,421,400	56.58
2007	4,518,871	2,116,600	6,635,471	11,135,800	59.59

Sources: Columns (1), (2), (3), (4) are from the 2008 Mortgage Finance Market Statistical Annual Vol. I, II (Inside Mortgage Finance); column (5) is calculated by the author.

Appendix B Non-agency MBS issuance by type

	(1) Prime	(2) Subprime	(3) Alt A	(4) Total	(5) [(2)+(3)]/(4)
1995	25,837.7	17,771.4	498.3	44,107.4	41.42%
1996	31,418.7	30,769.4	1,802.6	63,990.7	50.90
1997	49,974.9	56,920.7	6,518.0	113,413.7	55.93
1998	97,365.2	75,829.9	21,235.5	194,430.6	49.92
1999	74,630.9	55,851.5	12,022.8	142,505.2	47.63
2000	53,584.9	52,467.4	16,443.6	122,495.9	56.25
2001	142,202.5	87,052.9	11,373.6	240,629.0	40.90
2002	171,534.4	122,680.9	53,462.7	347,678.0	50.66
2003	237,454.6	194,958.5	74,151.0	506,564.1	53.12
2004	233,378.1	362,549.3	158,585.8	754,513.2	69.07
2005	280,703.7	465,036.3	332,323.2	1,078,063.2	73.96
2006	219,037.4	448,599.6	365,675.8	1,033,312.8	78.80
2007	180,462.4	201,546.7	249,610.0	631,619.1	71.42

Sources: Columns (1), (2), (3), (4) are from the 2008 Mortgage Finance Market Statistical Annual, Vol. I, II (Inside Mortgage Finance). Column (5) is calculated by the author. Columns (1), (2), (3) and (4) are billions of dollars.

CHAPTER 8

Real Estate Investors, the Leverage Cycle and the Housing Market Crisis*

by
Andrew Haughwout
Donghoon Lee
Joseph Tracy
Wilbert van der Klaauw

(Federal Reserve Bank of New York)

Abstract

We explore a mostly undocumented but important dimension of the housing market bubble and bust: the role played by real estate investors. Using unique credit report data, we document large increases in the participation of investors, especially in the states that experienced the largest housing boom-bust cycle, where at the peak almost half of purchase mortgage originations were associated with investors. Consistent with Geanakoplos' theory of the leverage cycle, we identify a shift toward optimistic buyers, identified here as "buy and flip" investors. During the house price boom these "flippers" showed an increased willingness and ability -- facilitated in part through apparent misreporting of their intentions to occupy the property -- to take on increased leverage. After home prices began to drop, investors defaulted at a much higher rate than single-home owners, representing over 30%

* We benefited from helpful comments and suggestions from participants at the April 2011 Housing Economics and Research Conference at UCLA. The views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System.

of aggregate delinquent mortgage balances. Our findings have important implications for the design of policies to address the deleterious consequences of the current crisis and to limit future occurrences of housing market bubbles.

The U.S. economy is still recovering from the financial crisis that began in the fall of 2007. The collapse of house prices across many markets was a precipitating factor in the financial crisis and adverse feedback effects between financial markets and the real economy led to the most severe recession in the post-war period. Extraordinary interventions by fiscal and monetary authorities both in the U.S. and abroad were required in order to prevent a complete collapse of global markets and the potential onset of another great depression.

Attention has shifted from containing the financial crisis to examining its causes and designing policies to limit both the likelihood and the severity of a similar crisis in the future. Given the central role that housing played as a catalyst to the crisis, it is important to better understand the determinants of the dynamics of house prices and of subsequent mortgage defaults over this recent cycle. While house prices were rising in many parts of the country over the period leading up to the crisis, these increases were particularly pronounced in four states-Arizona, California, Florida and Nevada (the "bubble" states). Figure 1 shows the path of house prices in the US, the bubble states as a whole, and in each of these states from 2000 Q1 to 2010 Q4.Over the period from 2000 to 2006 average house prices more than doubled in each of these states. The pace of house price appreciation accelerated starting in 2004. The peaks in prices across the four states occurred within a couple of months of each other in mid-2006. Following the turn in the markets, house prices declined rapidly in each state with much of the earlier gains given back within just two years.[1]

This rapid run-up and then crash in house prices exacted a terrible cost to homeowners, financial firms and to the economy. Current estimates are that around 23 percent of active mortgages are "under

1 California is a bit of an exception in that it appears that average house prices have stabilized at a level 50 percent higher than in 2000.

water” in that the balance on the mortgage exceeds the current value of the house.[2] As of 2010 Q4, nearly 2.8 million homes have gone through foreclosure, and another 2 million homes are in the process of foreclosure.[3] Serious delinquencies continue to add new homes to the foreclosure pipeline over time. Nationally distress sales represent around half of all repeat-sale transactions. These distress sales continue to exert downward pressure on house prices making it more difficult for housing markets to recover.

A focus on residential mortgage finance in order to understand what the determinants were of the house price and mortgage default dynamics generated over the recent cycle would inform efforts to enhance financial stability. A more robust system of residential mortgage finance should aim to limit the degree to which house prices rise and fall over a credit cycle. Reducing the amplitude of the house price swings will limit the potential for collateral damage created by housing markets for the real economy.

1. Related Literature

Given that housing is a durable asset, periods of rising prices are indicative of increasing demand for housing.[4] One strand of the literature on housing demand focuses on the determinants that affect the “user cost” of housing.[5] The user cost of housing (UC) is the annual flow cost to the owner per dollar of house price, taking into account after-tax financing costs, property taxes and insurance, maintenance and depreciation costs and the expected risk-adjusted return to owning the house. The value of the housing service flow is proxied by the annual

---

2 http://www.corelogic.com/About-Us/News/New-CoreLogic-Data-Shows-23-Percent-of-Borrowers-Underwater-with-$750-Billion-Dollars-of-Negative-Equity.aspx

3 http://www.ots.treas.gov/_files/490069.pdf

4 That is, with the exception of natural disasters and periods of armed conflict, the supply of housing in a market cannot contract significantly over a short period of time to drive up house prices. 6 See Hendershott and Slemrod (1983) and Poterba (1984) for early discussions.

5 See Hendershott and Slemrod (1983) and Poterba (1984) for early discussions.

rent (R). If we assume that there is arbitrage between owned and rental housing, then the annual rent should equate to the price of housing (*P*) times the user-cost.

$$R(Y, \gamma_m) = P * UC(\gamma_m, \tau, \delta, g^e)$$

where γ_m is the mortgage financing rate, τ describes the tax environment, δ the depreciation rate on housing net of that offset by maintenance expenditures, g^e the risk adjusted expected return to housing, and *Y* is the average income.

This framework suggests several possible candidates for explaining the rise in house prices in the early to mid-2000s. A rise in income in a housing market will increase area rental rates to a degree that reflects the elasticity of supply of rental housing in that local market. Higher rents will translate into higher house prices by a factor given by the reciprocal of the user-cost in that market. As a consequence, house prices will vary more with changes in rents in markets with low user-costs of housing.[6] The accommodative monetary policy following the bursting of the tech bubble lowered mortgage interest rates by over 300 basis points from mid-2001 to mid-2003, and facilitated a resumption of income growth after the end of the recession.[7] Lower financing costs for housing reduces the user-cost of housing which would lead to higher prices holding rents constant. However, if some of the benefits of lower financing costs to landlords are passed on to renters, then the impact of lower mortgage rates on house prices will be attenuated. The Bush tax cuts were enacted during this period which lowered marginal tax rates. These lower marginal tax rates would raise the user-cost by reducing the benefit from the mortgage interest deduction. These lower marginal tax rates would have led to lower house prices, all else the same, with the magnitude of the reduction reflecting in part expectations over whether the tax cuts would be made permanent.

While income, monetary policy and tax rates each underwent some changes in the first half of the 2000s, the term in the user-cost that has

---

6 See Himmelberg *et al* (2005) for a detailed discussion.

7 http://www.mortgage-x.com/general/historical_rates.asp

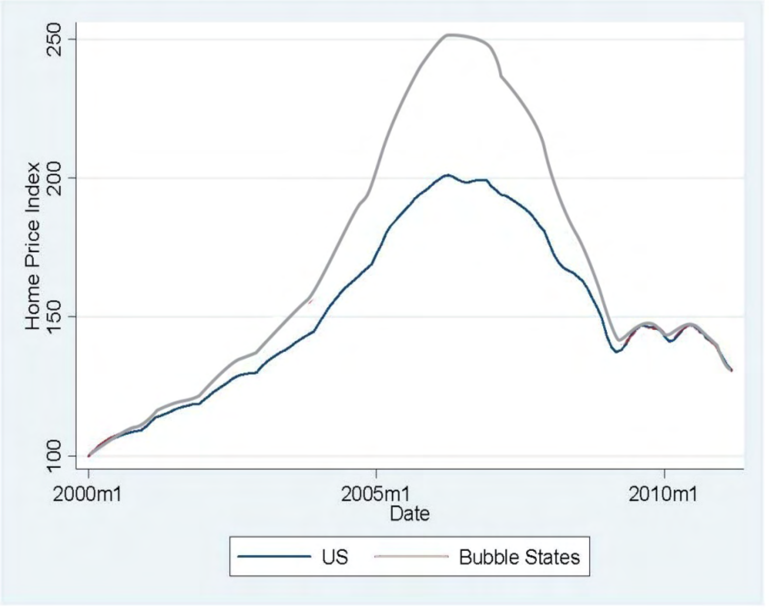
received the most attention in trying to explain the house price boom is the expected return to housing, g^e. The higher the risk-adjusted expected return, the lower the user-cost and the higher house prices will be in a market. As Himmelberg *et al* (2005) explain, the sensitivity of house prices to house price expectations increases with the degree to which house prices are expected to rise. The expected return to housing is the only forward-looking aspect to the user-cost of housing framework. The arbitrage condition listed above has a potential self-fulfilling characteristic. If owners expect house prices to rise in the future, then the user-cost of housing will fall and, given a constant rent, the value of houses will rise.[8] This rise in the value of housing can serve to confirm the earlier belief. This may lead to "irrational exuberance" in the housing market as argued by Shiller (2005).

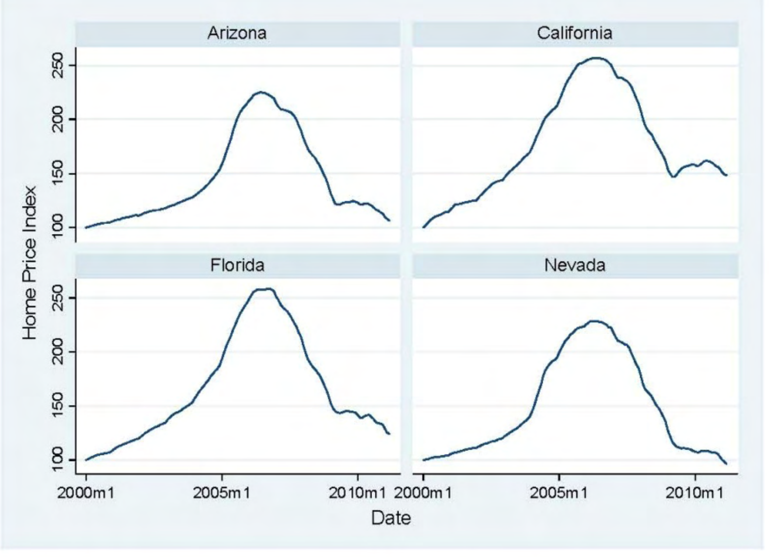
Himmelberg *et al* (2005) apply the user-cost formulation to assess the degree to which house prices dynamics track changes in fundamental demand determinants for housing. They calculate user-cost estimates for 46 metropolitan areas over a twenty-five year period ending in 2004. Their analysis identified only a few metropolitan areas where by 2004 house prices appeared to have risen significantly more than what would be predicted by average rents and estimated user-costs. It is unfortunate that their analysis ended in 2004 since the rapid acceleration in house price appreciation as shown in Figure 1 began in that year. Given their argument that the sensitivity of prices to user-costs increases at low values of the user-cost, it is possible that their methodology if extended through 2006 would have explained some of this acceleration in price appreciation. However, it is important to note that the average 30-year fixed-rate mortgage rate increased from 5.74 in January 2004 to 6.14 in December 2006, so that any further declines in the user-cost was not being driven by lower financing costs during this period.[9]

8 Rents would not be expected to rise since the value of the current flow of housing services has not changed.

9 Some authors pointed to the rise in price-rent ratios as likely to be followed by a reduction in subsequent price growth (Gallin 2008, Campbell, et al. 2009).

| Figure 1 | House Price Dynamics - US and AZ, CA, FL and NV





Notes: CoreLogic overall repeat-sale price indices. January 2000 = 100.

Glaeser *et al.* (2010) argue that the empirical connection between mortgage rates and house prices is not strong enough to explain the dynamics of house prices during the housing boom. On a conceptual level, they argue that the impact of any shift in housing demand on house prices depends on the housing supply elasticity in that market. For markets with inelastic housing supply, increases in housing demand will mainly result in higher house prices instead of increased production of new homes. In contrast, in housing markets with elastic housing supply, increases in housing demand will mainly result in the production of new homes. House prices in these markets are determined by the cost of building a new home.[10] Furthermore, they argue that expected future mortgage rates are important in addition to the current mortgage rate. If mortgage rates are expected to rise, then the effect of a low current mortgage rate on house prices will be attenuated. This argument can be captured in the user-cost arbitrage condition shown earlier by factoring the expected rise in financing costs into the expected house price appreciation term.[11]

Credit conditions enter into the standard user-cost formulation solely through the mortgage interest rate. However, a second important aspect is the required downpayment by the borrower. The interest rate and the required downpayment reflect the two underwriting constraints on a borrower when bidding on a property. The minimum downpayment percentage is also referred to as the "collateral rate" on the mortgage.[12] For a given mortgage balance, the mortgage interest rate impacts the monthly payment that the borrower will have to make. Underwriting standards will stipulate a maximum that the sum of the annual mortgage payments in addition to the taxes and insurance on the property can be as a fraction of the borrower's income.[13] We will refer to this as the

10 See Glaeser *et al.* (2010) for a detailed discussion.

11 An implication is that a reduction in interest rates that is perceived to be permanent would be expected to have a greater impact on house prices than a similar reduction in interest rates that is due to accommodative monetary policy and is expected to be transitory. This presents a challenge to those who hold the view that monetary policy was a primary determinant of the house price boom.

12 The collateral rate is also referred to as the "haircut" or "margin."

13 This is the called the front-end PITI (for principal, interest, taxes and insurance as a fraction of income) or DTI (debt-to-income ratio). There is also a back-end ratio that

"cash-flow constraint". A lender will also require the borrower(s) to make a minimum downpayment. The ratio of the downpayment to the sale or appraised value of the house determines the origination loan-to-value ratio (LTV). We refer to this as the "downpayment constraint." The maximum that a borrower may bid on a house will depend on which of these two constraints first becomes binding given the underwriting standards in use at the time.

The mortgage interest rate and the collateral rate are jointly determined in a credit market (see Fostel and Geanakoplos (2008) and Geanakoplos (2009)). For collateralized loans the collateral rate is determined by volatility of the asset price and the term of the loan. The higher the price volatility and the longer the term, the larger the collateral rate the lender will require (or equivalently the larger the minimum downpayment percentage). The purpose of the collateral rate is to safeguard the lender against defaults by the borrower when there are declines in the value of the collateral. Similarly, for a given level of price volatility, the higher the expected price appreciation for the asset the lower lenders may set the required collateral rate. In this case, the expected price appreciation acts as additional future collateral protecting the lender against losses in the event of a default.

There is an additional channel, not necessarily captured by changes in the average origination LTV and not explicitly addressed by Glaeser *et al* (2010), through which changes in credit conditions may have affected housing prices. Reduced loan documentation requirements may allow those who previously hit the collateral or cash-flow constraint to obtain more leverage, be it possibly at less favorable terms. This could occur when lenders no longer require income documentation or adopt more favorable ways of imputing the borrower's future income from wage earnings, bonuses and possibly rental income. Alternatively, in computing the DTI ratio, debts other than the mortgage loan under consideration may be ignored or incorporated differently.

---

adds to the numerator any recurring non-housing debt payments such as auto loans, student loans, and minimum credit card payments.

| Table 1 | Distribution of Mortgage Loan Documentation Level, by Year

Year	Full	Low	None
Subprime[1]			
2001	77.84	21.76	0.40
2002	71.13	28.30	0.57
2003	67.02	32.52	0.46
2004	65.37	34.34	0.29
2005	62.28	37.47	0.24
2006	61.71	38.00	0.29
2007	64.20	35.48	0.32
Alt-α			
2001	36.77	55.56	7.68
2002	40.64	51.96	7.40
2003	35.50	57.26	7.23
2004	37.75	55.72	6.53
2005	31.11	64.44	4.46
2006	18.92	76.56	4.53
2007	16.84	77.49	5.68

Notes: Authors calculations.
[1] *Source*: LoanPerformance data

As shown in Table 1, significant changes in documentation requirements for subprime and Alt-A loans took place during the run-up in house prices. The level of documentation for a new mortgage is reported as a data item on the origination file for that mortgage. The three values are full documentation, limited documentation and no documentation. While no-doc loans remained relatively uncommon, there was a sizeable shift from full-doc to low-doc loans for subprime and especially for Alt-A mortgage loans. By 2006 some 38% of newly originated subprime and 81% of new Alt-A loans were low- or no-doc loans.

We can incorporate changes in underwriting standards into the user-cost framework.

$$R(Y,\gamma_m) = P * UC(\gamma_m,\tau,\delta,g^e) * f(LTV^M,s)$$

where LTV^M is the maximum allowed origination loan-to-value ratio, s captures other prevailing underwriting standards at the time of the home purchase such as DTI and documentation, and *f* captures how changes in the degree of leverage and documentation impact house prices holding constant the user-cost.

Finally, there is a potentially important amplification effect of leverage on house prices which is not fully captured in our augmented user-cost arbitrage conditions. Geanakoplos (2009) posits that there is not a common house price appreciation expectation, g^e, that is shared by all potential buyers of an asset. Rather, he starts with the assumption that there is a distribution of expected appreciation rates across potential buyers. At the high end of the distribution are "optimistic" potential buyers with high values of g^e. Holding constant all of the other factors in our user-cost arbitrage condition, optimistic buyers will be willing to bid higher prices for housing since their user-costs are lower.

This distribution of buyers in terms of their opinions about the future value of housing can generate an amplification mechanism for house price dynamics. In normal times, optimistic buyers are infra-marginal participants in the housing market. At the prevailing house prices they would like to purchase additional housing but are prevented from doing so because the cash-flow constraint or the downpayment constraint is binding. However, during the early phase of a housing boom, lenders may reduce the required downpayment percentage on new mortgages and begin to relax other underwriting standards due to the strong performance of house prices and low delinquency rates. These actions enable the optimistic buyers to purchase additional housing. The increasing leverage allowed in the market, then, begins to shift the composition of new purchase transactions in the market toward more optimistic buyers who are willing to bid higher prices for houses. This is an additional channel by which higher leverage can amplify the upward pressure on house prices. Geanakoplos describes this dynamic as the upswing phase of a "leverage cycle".

Can increasing leverage help to explain the acceleration in house prices from 2004 to 2006? For leverage to have played an important role we need to establish at least two things. First, we need to show that leverage was increasing over these three years. Second, we need to demonstrate that the composition of purchasing activity was shifting toward more optimistic buyers. Table 2 summarizes changes in leverage from two different sources of information on housing transactions and mortgages. Glaeser *et al* (2010) report data on combined LTV ratios for purchases drawn from 89 metro areas and recorded by DataQuick. An advantage of this data is that they reflect all mortgage products that were used to finance these purchases and reflect up to three liens on the house. The second is changes in leverage on securitized nonprime purchase mortgages as recorded by LoanPerformance.[14]

| Table 2 | Mortgage Leverage During the Housing Boom

	All Housing Purchases[1]					Nonprime Purchase Mortgages[2]			
Year	25th	50th	75th	90th		25th	50th	75th	90th
2004	56	80	95	100		80	95	100	100
2005	64	86	99	100		80	95	100	100
2006	70	90	100	100		90	99	100	100

Notes: Percentiles of the distribution in each year of combined origination loan-to-value ratios
[1] *Source*: Glaeser *et al* (2010). DataQuick data from 89 metro areas
[2] *Source*: LoanPerformance data on securitized non-prime mortgages.

There are two observations that can be drawn from Table 2. First, as pointed out in Glaeser *et al* (2010) extreme leverage in the form of zero downpayment mortgages were available and used by at least 10 percent of borrowers.[15] Second, when we look below the 90th percentile, we see that leverage was increasing throughout the distribution of origination LTVs.[16] Glaeser *et at* (2010), however, conclude that the magnitude of the observed LTV changes do not appear to be large enough to be an

14 Our finding of increases in the median nonprime CLTV at origination is consistent with that of Mayer and Pence (2009). We use our matched sample, which represents a random sample of all LP loans as described below, for this table.

15 The Glaeser *et al* (2010) data indicate that the 90th percentile was at 100 going back to 1998.

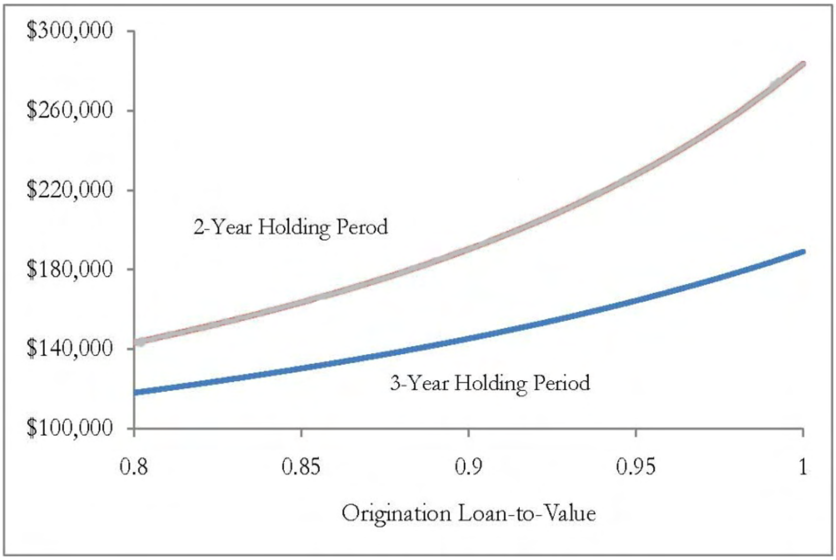
16 This is consistent with trends reported by Geanakoplos (2009, chart 1) for the average downpayment as a proportion of the purchase price. Among the 50% lowest downpayment ratios for subprime and ALT-A borrowers (based on CoreLogic data), he found a decrease in the average downpayment from 13% in the first quarter of 2000, to a low point of 2.7% in the second quarter of 2006.

important determinant of the acceleration in house prices. This conclusion may be model dependent. If Geanakoplos (2009) is correct that increasing leverage also affects prices through shifting the composition of buyers, then this additional amplification mechanism may imply that the observed changes in leverage are capable of explaining more of the acceleration in house prices. Glaeser *et al* (2010) in fact explicitly condition their conclusion (that credit market factors were not the main drivers of the housing boom) on the absence of significant composition changes in the population of buyers. They also raise the distinct possibility that the surge in the number of buyers during the boom may have been accompanied by an overall decline in credit quality of buyers not captured by LTVs, but were not able to find any evidence of large composition changes when measured by demographics.

One such potential shift in the composition of buyers during the housing boom and especially during the 2004-2006 period, concerns the number and activities of real estate investors. There are several reasons to expect credit conditions to have particularly affected investor activity in the buildup of the housing boom. In discussing these, we will distinguish between three different types of buyers in a housing market: buyers who want to live in the house (owner-occupiers), investors who want to keep the house as vacation or future retirement home or who want to rent the property and then resell at a future date (buy and hold), and investors who want to resell the property without living in or renting the house (buy and flip).

The first reason to expect a role for investors in bidding up prices concerns the impact of the previously discussed increase in average origination LTVs. For a given mortgage interest rate, reducing the required downpayment percentage can allow a borrower to bid more aggressively for a property, but this is especially so for investors. The easiest way to see the impact of variation in the allowed LTV on the maximum bid is to take the case of a "buy and flip" borrower. As an illustration, consider an investor who has $50,000 to invest in real estate. This money must cover the downpayment as well as the mortgage payments, property taxes and home insurance during the expected holding period. For simplicity, we assume that the house is financed with a 30-year fixed-rate mortgage with an interest rate of 5.5 percent. We assume that annual property taxes, insurance payments and any required maintenance expenditures equate to 2 percent of the house value. The investor will not be renting out the property during the time until resale. We consider two cases: in the first the investor plans to be able to finance the purchase for up to three years, and in the second the investor plans to be able to finance the project for up to two years.

Figure 2 Maximum House Bids and LTV



The relationship between the allowed level of leverage as indicated by the origination LTV and the maximum bid is shown in Figure 2. For the three year holding horizon, the maximum bid increases from $118 thousand with a twenty percent downpayment requirement to $189 thousand with no downpayment required - a sixty percent increase. The sensitivity of the maximum bid to changes in the origination LTV is increasing in the degree of leverage.[17] For example, reducing the

17 The nonlinearity is due to the fact that the origination LTV is determined by the ratio of the downpayment to the value of the house, which here equals the maximum house bid.

required downpayment from 20 percent to 19 percent raises the maximum bid by $2,260. Reducing the required downpayment from 5 percent to 4 percent raises the maximum bid by $4,189 - nearly double the earlier increase. For the same degree of leverage, investors with a shorter holding period have a higher maximum bid. We show how the schedule of maximum bids increases as the investor moves from a three to a two-year holding period. Shorter holding periods also increase the sensitivity of the maximum bid to changes in leverage at any given LTV. For example, with a two year holding period moving from a 5 percent to a 4 percent required downpayment raises the maximum bid by $9,335. Investors may have shortened their expected holding periods as the housing market heated up and the pace of house price increases accelerated. For a given maximum leverage, faster turn-around times for the investment properties would allow the investors to bid more aggressively.

A second mechanism through which investor behavior may have amplified the impact of changing credit conditions on house prices is through the loosening of loan documentation requirements. It has been speculated that the loosening of documentation standards may have facilitated the misreporting by borrowers of their true expected home-occupancy status.[18] This in turn may have enabled them to purchase homes under more favorable terms than they would have as investor. We explore the evidence for this possibility below.

A third channel affecting real estate investors concerns the use of second liens on existing mortgages to facilitate the down payment and meeting of loan requirements for purchasing additional investment properties. As documented in earlier work by Chakrabarti *et al* (2011), with house prices appreciating homeowners extracted home equity through higher balances on first mortgages, cash-out refinances, second mortgages and home equity lines of credit. In fact, on average for each 1% increase in home prices, homeowners increased their mortgage debt by 1%, so that proportionally their equity share in their homes actually remained relatively constant until the end of 2006. Equity extraction

---

18 See, for example, http://www.fincen.gov/news_room/rp/reports/html/mortgage_fraud112006.html.

may have been especially attractive to optimistic, but cash-constrained investors, by allowing them to use these funds to make downpayments on purchases of additional homes. Accordingly, we expect the combined LTV on existing mortgages to have increased for investors during the period in which they purchased additional properties.

Finally, we refer again to the amplification effects that result from shifts in the market toward more optimistic buyers. In the next section of the paper we explore the Geanakoplos hypothesis. We identify optimistic buyers as investors, and especially the "buy and flip" investors. We document the role of this class of investors over the past credit cycle both nationally as well in four boom states. We explore the extent to which the investor share of purchase transactions changes over the credit cycle. These changes are decomposed into both the extensive margin - more investors enter the market - and the intensive margin - existing investors increase the size of their portfolio of residential real estate exposures. We also examine the default behavior of investors as compared to owner-occupant borrowers. The final section of the paper discusses implications of our findings for current policy work on improving financial stability.

2. Investors and the Leverage Cycle

If Geanakoplos' description of the dynamics of the leverage cycle is applicable to the housing boom-bust cycle of the 2000s, we would expect to see changes in the characteristics of leveraged buyers of residential real estate over the period. In this section, we provide descriptive evidence of some major changes in the observable characteristics of mortgagors between 2000 and 2010.

While there has been some anecdotal evidence supporting the idea that investors played an important role in the boom, careful analysis of this issue has been impeded by lack of appropriate data.[19] For investors, the benefits of living in a house are immaterial to the decision of whether or not to keep making the mortgage payment, making default a

19 See, for example, http://www.metrotrends.org/commentary/mortgage-lending.cfm

less costly decision for investors than for owner-occupants. Of course, lenders are well aware of this difference, and typically require mortgagors to declare whether they will live in the collateral property, charging higher interest rates and requiring higher downpayments from those who acknowledge that they will not, *ceteris paribus*. But the interest rate penalty and limitations on leverage discourage borrowers from declaring their intention to live elsewhere, and self-reported "occupancy status" is thus considered a particularly unreliable piece of data. Haughwout *et al* (2008), for example, indicate their suspicion that miscoding of occupancy status in loan-level data may help to explain the large increase in early nonprime defaults that are unexplained by observable - i.e., reported - characteristics of loans and borrowers.[20] Fitch (2007) found evidence of occupancy misrepresentation in two-thirds of the small sample of subprime defaults they examined. It is thus desirable to identify a mortgage data source that allows the analysis of borrowers without relying on the information that is self-reported by the borrower on the mortgage application.

We bring two distinct kinds of data to the analysis of this important question. Our primary source is the FRBNY Consumer Credit Panel (CCP) which comprises an anonymous and nationally representative 5% random sample of US individuals with credit files and all of the household members of those 5%.[21] In all, the data set includes files for more than 15% of the population, or approximately 37 million individuals in each quarter from 1999-2011Q1.[22] The FRBNY CCP data allow us to overcome some of the difficulties with self-reported occupancy status. Unlike loan-level data, which focus on individual debt

---

20 Early defaults are defined to be defaults that occur within the first year.

21 The FRBNY CCP panel is based on Equifax credit report data. Lee and van der Klaauw (2010) provides further details on the data set. The analyses reported in this paper are solely based on the representative random sample and do not include the additional household members sampled.

22 In the balance of the paper, we use the term "mortgages" to refer to installment debt secured by residential real estate. Mortgage payments are typically determined so as to pay off the balance, plus interest, over a fixed time period, but some mortgages negatively amortize - the balance can grow over time. HELOCs are lines of credit, again with residential real estate as the collateral. HELOC borrowers may utilize credit up to some fixed limit.

contracts and the information used in underwriting them, credit reports are designed to give lenders (and potential lenders) dynamic credit information on individual borrowers, including the types and amounts of debt they have outstanding at any point in time. Our panel allows us to track individual *borrowers* over time, through refinances and moves, where at each point in time we observe all outstanding mortgage loans and non-mortgage debts.

We can use this information to separate mortgage borrowers based on how many distinct first-lien mortgage accounts appear on their credit reports. Since each property can secure at most a single first-lien mortgage, the number of such mortgages on a borrower's credit report is a reliable, non-self reported, indicator of the minimum number of properties a given individual has borrowed against.[23] This kind of information about individual borrowers is not available in loan-level data sets and thus the FRBNY CCP data provide a unique perspective into important questions about who is originating new mortgages at any point in time, as well as their subsequent behavior.

At this point, it is worth extending our earlier discussion of the relationship between the number of properties against which an individual has levered and what Geanakoplos describes in his leverage cycle theory. First, it is important to note that virtually all homeowners have some investment motivation in making a home purchase. While there is some debate in the academic literature about whether housing is a good investment relative to other assets, many buyers - whether they own only the home in which they live or own several units at a time - consider expected capital gains a part of their motivation for buying rather than renting (Case and Shiller 1988). However, some homebuyers differ from others in that some or all of their residential property portfolio does not also directly provide them with shelter: that is, they own multiple properties and do not live in all of them. While we recognize the investment motive of all homeowners, we will refer to these multiple property owners as "investors".

---

23 Because some properties may have no first-lien mortgages but do secure a HELOC, our count of properties is a minimum rather than an exact figure. In addition, some properties support no debt at all.

As we suggested above, some further differentiation among investors is in order. On the one hand, there exists a class of borrowers who buy properties in order to rent the housing units they contain. For these investors, the flow of rental income generated by real property is an important motivation for their investment, and the crucial consideration, as described above, is whether this income exceeds the cost of carrying the property (roughly speaking, principal, interest, taxes, insurance, maintenance net of any tax considerations) over a long period of time. Other investors may buy properties to use as a vacation or future retirement home. These "buy and hold" investors will thus be sensitive to changes in interest rates: a significant decline in rates can often offset the fixed cost of refinancing since they expect to hold the property for some time.

By comparison, the kind of investors portrayed in the popular television show "Flip This House" differs from those who hold assets for their income-generating potential. Indeed, in that program, a team of investors typically purchases a house, does some renovations and then re-sells the property to a new owner without ever receiving any rental income whatsoever. For these "buy and flip" investors, the primary motivation for the investment is capital gains, suggesting, for the reasons described above, that they will be both highly leveraged and will be considerably less sensitive to interest rate movements. In what follows, we will explore several dimensions of the behavior of investors in general, using the data on multiple first-lien mortgages as a way of distinguishing investors from owner-occupants. In our analysis we will distinguish between different categories of investors; by whether they are holding 2, 3 or 4 or more first mortgages. It is more difficult in our data to differentiate investor type - flippers vs. holders - within each category, although changes in many of the investor series as the boom unfolds are strongly suggestive of a change in composition of the investor group, as we shall see.[24]

---

24 Ex post we can differentiate by the average holding periods by different investors. However, the preferred classification would be based on the ex ante expected holding periods which we do not observe. Bayer et al (2011) use the former approach, defining flippers by the number of times they bought and sold a home in less than two years during the 1992-2005 period in Los Angeles. They then refer to

While the CCP data provides unique insights into the role of investors in the entire mortgage marketplace, it has some limitations. A specific drawback is the absence of information about the collateral property - its location and value - in the credit report data. To allow additional analyses, we have matched individual mortgage loans from the CCP data to loan-level data from CoreLogic's LoanPerformance ABS database. LoanPerformance ABS data provides detailed loan-level information on over 15 million securitized nonprime loans, including loans which were packaged into subprime and "Alt-a" private label securities, but excluding jumbo loans with balances that exceed the GSEs' conforming limits. The LoanPerformance (LP) data include detailed information on both the origination characteristics of the loans - such as level of documentation, interest rates, balance, and the value and location of the collateral property.[25] Interestingly, the data also include the borrower's self-reported occupancy intentions: indicating whether the property's purpose will be for owner-occupancy, for use as a second home, or for an investment property, which we can compare with our own definition of investors: CCP information on the number of first-liens reported contemporaneously on the borrower's credit report.[26] We will contrast the data on these "investor" definitions below; for now it is worth pointing out that the mortgage application refers to the reported use of a particular *property*, while the credit report refers to the extent of residential investment by *an individual*.

Our matched data are, of course, reflective of a subset of the entire market, albeit the part that changed most rapidly and noticeably during the boom. Even with the matched data, we are limited to analysis of individuals' credit reports: to the extent that residential real estate

---

individuals with two or three flips during the period as speculators, and those with over 10 flips as middlemen.

25 Our merge was provided by Equifax Corporation using servicers' loan numbering system. Given that the CCP constitutes a representative random sample of individuals, the matched sample represents a representative random sample of LP loans.

26 An alternative definition would define investors in a more static way, by identifying individuals who at *any point* in our sample period had more than one first-lien mortgage on their credit reports. Using that definition, our primary results remain similar to those reported here, although we discuss this distinction further below.

investment is conducted through incorporated businesses or partnerships, we will not capture that form of investor activity here. Notwithstanding these limitations, the FRBNY CCP and our matched dataset provides many unique benefits that will allow a much clearer picture of the kinds of borrowers holding, originating and defaulting on mortgages during the 2000s housing cycle. Moreover, the matched data overcome many of the limitations of the datasets used in isolation, and allow us to examine both the characteristics of the loans and details of the borrower's credit report simultaneously.[27]

3. Results

We begin by using the CCP data to provide a description of the part played by various types of buyers in the stock of outstanding mortgages. A fundamental stylized fact from the Geanakoplos model is that investors in their role as optimistic buyers ought to be playing an increasingly important role in borrowing during the upswing in the leverage cycle. Figure 3(a) shows the proportion of all new purchase mortgage balances originated by borrowers with 2, 3 and 4 or more first-lien mortgages on their credit reports in each quarter between 1999Q1 and 2010Q4. As can be seen in the figure, this investor proportion increased from around 20 percent in 2000 to a peak of nearly 35 percent in 2006. The purchase share for borrowers with 4+ first-lien mortgages increased by more than 5 percentage points over this period.

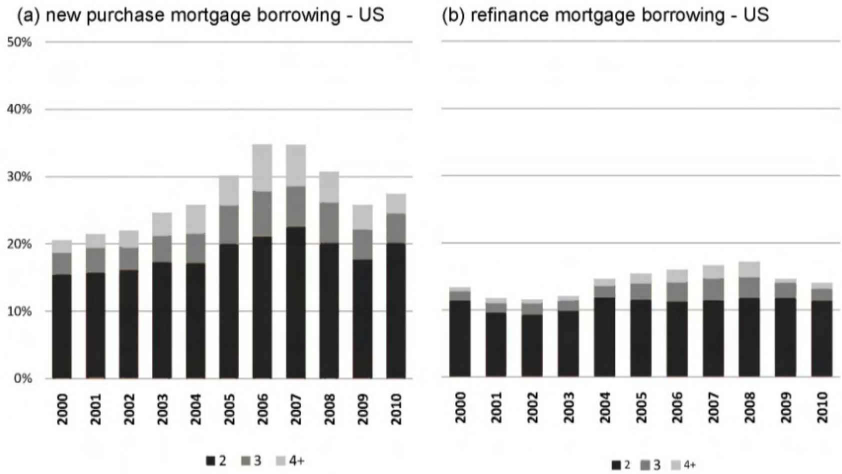
Meanwhile, investors make up a much smaller share of refinance originations (see Figure 3(b)), a result consistent with the view that investors hold properties for shorter periods.[28] For borrowers with short

---

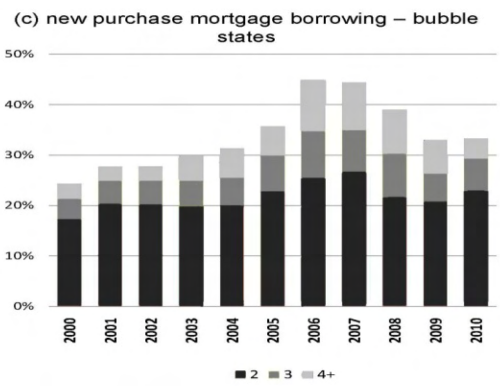
27 We observe borrowers' credit reports on the final day of each quarter. Because there can be delays in credit reporting such that a mortgage that has been paid off may stay on the credit report for a period of time, we use the data's panel structure to correct for these delays. Throughout this section of our analysis investor status is determined based on the maximum number of first mortgages that appear in both of the two most recent quarters. Thus we can be more confident that each first-lien we consider is in fact associated with a unique property.

28 Unlike loan-level data, borrower-level credit report data indicates the closing and

| Figure 3 | Investor shares in new mortgage borrowing(cont) (by number of first mortgages)



Source: FRBNY Consumer Credit Panel, 1% of population.



Source: FRBNY Consumer Credit Panel, 1% of population

---

opening of mortgage loans over time but do not include an indicator for whether a new mortgage loan represents a new purchase origination or a refinance. We identify refinances as a closing and opening of a new mortgage loan within a 6 month period during which the loan holder did not change address. Our refinance measure therefore may include some purchase loans associated with cases where an investor sold and bought a new property within a relatively short period of time. Nonetheless as shown later the patterns displayed in Figure 3 are mirrored in our matched sample, where loan purpose (purchase vs refinance) is explicitly measured. See Figure 7.

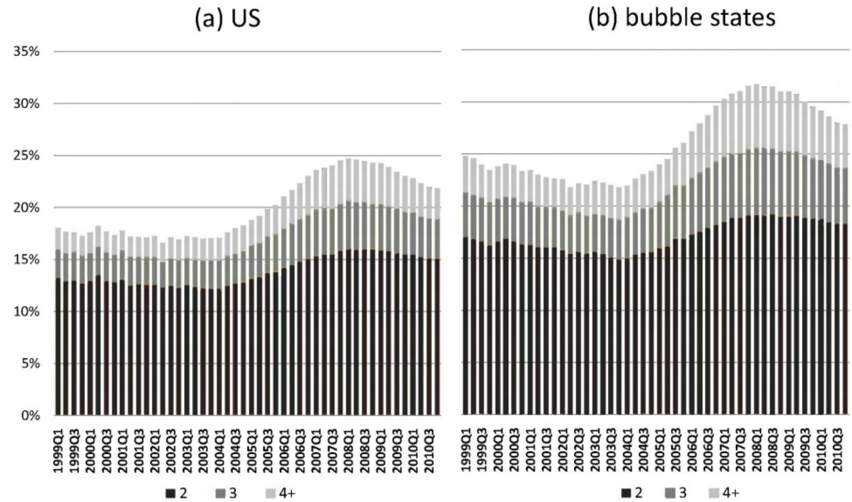
time horizons, the fixed costs of refinancing can make the option to refinance uneconomic.

Previous research has indicated that there was significant variation in the timing and, most importantly, the amplitude of the housing cycle over space (see, for example, Himmelberg *et al* 2005). If investors were playing an important role in fueling the growth of house prices in those states which experienced the greatest increases, we would expect to see differences over space in investors' share of the mortgage market. Our data confirm this conjecture. Figure 3(c) displays the same information focusing on the four states that experienced an especially pronounced housing cycle: Arizona, California, Florida and Nevada. Multiple lien holders of all types (2, 3, and 4+) were more prevalent in these "bubble" states than they are for the nation as a whole. The investor share of purchase mortgages also increased faster in these states as the housing boom peaked, rising from almost 25 percent in 2000 to 45 percent in 2006. The purchase share for borrowers with 4+ first-lien mortgages increased by more than 7 percentage points (or 350%) over this period.

Given this evolution of the flow of mortgage borrowing, it is unsurprising that we find investors increasing their share in the *stock* of mortgage debts. Figure 4(a), shows the share of outstanding mortgage *balances owed* by the number of first-liens reported on the borrowers' credit report.[29] Beginning in 2004, we see a pick-up in the share of all mortgage debt owed by borrowers with multiple first-liens, and this figure reached 24.7% by early 2008. Figure 4(b) displays the same information for the four "bubble" states. At the peak nearly one-third of all first-lien balances in these four states were owed by borrowers with at least two first-liens. By the peak in early 2008, first-lien mortgage debts owed by bubble state borrowers with four or more first-liens had risen to nearly $170 billion, over three and a half times their levels of

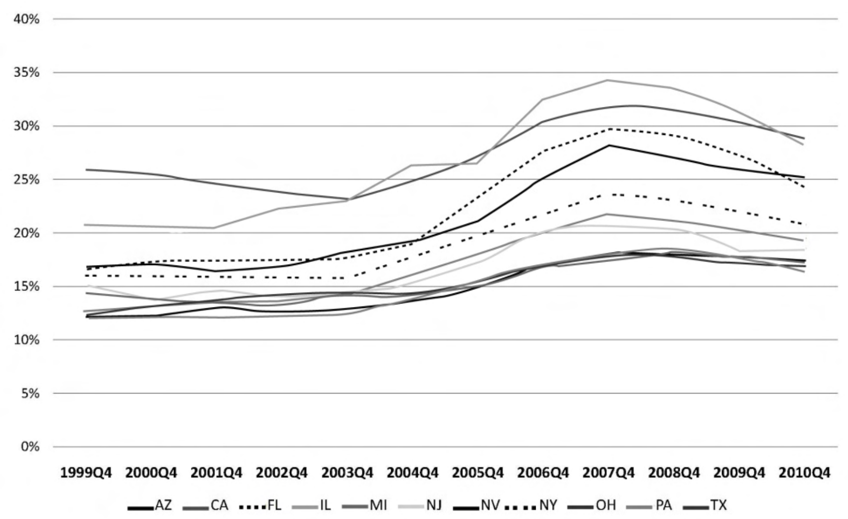
29 Between 2004Q1 and 2008Q1 the share of borrowers who had multiple first-liens increased from 7.3% to nearly 10%. During this four year period, the share of all mortgage borrowers with four or more first-lien mortgages on their credit reports increased by more than 50% (from 0.43% to 0.70%). In bubble states the investor share increased from 10% to 14% during this period, while the share of borrowers with four or more first-liens increased from 0.65% to 1.22%

| Figure 4 | Investor share in aggregate first mortgage balance (by number of first mortgages)



Source: FRBNY Consumer Credit Panel, 1% of population

| Figure 5 | Balance share of 2+ first mortgage holders by state



Source: FRBNY Consumer Credit Panel, 1% of population.

early 2004. Figure 5 shows the investor share over time by selected states. For all states listed, their share of mortgage balances increases over the boom but with less amplitude than in the bubble states.

Table 3 Average balance per mortgage, by investor status[1]

a. US

Year	Number of First Mortgages Held			
	1	2	3	4+
1999	95,160	102,636	105,917	100,489
2000	98,317	105,618	109,636	102,624
2001	102,664	113,750	120,447	110,290
2002	110,227	122,222	130,521	117,925
2003	123,335	139,651	146,173	128,873
2004	131,992	153,712	164,099	142,328
2005	141,977	171,042	187,087	158,730
2006	150,840	186,609	207,801	185,337
2007	157,965	197,464	218,394	201,147
2008	160,818	198,283	217,812	197,969
2009	161,194	197,127	212,477	193,175
2010	160,331	193,429	203,860	183,640

b. Bubble States

Year	Number of First Mortgages Held			
	1	2	3	4+
1999	116,037	120,670	125,548	120,033
2000	120,534	124,715	132,297	122,919
2001	125,809	133,262	141,287	127,844
2002	136,490	145,598	152,644	139,575
2003	155,907	167,779	171,582	162,933
2004	172,227	190,560	201,527	185,925
2005	193,432	220,128	232,655	211,273
2006	212,578	242,978	262,601	243,702
2007	225,048	258,618	278,413	262,718
2008	226,177	256,374	272,984	258,179
2009	223,594	252,630	266,031	254,578
2010	219,407	245,778	255,779	238,709

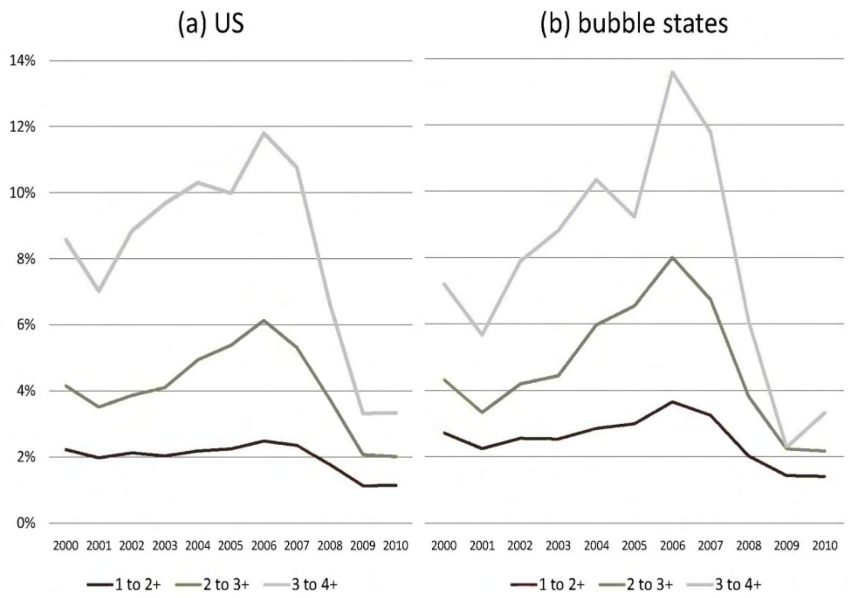
[1] *Notes*: Q4 values of average mortgage balance.
Source: FRBNY Consumer Credit Panel.

In the upswing phase of a leverage cycle that is unfolding as described by Geanakoplos, we would expect to see increases in both the extensive margin (reflected here as increases in the share of buyers who have multiple first-liens) and in the intensive margin (increases in the number of first-liens held, conditional on having more than one and increases in the average balances for each mortgage). The data demonstrate both increasing prevalence of investors in the housing marketplace and an increase in their share of outstanding and newly originated debt. An increased share of borrowers with multiple first-liens is evidence of an increase in the *extensive* margin: investor status became more widespread during the boom, especially in those markets where prices rose the most sharply. The fact that the share of new and existing purchase mortgages were owed by investors reflects this fact, but also may reflect the *intensive* margin, as investor-types increased their exposure to the housing market by borrowing more against residential property.

In order to discriminate between the two, we examine the intensive margin more carefully. In Table 3, we show the average balances on mortgages owed by the number of first-liens reported on the borrower's credit report for the US and the bubble states. We observe a change in the relative size of first-liens owed by highly leveraged borrowers. Because property values were rising sharply between 2000 and 2006, it is no surprise that the average balance on outstanding first-lien mortgages rose as well. But balances owed by investors rose even more sharply than those owed by owner-occupants. In 1999 for the US as a whole (panel (a) of Table 3), the average balance on first-liens owed by borrowers with debt secured by more than three properties was 13% higher than that of owner-occupants. [30] By 2006, the average investor balance (in the 4+ line) was nearly 50% larger than the corresponding owner-occupant figure. Interestingly, a more muted version of the same pattern obtains in the bubble states, shown in panel (b).

30 A similar relative increase is observed for average origination balances.

| Figure 6 | Average quarterly transition rates(t-1to t) in # of first mortgages, by year



Source: FRBNY Consumer Credit Panel, 1% of population.

Further evidence of this increase in the intensive margin is found in Figure 6. Here we track the transitions into and out of the various first-lien mortgage count categories. In both panels (a) and (b) we find what we will refer to as "up-leveraging" transitions: the proportion of investors in year t-1 who have additional mortgages in *t*.[31] For example, Figure 6(a) shows that approximately 6% of all investors with two first-lien mortgages in 2005 had added a third by 2006. As can easily be seen in the figure, the proportion of all mortgage borrowers who added additional properties to their portfolios grew between 2000 and 2006, with the sharpest increases found among those who already had the highest residential real estate holdings. Around 12% of US borrowers with exactly three first-liens in 2005 added additional properties to their portfolios during 2006; in the bubble states this figure exceeded 16%.

31 The transition rates are based on the maximum number of first mortgages held during the two most recent quarters at time t, and the maximum number of first mortgages held during the two most recent quarters at time *t*-1.

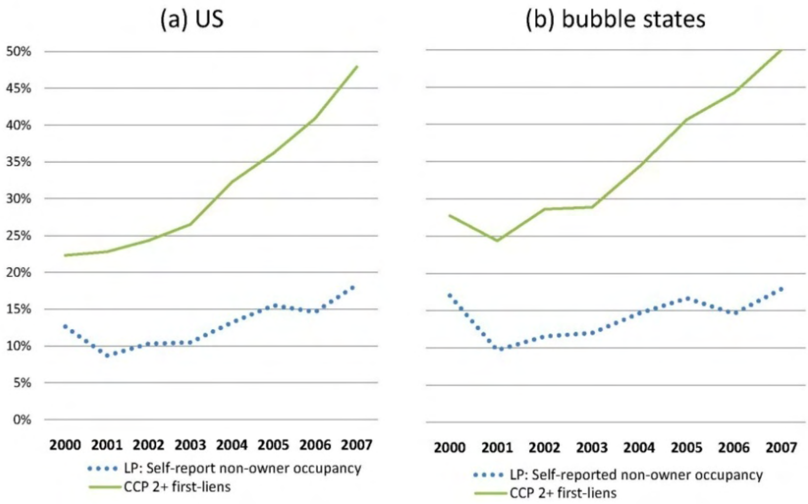
We conclude from our analysis that mortgage borrowing by investors - defined as those with multiple properties in their portfolios - increased substantially during the boom, especially in those markets where house price increases were particularly pronounced. We find evidence of increases in both the extensive margin - new investors entering the marketplace - and the intensive margin. increased exposure to residential real estate among previous investors.

These results contrast with previous discussions of the role of investors in the mortgage marketplace, and underscore the benefits of the FRBNY CCP for analyzing these questions. Mayer, Pence and Sherlund (2008), for example, conclude "because our data show that [self-identified] investors were a small or declining share of overall originations [of non-prime mortgages], it seems unlikely that they accounted for much of the rise in the overall delinquency rate *unless they increasingly misrepresented themselves as owner-occupiers* or their unobserved characteristics deteriorated over time." (2009, pg. 44, emphasis added). Our data allow us to "see through" the self-reported information captured on the mortgage application, and show precisely this - an increasingly large discrepancy between mortgage application occupancy self-reports and the number of first-liens on the credit report during the crucial 2004-2006 period. These results thus leave open the question of the role of these investors in the subsequent increase in defaults and delinquencies.

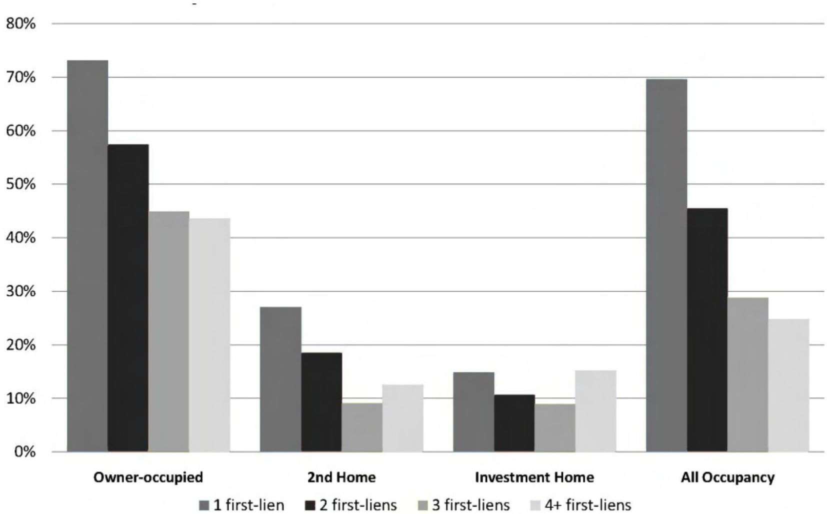
Figure 7 contains three panels which explore the relationship between occupancy self-reporting on mortgage applications and borrowers' first-lien counts for our matched CCP-LP sample. In panel (a), we plot the proportion of new nonprime purchase originations by self-reported occupancy status (from LP) and number of first-liens (from CCP). The dashed line plots the proportion of balances taken out by borrowers who checked either "2nd home" or "investor property" on the mortgage application, while the solid line shows the proportion of balances originated by these same borrowers who, after closing this mortgage, simultaneously have two or more first-liens.[32]

32 Recall we count only those first-liens that remain on the credit report for at least two calendar quarters.

| Figure 7 | Investor share of purchase origination balances, LP and CCP definitions



(c). Share of new purchasers moving to property Zip code, LP and CCP investor definitions



Source: FRBNY Consumer Credit Panel, 1% of population.

Comparing the two series provides some insight into the value of self-reported occupancy status. First, at the beginning of the period the two series are reasonably close together, but even in 2000 there is a significant discrepancy between what borrowers report on the mortgage application and the number of properties they own. Over time, the proportion of new originations by borrowers who acknowledge that they will not use the home as their primary residence (the dashed line) increases slowly, and is fairly flat at 13-15% for the crucial 2004-2006 vintages. While we are including second homes, balance weighting and using only purchase mortgages, this pattern is similar to the results found by Mayer, Pence and Sherlund (2008). Meanwhile, however, the proportion of borrowers who have 2 or more first-lien mortgages rises much more quickly, and approaches 41% by 2006. The bubble states, shown in Figure 7(b), exhibit the same pattern, although in somewhat more extreme, where in 2006 the gap between self-reported occupancy status and the number of first-liens reached 30 percentage points. In other words, many of the borrowers who claimed on the mortgage application that they planned to live in the property they were purchasing had multiple first-lien mortgages when the transaction was complete. Mayer, Pence and Sherlund (2008) accurately report that borrowers' self-reported occupancy status was not changing dramatically during this period, but they are unable to observe the change in the characteristics of borrowers who report themselves as owner-occupants. In fact, the importance of investors as defined in the CCP - borrowers who have 2 or more first-liens on their credit reports - expanded sharply during this period, especially in the bubble states. Also note that this increase in the share of investors in non-prime purchase originations is very similar to that shown earlier for all purchase originations in Figure 5 (based on the entire CCP), which is reassuring.

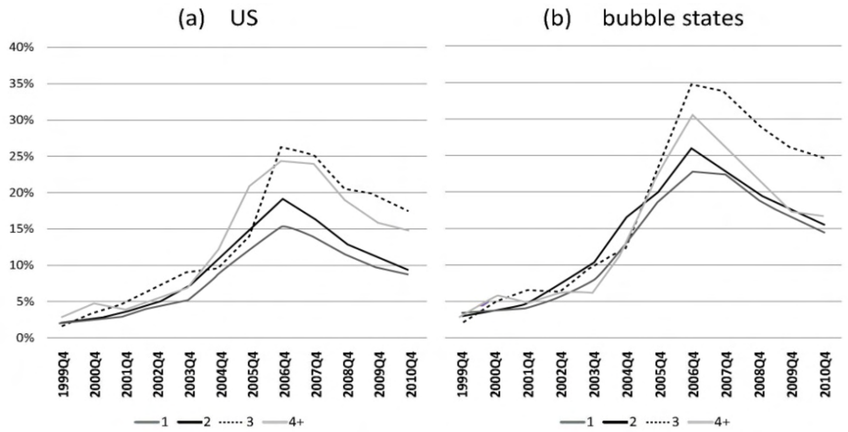
While it is possible that all of these borrowers intended to live in the purchased property, it seems unlikely. In addition, the matched data allow us to track whether the individual changed addresses after closing the mortgage, and whether they moved to the same zip code recorded for the property. Figure 7(c) shows, by borrowers' self-reported occupancy status and the number of first-liens on their credit reports, the percentage who changed addresses to the zip code containing the property within two years of originating a nonprime purchase mortgage. Unconditional on self-reported occupancy status we find respectively 70% and 25% of single and four first-lien holders to have moved to the property zip code within two years of the new purchase. The data indicate further that 73% of those who claim owner occupancy while holding a single mortgage changed addresses and their new zip code matches that of the property. By contrast, only 43% of those who claimed owner occupancy on the mortgage application while carrying four or more first-liens prior to closing moved to the property zip code within two years. Unsurprisingly, relatively low shares (under 30%) of those who reported the property as a second home or investor property moved to the property zip code within two years. While the evidence cannot be definitive, we take this as suggestive of significant occupancy misrepresentation in nonprime mortgages during the boom.

4. Mortgage Products and Leverage

Since the Geanakoplos theory focuses on highly-leveraged positions taken by optimistic buyers here identified as investors, a natural next step in our discussion is to explore the leverage obtained by investors relative to owner-occupants. Our discussion proceeds on two fronts. We are able to provide some insight into this issue by using the CCP-LP matched sample to examine the mortgage products used by investors. Leverage theory suggests that "buy and flip" investors will want to use as much leverage as lenders will allow. We would expect to observe investors using non-prime mortgages - which allow for higher leverage than conforming mortgages - relatively more than non-investors as the boom progressed. As noted above, investors are less sensitive to the higher interest rates charged on non-prime loans than owner-occupants due to the shorter expected holding period. Buy and flip investors are more willing to pay higher rates in order to increase leverage.

Figure 8 provides some insight into the mortgage products chosen by investors and non-investors. In panel (a) we plot the national proportion of first mortgage balances that were securitized by private ABS issuers,

| Figure 8 | Share securitized non-prime in aggregate mortgage balance, by investor type



Source: FRBNY Consumer Credit Panel, 1% of population.

by the number of first-liens on the borrower's credit report. [33] These are essentially market shares for the nonprime mortgage lenders for each group of first mortgage holders, respectively. The temporal patterns are interesting: while the total nonprime share rose sharply in 2004 and 2005, borrowers with multiple first-liens were even more likely than other borrowers to obtain credit from the nonprime part of the market. By 2006, 26.0% and 24.4% of first mortgage balances associated with borrowers with three first-liens and four first-liens on their credit reports respectively, was nonprime, compared with 15% for those with a single first-lien. Panel (b) reports the same information for the bubble states, and demonstrates a more significant increase for all borrowers, as well as providing a similar picture of investors' preferences for nonprime credit.

The second piece of evidence we can bring to bear on the leverage issue is also from the matched sample. For securitized subprime and alt-a mortgages we observe lender-reported combined LTVs at the origination of each first-lien. Table 4 extends Table 2 by reporting the

33 We calculated these figures by comparing the total originations in our matched sample with all purchase originations in our CCP data.

median combined LTV and the share exceeding 90 for our matched sample, focusing on 2002, 2004 and 2006 purchase originations only, drawing on the loan-level LP data. The table shows some striking features of the data. First, note that reporting an intention to live in the purchased property is consistent with higher leverage: in all years, self-declared owner occupants have higher median LTVs, and are much more likely to have LTVs above 90. Second, conditional on their self-reported status, borrowers' property ownership, reflected by the number of first-liens on the credit report, does not have a consistent relationship with LTV.

Table 4 Combined Loan to Value ratios at Origination, Securitized Non-prime Mortgages

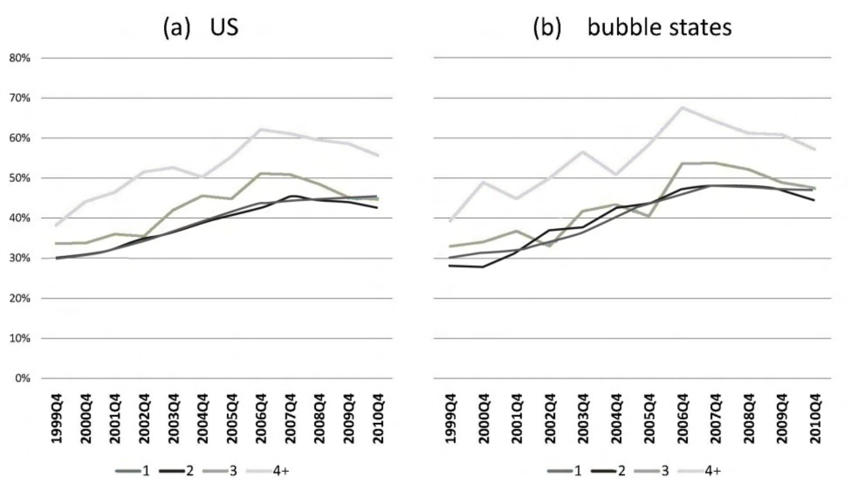
	2002		2004		2006	
	Median	% > 90	Median	% > 90	Median	% > 90
Self-report owner-occupied . . .						
. . .with 1 first-lien	85	31.6%	95	58.8%	100	70.3%
. . .with 2 first-liens	90	31.6%	95	57.4%	100	67.6%
. . .with 3 first-liens	80	25.0%	95	51.1%	100	65.1%
. . .with 4+ first-liens	85	15.8%	90	44.5%	100	70.5%
Self-report second home . . .						
. . .with 1 first-lien	80	11.1%	80	10.0%	90	37.8%
. . .with 2 first-liens	83	0.0%	90	19.0%	90	40.9%
. . .with 3 first-liens	80	0.0%	90	13.3%	90	29.8%
. . .with 4+ first-liens	73	0.0%	80	6.3%	90	35.4%
Self-report investment home . . .						
. . .with 1 first-lien	85	12.8%	90	27.4%	90	38.3%
. . .with 2 first-liens	85	5.8%	90	18.8%	90	31.2%
. . .with 3 first-liens	85	5.0%	90	16.2%	90	36.3%
. . .with 4+ first-liens	80	0.0%	90	16.7%	90	31.1%

Source: FRBNY Consumer Credit Panel.

We conclude that, given down payment requirements in the prime market, investors were able to increase their leverage by disproportionately using nonprime securitized mortgages, and were a major driver of growth in that important market segment. By declaring an intention to live in the properties collateralizing these loans, investors were able to reduce both the interest rates and the minimum downpayments, with the latter being the most valuable for our buy and flip investors.

An additional development of interest in the type of mortgage loans chosen by investors and non-investors is shown in Figure 9(a) for the nation and 9(b) for the bubble states. Here we consider whether the account was an individual or joint account. This distinction is interesting for two reasons. First, we do not observe debt-to-income ratios in our data, but it is a reasonable presumption that individual accounts carry higher ratios since they depend on the income of a single borrower; of course the narrower support makes these mortgages riskier as well.[34] Second, if the borrower is making a speculative leveraged investment, it is presumably a dominant strategy to expose only one credit account to the risk of a foreclosure. The figure documents a general increase in the use of individual as opposed to joint mortgage accounts that began after 2000 and finally began to taper off at the end of 2007. However, the shift from joint to individual mortgage accounts since 2001 was much greater for investors, especially for investors with 4 or more first

Figure 9 Share of aggregate mortg age balance in individual accounts, By investor type



Source: FRBNY Consumer Credit Panel, 1% of population

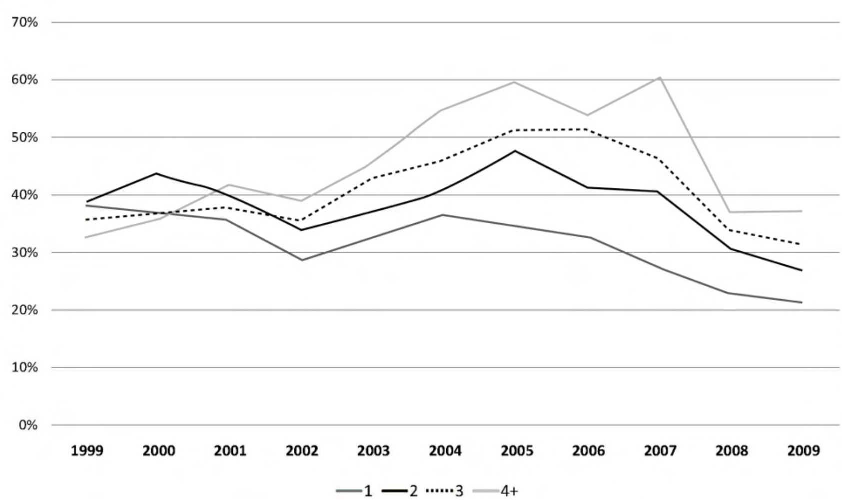
---

34 Indeed, in the hazard analysis described below, we find that individual accounts are more likely to transition into 90+ day delinquency, cet. par.

mortgage accounts. By 2007 over 60% of the total outstanding first mortgage balance in the US was associated with individual accounts. Interestingly, unlike for single mortgage-holders, for investors the balance share associated with individual accounts began to drop after 2007. As shown in Panel 9(b), the same trends apply to the bubble states, except that the increase in the balance-weighted share of debt in individual accounts was more pronounced.

The Geanakoplos leverage cycle theory predicts that as an asset price boom unfolds, buy and flip investors will become a more important share of the investment property marketplace. As hypothesized earlier, as the housing boom intensifies investors are likely to reduce their expected and actual holding periods. We can provide insight into whether this hypothesis is consistent with housing market developments during the 2000s boom by examining the holding periods for mortgages originated during the boom. Figure 10 provides some of this evidence. Here we plot the share of all purchase mortgages securing property sold in year t that had been held for less than three years. As an example,

Figure 10 Proportion of mortgage pay offs that were originated within The past three years, by investor type



Source: FRBNY Consumer Credit Panel, 1% of population during 2006 60% of mortgages paid off (excluding re-finances) by borrowers with at least four first-liens had been originated less than 36 months earlier.

In the figure, we see evidence of several interesting phenomena.[35] First, fairly large shares of first-liens are held for a short period of time. Even in the early part of our sample, between 30 and 40 percent of pay-offs are for mortgages on properties held for less than three years, regardless of the investor status of the seller. Second, as a group, borrowers with multiple first-liens initially look quite similar to owner-occupants in their mortgage durations. As the boom unfolds, however, we see increasing shares of properties held by investors (see especially the 3 and 4+ lines) being sold quickly. By the peak of the market, a large share of sales by investors complete relatively short holding periods. Our interpretation of this phenomenon is that the composition of those in the multiple first-lien categories is shifting from “holders” to “flippers”.

5. Delinquencies and de-leveraging during the bust

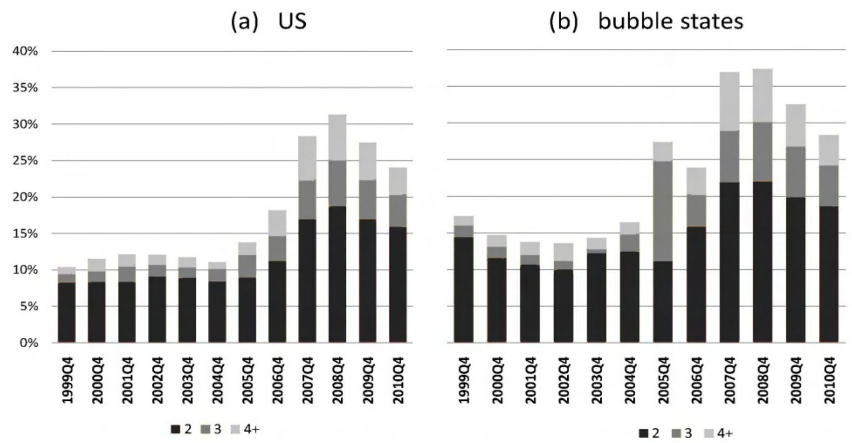
We have shown that short time-horizon multiple first-lien holders became an increasingly important part of the mortgage marketplace during the boom between 1999 and 2006, thus confirming that several elements of the Geanakoplos leverage cycle model are applicable to the US housing market. The second stage of the cycle is the bursting of the bubble, reflected in this case by the collapse in housing prices and sharp increases in delinquencies and defaults after 2006. Here again the model contains several implicit predictions that we can examine with our data.

One such prediction is that investors will stop increasing their exposure to real estate and will rapidly begin to divest themselves of their positions. We saw clear evidence of this in Figure 6, particularly in

35 A chart showing the share of mortgage closings that were originated within the past two years showed very similar trends, but with levels peaking during the 2004-2007 period at around 35 percent for 4+ investors while the rate during the period fell from about 20 to 15 percent for single homeowners during that period.

the bubble states (panel (b)), where 17% of three property owners had increased their exposure to housing during 2006; by 2009 that figure was just over 1%. This sharp reduction in additions to the intensive margin is consistent with a rapid retrenchment among investors. Figures 3 and 4 and Table 3 all contain evidence consistent with the conclusion that investors reduced their role in the market after prices peaked in 2006-2007, including reductions in both the extent and intensity of investor activity.

Figure 11 Investor share of 90+ DPD first-mortgage balances, by investor type



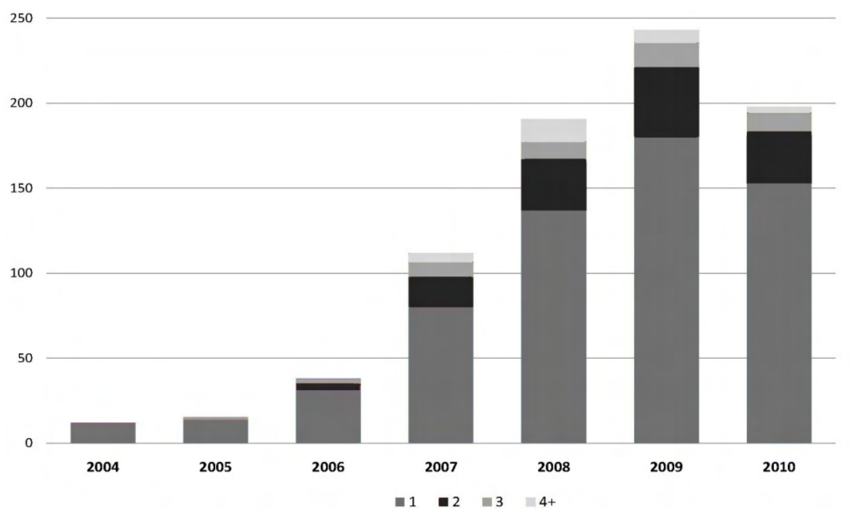
Source: FRBNY Consumer Credit Panel, 1% of population

A second prediction, not only from Geanakoplos but also from the previous literature on mortgage defaults, is that investors will be quite influenced by house price changes in their repayment behavior.[36] Figure 11 provides some evidence for this hypothesis. In panel (a), which depicts the severe delinquencies contributed by multiple first-lien holders in the nation as a whole, we see an extraordinarily rapid increase in the investor share. Early in the period, as house prices were rising, severe delinquencies by investors, especially those with three or more

36 This is consistent with the findings in Mayer et al (2009) which found the decline in house prices to be a key factor in explaining the big increase in mortgage delinquencies.

first-liens, were quite rare, and considerably below their proportionate share of outstanding first-lien balances. Beginning in early 2006, however, as the housing market peaked, serious mortgage delinquencies by investors rose sharply, and by 2007 investors' delinquency share exceeded their share of outstanding mortgage debt. This period was marked by especially large, disproportionate delinquencies by borrowers with three or more first-liens. A similar, even more dramatic, version of this dynamic is present in the bubble states, depicted in Figure 11(b). Here, the data indicate a virtual explosion in delinquencies among multiple first-lien borrowers, especially those with more than two properties.

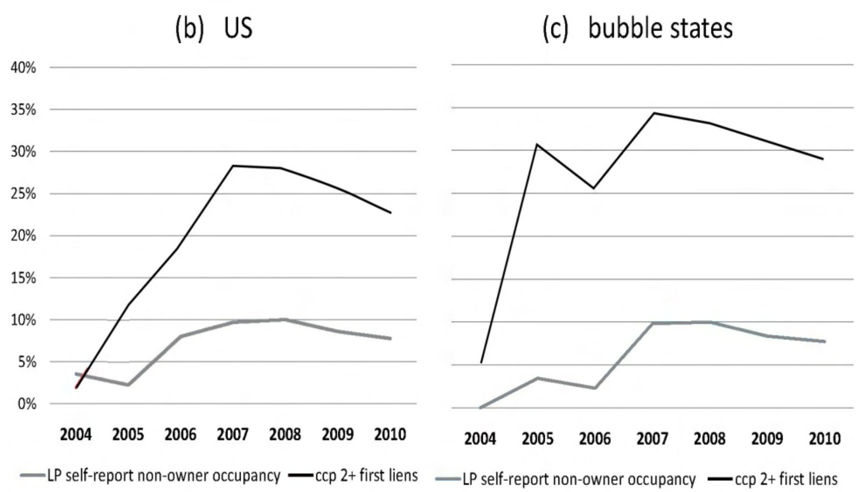
Figure 12-1 (a) Aggregate 90+DPD securitized non-prime mortgage balances ($B), by investor type



Source: FRBNY Consumer Credit Panel, 1% of population

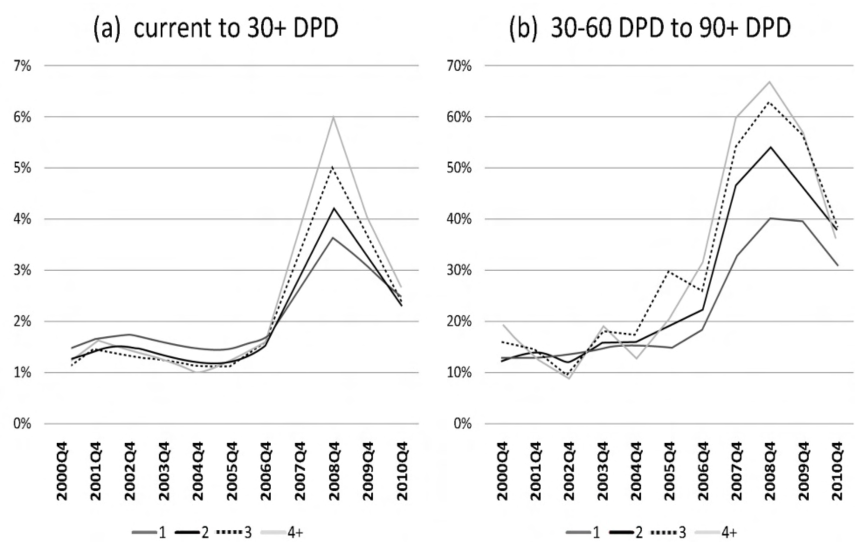
We can also investigate the relationship between investor status and delinquencies in the securitized non-prime sector using our matched sample. Moreover, we are able to do so for both the multiple first-lien and declared owner-occupancy measures of investor status. In Figure 12, panel (a) shows the contributions, in billions of dollars, to serious delinquencies for the nation for borrowers with single and multiple first

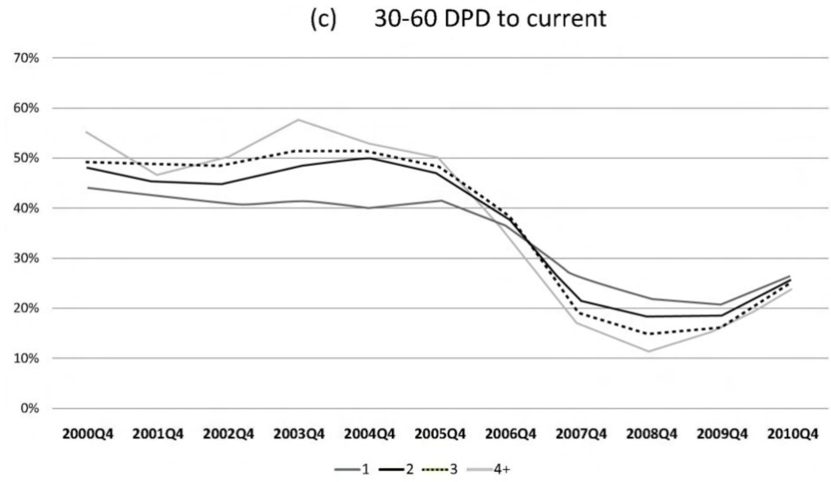
| Figure 12-2 | Investor share 90+DPD securitized non-prime balances, LP and CCPdefinitions



Source: FRBNY Consumer Credit Panel, 1% of population

| Figure 13 | Average quarterly transition rates by investor type





Source: FRBNY Consumer Credit Panel, 1% of population

liens. Panels (b) and (c) instead show the balance-weighted shares in serious delinquent debt for the nation and the bubble states for borrowers who had multiple first-liens and those who reported that they would not be using the home as their primary residence. What is evident from these figures is, as noted previously, a huge increase in serious would not be using the home as their primary residence. What is evident from these figures is, as noted previously, a huge increase in serious delinquencies in the non-prime sector after the house price peak in 2006, with a large share coming from the bubble states. Perhaps more important for our purposes is that reliance on the self-reported occupancy status to understand the increase would lead researchers to conclude that investors had relatively little to do with the rise in delinquencies, whereas in fact the contribution from borrowers with multiple first-liens (the CCP measure) is very large, reaching almost $250 billion by 2009 in the securitized non-prime sector alone.

Among the underlying forces behind the increase in delinquencies among investors are (a) a sharp increase in the rate of initial delinquency among investors, (b) a large increase in the rate at which initial delinquencies transition into a severe delinquency and (c) a large decrease in the rate at which initial delinquencies cure. Investors not living in houses they own will make their default decisions purely based on investment motives, as opposed to consumption motives. This suggests more ruthless or strategic behavior on the part of investors where conditional on an initial delinquency, loans would transition more quickly into defaults. As shown in Figure 13(a), while transition rates into early delinquency were lower among investors before 2007, they were much higher in the subsequent period, especially for those with 4 or more first mortgages. Figure 13(b) shows that such early delinquencies after 2006 also transitioned into defaults at a much higher rate for investors, with fewer early delinquent loans curing as seen in Figure 13(c).[37]

To obtain some further insight into the sharp increase in delinquencies among investors, we next investigate the role of various investor characteristics. First, as documented earlier, the investor share of mortgage holders was much greater in the bubble states, states which subsequently experienced the sharpest house price declines. Second, reflecting their growing share in real estate transactions, mortgages held by investors were more likely to have been originated in more recent years. Unlike homes purchased in earlier years, homes bought after 2005 experienced little or no price appreciation and their buyers therefore saw no gains in home equity. The subsequent drop in house prices was therefore more likely to cause these mortgages to go underwater, a necessary condition for default. Third, as shown earlier, investors are more likely to use non-conforming loans, which generally carry higher interest rates, and to use individual rather than joint mortgage loans. Moreover, average origination balances generally were higher among investors. All these factors could put mortgages held by investors at greater risk of default.

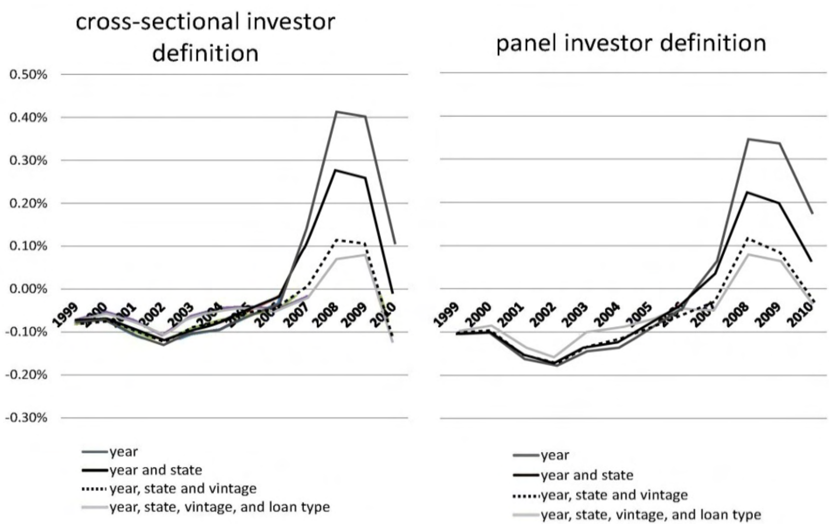
To analyze the respective importance of these factors, we estimated a set of loan-level delinquency hazard models, relating the quarterly rate of entry into 90+ day delinquency to loan and borrower characteristics. Linear probability model estimates of the year-specific impacts of investor status on the delinquency rate are presented in Table 5. The models underlying the estimates in the first panel of the table impose a

37 In additional analyses, not reported here, we found very similar trends in transition rates for the subset of conforming loans.

linear effect of number of first-mortgages held, while the second panel estimates separate effects for investors holding 2 and 3+ first mortgages. For each, we estimated four different models. The first includes only includes year fixed effects as controls. The second specification adds state fixed effects, while the third specification in addition includes loan vintage-year dummies. Finally, the fourth also includes controls for loan characteristics including loan origination amount, loan type (whether guaranteed by Government Sponsored Enterprises Fannie Mae and Freddie Mac, FHA/VA, other) and whether the mortgage account was individual or joint.[38]

The estimates for specification (1) mirror those in Figure 7, showing lower average delinquency rates for investors up to 2006, and higher rates since then, especially among those with 4 or more first mortgages. Adding controls for state fixed effects in specification (2), vintage effects in specification (3), and loan characteristics, in specification (4), leads to subsequent declines in the estimated remaining investor effect,

| Figure 14 | Investor effect on quarterly flow into 90+ delinquency



Source: FRBNY Consumer Credit Panel, 1% of population

38 For a subset of GSE mortgage loans in our database, the GSE identifier was missing. Therefore the included measure is only a rough proxy of true loan type.

▌Table 5 ▌ Annual Investor Effects on 90DPD Delinquency Rate[1]

Linear Investor Effect

Year	Specification (1)	(2)	(3)	(4)
1999	-0.08%	-0.08%	-0.08%	-0.07%
2000	-0.07%	-0.07%	-0.07%	-0.05%
2001	-0.11%	-0.09%	-0.10%	-0.07%
2002	-0.13%	-0.12%	-0.12%	-0.11%
2003	-0.11%	-0.09%	-0.09%	-0.06%
2004	-0.09%	-0.08%	-0.07%	-0.04%
2005	-0.06%	-0.05%	-0.06%	-0.04%
2006	-0.03%	-0.02%	-0.05%	-0.05%
2007	0.14%	0.11%	0.01%	-0.02%
2008	0.41%	0.28%	0.12%	0.07%
2009	0.40%	0.26%	0.11%	0.08%
2010	0.10%	-0.01%	-0.12%	-0.12%

Separate Investor Effects

	Specification (1)		(2)		(3)		(4)	
	Investor Type		Investor Type		Investor Type		Investor Type	
Year	2	3+	2	3+	2	3+	2	3+
1999	-0.06%	-0.20%	-0.06%	-0.21%	-0.06%	-0.21%	-0.04%	-0.20%
2000	-0.15%	-0.13%	-0.14%	-0.13%	-0.14%	-0.13%	-0.10%	-0.11%
2001	-0.08%	-0.30%	-0.06%	-0.27%	-0.07%	-0.27%	-0.02%	-0.23%
2002	-0.15%	-0.30%	-0.14%	-0.28%	-0.14%	-0.29%	-0.11%	-0.25%
2003	-0.14%	-0.23%	-0.12%	-0.20%	-0.12%	-0.19%	-0.08%	-0.13%
2004	-0.15%	-0.15%	-0.12%	-0.12%	-0.12%	-0.11%	-0.07%	-0.05%
2005	-0.18%	-0.06%	-0.16%	-0.04%	-0.17%	-0.06%	-0.13%	-0.03%
2006	-0.04%	-0.06%	-0.02%	-0.03%	-0.06%	-0.10%	-0.03%	-0.10%
2007	0.28%	0.33%	0.24%	0.26%	0.09%	0.02%	0.11%	-0.05%
2008	0.41%	1.03%	0.27%	0.71%	0.03%	0.34%	0.04%	0.20%
2009	0.34%	1.15%	0.18%	0.81%	-0.02%	0.45%	0.02%	0.35%
2010	0.19%	0.23%	0.08%	-0.04%	-0.03%	-0.30%	0.03%	-0.34%

[1]*Source*: FRBNY Consumer Credit Panel. Linear probability model estimates of year-specific impacts of the current number of first mortgages held on the quarterly entry rate into 90+ delinquency. The first panel imposes a linear effect, while the second panel estimates separate effects for investors holding 2 and 3+ first mortgages. Specification (1) controls for year fixed effects. Specification (2) adds state fixed effects, and specification (3) in addition includes loan vintage-year dummies. Specification (4) adds controls for loan characteristics such as the loan origination amount, whether loan was guaranteed by a GSE, and whether mortgage account was individual or joint.

indicating that each set of controls can explain a piece of the higher overall delinquency rates of investors. A graphical depiction of the year-specific investor effects are shown in Figure 14. The estimates imply that slightly more than half of the change in the relatively delinquency rates of investors versus non-investors can be accounted for by differences in the timing and location of home purchases and differences in the types of mortgages used to finance these purchases. However, substantial investor effects remain, suggesting that there were additional unmeasured differences between investors and non-investors that put mortgage loans of the former at higher risk of default.

The second panel in Table 5 repeats the same analysis but using a specification that allows for year-specific effects of investors with 2 or 3+ first mortgages. The estimates indicate that the difference between delinquency rates for investors with 3+ mortgages and single home-owners was much larger than for investors with 2 mortgages - they were much safer before 2006 and much riskier after 2006, when prices had begun to decline.

Finally, we repeated the loan-level delinquency hazard models using a different definition of investor. Instead of a cross-sectional definition, where investor status can change over the life of a loan as loans are added or closed, we adopt a panel definition, where investors are defined by the maximum number of first mortgage loans held during the lifetime of the loan. Such a definition allows us to identify loans as associated with individuals who previously were investors but closed some of their other mortgages. This may occur, for example, where other properties in an investor portfolio are sold or foreclosed on. As shown in Table 6, investor effect estimates both before and after 2006 are generally somewhat larger in absolute magnitude. The biggest difference in estimates when compared to Table 5 are for 2010 representing the extent of deleveraging by investors. Figure 14(a) and 14(b) summarize the investor effects for each analysis.

❙ Table 6 ❙ Annual Investor Effects on 90DPD Delinquency Rate . Panel Definition of Investor[1]

Linear Investor Effect

	Specification			
Year	(1)	(2)	(3)	(4)
1999	-0.10%	-0.10%	-0.10%	-0.09%
2000	-0.10%	-0.10%	-0.10%	-0.08%
2001	-0.16%	-0.15%	-0.15%	-0.13%
2002	-0.18%	-0.17%	-0.17%	-0.16%
2003	-0.14%	-0.13%	-0.13%	-0.10%
2004	-0.13%	-0.12%	-0.12%	-0.09%
2005	-0.09%	-0.08%	-0.09%	-0.06%
2006	-0.05%	-0.03%	-0.05%	-0.04%
2007	0.06%	0.03%	-0.03%	-0.04%
2008	0.35%	0.23%	0.12%	0.08%
2009	0.34%	0.20%	0.09%	0.07%
2010	0.17%	0.07%	-0.03%	-0.03%

Separate Investor Effects

	Specification							
	(1)		(2)		(3)		(4)	
	Investor Type		Investor Type		Investor Type		Investor Type	
Year	2	3+	2	3+	2	3+	2	3+
1999	-0.13%	-0.22%	-0.13%	-0.23%	-0.13%	-0.22%	-0.12%	-0.21%
2000	-0.17%	-0.20%	-0.17%	-0.20%	-0.17%	-0.20%	-0.14%	-0.17%
2001	-0.20%	-0.36%	-0.19%	-0.34%	-0.20%	-0.35%	-0.16%	-0.29%
2002	-0.27%	-0.37%	-0.27%	-0.36%	-0.27%	-0.37%	-0.24%	-0.33%
2003	-0.20%	-0.30%	-0.19%	-0.28%	-0.19%	-0.28%	-0.14%	-0.21%
2004	-0.22%	-0.26%	-0.20%	-0.23%	-0.20%	-0.22%	-0.14%	-0.15%
2005	-0.21%	-0.17%	-0.20%	-0.14%	-0.20%	-0.16%	-0.16%	-0.11%
2006	-0.19%	-0.06%	-0.17%	-0.02%	-0.18%	-0.07%	-0.15%	-0.06%
2007	0.01%	0.18%	-0.03%	0.11%	-0.08%	-0.05%	-0.05%	-0.10%
2008	0.14%	1.01%	0.00%	0.71%	-0.09%	0.43%	-0.07%	0.33%
2009	0.10%	0.97%	-0.07%	0.65%	-0.16%	0.36%	-0.10%	0.29%
2010	0.13%	0.51%	0.01%	0.26%	-0.07%	0.03%	-0.01%	0.01%

[1] *Source*: FRBNY Consumer Credit Panel. Linear probability model estimates of year-specific impacts of the maximum number of first mortgages held during the lifetime of the loan on the quarterly entry rate into 90+ delinquency. The first panel imposes a linear effect, while the second panel estimates separate effects for investors holding 2 and 3+ first mortgages. Specification (1) controls for year fixed effects. Specification (2) adds state fixed effects, and specification (3) in addition includes loan vintage-year dummies. Specification (4) adds controls for loan characteristics such as the loan origination amount, whether loan was guaranteed by a GSE, and whether mortgage account was individual or joint.

6. Conclusion

The effects of boom-bust cycles in asset prices are nowhere more potentially dangerous than in housing, which makes up about 80% of the debts owed by households. While changes in underwriting standards have been the focus of many studies trying to understand housing cycles, less attention has been paid to how these standards interact with the distribution of borrowers in the marketplace. Our exploration of the 2000s housing cycle suggests that this interaction was an important, but poorly understood, dynamic. Our analysis reveals patterns consistent with Geanakopolos's theory of the leverage cycle. Possibly house price-driven relaxation of down payment and documentation standards induced or facilitated a change in the composition of mortgage borrowers toward more optimistic buyers, here identified as short time horizon investors. Giver their willingness to bid more aggressively, the large influx of investors is likely to have amplified the upward pressure on house prices during the boom. As they represented almost half of all buyers in the bubble states during the boom, we can expect an impact on the appraisals and purchase prices of homes bought by non-investors. Our analysis also indicates that these marginal borrowers appear to have contributed substantially to both the increasing amount of real estate-related debt during the boom, and to the rapid deleveraging and delinquency that accompanied the bust.

The findings in our paper so far have important implications for the design of future policies to reduce the likelihood and deleterious consequences of future house price bubbles. While investors in the role of 'middlemen' can provide important liquidity to the housing market (Bayer et al, 2011), investors as speculators can generate amplifications of house price movements. There is thus scope for policy instruments that target the activities of speculative investors. To dampen speculation and to cool down the nation's housing market, the Chinese government during the past few years has implemented a number of successive tightening measures that include higher down-payments and mortgage rates on second and additional investment homes.[39] Some cities in

39 Down payment requirements for the purchase of second homes and additional

China have also introduced a new real estate tax on such properties as well as limits and freezes on the purchasing of second and additional investment homes. Such explicit management of the use of leverage by optimistic buyers may serve to dampen upswings in asset markets, thereby ameliorating the effects of the decline if and when it occurs.

Our findings regarding the role of investors in the housing boom and bust and the high rate at which they defaulted after 2007 also has important implications for the design of effective, equitable and targeted assistance programs. While the majority of home-owner assistance programs developed over the past several years have been targeted to owner-occupants, many have experienced relatively low take-up rates. If, as indicated here, a large share of defaulters are not living in the collateral home, then programs such as HAMP may not be effective in stemming foreclosures. On the other hand, less sensitive policies, like blanket modifications offered regardless of occupancy status might be more efficient, but would provide assistance to a large class of multiple property owners. No one's first priority for receiving taxpayer dollars.

---

investment properties were increased to 30% of the property price in January 2010, to 50% in April 2010 and 60% in February 2011.

References

Bayer, Patrick, Christopher Geissler and James W. Roberts, "Speculators and Middlemen: The Role of Flippers in the Housing Market." Duke University, February 2011.

Campbell, Sean, Morris A. Davis, Joshua Gallin and Robert F. Martin. "A Trend and Variance Decomposition of the Rent-Price Ratio in Housing Markets." *Journal of Urban Economics* 66(2), 90-102. 2009.

Case, Karl E., and Robert J. Shiller. "The Behavior of Home Buyers in Boom and Post-Boom Markets." *New England Economic Review*, November/December 1988, 29-46.

Chakrabarti, Fajashri, Donghoon Lee, Wilbert van der Klaauw, and Basit Zafar. "Household Debt and Saving During the 2007 Recession." Federal Reserve Bank of New York Staff Report No. 482, January, 2011.

Fitch Ratings. *The Impact of Poor Underwriting Practices and Fraud in Subprime RMBS Performance*. U.S. Residential Mortgage Special Report, 2007.

Fostel, Ana, and John Geanakoplos. "Leverage Cycles and the Anxious Economy." *American Economic Review* 98 (September 2008): 1211-1244.

Gallin, Joshua. "The Long-Run Relationship between House Prices and Rents." *Real Estate Economics* 36: 635-658. 2008.

Geanakoplos, John. "The Leverage Cycle." Cowles Foundation Discussion Paper No. 1715, July, 2009.

_____. "Solving the Present Crisis and Managing the Leverage Cycle." *Economic Policy Review* 16 (August 2010): 101-131.

Glaeser, Edward L., Joshua D. Gottlieb, and Joseph Gyourko. "Can Cheap Credit Explain the Housing Boom?" Working Paper No. 16230. National Bureau of Economic Research, July, 2010.

_____, Joseph Gyourko, and Albert Saiz. "Housing Supply and Housing Bubbles." *Journal of Urban Economics* 64 (September 2008): 198-217.

Haughwout, Andrew, Richard Peach, and Joseph Tracy. "Juvenile Delinquent Mortgages: Bad Credit or Bad Economy." *Journal of Urban Economics* 64

(September 2008): 246-257.

Hendershott, Patric, and Joel Slemrod. "Taxes and the User Cost of Capital for Owner-Occupied Housing." *Journal of Americal Real Estate & Urban Economics Association* 10 (Winter 1983): 375

393. Himelberg, Charles, Christopher Mayer, and Todd Sinai. "Assessing High House Prices: Bubbles, Fundamentals, and Misperceptions." *Journal of Economic Perspectives* 19 (Fall 2005): 67-92.

Lee, Donghoon, and Wilbert van der Klaauw. "An Introduction to the FRBNY Credit Panel," Federal Reserve Bank of New York Staff Report No. 479, November, 2010.

Mayer, Christopher, Karen Pence and Shane M. Sherlund, "The Rise in Mortgage Defaults," *Journal of Economic Perspectives* 23 (Winter 2009): 27-50.

Poterba, James. "Tax Subsidies to Owner-Occupied Housing: An Asset Market Approach." *Quarterly Journal of Economics* 99 (November 1984): 729-752.

Shiller, Robert. *Irrational Exuberance: 2nd Edition*. Princeton University Press, 2005.

CHAPTER 9

Optimal Provision of Loans and Insurance against Unemployment from Lifetime Perspective

by
Joseph Stiglitz
(Columbia University)
Jungyoll Yun
(Ewha University, Seoul, Korea)

Abstract

In an earlier paper, we showed that integrated individual accounts, allowing individuals to borrow against future pensions when they are unemployed, can be welfare increasing, because it allows increased intertemporal consumption smoothing without attenuating incentives to search. Here, we examine from a lifetime perspective how the optimal mix between publicly provided unemployment insurance (UI) and loans against pension accounts changes over time in a model where unemployment may occur in any period. We show that, if the incidence of unemployment is relatively low when old, i) the optimal mix for the young entails a positive amount of loans regardless of its incentive costs; ii) the amount of consumption for those unemployed young is greater than for those unemployed old while the converse may be true in the absence of borrowing; iii) the optimal mix entails more loans and a smaller UI benefit for the young than for the old. We demonstrate that there will be incentives to save excessively in good states as well as to borrow excessively from the market when unemployed. Individuals and markets do not take into account the externalities such actions: they affect search, and thus the magnitude of UI payments and loan defaults. Finally, we show how non-market groups can improve welfare through loan-cosigning, which may be voluntarily provided within the group, as

it allows them to smooth their incomes with lower incentive costs, and while the income sharing is less effective than market pooling, the incentive benefits dominate. Current UI programs have benefits that are typically dependent only on recent employment history, and do not have any loan provisions. Thus, this paper suggests that there is considerable scope for reforms that allow better intertemporal smoothing and risk mitigation while simultaneously improving search incentives.

1. Introduction

Unemployment insurance has been criticized because of its adverse incentive effect.[1] Stiglitz and Yun (2005) show that under seemingly fairly weak conditions, provided that the duration of unemployment is limited, self-insurance through borrowing e.g. against future retirement benefits, could enhance welfare of workers by providing them with intertemporal income smoothing.[2] For most individuals, the fraction of life time income that is lost as a result of episodic unemployment is small, so that individuals are close to risk neutral with respect to such losses. The problem is that, with imperfections in capital markets, temporarily unemployed individuals are forced to cut consumption. Several studies (Chatty(2008)) have shown that the liquidity constraint is one of the most serious difficulties facing unemployed individuals. But by borrowing against one's retirement savings, capital market constraints are resolved. Self-insurance has the advantage that there are no moral hazard (adverse incentive) effects. More specifically, the loan provides inter-temporal consumption smoothing with little incentive costs, while unemployment insurance provides inter-state consumption smoothing (insurance) with some incentive costs (i.e. between states of the world where unemployment does not occur and those where it does). This suggests that a desirable form of income support for unemployed individuals may involve a combination of loans and UI benefit. Analyzing the optimal mix in a model where unemployment shock is

1 See Flemming (1978), Hopenhayen and Nicolini (1991).
2 See also Altman and Feldstein (1998), Costain (1999).

small so that it may not cause default associated with loans, Stiglitz-Yun(2005) showed that the optimal mix always entails a positive amount of loans, collateralized by pension savings. There should not be complete reliance on unemployment insurance (UI).

When there is a chance of more extended unemployment, however, the benefits of loans may be limited while an incentive costs associated with loans may arise. With the chance of extended unemployment, there is the risk that the requisite borrowing against retirement savings results in individuals depleting their pension accounts. The possibility of long periods of unemployment means that intertemporal smoothing through loans may not suffice in limiting the costs of the risks posed by unemployment. Also, as there is a chance that those unemployed in later periods cannot repay what they borrowed in the earlier periods and thus, in some sense, have to be bailed out, the provision of loan to the unemployed may not give them right incentives to get reemployed. Furthermore, if a certain amount of UI benefit has to be paid to those bailed-out, the additional adverse incentive effects associated with the UI benefit may interact with those associated with loans, implying a reduction in the scale loans.

This paper analyzes the optimal combination of loans and UI in a model where unemployment may occur at any point in time. The precise mix depend upon the relative benefits and costs of the one compared to the other, which vary with one's employment history and the point an individual is in his lifetime. Since the amount of loans at any point is based upon one's lifetime income expected at the time of unemployment, the introduction of loans necessitates designing social insurance against unemployment and retirement from a lifetime perspective.[3]

While our earlier analysis suggested that loans were preferable to UI (income was smoothed, but there was no attenuation of incentives), in a life-time model, there is a risk of extended unemployment, and

---

3 The analysis below will make it clear that even without a formal government program of lending against pensions (as is considered here), the fact that individuals borrow and save affects the optimal design of UI programs, and necessitates taking a life-time perspective.

hence a risk that individuals will not be able to repay what they have borrowed. The "bailout" that then results can be thought of as a form of UI for extended bouts of unemployment, but that means that there are incentive effects associated with loans as well as with UI. Given this, the question is, is it still the case that loan provision should be a part of optimal package of benefits for unemployed individuals. A critical factor is the amount of UI benefit paid to them in the future when they are unemployed again, as it exerts a negative externality upon the loans. It induces less search, and thus a higher likelihood of unemployment, and hence a higher likelihood of non-repayment. While both the UI benefit for the long-term unemployed and the loans for the young unemployed decrease as the incentive problem for the old unemployed gets more serious, the former may decline faster than the latter when the probability of getting re-unemployed is small. The main reason for this is that the incentive problem for the loans arises with a certain probability whereas that for the UI benefit does with probability 1. This suggests that the optimal package will entail a positive amount of loans for the unemployed young, regardless of the incentive costs, unless the probability of long-term unemployment is high.

Not surprisingly, the optimal mix of loans and UI changes over time: Unless the probability of being long-term unemployed is high, it should entail more of loans and less of UI benefit when unemployed young than when old, while the amount of consumption for those unemployed young should be greater than for those unemployed old. After all, those unemployed when they are young anticipate that the losses are likely to be made up over the rest of their lives; while those who are unemployed when they old know that that cannot be the case. The latter result suggests a role of loans, because, without loan provision, the UI benefit would be smaller in the earlier periods than in the later periods . It is also shown that the optimal mix depends upon the incidence of unemployment and its (expected) duration: it entails more loans with a small UI benefit when more individuals are unemployed for a shorter period of time.

One interesting implications of our analysis is that, even with optimal UI, private incentives for (precautionary) savings may be excessive, which can affect the optimal mix of unemployment benefits in later periods, as the excessive savings aggravate incentive costs while increasing the to smooth out consumption on their own. It is shown that the excessive private savings, if not controlled by the government, implies that, for the optimal package of benefits, greater reliance on loans and smaller UI benefits than would be the case if it can be controlled by the government.

We noted that, with the possibility of the extended unemployment, the loan may not only smooth consumption across time, but serve as insurance (as a result of the bail-out.) This aspect of loans raises a couple of other important issues we deal with in this paper: the externality associated with private loan markets and welfare effects of loan-cosigning. As for the first issue, as the government tries to discourage the market from offering excessive loans for the unemployed because of the externality associated with private loans, it has to offer more UI to reduce the scope of inefficient provision of loans by the market, resulting in the excessive unemployment benefits for the unemployed.[4] This paper thus uncovers a market failure—the risk that the market provides too much income smoothing, thereby attenuating incentives to search and imposing additional costs on the government (as the provider of unemployment insurance.)[5] In short, unrestrained loan markets are socially dysfunctional: It leads to too little efforts at job search. It is markets, not government, that, in some sense, is responsible for excessively high unemployment.[6]

On the other hand, we show that a *non-market* group, such as family, village, etc., that has a superior monitoring ability (to that of government or markets) and a sense of peer pressure among its

---

4 Chetty and Saez(2010) discuss, in a general framework of insurance provision, how the presence of private insurance market affects the optimal social insurance.

5 In a sense, this market failure is related to that analyzed by Arnott and Stiglitz (1991), who point out that the provision of insurance against one risk may affect risk taking affecting other insurance contracts. This, in turn, is related to the fundamental non-decentralizatibility theorem of Greenwald and Stiglitz [1986).

6 Of course, some governments may have provide UI benefits in excess of the optimum. Our analysis shows that "excessive" unemployment that to obtain the optimal level of unemployment, restrictions have to be placed both on the amount of insurance that government provides and the amount of borrowing that individuals can undertake.

members, can be used to improve welfare as they supplement publicly provided UI benefits through loan-cosigning.[7] We show that the informational advantage and the peer pressure associated with a non-market group can interact with each other, leading to an equilibrium where loans for one member are voluntarily cosigned by another member and can be Pareto superior.

This paper is somewhat related to the literature on optimal unemployment insurance (Hopenhayn-Nicolini (1997), Kocherlakota (2004), Shimer-Wernings (2005)). The existing literature focuses, however, on how consumption changes during the unemployment tenure of an individual, while this paper analyzes how the optimal consumption for the unemployed changes with the timing of unemployment in one's lifetime career. The rest of this paper is organized as follows. The next section describes the basic model that characterizes the optimal mix of UI benefits and loans from the lifetime perspective and analyzes how it varies with changes in the probability of unemployment in later periods and the possibility of the extended unemployment. Section III explores how the externality associated with private loans affects the optimal program for the unemployed by the government, while Section IV addresses the welfare implications of loan-cosigning. Section VI collects the main results of this paper with some concluding remarks.

2. The Model and Baseline Optimum

Consider a 3-period model in which an individual may work for period 1 and 2 at the wage w per period, and then retires in period 3 (Fig.1). For simplicity, we assume w is fixed and there is no discounting (the safe rate of interest is zero). The worker may be confronted with an unemployment shock in each of the two periods. The probability of an unemployment shock occurring to an individual in period 1 is q, while

7 This can be compared to Arnott-Stiglitz(1991), who argues that the presence of a non-market group may not be welfare-increasing in the provision of insurance unless it has perfect control of the actions taken by its members.

that in period 2 depends upon whether or not he is unemployed in period 1. The probability of a shock in period 2 for a worker who was previously unemployed is p_N, while that for a worker who was not unemployed is p_U.[8]

There are three different unemployment shocks in the model: unemployment shock in period 1 (called unemployment shock 1), unemployment shock in period 2 for those who have not been unemployed (unemployment shock 2), and unemployment shock in period 2 for those who have previously been unemployed (unemployment shock 3). Each unemployment shock occurs at the beginning of the period. After each shock, a worker may choose to search or not to search for a job. If he expends search effort e, then he finds a job; if he doesn't search, he is unemployed that period.[9]

Search cost may differ across the three shocks: we denote e_1, e_2, e_3, for the unemployment shock 1, 2, 3, respectively. With expenditure of e, he gets reemployed in the period.[10] The search cost $\{e_1, e_2, e_3\}$ are independent random variables with distribution functions F_1, F_2, F_3, respectively. The individual finds out his search costs before committing to search.

It is easy to show that there exists a threshold level $\bar{e}_i$ $(i = 1,2,3)$, such that when the realized search cost for an individual in a period is lower (higher) than a threshold level, he will choose to search (not to search). Hence, if the threshold for the unemployment shock 1 is $\bar{e}_1$, then the first period unemployment rate *among those who have faced*

---

8 A couple of different interpretations of the parameter Pu should be noted. It could be interpreted as the correlation coefficient between the two unemployment shocks in period 1 and 2, i.e. $p_U > p_N$ implies that an individual who is unemployed today has a higher probability of facing an "unemployment shock" next period than an individual who is employed. $p_U = 1$ means that an individual who is unemployed today will be unemployed next period (in the absence of search) . Also, Pu could also refer to the duration of unemployment relative to one's lifetime income.

9 An individual who is unemployed in period 1 and does not experience an unemployment shock in period 2 can be thought of as having been laid off for one period.

10 The search activity in the model takes no time and guarantees a job for the worker with probability 1. Thus, a worker choosing to search will not be unemployed in the period, like the one with no unemployment shock.

the shock is $1 - F_1(\bar{e}_1)$. We denote by $h_i \equiv \frac{f_i(e)}{1-F_i(e)}$ $(i = 1,2,3)$ the search elasticity of unemployment, i.e., the degree of sensitiveness of unemployment with respect to search activity, for a shock i, and assume that h_i is constant over e.

An unemployed worker (under shock i) receives unemployment compensation from the government consisting of two components: an unemployment insurance (UI) benefit r_i $(i = 1,2,3)$, which does not have to be repaid), and a loan in the amount of R_i $(i = 1,2,3)$. The loan for a worker is to be repaid out of his retirement income, which consists of the savings he has made out of his wage income during his working career. If an individual gets unemployed again in period 2 as well as in period 1, his retirement income will become zero so that he may be unable to repay what he borrowed in period 1 when he was unemployed. In this case the government bails him out. But at the time the government provides the "loan" in period 1, it does not know whether the individual will have to be bailed out.[11]

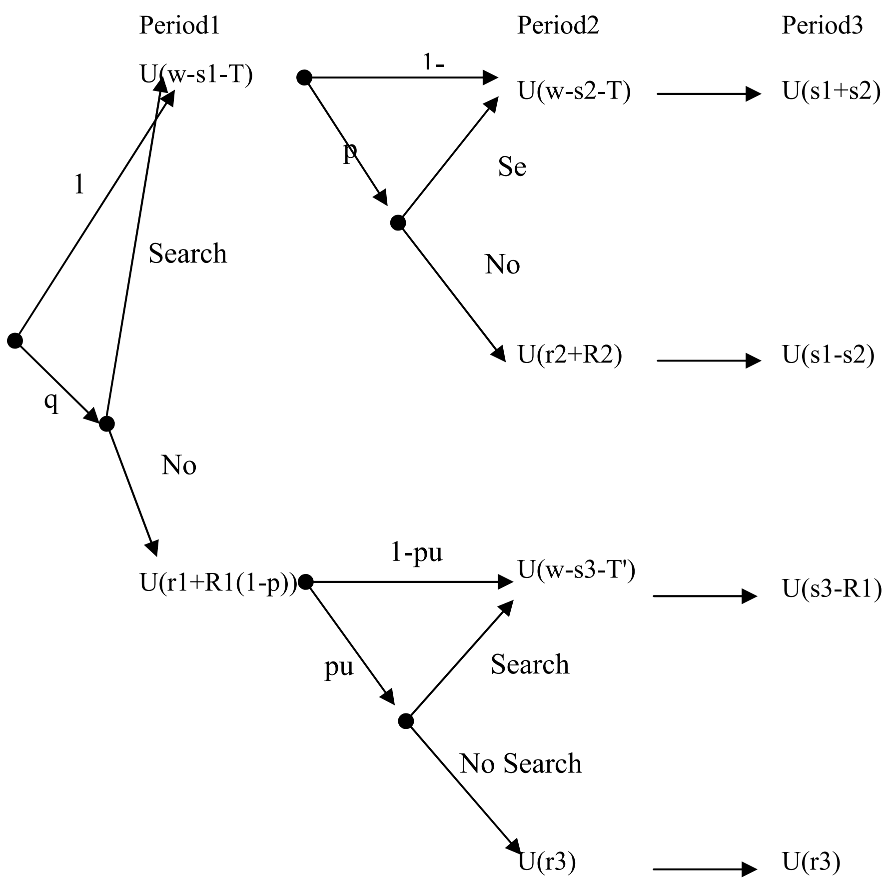
The UI benefit provided to the unemployed workers is financed by an (unemployment) tax imposed upon employed workers. In particular, the UI benefits for shock 1 and 2, r_1 and r_2, are financed by the tax T imposed upon those who get employed the first period, while the UI benefit r_3 for shock 3 is financed by the tax T' imposed upon those who get employed the second period who were not employed the first (i.e. who either don't experience shock 3 or who avoiding the resulting unemployment by searching). (see Figure 1)[12] Also, we assume that the *expected* cost of the bail-out for defaulted loans is, in effect, borne by

---

11 As for the loan to be offered in period 2, the government would not provide it to those who were unemployed in period 1 as the government knows that they would not have any income left for retirement. This point will be made clear later in this section.

12 We separate financing UI benefit r_3 against shock 3 from financing the other UI benefits in this paper, as we would like to avoid the analytical complexity caused by the inter-shock moral hazards, i.e., caused, for example, by the effect upon unemployment incidence under shock 1 of UI benefit r_3 against shock 3. But we could think of lump-sum taxes for different periods and states that can equalize the UI taxes in this case without affecting the nature of the optimal set of UI benefits and loan in this model.

| Figure 1 | Model



$T = \frac{1}{2-\bar{p}N}\left(\frac{\bar{q}}{1-\bar{q}} r1 + \bar{p}N\ r2\right)$ $T' = \frac{\bar{p}u}{1-\bar{p}u} r3$ p: price of loan

borrowers at the time of borrowing as part of the price for the loan. Heuristically, we can think of the loan as facilitating life time consumption smoothing while the UI benefit focuses on inter-state smoothing. Note, however, that a loan with the possibility of bail-out provides insurance against an unemployment shock in period 2.

We will consider a case when the government is the only provider of unemployment insurance and loans for the unemployed workers[13] and individuals are homogeneous. We will explore later in this paper the desirability of the government provision of loans compared to the private provision by considering a case when individuals are heterogenous and they are privately informed of their types: they differ in the probability of getting re-unemployed in period 2 after being unemployed in period 1.

Let $V_i (i = N, U)$ denote the lifetime expected utility for a worker who is employed or unemployed in period 1. The lifetime expected utility V for an individual at the beginning of period 1 will be

$$V = \mathrm{Max}_{\bar{e}_1, s_1} (1 - \bar{q})V_N + \bar{q}V_U - q\int_0^{\bar{e}_1} e dF_1 \quad (1)$$

where q is the probability that the shock 1 occurs, and

$$\bar{q} \equiv q(1 - F_1(\bar{e}_1)),$$

giving the probability of being unemployed under shock 1. $\bar{e}_1$ is the threshold search cost in period 1, above which individuals do not search; it is determined as follows:

$$\bar{e}_1 = V_N - V_U \quad (2)$$

where V_N and V_U are the life-time expected utility of an individual who is employed (not employed) in period 1. The individual searches if the lifetime benefits from search are worth the costs. We will not calculate these lifetime values. Let V_{ij} $(i, j = N, U)$ indicate the lifetime utility expected at the beginning of period 2 for a worker who is employed or unemployed in period 2 after having been employed or unemployed in period 1. Normalizing the constant wage to 1 for

13 Or, we can assume that the government can perfectly control provisions of benefits and loans from private sectors so that it may implement optimal package of unemployment compensations.

simplicity, we then have the individual maximizing with respect to e_i, given s, and the government maximizing with respect to s:

$$V_N = \mathrm{Max}_{s_1}\{U(1 - s_1 - T)$$
$$+\mathrm{Max}_{\bar{e}_2}\{[1 - p_N(1 - F_2(1 - \bar{e}_2))]V_{NN} + p_N(1 - F_2(1 -$$
$$\bar{e}_2))V_{NU} - \int^{\bar{e}_2} edF\}\}$$
$$V_{NN} = \mathrm{Max}_{s_2}\{U(1 - s_2 - T) + U(s_1 + s_2)\} = 2U(\tfrac{1+s_1-T}{2})$$
$$V_{NU} = \mathrm{Max}_{R_1}\{U(r_2 + R_2) + U(s_1 - R_2)\} = 2U(\tfrac{r_2+s_1}{2})$$
$$V_U = U(r_1 + R_1(1 - P))$$
$$+\mathrm{Max}_{\bar{e}_3}\{(1 - p_U(1 - F_3(\bar{e}_3)))\, V_{UN} + p_U(1 - F_3(\bar{e}_3))V_{UU} -$$
$$\int^{\bar{e}_3} edF\}$$
$$V_{UN} = \mathrm{Max}_{s_3}\{U(1 - s_3 - T') + U(s_3 - R_1)\} = 2U(\tfrac{1-R_1-T'}{2})$$
$$V_{UU} = 2U(\tfrac{r_3}{2})$$

where s_1, s_2 , s_3 denote savings for an individual and $\bar{e}_2, \bar{e}_3$ indicate the threshold search cost upon unemployment shock 2, 3, respectively. (In this formulation, the government gets to set the savings rate s_i; later, we will deal with the more general case.) We can then see that

$$\bar{e}_2 = V_{NN} - V_{NU} \tag{3}$$
$$\bar{e}_3 = V_{UN} - V_{UU} \tag{3'}$$

and

$$T = \frac{1}{(2-\bar{p}_N)}(\frac{\bar{q}}{(1-\bar{q})} r_1 + \bar{p}_N r_2)$$

$$T' = \frac{\bar{P}_U}{1-\bar{P}_U} r_3$$

where

$$\bar{p}_N \equiv p_N\big(1 - F_2(1 - \bar{e}_2)\big)$$
$$\bar{p}_U \equiv p_U\big(1 - F_3(1 - \bar{e}_3)\big)$$

The price of loan, P, refers to the price of public loans provided to those unemployed in period 1, which is equal to the probability of default, $\bar{p}_U$. Note also that individual savings s_2, s_3 are determined so as to equalize consumption across periods.

The optimal savings s_1 in period 1 is determined so as to balance inter-temporal consumption smoothing against the disincentive associated with savings:

$$-U'(1 - s_1^* - T) + (1 - \bar{p}_N)U'\left(\frac{1+s_1^*-T}{2}\right) + \bar{p}_N U'\left(\frac{r_2+s_1^*}{2}\right)$$
$$-h_2 \frac{\bar{p}_N}{(2-\bar{p}_N)^2}\left(\frac{\bar{q}}{1-\bar{q}} r_1 + 2r_2\right)\left\{U'\left(\frac{r_2+s_1^*}{2}\right) - U'\left(\frac{1+s_1^*-T}{2}\right)\right\}\{U'(1 - s_1 -$$
$$T) + (1 - \bar{p}_N)U'\left(\frac{1+s_1^*-T}{2}\right)\} = 0. \qquad (5)$$

In this paper we treat the savings in period 1 as the mandatory savings for retirement. This can be justified by the possible moral hazard behavior in the choice of savings that can be caused by the government subsidy r_3 to those with no income for retirement, which is a part of government program in this paper.[14]

On the other hand, the private savings s_1^o that individuals would like to make, which is different from s_1^*, is determined as

$$-U'(1 - s_1^o - T) + (1 - \bar{p}_N)U'\left(\frac{1+s_1^o-T}{2}\right) + \bar{p}_N U'\left(\frac{r_2+s_1^o}{2}\right) = 0. \qquad (5')$$

We can then prove the following Proposition.

14 Without the mandatory retirement savings in period 1, the loan for the unemployed in period 2 may not be necessary, because individuals can consume a part of their savings when unemployed.

Proposition 1

$s_1^* < s_1^o$.

This is clear from the comparison of (5) with (5'). That is, because private individuals do not take into account the incentive costs caused by their savings, the level that individuals would save on their own account (private savings) s_1^o is greater than the socially optimal level s_1^*.

This implies that individuals will make saving beyond the optimal level s_1^* mandated by the government. We will assume for the remaining part of this paper that, whenever applicable, the government can implement the optimal savings s_1^* on the part of individuals by imposing taxes upon the extra (private) savings individuals make in period 1. We will check later, however, how, if this is not the case, the uncontrolled private savings may affect the government program for unemployed individuals.

The interior solution for the optimum unemployment insurance/loan program $\{r_i^*, R_i^*\}(i = 1,2,3)$, which maximizes the lifetime expected utility V, will then satisfy the following conditions (by the envelope theorem on the savings s_1):

r_1^* :

$$U'\left(r_1^* + R_1^*(1-\bar{p}_U)\right)(2-\bar{p}_N) - M(T)$$
$$-\frac{h_1}{1-\bar{q}} r_1^* \left\{U'\left(r_1^* + R_1^*(1-\bar{p}_U)\right) + \frac{1}{2-\bar{p}_N}\frac{\bar{q}}{(1-\bar{q})} M(T)\right\} \cdot M(T) = 0 \quad (6)$$

R_1^*:

$$U'\left(r_1^* + R_1^*(1-\bar{p}_U)\right) - M(T')$$
$$-\frac{h_3\bar{p}_U}{1-\bar{p}_U}\left\{R_1^* U'\left(r_1^* + R_1^*(1-\bar{p}_U)\right) + \frac{1}{1-\bar{p}_U} {r_3^*}' M(T')\right\} \cdot M(T') = 0 \quad (7)$$

r_2^* :

$$U'(r_2^* + R_2^*)(2 - \bar{p}_N) - M(T)$$
$$-h_2 \frac{1}{(2-\bar{p}_N)}\left(\frac{\bar{q}}{1-\bar{q}} r_1^* + 2r_2^*\right)\{U'(r_2^* + R_2^*) + \frac{\bar{p}_N}{2-\bar{p}_N} M(T)\} \cdot M(T) = 0 \quad (8)$$

R_2^*:

$$R_2^* = \frac{s_1^* - r_2^*}{2} \quad (9)$$

r_3^* :

$$U'\left(\frac{r_3^*}{2}\right) - M(T') - h_3\{R_1^* U'\left(r_1^* + R_1^*(1 - \bar{p}_U)\right) + \frac{1}{1-\bar{p}_U} r_3^* M(T')\}$$
$$\cdot \left\{U'\left(\frac{r_3^*}{2}\right) + \frac{\bar{p}_U}{1-\bar{p}_U} M(T')\right\} = 0 \quad (10)$$

R_3^*:

$$R_3^* = 0 \quad (11)$$

where $M(T) \equiv U'(1 - s_1^* - T) + (1 - \bar{p}_N)U'\left(\frac{1+s_1^*-T}{2}\right)$

and $M(T') \equiv U'\left(\frac{1-R_1^*-T'}{2}\right)$, is the marginal disutility of tax T (for an individual employed in period 1) and of tax T' (for an individual unemployed in period 1), respectively. The above conditions imply that the optimal UI benefit and loan in each of the optimal government programs is determined so as to balance its benefit of consumption smoothing across states or periods with the moral hazard (adverse incentive) costs associated with them.[15]

15 In this paper we confine ourselves to the case when an unemployed individual is liquidity-constrained so that he would not saving out of his UI benefit. That is, we assume for the sake of non-negativity of R_1 or R_2 that $r_1 < \frac{1-T}{2}$ and that $r_2 < s_1^*$.

We will first explore whether or not it is desirable to introduce loans when there is some chance that an individual remains unemployed for a long time and thus may not repay what he has borrowed.[16]

Proposition 2

R_1^* (or r_3^*) > 0 whenever r_3^* (or R_3^*) = 0. In general, there exist p_U^o such that $R_1^* > 0$ for any h_3 when $p_U < p_U^o$.

Proposition 2 demonstrates that unless the probability of long-term unemployment is high, it is optimal to introduce some loans as a part of the benefit package for unemployed individuals despite some chance of extended unemployment and the resulting loan-default. The intuition is that loans and UI benefit r_3 exert an externality, and the magnitude of the effect of loans depends on the probability $\bar{p}_U$. This implies that, the presence of loans decreases the optimal level of UI and vice versa. Now, as the sensitivity of search to effort increases (h_3) increases, overall benefits decrease. The question is, which of the two – loans and r_3 - becomes zero earlier than the other as h_3 increases. Proposition 2 shows that r_3 does, unless the probability p_U of extended unemployment is high. The main reason is the following. Although consumption-smoothing effect is larger for r_3 than for loans, the incentive effect is also larger. Furthermore, while the incentive effect for loans is limited by the probability p_U of extended unemployment, that is not the case for r_3, so that a lower p_U increases the difference in the incentive effect between the two.

Now let us examine how the government loan program and the resulting unemployment benefit will be affected by the timing of unemployment by analyzing $\{r_i^*, R_i^*\}$ (i=1,2,3). We can state the following proposition.

16 Recall that for this part of the paper, , in analyzing the optimal government program $\{r_i^*, R_i^*\}$, we will assume that the government can control private savings so that $s_1 = s_1^*$.

Proposition 3

i) $r_1^* + R_1^*(1 - \bar{p}_U) > \frac{r_3^*}{2}$.

ii) Suppose that $F_1 = F_2$ and that $p_N \leq q$. In the absence of loan provisions, $r_1^* < r_2^*$. In the presence of loan provisions, there exists $p'_U(> 0)$ such that, for $p_U \leq p'_U$, the followings are true: (a) $r_1^* < r_2^*$ and $R_1^* > R_2^*$, and (b) there exists $p'_N(> 0)$ such that $r_1^* + R_1^*(1 - \bar{p}_U) > r_2^* + R_2^*$ for $p_N \leq p'_N$.

The proof is delegated to the Appendix. Proposition 3i says that there is imperfect consumption smoothing: those who are unemployed the first period and the second have a higher level of consumption the first period than they do in the second and third. This is perhaps not surprising: the first period, their expected income is still high because of the expectation of a second period job, and they wish to smooth consumption over their lifetime. Insurance does not fully replace lost income, and so when the adverse shock occurs in the second period, consumption has to decrease.

Proposition 3ii looks at the central case where the search cost distributions in the first and second period (in the second period, for those who have been employed in the first) are equal and where the probability of an unemployment shock the second period to someone who was unemployed the first is equal to (or less than) the probability of an unemployment shock the first period. If there are no loans, then first period unemployment insurance is less than the second, and that means consumption of those unemployed only in the second period is greater than consumption of those only unemployed in the first period. But if there are loans in the first and second period, if the probability of a person who is employed the first period facing an unemployment shock the second period is small, then consumption of the person who is unemployed the first period is greater than that of the person who is unemployed (only) in the second, and the mix of benefits in the first period is weighted more towards loans and away from UI.

In comparing consumption of those who are unemployed in only one period—when young and when old—there are three effects: (a) an unemployed young person has a longer time over which to smooth consumption; (b) the person who is unemployed the first period *on average* will be worse off than the person who is only unemployed in the second period (and in that sense, his lifetime expected income is lower); (c) one can only smooth going forward, not going backward and because the young spent so much of his lifetime income (rationally) in the first period, in the expectation that he would not be unemployed, he has to take a big hit the second period. How these effects weigh out depends on the usage of loans. When there are no loans, consumption of the young unemployed is lower; when there are loans and the probability of future unemployment (of those already unemployed) is low enough, consumption of the young unemployed is higher.

Unrestricted private savings

We have thus far assumed that the government can ensure that individuals' savings s_1 in period 1 are at the optimal level s_1*. If the government cannot control individual savings, however, it has to take into account the change in private savings in response to the government unemployment insurance program. In particular we can prove the following Proposition on the government program (r_2^o, R_2^o) in the presence of uncontrolled private savings.

Proposition 4

If $U''' \approx 0$, then there exists p_N^o such that, for $p_N < p_N^o$, $r_2^o < r_2^*$ and $R_2^o > R_2^*$.

The proof can be found in the Appendix. The proposition says that, provided that the probability of someone employed the first period facing an unemployment shock the second period is small enough, private savings induces government to provide smaller UI benefits and a larger loan.

Recall that Proposition 1 says that in making savings for retirement, an individual would not take into account its effect upon the UI tax T

through the change in search incentive in period 2. If the government cannot intervene to cure the distortion in savings caused by the externality, therefore, the government program needs to be modified as implied by (5') and (6)-(10). In particular, it involves reduction in UI and increase in loans because the larger amount of savings increase the incentive cost of UI benefit under shock 2 while increasing the retirement income .[17] [18]

Comparative statics: effects of changes in the incidence of unemployment and its expected duration

Lastly we will examine how the optimal mix of unemployment benefits changes as the incidence of unemployment or its (expected) duration varies. In doing this we will focus upon the impact on $\{r_1^*, R_1^*\}$ [19]. In particular, we assume for simplicity that $P_N = 0$ and that that the UI benefit r_3 for shock 3 is exogenously given as a socially acceptable minimum level $\bar{r}$ of consumption.[20] In this case individuals choose the socially optimal level of savings s_1 in period 1:

$$s_1^* = \frac{1-T}{3} = s_1^o. \tag{12}$$

17 This result is not so obvious as it looks, because, in the presence of uncontrolled private savings, the government may want to reduce savings by *increasing the UI benefit* in period 2 (reducing the need for precautionary savings). But this effect gets relatively small when P_N is small.

18 This result can be compared to the one reported by the existing literature, which analyzes how unobservable savings affects the inter-temporal consumption structure for the unemployed during the unemployment duration (Kocherakota(2004), Shimer-Wernings(2005)), who found the optimal UI structure to be non-decreasing in unemployment tenure, as opposed to Hopenhyen-Nicolini (1998) (who argue for decreasing UI), as the amount of savings gets depleted during the unemployment duration . In this paper we focus upon how the unobservable savings affects the optimal mix of benefits in the government program for the next unemployment bout in one's future career.

19 since $R_3^* = 0$ while R_2^* is uniquely determined by r_2^* through inter-temporal smoothing (as R_2^* does not entail any bailout costs).

20 Those who get unemployed for both of the two periods should be the ones who need to be assisted by the government through the various social assistant programs. In many cases the level of assistance for these individuals tend to be determined by social and political factors, as well as economic ones.

Confining ourselves to the case when the optimal mix of UI benefit and loans involves positive amount of each type of unemployment benefit, we can establish the following Proposition:

Proposition 5(Comparative statics of unemployment insurance)

As the two unemployment shocks in period 1 and 2 get more highly correlated (as p_u increases), the optimal mix of unemployment benefits involves a greater UI benefit, r_1^*, and smaller loan-based self-insurance, $R_1^*(1-\bar{p}_U)$. The converse will be true, however, when the probability q of the period 1 unemployment shock 1 gets higher. The total unemployment benefits $(r_1^* + R_1^*(1-\bar{p}_U))$ is decreasing in q or in p_U.

The results of Proposition 5 are intuitive. First, as an unemployed worker is more likely to experience a longer duration of unemployment, i.e., as p_u gets higher, the welfare benefit from the loan-based self-insurance decreases, because loans provide limited inter-temporal smoothing and suffer from higher incentive costs. This leads to a partial substitution of UI benefits for loans, while the total amount of unemployment compensation decreases. Second, when the probability q of the period 1 unemployment shock gets higher, the burden of financing UI benefit gets larger, leading to an optimal mix involving a smaller total amount of unemployment compensation, with a smaller fraction of the benefits in the form of UI benefits.

The possibility of extended unemployment and the resulting default on loans means that loans act as insurance. For the rest of this paper we will deal with some issues related to this property of loans: the externality associated with private provision of loans and welfare effects of loan-cosigning in the presence of default possibility (i.e., when $p_U > 0$). In this analysis, for simplicity, we will focus upon $\{r_1^*, R_1^*\}$ by assuming that $p_N = 0$ and that the UI benefit r_3 for shock 3 is set exogenously to be $\bar{r}$.

3. Presence of Private Loan Market

So long as the optimal provision of loans entails incomplete inter-temporal smoothing due to the moral hazard, private lenders would have an incentive to provide additional loans, as they would not take into consideration the effect of their loan provision upon search and losses under the government loan and UI programs. This suggests that the market provision of loans is not desirable.

Obviously, if the government can observe and control the provision of loans by private lenders, the private provision of loans would not matter. Suppose, however, that the government cannot observe the private provision of loans by the market, but, of course, it can *infer* what the market will do. In this case the presence of the market would affect the government program in an important way.[21] That is, the presence of the loan market may affect the government provision. Here we will discuss this issue, together with the welfare consequence that the presence of the market may bring about.

Because under shock 2, government loans achieve complete inter-temporal consumption smoothing, there is no incentive for private markets to provide loans then. But the presence of the market would affect the loan provision under shock 1. We will first analyze the set of sustainable loan contracts in the presence of a competitive loan market and then analyze the optimal response on the part of the government. Private lenders always have incentives to provide loans to individuals whenever additional loan can improve intertemporal consumption smoothing. That is, for a given r_1, R_1 and P, additional loans will be offered whenever

$$U'\left(r_1 + R_1(1 - P)\right) - U'\left(\frac{1 - R_1 - T'}{2}\right) > 0,$$

because the price P of additional loan in a competitive loan market is equal to the (expected) probability that an individual is unemployed in

21 This is the case, even if, after the government adjusts its loan and UI program, the market chooses not to make any loans.

period 2, that is, $P = p_U(1 - F(\bar{e}_3))$, where $\bar{e}_3 = U\left(\frac{1-R_1-T\prime}{2}\right) - U(\frac{\bar{r}}{2})$. The amount of loan $R_1(P, r_1)$ for an individual that is sustainable in the market should then satisfy

$$r_1 + R_1(P, r_1)(1 - P) = \frac{1-R_1(P,r_1)-T\prime}{2} \quad (12)$$

In other words, the sustainable loan contract for the individual has to entail complete inter-temporal consumption smoothing in the presence of private loan market. Note, however, that it does not imply that the contract entails complete insurance against shock 3. In other words, an individual choosing the contract would have some (albeit insufficient) incentive to search under shock 3.

Suppose that the government delegates the provision of loans to the private market as it cannot control private loans. That is, once the government offers UI r_1^o, the private market responds to it by offering loans. We can see from (12) that the package (r_1^o, R_1^o) has to satisfy

$$R_1^o = \frac{1-2r_1^o-T\prime}{3-2\bar{p}_U}, \quad (13)$$

where $\bar{p}_U = p_U(1 - F(\bar{e}_3^o))$ and $\bar{e}_3^o = U\left(\frac{1-R_1^o-T\prime}{2}\right) - U(\frac{\bar{r}}{2})$.

The UI benefit r_1^o should then be the one that satisfies the following condition:

$$U'\left(r_1^o + R_1^o(1 - \bar{p}_U)\right) - U'\left(\frac{2-T}{3}\right) - \frac{h_1}{1-\bar{q}} r_1^o U'\left(r_1^o + R_1^o(1 - \bar{p}_U)\right)U'\left(\frac{2-T}{3}\right)$$
$$- \frac{\partial R_1}{\partial r_1}\frac{\bar{p}_U}{1-\bar{p}_U} h_3\{R_1^o U'\left(r_1^o + R_1^o(1 - \bar{p}_U)\right) + \bar{r}U'(\frac{1-R_1^o-T\prime}{2})\}U'(\frac{1-R_1^o-T\prime}{2}) \quad (14)$$

where, by (13),

$$\frac{\partial R_1}{\partial r_1} = \frac{-2}{3-2\bar{p}_U} < 0.$$

Condition (14) shows that (a) because in the presence of the market, loan provision is increased; and (b) the government must take this into account in setting the UI benefit. We can establish the following proposition on the optimal provision of UI benefit in response to the market provision of loans.

Proposition 6

In the presence of a private market for loans that cannot be controlled by the government, compared to the case of its absence, the resulting amount of unemployment benefits under shock 1 is greater than that in the absence of private market, i.e., $r_1^o + R_1^o(1 - \bar{p}_U) > r_1^* + R_1^*(1 - \bar{p}_U)$. Welfare is lower than in the case without private loans.

The proof is delegated to the Appendix. As the market provide loans so as to secure complete inter-temporal smoothing, which is costly in terms of search incentives, the government may want to substitute unemployment insurance for loans offered by the market. As a result, it may be the case that the government offers larger amount of UI to reduce the amount of loans that the market offers.[22] On the other hand, total unemployment benefits increase. The excessive amount of total unemployment benefits the government provides in response to the market comes from the two sources. The first is the pressure of the market (that does not care about the adverse incentive effects induced by its loans), which offers loans that secure full inter-temporal consumption smoothing. The second arises from the fact that it is then desirable for the government to provide more UI benefit, even though UI itself has adverse incentive effects .

The excessive amount of unemployment compensation reduces welfare. The amount of reduction in welfare depends upon the seriousness of moral hazard associated with loans. This suggests that the government should not delegate the provision of loans to the market

---

22 This can be compared to the arguments of the literature (Bailey(1978), Crossley-Low(2011)) that the presence of borrowing constraint increases optimal UI. Here the presence of uncontrolled market for loans reduces optimal UI. It is the possibility of default associated with loans and of excessive borrowing in the private market that plays a role in the model, whereas it is not considered in the literature.

whenever it can control its provision of loans, so long as there is some positive probability of default associated with loans.

4. Welfare Effects of Loan-Cosigning in the Provision of Loans

Faced with possibility of default associated with the loans on the part of the unemployed, the government may want the loans to be cosigned by other employed workers. The issue we will deal with in this section is whether the introduction of a loan-cosigning program would increase welfare and whether such a program could be made to work, i.e. would a potential cosigner have an incentive to cosign the loans. In this analysis we will use the original model in which individuals are homogenous.

We first note that an individual who is close to the cosignee, such as member of the same informal group (like the same family, close friends, etc.), is in a superior position for loan-cosigning t for a couple of reasons. First, he may be in a position to monitor the actions taken by the cosignees more effectively than the others. Second, the cosignee may be subject to so-called peer pressure from an individual within the same informal group.[23] How much an individual cares about the peer pressure within a group and how much one can effectively control the actions taken by the others would depend upon many other factors (cultural ones, for example) exogenous to this model.[24] The peer pressure may affect the behavior of both cosigner and cosignee, as will be discussed below.

In the model below examining the welfare effect of loan-cosigning weformulate a precise specification of both the informational advantage and the effect of peer pressure. When the search cost for an unemployed individual is realized, the cosignee can notice it with some probability

23 The incentive effects created by the peer pressure in a group have been discussed by Lazear (1990).

24 Here we suppose that an individual does not care about the well-being of the others within a group, implying that there is no need for intra-group transfers, *ex* post.

γ (<0). The probability γ thus indicates the degree of informational advantage that a cosigner has over the government, who cannot identify the realized search cost at all. We will next suppose that the co-signing group determines the target threshold cost of search, $\bar{e}_3'$, which is the one that is optimal from the group's point of view, as will be specified below. We will also suppose that whenever the search cost is found to be lower than the target level $\bar{e}_3'$, the cosignee must choose to search, because of the peer pressure.[25] Then, the probability that an unemployed individual chooses to search under the loan-cosigning, $\bar{F}_3$, would be

$$\bar{F}_3(\bar{e}_3, \bar{e}_3') = F_3(\bar{e}_3) + \gamma\{F_3(\bar{e}_3') - F_3(\bar{e}_3)\} = (1-\gamma)F_3(\bar{e}_3) + \gamma F_3(\bar{e}_3'), \tag{15}$$

where $\bar{e}_3$ is the threshold search cost that is chosen by an unemployed individual in the absence of loan-cosigning.

Let us consider a "group" of two individuals, within which an employed individual is supposed to cosign the loans provided to unemployed individual.[26] The government sets the rate of loan-cosigning, c, as well as UI benefit and loan for the unemployed, to maximize the expected utility of an individual; if the borrower defaults, the co-signee must repay a fraction c of the amount owed.

Suppose that an employed individual cosigns the loans R_1 for his colleague (within the same group) who is unemployed in period 1. If both of the two individuals in a group are employed or unemployed, there would be no loan-cosigning. In general, an individual who is unemployed in period 1 may have different packages of UI benefit and loans, depending upon whether his colleague is employed or not. In this section, however, for expositional convenience we will assume that the same amount of UI benefit and loans are provided to unemployed individuals in period 1 regardless of whether or not the loan can be

25 We assume here that the amount of peer pressure for a cosignee choosing not to search is greater than his benefit associated with pursuing a no-search strategy.

26 The results can be easily generalized.

cosigned.[27] But we will let the price of the loan vary with the rate of loan-cosigning.

Using this formulation we can write the expected payoff of an individual in the group as follows:

$$V = (1-\bar{q})V_N + \bar{q}V_U - \int^{\bar{e}_1} edF_1$$

where

$$V_N = U(1-s-T) + (1-\bar{q}(1-\bar{F}_3)2U\left(\frac{1+s-T}{2}\right) + \bar{q}(1-\bar{F}_3)2U\left(\frac{1+s-T-cR_1}{2}\right)$$

$$V_U = \bar{q}\left\{U\left(r_1 + R_1(1-P(0))\right) + F_3(\bar{e}_3)2U\left(\frac{1-R_1-T'}{2}\right) + (1-F_3(\bar{e}_3))2U\left(\frac{\bar{r}}{2}\right) - \int^{\bar{e}_3} edF_3\right\} + (1-\bar{q})[U\left(r_1 + R_1(1-P(c))\right) + \bar{F}_3 2U\left(\frac{1-R_1-T\prime}{2}\right) + (1-\bar{F}_3)\{2U\left(\frac{\bar{r}}{2}\right)\} - \int^{\bar{e}_3\prime} edF_3] \ ,$$

While s is the savings in period 1 and

$$P(c) = (1-\bar{F}_3)(1-c) \quad \text{for } c \geq 0,$$

where $P(c)$ is the price of loan that is cosigned at the rate c.[28] This suggests that with the loan-cosigning cR_1 the price of loan decreases by $c(1-\bar{F}_3)$ as the expected default cost gets lowered. In that respect the loan-cosigning provides an inter-personal (and inter-state) income transfer within the group. In particular, we can notice that it plays a similar role to UI benefit, in that the cost is to be borne by an employed individual and that the total resources available to the unemployed

---

27 This assumption does not matter for the analysis below, because we will be examining whether or not the *introduction* of loan-cosigning is welfare-increasing.

28 Note that $\bar{F}_3 = F_3$ when $c = 0$.

increases. The loan-cosigning is different from the UI benefit, however, in that the risk for an individual is effectively pooled in the latter case whereas it is not in the former case. On the other hand, loan-cosigning can improve individual search incentives (especially in period 2) due to the informational advantage on the part of cosigners.

The threshold search costs $\bar{e}_1$, $\bar{e}_3$ and $\bar{e}_3{}'$ will be determined as follows:

$$\bar{e}_1 = V_N - V_U$$
$$\bar{e}_3 = 2U\left(\frac{1-R_1-T'}{2}\right) - 2U(\bar{r})$$
$$\bar{e}_3{}' = 2U\left(\frac{1-R_1-T'}{2}\right) - 2U(\bar{r}) + 2\{U\left(\frac{1+s-T}{2}\right) - U\left(\frac{1+s-T-cR_1}{2}\right)\}, \tag{16}$$

The target threshold search cost $\bar{e}_3{}'$ set by the group maximizes the welfare of the group, i.e., the sum of utilities of cosigner and cosignee.[29] We can also notice from (15) and (16) that the incentive effect of loan-cosigning hinges upon the amount cR_1 of loan cosigned, as well as the degree of informational advantage indicated by γ.

Using the above conditions for the threshold search costs, we can express the incentive effects of the loan-cosigning as follows:

$$\frac{\partial \bar{e}_1}{\partial c} = -(1-F_3)R_1[\bar{q}U'\left(\frac{1+s-T-cR_1}{2}\right) + (1-\bar{q})U'$$
$$(r_1 + R_1(1-P))] < 0$$
$$\frac{\partial \bar{e}_3{}'}{\partial c} = R_1 U'\left(\frac{1+s-T-cR_1}{2}\right) > 0$$

That is, the loan-cosigning induces positive incentive effect under shock 3, while it causes negative incentive effect under shock 1.[30]

29 We can think of the group as agreeing on this ex ante, before they know which (if any) of the members of the group will be unemployed.

30 This negative incentive effect is due to the risk sharing of cosigning which lowers the difference in expected utility between getting a job and getting unemployed under shock 1.

On the other hand, the savings by an employed individual will be determined by inter-temporal smoothing[31]:

$$-U'(1-s-T_1)+\bar{F}_3U'\left(\frac{1+s-T}{2}\right)+(1-\bar{F}_3)U'\left(\frac{1+s-T-cR_1}{2}\right)=0 \tag{17}$$

Given this behavior of individuals in a group, the government sets the amounts of UI benefit and loan, and the ratio γ of cosigned loan (before an individual in the group turns out to be employed or not in period 1) to maximize the payoff $\mathbf{V}$. To see whether or not the loan-cosigning enhances welfare we will evaluate the expected payoff $\mathbf{V}$ of an individual at $c=0$. Differentiating $\mathbf{V}$ with respect to $\mathbf{c}$ at $c=0$, we have

$$\begin{aligned}\frac{\partial V}{\partial c} &= (1-\bar{q})\bar{q}(1-F_3)R_1[U'\big(r_1+R_1(1-P)\big)-U'\left(\frac{1+s-T}{2}\right)] \\ &\quad -h_1\frac{r_1}{1-\bar{q}}U'(1-s-T)\{(1-\bar{q})U'\big(r_1+R_1(1-P)\big) \\ &\quad +\bar{q}U'\left(\frac{1+s-T}{2}\right)\} \\ &\quad +h_3\frac{1}{1-\bar{q}}\{R_1U'\big(r_1+R_1(1-P)\big)+\bar{r}U'(\frac{\bar{r}}{2})\}\gamma U'\left(\frac{1+s-T}{2}\right)]\end{aligned}$$

Since $1-s-T=\frac{1+s-T}{2}$ when $c=0$ by (17), we can see from (6) that, if $\gamma>0$,

$$\frac{\partial V}{\partial c}>0.$$

In other words, as long as a cosigner has any informational advantage compared to the government, the introduction of loan cosigning is welfare-enhancing. Because at $c=0$, the risk pooling effect is trivial by the envelope theorem, the benefits of co-signing arise

31 Notice that we have allowed peer pressure to effect only search costs the second period for an individual who is unemployed the first period. The group could exercise peer pressure to effect search in other contingencies and savings. If so, the welfare benefits of group co-signing would be enhanced.

solely from the improved incentives. This establishes the following proposition:

Proposition 7

The introduction of loan-cosigning in the government provision of loans for the unemployed increases welfare so long as an individual cosigner is better informed of the realized search cost for the cosignee than the government.

Although loan-cosigning is similar to UI benefit in its risk sharingproperties, the former would not be able to fully replace the latter, as the risk is not perfectly pooled. Co-signing is only advantageous because of the informational advantage compared to the government.,. The mix of UI and loan-cosigning balances the imperfect risk-pooling and the informational advantage associated with the loan-cosigning determine.[32]

Finally, we argue that peer pressure provides the motivation for members of the group to co-sign loans. Note that the loan cosigned by an employed member of a group is priced lower than it would otherwise be because the loan cosigned is subject toa lower default probability, and all members of the group benefit in expected utility. Thus, ex ante, it makes sense for them to agree to co-sign, and it is reasonable to assume that peer pressure is sufficiently great (at least for small c) that they would not renege from such an agreement.[33]

---

32 The question arises over the possibility of private incentive for loan-cosigning beyond the rate mandated by government. As long as the group does not take into account the effect of its choice upon the UI tax or the price of loan, it is easy to imagine circumstances under which the private group might have an incentive to go beyond the level required by the government. But unless the cosigners fully control the action of unemployed cosignee, showed this may decrease welfare, because the additional risk sharing attenuates first period search effort (The argument is analogous to Arnott-Stiglitz(1991).) In this paper, however, there is no incentive for additional loan-cosigning on the part of a group once unemployment risk is realized in period 1, because there is no interdependence of utility functions by members of the group.

33 More formally, we can assume that the marginal peer pressure associated with the initial cosigning (i.e., at $c = 0$) is greater than the marginal cost (for the

5. Conclusion

This paper analyzes the optimal combination of UI benefit and loan-based self-insurance for unemployed individuals from the lifetime perspective, i.e., examines how it changes over time in a model where unemployment may occur in any period. This paper demonstrates that providing loans should be part of any unemployment system, unless the probability of long-term unemployment is high: i) the optimal package of unemployment benefit should involve a positive amount of loans regardless of search elasticity, ii) the optimal mix should entail more loans and smaller UI benefits for young unemployed individuals than for the old, iii) when the probability of being unemployed when old is low, the amount of consumption for those unemployed when young is greater than those unemployed when old, while the opposite is true in the absence of the provision of loans. This paper also points out the externality associated with private savings, given the government program—precautionary savings are higher than is optimal and search is lower-- and shows that, when the government cannot control individual savings, the UI benefit should be lower and loans should be larger \ in later periods.

The possibility of extended unemployment and the resulting default on loans means that loans act as insurance. This suggests that the market may offer excessive amount of loans because of the externality associated with private provision. When the government cannot control the market provision of loans, however, the government may offer more UI to reduce the amount of inefficient provision of loans by the market, compared to the case of no private market, resulting in the excessive unemployment benefits for the unemployed.

Finally, this paper shows that a government program of co-signing can be welfare enhancing, and can be implemented voluntary by members of a group, if members of the co-signing group have an informational advantage over the government and can exert peer pressure to ensure that individuals with low search costs search for a job.

cosigner) of cosigning at $c = 0$, so that the cosigner would voluntarily accept some level of cosigning for his (unemployed) colleague.

References

R. Arnott, J Stiglitz (1991), “Moral Hazard and Non-market Institutions: Dyfunctional Crowding Out or Peer Monitoring,” American Economic Review, 81, pp. 179-190

R. Arnott, J Stiglitz (1991), “The Welfare Economics of Moral Hazard,” NBER Working Paper, w3316

M. Bailey (1978), “Some Aspects of Optimal Unemployment Insurance,” Journal of Public Economics, 10(3), 379-402

R. Chetty (2008), “Moral Hazard vs. Liquidity, Optimal Unemployment Insurance,” Journal of Political Economy, 116(2), 176-233

R.Chetty, E. Saez (2010), “Optimal Taxation and Social Insurance with Endogenous Private Insurance,” American Economic Journal: Economic Policy, 2(2), 85-114

T. Cosley, H. Low (2011), “Borrowing Constraint, The Cost of Precautionary Savings, and Unemployment Insurance,” mimeo

J. Costain (1999), “Precautionary Savings,” Working Paper, University Pompeu Fabra

M. Feldstein, D. Altman (1998), “Unemployment Insurance Savings Account,” NBER Working Paper 6860

J. Flemming (1978), “Aspects of Optimal Unemployment Insurance: search, leisure, savings, and capital market imperfections,” Journal of Public Economics, 10, pp. 403-25

B. Greenwald, J. Stiglitz (1986), “Externalities in Economies with Imperfect Information and Incomplete Market,” Quarterly Journal of Economics, 102, pp. 229-64

H. Hopenhayn, J. Nicolini (1997), “Optimal Unemployment Insurance,” Journal of Political Economy, 105, 312-438

E. Kandel, E. Lazear (1992), “Peer Pressure and Partnership,” Journal of Political Economy

N. Kocherlakota (2004), “Figuring out the Impacts of Hidden Savings on Optimal Unemployment Insurance,” Review of Economic Dynamics, 7. 541-554

R. Shimer, I. Wernings (2005), “Liquidity and Insurance for the Unemployed,” Federal Reserve Bank of Minneapolis, Staff Report, 366

M. Whinston (1983), “Moral Hazard, Adverse Selection, and Optimal Provision of Social Insurance,” Journal of Public Economics, Vol. 22, No. 1, pp 49-71

M. Rothschild, J. Stiglitz (1976), “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Information,” Quarterly Journal of Economics, Vol. 90, pp 629-650

J. Stiglitz, J. Yun (2005), “Integration of Unemployment Insurance with Retirement Insurance,” Journal of Public Economics, 89, pp. 2037-67

J. Stiglitz, J. Yun (2009), “Optimality and Equilibrium In a Competitive Insurance Market under Adverse Selection and Moral Hazard,” mimeo

Appendix

<Proofs of Propositions>

<Proposition 1>

Comparing (5) with (5'), we can see that $s_1^* < s_1^o$.

<Proposition 2>

Let r_1^o and r_3^o be the solution to (6) and (10) when $R_1 = 0$, respectively. Since $r_1^o < \frac{1-T\prime}{2}$, we can see that (7) or (10) cannot hold if $R_1 = r_3 = 0$. This proves the first part of Proposition 2. Let us substitute $R_1 = 0$ into (7) and (10) to see if $R_1^* = 0$. We can then see that $R_1^* > 0$ for any h_3 if

$p_U < \frac{U'(\frac{r_3^o}{2})}{U'\left(\frac{r_3^o}{2}\right) - U'(\frac{1-T'}{2})} \frac{U'\left(\frac{r_1^o}{2}\right) - U'(\frac{1-T'}{2})}{U'\left(\frac{r_1^o}{2}\right)}$. If we let $\frac{U'\left(\frac{r_1^o}{2}\right) - U'(\frac{1-T'}{2})}{U'\left(\frac{r_1^o}{2}\right)} \equiv p_U^o$,

then $R_1^* > 0$ for any h_3 for $p_U < p_U^o$, proving the second part.

<Proposition 3>

i) $U'\big(r_1 + R_1(1 - \bar{p}_U)\big) < U'(\frac{r_3}{2})$ by (7) and (10) because $\frac{\bar{p}_U}{1-\bar{p}_U} < 1$ and $U'\left(\frac{r_3}{2}\right) > U'(\frac{1-R_1-T'}{2})$ (from (10)), implying that $r_1^* + R_1^*(1 - \bar{p}_U) > \frac{r_3^*}{2}$.

ii) Suppose that $R_1 = R_2 = 0$. Then, (6) and (8) implies that, since $q \geq p_N$ and $\bar{e}_1 = V_N - V_U < V_{NN} - V_{NU} = \bar{e}_2$, $\frac{1}{1-\bar{q}} r_1 < \frac{1}{(2-\bar{p}_N)}(\frac{\bar{q}}{1-\bar{q}} r_1 + 2r_2)$, which in turn suggests that $r_1^* < r_2^*$. Now suppose that $R_1, R_2 > 0$. When $p_U \cong 0$, $r_1 + R_1(1-\bar{p}_U) \cong \frac{1+r_1}{3}$ by (7). while $r_2 + R_2 = \frac{s_1^*+r_2}{2}$. Suppose $r_1^* > r_2^*$. Since $s_1^* < \frac{2}{3}$, $r_1 + R_1(1-\bar{p}_U) > r_2 + R_2$, which contradicts to (6) and (8). Suppose that $r_1^* < r_2^*$. Since $R_1 = \frac{1-2r_1}{3}$ and $R_2 = \frac{s_1-r_2}{2}$, and since $s_1 < \frac{2-r_2}{3}$, $R_1^* > R_2^*$. Because $s_1^* \geq r_2$ (by (9)) and $s_1^* \leq \frac{1}{3}$ when $p_N \cong 0$, $r_1^* + R_1^*(1-\bar{p}_U) > r_2^* + R_2^*$

<Proposition 4>

Knowing from (9) that $r_2^* + R_2^* = \frac{r_2^*+s_1^*}{2}$, we can rewrite (8) as

$$U'\left(\frac{r_2^*+s_1^*}{2}\right)\{(2-\bar{p}_N) - h_2 \frac{1}{(2-\bar{p}_N)}\left(\frac{\bar{q}}{1-\bar{q}} r_1^* + 2r_2^*\right)\}M(T) - M(T)\left\{1 + h_2 \frac{1}{(2-\bar{p}_N)}\left(\frac{\bar{q}}{1-\bar{q}} r_1^* + 2r_2^*\right)\frac{\bar{p}_N}{(2-\bar{p}_N)} M(T)\right\} = 0 \quad (8')$$

where $M(T) \equiv U'(1-s_1^*-T) + (1-\bar{p}_N)U'\left(\frac{1+s_1^*-T}{2}\right)$. Note that $(2-\bar{p}_N) - h_2 \frac{1}{(2-\bar{p}_N)}\left(\frac{\bar{q}}{1-\bar{q}} r_1^* + 2r_2^*\right) > 0$ (from (8)), and $\frac{\partial \bar{p}_N}{\partial s_1} > 0$, and that $\frac{\partial M(T)}{\partial s_1} > 0$ (because U'''=0). Differentiating (8') with respect to r_2^* and s_1^*, we have $\frac{\partial r_2^*}{\partial s_1} < 0$. When $s_1 = s_1^o$, the choice of r_2 by the government, r_2^o, will satisfy

$U'\left(\frac{r_2^o+s_1^o}{2}\right)\{(2-\bar{p}_N)$ - $h_2\frac{1}{(2-\bar{p}_N)}\left(\frac{\bar{q}}{1-\bar{q}}r_1^*+2r_2^o\right)\}M(T)-M(T)\Big\{1+h_2\frac{1}{(2-\bar{p}_N)}\left(\frac{\bar{q}}{1-\bar{q}}r_1^*+2r_2^o\right)\frac{\bar{p}_N}{(2-\bar{p}_N)}M(T)\Big\}$ $-h_2\frac{1}{(2-\bar{p}_N)}\left(\frac{\bar{q}}{1-\bar{q}}r_1^*+2r_2^o\right)\}$

$\Big\{U'\left(\frac{r_2^o+s_1^o}{2}\right)-U'\left(\frac{1+s_1^o-T}{2}\right)\Big\}M(T)\left(\frac{\partial s_1}{\partial r_2}\right)=0$, where —$\left|\frac{\partial s_1}{\partial r_2}\right|<\frac{\bar{p}_N}{3}$ (by differentiating (5') with respect to s_1 and r_2 (taking as given T)). Thus, if p_N is small enough, $r_2^o<r_2^*$ since $\frac{\partial r_2^*}{\partial s_1}<0$. This implies by (9) that $R_2^o>R_2^*$.

<Proposition 5>

If $p_N=0,\ 1-s_1-T=\frac{1+s_1-T}{2}=\frac{2(1-T)}{3}$, so that the condition (6) can be rewritten as

$$U'\big(r_1+R_1(1-\bar{p}_U)\big)-U'\left(\frac{2(1-T)}{3}\right)\qquad-\frac{h_1}{1-\bar{q}}r_1U'\big(r_1+R_1(1-\bar{p}_U)\big)U'\left(\frac{2(1-T)}{3}\right)=0 \qquad (6')$$

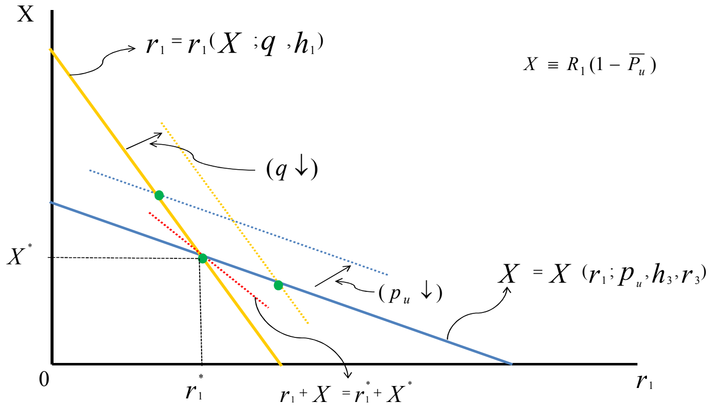
Also, let $X\equiv R_1(1-\bar{p}_U)$. We can then rewrite (7) as

$$U'(r_1+X)-U'\left(\frac{1-\frac{X}{1-\bar{p}_U}-T'}{2}\right)-\frac{h_3\bar{p}_U}{1-\bar{p}_U}U'\left(\frac{1-\frac{X}{1-\bar{p}_U}-T'}{2}\right)\{\frac{X}{1-\bar{p}_U}U'(r_1+X)+\frac{1}{1-\bar{p}_U}r_3U'\left(\frac{1-\frac{X}{1-\bar{p}_U}-T'}{2}\right)\}=0 \qquad (7')$$

Let $r_1^*=r(X;q,h_1)$ or $X={r_1}^{-1}(r_1^*;q,h_1)$ from (6'), and let $X^*=X(r_1;\ p_U,h_3,r_3)$ or $r_1=X^{-1}(X^*;\ p_U,h_3,r_3)$ from (7'). See Figure 2. We can see from (6') and (7') that $\frac{\partial r^{-1}}{\partial r_1}<\frac{\partial X}{\partial r_1}<0$, because

$\frac{\partial \bar{q}}{\partial r_1} > 0$, $\frac{\partial \bar{p}_U}{\partial r_1} = 0$ and $\frac{\partial \bar{q}}{\partial X} = 0$, $\frac{\partial \bar{p}_U}{\partial R_1} > 0$. Since $\frac{\partial X}{\partial R_1} > 0$ (from (7')), $\frac{\partial \bar{p}_U}{\partial X} > 0$. Also, from (6') and (7,), $r_1^{-1}(0; q, h_1) > X(0;\ p_U, h_3, r_3)$ because, for $p_N = 0$, $r_1^{-1}(0; q, h_1) = 1 - s_1 = \frac{2}{3} > \frac{1-T'}{3} > X(0;\ p_U, h_3, r_3)$. Suppose that $r_1(0; q, h_1) < X^{-1}(0;\ p_U, h_3, r_3)$ so that an interior solution (r_1^*, X^*) for (6') and (7') exists. Since $\frac{\partial r_1^{-1}}{\partial r_1} < \frac{\partial X}{\partial r_1} < 0$ and $r_1(0; q, h_1) < X^{-1}(0;\ p_U, h_3, r_3)$, the solution, which is unique, can be illustrated by the intersection A of $r_1(R_1; q, h_1)$ and $X(r_1;\ p_U, h_3, r_3)$ curves as in Figure 2. As q increases, $r_1(R_1; q, h_1)$ curve shifts leftward, so that r_1^* decreases while X^* increases. On the other hand, as p_U increases, $X(r_1;\ p_U, h_3, r_3)$ curve shifts leftward, so that r_1^* increases while X^* decreases. Note also from (6') and (7') that $\frac{\partial r_1^{-1}}{\partial r_1} < -1 < \frac{\partial X}{\partial r_1} < 0$, suggesting that $\frac{\partial (r_1^* + X^*)}{\partial q}$, $\frac{\partial (r_1^* + X^*)}{\partial p_U} < 0$.

| Figure 2 | Optimal Mix (r_1^*, R_1^*)

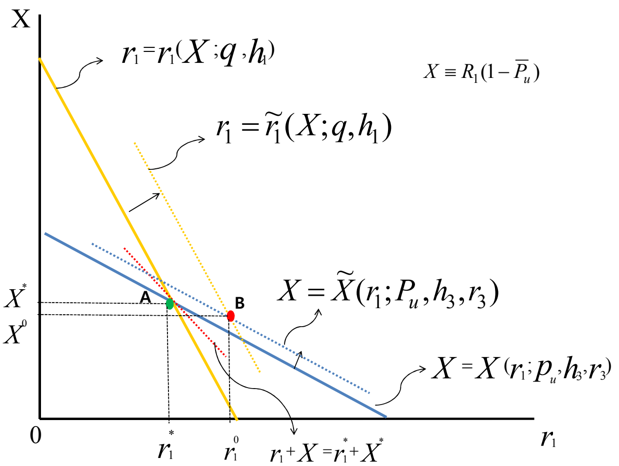


$X \equiv R_1(1 - \overline{p_u})$

<Proposition 6>

Figure 3 depicts (by A) how the original optimal mix $(r_1^*, R_1^*(1 - \overline{P}_U))$ is determined by (6)($r_1 = r_1(X,;q,h_3)$) and (7) ($X = X(r_1;p_U,h_3,r_3)$). Let (12) and (14) be represented by the curve $X = \widetilde{X}(r_1;p_U,h_3,r_3)$, and by the curve $r_1 = \tilde{r}_1(X,;q,h_3)$ in Figure 2, respectively. Comparing (14) and (12) with (6) and (7), respectively, we can see that the curve $r_1 = \tilde{r}_1(X,;q,h_3)$ and the curve $X = \widetilde{X}(r_1;p_U,h_3,r_3)$ are on the right hand side of the curve $r_1 = r_1(X,;q,h_3)$ and of the curve $X = X(r_1;p_U,h_3,r_3)$, respectively. Since $\frac{\partial {r_1}^{-1}}{\partial r_1} < -1 < \frac{\partial X}{\partial r_1} < 0$, we can see (from A and B in Figure 3) that $R_1^o(1 - \overline{P}_U) + r_1^o > R_1^*(1 - \overline{P}_U) + r_1^*$.

Figure 3 | Optimal Mix (r_1^0, R_1^0)



$X \equiv R_1(1 - \overline{p_u})$

CHAPTER 10

The Effects of Free Trade Agreements and Fair Trade for All

by
*Junsoo Lee**
(University of Alabama)

Abstract

The goal of this paper is to obtain more reliable quantitative estimates of the treatment effects of Free Trade Agreements (FTA) on trade flows. In particular, we are interested in examining whether the benefits of FTAs will be greater for rich countries than poor countries. In previous studies, FTAs were often treated as an exogenous variable, and a static model specification has been employed. However, as member countries choose to participate in an FTA, the endogeneity issue of FTAs must be addressed empirically. Furthermore, the persistence effect of trade flows can be modeled more effectively in a dynamic specification. We found that FTAs can increase bilateral trade flows by 174% ~ 220% in the long-run when we adopt a dynamic panel model with proper estimation methods while controlling for the contemporaneous endogeneity and the feedback effect. In addition, we found that the difference in the benefits of FTAs on trade flows between rich (bigger) and poor (smaller) countries is rather small or insignificant.

* Junsoo Lee, Professor, and Rick and Elaine Horsley Faculty Fellow, Department of Economics, Finance and Legal Studies, University of Alabama, Tuscaloosa, AL 35486, USA, Phone: 205-348-8978, Email: jlee@cba.ua.edu

1. Introduction

The recent trend of increased free trade agreements (FTAs) is expected to have a marked impact on the world economy. Through FTAs, the members are granted exclusive favoritism in trade policy. FTAs are easier and quicker to establish than multinational agreements and the number of nations involved in FTAs has been growing rapidly. As such, the issue on the effects of FTAs has attracted much attention in the literature.

One popular view of FTAs is that they are a means of increasing bilateral trade flows by removing tariffs or non-tariff barriers. It is often argued that FTAs can induce trade creation, implying that trade flows are redirected due to the formation of a free trade area. According to this view, FTAs can contribute to increasing productivity by inducing foreign investment and opening up markets, thereby further increasing trade flows. Comparative advantage in different phases of production can yield higher productivity for both partner countries. The cost of producing goods will be decreased, and FTAs can contribute to increase social welfare as lowered prices can increase real income in the signed countries. Trade flows in the partner countries will keep improving according to this view.

On the other hand, the opposite view is also possible. Trade flows can be diverted from a cost-efficient country to a less efficient country, which may produce goods at a higher cost than non-member countries. Also, it is possible that FTA will have a detrimental effect for the efficiency of certain countries. For example, removing trade barriers and inducing outside competition can yield a loss of jobs. If FTAs are not set up properly, they can decrease rather than uphold economic welfare. Opponents of FTAs argue that free trade may not necessarily induce poor countries to improve their standard of living. Instead, critics sometimes advocate the view that protectionism can help less developed countries. They tend to support the notion that rich countries have numerous advantages in negotiating trade agreements by disproportionately controlling the agenda of agreements. For example, FTAs can give more benefits to rich countries if the focus is given to removing barriers for high-value goods, service and investment and not farm products or low-tech goods. Then, the benefit of FTAs will flow mostly to rich countries but not poor countries and trade flows (exports) of poor countries will suffer in the long-run.

It is an empirical matter to see which of these two different viewpoints on the effects of FTAs on trade flows is supported from the data. Actually, previous empirical studies on the effect of FTAs on bilateral trade provided somewhat mixed results. The estimates of the effects of FTAs on trade flow can be either positive or negative, depending on the employed estimation method. For example, a recent and perhaps, the most influential work by Baier and Bergstrand (2007, *Journal of International Economics*) shows that the average treatment effect of FTAs can vary significantly if the effects of unobserved heterogeneity of bilateral factors are not controlled for. Although many other papers report that the estimates of the treatment effects of FTAs are rather overwhelmingly positive, empirical results are not yet clear. Baier and Bergstrand (2007) note potential bias of using cross-sectional models as they fail to control for the effects of unobserved heterogeneity. This key reference paper motivates the present paper. However, it seems clear that their estimation methods are also incomplete, and there are some important econometric issues to resolve, as we will discuss in more detail in the next section.

In this paper, we attempt to obtain more reliable quantitative estimates of the average effect of an FTA on bilateral trade. In particular, we note that Baier and Bergstrand (2007) and others have used a static model specification in a panel model setting, despite the existence of a persistent effect in trade flows. Moreover, it is clear that additional sources of endogeneity still remain. As such, we adopt a dynamic panel data model to revisit the issue, while we control for endogenous bias that is often ignored in previous studies. We are particularly interested in examining the issue of who benefits more from joining an FTA. Will rich countries benefit more than poor countries? Are there any significant differences in the benefits of FTAs between rich and poor countries? Is trade fair to all? Some authors, including Stiglitz (2006) and Jang (2010), argue that poor countries tend to suffer from FTAs while rich countries benefit greatly from FTAs. However, it remains to be seen if this argument holds empirically. To our surprise, this important issue has not been examined fully in the literature. The goal of this paper is to provide empirical evidence on this question.[1]

The rest of the paper is organized as follows. In section 2, we discuss various econometric issues in examining the treatment effects of FTAs on trade flows. Section 3 explains the estimation models, data, and econometric techniques. The estimation results are then provided and discussed. Section 4 concludes.

2. Econometric Issues

Here, we wish to discuss in more detail the various econometric issues to obtain reliable estimates of the treatment effects of FTAs on trade flows. Indeed, there are a few important econometric issues to clarify.

First, as pointed by Baier and Bergstrand (2007), trade policy is not an exogenous variable, although many previous studies in international trade typically assume that the dummy variable denoting a FTA is exogenous. Baier and Bergstrand (2007) show that the effect of FTA on trade flows would be underestimated considerably if endogeneity is not adjusted properly. The pivotal issue lies around the endogeneity bias caused by ignoring the unobserved bilateral fixed effects that exist between two trading partners. It is plausible to believe that countries likely select endogenously into FTAs for various reasons but this choice may not be explained fully by the observable factors. The choice can be correlated with the level of trade. Clearly, it is not a random assignment when a country chooses to seek an FTA with other countries.

1 For example, in his best-seller book, Jang, Ha-Joon (2010) argues that free-trade, free market policies are policies that have rarely, if ever, worked, and that these policies have slowed down growth and increased income inequality in the developing countries (Thing 7 out of 23 things about capitalism). However, Stiglitz and Charlton (2005, p. 35) explain that the relationship between trade liberalization and growth is much more controversial due to econometric difficulties. In this paper, we do not examine the effects of FTAs on growth or income inequality. Instead, we focus on bilateral trade flows and examine if the long-run treatment effects of FTAs on trade flows are significantly different.

Nonetheless, many of these studies treat the FTA as an exogenous variable. Baier and Bergstrand (2007) compellingly argue that trade policy is an endogenous variable because countries voluntarily choose to participate in an FTA. Thus, the so-called selection bias reflecting contemporaneous simultaneity should be handled carefully. The contemporaneous endogeneity issue can be handled by using instrumental variables estimation. These approaches require instrumental variables related to the selection of FTAs. There may be some criticisms of instruments as weak. However, it is still helpful to employ instruments to control for contemporaneous endogeneity rather than not using them at all. In this regard, we employ peer effects that lead to FTAs as instrumental variables. Specifically, the number of FTAs formed by each of the countries is expected to be correlated with FTAs; see Buther and Milner (2008). Regarding the exclusion restriction requirement, these variables can be less critical than the other possible instruments that denote institutional and political conditions.

Second, there is an issue of unobserved heterogeneity of trade flows. When the issue of selection bias can be understood as simultaneity between trade flows, it is reasonable to argue that FTAs and trade flows can be explained by common factors that affect trade policy of two member countries. Specifically, the reasons why member countries select into FTAs can be possibly correlated with unobserved factors that are also correlated with trade flows. In such cases, such common factors can be unobserved but can adversely affect the estimation results if not controlled for. As a matter of fact, we will show in the next section that when unobserved bilateral fixed effects are not controlled for, the estimates of the effects of FTAs on trade flows are fairly unstable. The effects of FTAs on trade flows can be positive or negative, and the magnitudes vary significantly over different time-periods and different model specifications. Thus, this finding illustrates the peril of using cross-sectional data in the analysis of the treatment effects of FTAs. In this regard, we note the importance of controlling for country specific fixed effects, bilateral unobserved fixed effects, time fixed effects and country-specific time effects. Indeed, time varying variables cannot fully explain the patterns of trade flows and unobserved factors could be important.

Third, it seems important to recognize that trade flows can be modeled more effectively as a dynamic process. Indeed, past trade activities can be associated strongly with current and future trade flows. Moreover, it might take much time to have the effects of FTAs realized in actual trade flows. When the relationship is dynamic, the usual static panel estimates can be inconsistent, depending on the length of the time period. In this regard, a dynamic specification might well be needed. In spite of this, the pioneering work by Baier and Bergstrand (2007) adopts a static specification, which can possibly yield biased estimates. Moreover, there are additional econometric issues regarding dynamic panel models. Recent developments in econometrics explain clearly the perils of using the usual panel data techniques in dynamic panel data models. In particular, the usual fixed effect estimates (also random effects estimates) can be biased and inconsistent.[2] In such cases, the literature often employs the GMM estimation method using the lagged dependent variable as an instrument as suggested by Arellano and Bond (1991). However, there is an additional issue of controlling for contemporaneous endogeneity as described above even with GMM estimation. Moreover, it is helpful to recognize the importance of controlling for both dynamic persistence effects and unobserved heterogeneity effects, since an estimator allowing for only one of these factors can be biased. For example, Angrist and Pischke (2009, p. 246) show that estimates of a positive treatment effect will tend to be large when a dynamic specification is proper but one mistakenly uses fixed effects. On the other hand, if fixed effects are present but one mistakenly estimates an equation with a lagged dependent variable without fixed effects, estimates of a positive treatment effect will tend to be too small. Thus, it seems important to control for both of these

2 One may consider using the difference-in-difference (DiD) estimates to overcome the endogeneity issue and obtain consistent estimates. Adopting the DiD procedure can control for all time invariant unobserved effects. However, one major difficulty lies in choosing the proper two time periods for pre- and post-treatment, which has to be of equal length for both the treated and control groups. In addition, it is necessary to assume the absence of autocorrelation in the data; see Egger et al. (2008). We do not adopt this procedure in our paper but leave it as a tool for future research.

effects at the same time. In this paper, we take care of this point in our estimation methods.

Fourth, another complication arises when there is a possibility of a feedback effect between trade flows and FTAs. The feedback effect occurs if current or past trade flows affect future FTAs, or current or future FTAs can be explained by past trade flows. This issue is related to the assumption of strict exogeneity, which may not hold in practice. Then, in addition to contemporaneous endogeneity, this other source of bias exists if FTAs are associated with past errors in the trade equation, thus creating a feedback effect. Obviously, the static version of the treatment effect model will not guarantee consistent estimates. In this paper, we examine the possibility of feedback effects in our analysis. We note in passing that Baier and Bergstrand (2007) acknowledge and test for a feedback effect, but find no significant evidence for such.

3. Estimation Methods and Empirical Results

In this paper, we attempt to address almost all of the econometric issues discussed in the previous section in order to examine our main empirical issues. As a main equation to explain the determinants of trade flows, we employ the gravity equation. Perhaps, the gravity equation is the most popular in the empirical literature on trade. Levi Yeyati et al. (2003) discuss the theoretical foundations of the gravity equation. The gravity model states that bilateral trade depends on the GDPs of the partner countries and on the distance between them. We also add some other variables including indicators for common languages, and common land boarders. Then, the dummy variable indicating the presence or absence of an FTA is added to the pair of two countries. One potential issue in applying the gravity model is the possible endogeneity of GDP. Baier and Bergstrand (2007) discuss reasons to ignore the potential endogeneity of incomes in the equation for trade, but this does not preclude potential endogeneity of GDP. We deal with this potential issue along with IV and GMM estimation in a dynamic model frame.

As noted previously, we employ panel data models using the following gravity equation:

$$y_{ijt} = c + \alpha_{ij} + \gamma_{+} + c_{it} + d\,FTA_{ijt} + \beta\, x_{ijt} + e_{ijt}, \tag{1}$$

where y_{ijt} denotes trade flows between a pair of two countries; i and j indicates the export and import country in a specific year, t; FTA_{ijt} is the dummy variable for a pair of two countries that have established a FTA at year t; and x_{ijt} represents the control variables in the gravity equation. Here, α_{ij} denotes the unobservable bilateral heterogeneity effect between two countries, and γ_t denotes time fixed effects. Unobserved heterogeneity may result from unobserved time invariant bilateral random variables that influence the presence of an FTA and the volume of trade. Following Baier and Bergstrand (2007), we also include country specific time effects, which are captured by c_{it}; see also Anderson and van Wincoop (2003) and Feenstra (2005). These terms are expected to capture the effects of multilateral price variables in a panel setting. Multilateral price variables are time varying and bilateral fixed effects may not control for these effects. In this standard gravity equation, the dependant variable is bilateral trade between country i and county j and is specified as exports from the ith country to the jth country in time period t. The control variables include GDP_i and GDP_j (the level of gross domestic product in country i and j, respectively), $DIST_{ij}$ (the distance between two trading countries), $LANG_{ij}$ (a binary variable with 1 representing if the two trading countries share a common language and 0 otherwise), and ADJ_{ij} (a binary variable with 1 meaning the two trading countries share a common boarder and 0 otherwise). We use the same panel data set of Baier and Bergstrand (2007) who provide the source data on the web site, http://people.clemson.edu/~sbaier/. We take logs of all variables that take positive values. In particular, for the dependent variable, we omit the observations of the pair of countries that do not have a record of positive trade volumes. We provide the description of these variables in Appendix Table 1 along with descriptive statistics.

To begin with, we first consider cross-sectional regressions for each of 1960, 1970, 1980, 1990 and 2000, using the gravity model in (1).

❙ Table 1 ❙ Estimation Results of Cross-sectional Regression

(a) Baseline Estimation Results

	1960	1970	1980	1990	2000
fta	0.614***	1.374***	-0.129	-0.142	0.294***
	(3.40)	(6.64)	(-0.76)	(-0.95)	(2.85)
ldist	-0.648***	-0.845***	-1.061***	-1.070***	-1.168***
	(-16.81)	(-21.10)	(-28.17)	(-28.82)	(-32.57)
adj	0.148	0.139	0.369**	0.589***	0.743***
	(1.00)	(0.85)	(2.35)	(3.72)	(4.88)
lang	0.051	0.341***	0.551***	0.796***	0.715***
	(0.57)	(3.48)	(5.80)	(8.16)	(7.71)
lgdpexp	0.755***	0.883***	1.011***	1.085***	1.180***
	(47.06)	(57.55)	(69.75)	(85.13)	(104.13)
lgdpimp	0.760***	0.917***	1.005***	0.967***	0.976***
	(50.16)	(63.95)	(73.92)	(78.08)	(87.39)
constant	-10.162***	-14.027***	-17.083***	-18.330***	-19.689***
	(-21.81)	(-30.16)	(-37.52)	(-43.33)	(-51.41)
N	2789	4030	5494	6474	7302
R^2	0.604	0.633	0.645	0.665	0.714

Dependent: log(trade flow); t-statistics in parentheses, * p<.10, ** p<.05, *** p<.01

(b) Additional Estimation Results

	1960	1970	1980	1990	2000
(lagged fta is added)					
fta		1.263***	-2.825***	-0.158	0.507***
		(3.13)	(-5.85)	(-0.60)	(3.51)
fta(-1)		0.120	2.754***	0.022	-0.376**
		(0.32)	(5.96)	(0.07)	(-2.10)
(interaction term fta*big is added)					
Fta	0.511**	1.298***	-0.132	-0.062	0.338**
	(2.09)	(4.69)	(-0.57)	(-0.31)	(2.52)
Fta_big	0.204	0.157	0.005	-0.161	-0.087
	(0.62)	(0.41)	(0.02)	(-0.63)	(-0.51)
(interaction term fta*rich is added)					
fta	0.495**	1.233***	-0.152	-0.201	0.143
	(2.00)	(4.36)	(-0.66)	(-1.02)	(1.07)
fta_rich	0.230	0.279	0.045	0.118	0.304*
	(0.70)	(0.73)	(0.15)	(0.46)	(1.79)

Note: Control variables were added but their coefficients are not shown here.
Note: Other results are also available when real variables or other specifications are used.

Table 1 provides the estimation results. In this regression, nominal variables are used, as in Baier and Bergstrand (2007), while we use logged real variables for the estimation results in Table 2 and others. It is obvious that the estimated coefficients of FTAs do not show a clear pattern. The coefficients are positive for the estimates for 1960, 1970 and 2000, but negative for the estimates for 1980 and 1990. These results are not very reliable since various econometric issues discussed in the previous section are not taken care of. In panel (b) of Table 1, we report the results of the model that use the lagged FTA as an additional regressor. We also examine the results with the dummy variables BIG and RICH and the interaction terms with FTA, where BIG (or RICH) take a value of 1 if real GDP (or per-capita real GDP) of the ith country is greater than that of the jth country. The results using these interaction dummy variables are also unstable in these cross-sectional regression models.

As the theoretical and empirical literature has noted, a proper estimation procedure should account for the problem of endogeneity of FTA and the fact that the relationship between FTA and trade flows is dynamic. Thus, we consider the estimation results using panel data techniques with bilateral fixed effects and country-specific time effects. Using the gravity model as our main empirical specification, we now use a balanced panel data set of the bilateral trade flows of 135 countries for the time-period of every 5 years from 1960 to 2000.

To capture the dynamic relationship between trade flows and FTAs, we also consider models where the lagged dependent variable, the lagged variables of FTA, and the forward variable of FTA are added as regressors along with the interaction terms of the two different dummy variables, BIG and RICH. Thus, our model specification is modified to

$$\begin{aligned} y_{ijt} = {} & c + \alpha_{ij} + \gamma_t + \rho y_{ijt-1} + d_0\mathrm{FTA}_{\mathrm{ijt}} + d_1\mathrm{FTA}_{\mathrm{ijt\text{-}1}} + d_2\mathrm{FTA}_{\mathrm{ijt\text{-}1}} + \\ & d_3\mathrm{FTA}_{\mathrm{ijt+1}} + f_0(\mathrm{FTA}_{\mathrm{ijt}}\mathrm{BIG}_{\mathrm{ijt}}) + f_1\left(\mathrm{FTA}_{\mathrm{ijt\text{-}1}}\mathrm{BIG}_{\mathrm{ijt\text{-}1}}\right) + \\ & f_2\left(\mathrm{FTA}_{\mathrm{ijt\text{-}2}}\mathrm{BIG}_{\mathrm{ijt\text{-}2}}\right) + f_3\left(\mathrm{FTA}_{\mathrm{ijt+1}}\mathrm{BIG}_{\mathrm{ijt+1}}\right) + \beta \mathrm{x}_{\mathrm{ijt}} + \mathrm{e}_{\mathrm{ijt}} \qquad (2) \end{aligned}$$

Then, the long-run effect of FTAs can be given as the sum of these dummy coefficients. However, given the persistent effect captured by the coefficient (ρ) of the lagged dependent variable, this sum needs to be divided by ($1-\rho$). Thus, it is given as $(d_0+d_1+d_2+d_3+f_0+f_1+f_2+f_3)/(1-\rho)$. The sum of these coefficients of the dummy variables excluding the interaction dummy variables denotes the long-run effects of FTAs on the smaller (poorer) countries when BIG = 0 (or RICH = 0), and the sum of the coefficients of the interaction dummy variables reflects the difference in the long-run effects of FTAs between the bigger (richer) and smaller (poorer) countries. These long-run effects are thus $(d_0+d_1+d_2+d_3)/(1-\rho)$, and $(f_0+f_1+f_2+f_3)/(1-\rho)$, respectively. For example, if the sum of the coefficients of the interaction dummy variables (fta_big, fta_big(-1), fta_big(-2) and fta_big(+1)) is negative, it implies that the long-run effects of FTAs for bigger countries (BIG = 1) are smaller than those of smaller countries (BIG = 0).

The results in the left panel in Table 2 show that the long-run effects of FTAs are actually lower as much as 17.3% (=100%($e^{-0.191}$-1)), 18.1% and 20.1% for the bigger countries than for the smaller countries. However, the differences are statistically insignificant. The results in the right panel in Table 2 show that the long-run effects of FTAs are actually higher by as much as 13.1% (=100%($e^{0.123}$-1)), 8.0% and 9.1% for the richer countries than those of the relatively poorer countries. However, these differences are also statistically insignificant. We note that the results of the regression without the lagged dependent variable are rather smaller. For these results, we tested the significance of a dynamic relationship by adding the lagged residuals and confirmed that a dynamic relationship should be considered. This analysis allows us to understand if there is a dynamic effect between the two variables. When included, the coefficient of the lagged dependent variable is highly significant, which implies that the static model is mis-specified.

As explained previously, the country specific time effects could reflect the effects of multilateral price variables as argued by Anderson and van Wincoop (2003) and Feenstra (2005). Thus, it can be argued that the treatment effects of FTAs can be smaller after controlling for the effects of multilateral price variables. Also, the elasticities of real

| Table 2 | Panel Estimation Results with Bilateral and Country-time FE

	With the dummy BIG			With the dummy RICH		
	Bilateral FE	Bilateral FE	Country-Time FE	Bilateral FE	Bilateral FE	Country-Time FE
lagged dep		0.140***	0.113***		0.141***	0.113***
		(21.20)	(16.69)		(21.23)	(16.73)
fta	0.776***	0.322***	0.403***	0.619***	0.269**	0.358***
	(11.82)	(2.89)	(3.62)	(9.71)	(2.41)	(3.21)
fta(-1)		0.278**	0.228*		0.255*	0.190
		(2.14)	(1.75)		(1.94)	(1.44)
fta(-2)		0.170	0.145		0.152	0.117
		(1.36)	(1.16)		(1.22)	(0.94)
fta(+1)		0.150*	-0.089		0.127	-0.116
		(1.80)	(-1.05)		(1.56)	(-1.40)
fta_big	-0.191**	-0.126	-0.115	0.123	-0.021	-0.027
(fta_rich)	(-2.12)	(-0.81)	(-0.77)	(1.45)	(-0.13)	(-0.18)
fta_big(-1)		-0.174	-0.190		-0.126	-0.112
(fta_rich(-1))		(-0.95)	(-1.09)		(-0.68)	(-0.64)
fta_big(-2)		0.005	-0.004		0.041	0.052
(fta_rich(-2))		(0.03)	(-0.02)		(0.25)	(0.32)
fta_big(+1)		0.124	0.109		0.172	0.165
(fta_rich(+1))		(1.07)	(0.98)		(1.56)	(1.56)
lrgdpexp	1.267***	0.952***	0.172*	1.268***	0.951***	0.172*
	(47.16)	(25.83)	(1.81)	(47.17)	(25.83)	(1.81)
lrgdpimp	1.223***	1.129***	0.073	1.222***	1.129***	0.073
	(41.60)	(25.79)	(0.69)	(41.58)	(25.79)	(0.69)
Time FE	Yes	Yes	No	Yes	Yes	No
Country-Time FE	No	No	Yes	No	No	Yes
N	47081	27575	27575	47081	27575	27575
R^2	0.227	0.179	0.262	0.227	0.179	0.262
Sum (all dummy coefficients)	0.585	0.871	0.549	0.742	1.012	0.707
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sum (no interaction terms)	0.776	1.070	0.775	0.619	0.935	0.619
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sum (interaction terms)	-0.191	-0.199	-0.225	0.123	0.077	0.088
	(0.034)	(0.374)	(0.297)	(0.146)	(0.643)	(0.574)

Dependent: log(real trade flow), t-statistics in parentheses; * p<.10, ** p<.05, *** p<.01

GDP in export and import countries become much smaller in these cases. However, we refrain from using country specific time effects from the additional estimation strategies below since the estimation often fails in these cases.[3] In addition, since there is the possibility of a

3 One may consider random effects estimators, but we need to satisfy the assumption of

feedback effect, we test for this introducing, according to Wooldridge (2010), the future levels of FTA. The coefficient of the forward variable of FTA is insignificant in the results on the left panel of Table 2, but significant in the results on the right panel, implying that strict exogeneity fails because the existence of future an FTA is correlated with the present flow of FDI.

We next consider IV estimation for the panel models in (2). We wish to correct for the bias caused by contemporaneous endogeneity of FTAs. To do so, we use our peer effects instruments, which are the number of FTAs signed by each of two countries at year t. It can be reasonably argued that these instruments are correlated with the tendency of seeking FTAs in each of the pair countries but are exogenous to trade flows. We refer to these results as FE-IV1. Another approach is a two-step IV estimation where we first estimate a probit model and use the predicted probability of having FTAs as an additional instrument for FTA.[4] We denote such results with FE-IV2. These results are shown in Table 3. The results of FE-IV1 and FE-IV2 are not much different, but the magnitudes of the sum of the coefficients are somewhat different. When the dummy variable BIG is used, we find that the sum of the interaction dummy variables is negative and significant, implying that the benefit of FTAs is greater for smaller countries. However, when the dummy variable RICH is used, we find that the sum of the interaction dummy variables is mixed and insignificant, implying that there is no difference in the benefit of FTAs for rich and poor countries.

Utilizing the panel dimension of our data is not sufficient to fix the problem of endogeneity. Although such results can resolve the bias caused by ignoring unobserved heterogeneity and contemporaneous endogeneity, they may not fully resolve the feedback effects as explained in the previous section. In particular, it is now well known that the estimates of dynamic panel data models with lagged dependent variables can be biased and inconsistent. Therefore, we suggest a further

---

no correlation between unobservable invariant bilateral factors and any regressors including FTA. This assumption does not seem realistic. Egger (2000) also provides strong empirical evidence against a random effects gravity model.

4 We report the estimation results of the probit model in Appendix Table 2. 13

| Table 3 | Panel IV Estimation Results

	No Interaction Dummy Variable		With the dummy BIG		With the dummy RICH	
	FE-IV1	FE-IV2	FE-IV1	FE-IV2	FE-IV1	FE-IV2
lagged	0.141***	0.141***	0.140***	0.140***	0.140***	0.141***
dep	(12.43)	(12.43)	(12.39)	(12.41)	(12.37)	(12.42)
fta	0.414	0.292***	0.992**	0.439***	1.615***	0.480***
	(1.62)	(4.50)	(1.98)	(3.82)	(3.28)	(4.30)
fta(-1)	0.107	0.173**	-0.087	0.214**	-0.492*	0.138*
	(0.73)	(2.53)	(-0.31)	(2.38)	(-1.73)	(1.67)
fta(-2)	0.181***	0.175***	0.200**	0.175**	0.200**	0.159**
	(2.80)	(2.76)	(2.31)	(2.13)	(2.51)	(2.17)
fta(+1)	0.166*	0.203***	-0.064	0.113*	-0.310*	0.059
	(1.93)	(4.89)	(-0.38)	(1.74)	(-1.83)	(0.93)
fta_big			-0.792	-0.242*	-1.362***	-0.231*
(fta_rich)			(-1.58)	(-1.91)	(-2.76)	(-1.89)
fta_big(-1)			0.191	-0.110	0.627**	-0.008
(fta_rich(-1))			(0.64)	(-0.83)	(2.05)	(-0.06)
fta_big(-2)			-0.017	0.001	0.018	0.037
(fta_rich(-2))			(-0.14)	(0.01)	(0.15)	(0.33)
fta_big(+1)			0.344*	0.162**	0.633***	0.244***
(fta_rich(+1))			(1.94)	(2.11)	(3.50)	(3.39)
lrgdpexp	0.954***	0.952***	0.956***	0.952***	0.959***	0.953***
	(20.84)	(20.90)	(20.85)	(20.90)	(20.92)	(20.91)
lrgdpimp	1.130***	1.128***	1.132***	1.129***	1.136***	1.130***
	(24.06)	(24.08)	(24.12)	(24.11)	(24.16)	(24.12)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	26791	26791	26791	26791	26791	26791
R^2	0.179	0.179	0.177	0.179	0.173	0.179
Sum (all dummy	0.868	0.843	0.874	0.874	1.080	1.022
coefficients)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sum (no			1.210	1.094	1.178	0.973
interaction terms)			(0.00)	(0.00)	(0.00)	(0.00)
Sum (interaction			-0.319	-0.220	-0.098	0.049
terms)			(0.028)	(0.052)	(0.372)	(0.527)

Dependent: log(real trade flow), t statistics in parentheses; * p<.10, ** p<.05, *** p<.01

step. For this, we first use IV estimation based on the first difference data to remove the bilateral fixed effects represented by α_{ij} in equation (2), and then use the second step of the lag of the dependent variable as an instrumental variable since this variable is uncorrelated with the error term. This approach was initially suggested by Anderson and Hsiao (1982). The results reported in Table 4 show that the sum of the coefficients of FTA dummy variables is generally greater than the

Table 4 Dynamic Panel Hsiao-Anderson and IV Estimation

	No Interaction Dummy Variable		With the dummy BIG		With the dummy RICH	
	AH-IV1	AH-IV2	AH-IV1	AH-IV2	AH-IV1	AH-IV2
lagged	-0.036***	-0.031***	-0.036***	-0.032***	-0.036***	-0.032***
dep	(-8.75)	(-7.42)	(-8.74)	(-7.51)	(-8.75)	(-7.64)
fta	0.332***	0.067	0.343***	-0.050	0.324***	0.133
	(11.30)	(0.43)	(8.70)	(-0.17)	(8.40)	(0.50)
fta(-1)	0.365***	0.372***	0.403***	0.442***	0.415***	0.419***
	(8.87)	(11.35)	(7.33)	(8.81)	(9.41)	(8.51)
fta(-2)	0.342***	0.291***	0.305***	0.287***	0.245***	0.237***
	(4.20)	(4.56)	(3.52)	(3.94)	(3.21)	(3.77)
fta(+1)	0.204***	0.197***	0.188***	0.202***	0.171***	0.191***
	(5.52)	(5.30)	(3.85)	(3.97)	(3.88)	(3.83)
fta_big			-0.022	0.327	0.016	0.166
(fta_rich)			(-0.44)	(1.21)	(0.32)	(0.69)
fta_big(-1)			-0.077	-0.139**	-0.097	-0.088
(fta_rich(-1))			(-1.03)	(-2.08)	(-1.30)	(-1.10)
fta_big(-2)			0.071	0.009	0.184	0.111
(fta_rich(-2))			(0.52)	(0.08)	(1.58)	(1.11)
fta_big(+1)			0.031	-0.007	0.065	0.019
(fta_rich(+1))			(0.47)	(-0.10)	(1.10)	(0.24)
lrgdpexp	0.963***	1.000***	0.963***	1.001***	0.962***	1.003***
	(15.13)	(15.05)	(15.13)	(15.04)	(15.12)	(15.09)
lrgdpimp	1.197***	1.294***	1.196***	1.295***	1.198***	1.297***
	(17.06)	(17.71)	(17.05)	(17.73)	(17.07)	(17.73)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	20902	17626	20902	17626	20902	17626
R^2	0.067	0.073	0.067	0.073	0.067	0.074
Sum (all dummy coefficients)	1.200 (0.00)	0.899 (0.00)	1.199 (0.00)	1.038 (0.00)	1.277 (0.00)	1.151 (0.00)
Sum (no interaction terms)			1.196 (0.00)	0.854 (0.00)	1.115 (0.00)	0.950 (0.00)
Sum (interaction terms)			0.003 (0.987)	0.184 (0.377)	0.162 (0.064)	0.202 (0.071)

Dependent: log(real trade flow), t-statistics in parentheses; * p<.10, ** p<.05, *** p<.01

results in Table 3, and that the sum of the coefficients of the interaction dummy variables is not significant in all the specifications. Thus, there is no significant difference in the long-run treatment effects of FTAs between bigger and smaller countries and between richer and relatively poor countries.

However, the estimation performed using the lags as instrumental variables is inefficient since it does not make use of all available orthogonality restrictions and it does not consider the structure of the disturbance. As such, we adopt the Arellano-Bond GMM estimation method, which can eliminate the time-invariant fixed effect by taking the first difference of equation (2). In addition, this method uses all possible lags of the independent variables to generate orthogonality restrictions. Following recent developments on the use of weak

| Table 5 | Dynamic Panel Results Using GMM Estimation

	No Interaction Dummy Variable		With the dummy BIG		With the dummy RICH	
	GMM	GMM-IV	GMM	GMM-IV	GMM	GMM-IV
lagged dep	0.291***	0.386***	0.290***	0.386***	0.291***	0.386***
	(9.61)	(14.03)	(9.60)	(14.03)	(9.61)	(14.01)
fta	0.232***	0.200***	0.253***	0.223***	0.228***	0.200***
	(7.49)	(6.41)	(5.86)	(5.08)	(5.45)	(4.73)
fta(-1)	0.209***	0.182***	0.229***	0.195***	0.254***	0.224***
	(4.60)	(3.97)	(3.53)	(2.95)	(5.52)	(4.84)
fta(-2)	0.139*	0.084	0.085	0.018	0.040	-0.013
	(1.65)	(0.99)	(0.81)	(0.17)	(0.52)	(-0.17)
fta(+1)	0.174***	0.154***	0.157***	0.125**	0.156***	0.130**
	(4.06)	(3.53)	(2.77)	(2.15)	(3.04)	(2.48)
fta_big			-0.043	-0.047	0.007	-0.001
(fta_rich)			(-0.74)	(-0.79)	(0.12)	(-0.01)
fta_big(-1)			-0.040	-0.025	-0.088	-0.083
(fta_rich(-1))			(-0.45)	(-0.27)	(-0.99)	(-0.91)
fta_big(-2)			0.106	0.128	0.192	0.189
(fta_rich(-2))			(0.72)	(0.87)	(1.54)	(1.50)
fta_big(+1)			0.034	0.057	0.035	0.050
(fta_rich(+1))			(0.44)	(0.73)	(0.51)	(0.71)
lrgdpexp	0.878***	0.702***	0.878***	0.702***	0.878***	0.702***
	(12.58)	(10.88)	(12.57)	(10.88)	(12.57)	(10.88)
lrgdpimp	0.955***	1.004***	0.955***	1.004***	0.955***	1.005***
	(12.49)	(13.24)	(12.48)	(13.23)	(12.49)	(13.24)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	20902	20902	20902	20902	20902	20902
Sum (all dummy coefficients)	1.063	1.010	1.100	1.098	1.162	1.134
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sum (no interaction terms)			1.020	0.914	0.956	0.881
			(0.00)	(0.00)	(0.00)	(0.00)
Sum (interaction terms)			0.080	0.184	0.206	0.252
			(0.702)	(0.441)	(0.096)	(0.080)

Dependent: log(real trade flow), t statistics in parentheses; * p<.10, ** p<.05, *** p<.01

instruments, we refrain from using all possible lagged instruments. Instead, we use three lags using the second, third and fourth lagged dependent variables as instruments for the endogenous regressor of the lagged dependent variable. In addition, we also use a modified GMM estimator where we use our peer effects instruments as an additional orthogonality condition in the GMM estimation. The motivation of this approach is to further control the effect of the contemporaneous selection effect of FTAs as in the usual IV estimations. These results are denoted as GMM-IV. We find that the results in Table 5 using GMM estimation are fairly close to the results in Table 4. Specifically, there is no significant difference in the long-run treatment effects of FTAs between bigger and smaller countries, and between richer and relatively poor countries. The magnitudes of the long-run treatment effects of FTAs range from 175% to 220%, when we employ the results in Table 5; these are given as $100\%(e^{1.01}-1) = 175\%$ and $100\%(e^{1.162}-1) = 220\%$. These results are much bigger than the results of Baier and Bergstrand (2007), who noted that after 10 years an FTA essentially doubles the level of members' international trade.

We wish to note that there are important differences between our approach and their approach. First, our results reflect long-run estimates of the treatment effects but the effects are not limited to the span of 10 years. Instead, it captures the eventual effects in the long-run. Second, their estimates might be still under-estimated. This is so since they used a static model, which can yield biased estimates. Third, our analysis permits us to examine the differences in two different groups of different size of the economy or different per-capita income levels. The results in the last panel of Table 5 show that there are differences in the benefits of FTAs and richer countries could benefit more from FTAs than relatively poor countries, although the differences are not significant at the 5% level but are significant at the 10% level. Fourth, in our results, all sources of endogeneity bias are controlled for, in addition to the effects of unobserved heterogeneity. These additional sources include contemporaneous endogeneity, the feedback effect and, more importantly, the persistent effect from a dynamic model specification.

We next conduct robust-checks using the sub-samples of the data where the dummy variables BIG and RICH are defined differently. We have also examined the cases where the sub-samples of BIG =1 and BIG = 0 are separately estimated, and compared the estimated coefficients. These results are not much different from the major findings in Tables 4 and 5 and are omitted here to conserve space.

Finally, we employ matching techniques where we estimate the probability of belonging to an FTA conditional on observable characteristics, and match the samples of the treatment group with the samples in the control group before we examine the differences of matched samples. From the probability model we have an estimate of the propensity score, a unit-free metric that we use to construct the control group. The control group includes non-FTA-member country pairs that have a probability of signing an FTA equal to the ones in the treated group that are the country pairs that already belong to an FTA. These results vary depending on how we choose two different time-periods and we chose not to report them.

4. Concluding Remarks

The goal of this paper is to determine how much free trade agreements (FTAs) affect members' international trade when correcting for various sources of bias in a dynamic specification. Various econometric models were used to compare the results for the relationship between FTAs and international trade. The treatment effects of Free Trade Agreements (FTAs) have been extensively examined in the empirical literature. However, FTAs have been often treated as exogenous and some important sources of endogeneity bias were not fully examined. In this paper, we revisit the issue of obtaining more reliable estimates of the treatment effects of FTAs. To do so, we adopt a dynamic model specification along with other measures of correcting for various endogeneity biases. In particular, we were interested in whether bigger or rich countries could benefit more from joining FTAs than smaller or poor countries. Our findings show that the evidence of greater benefits for rich countries is rather weak. Instead, we find that both signatory countries can benefit significantly from joining a FTA. Overall, the long-run treatment effects on increased trade from joining a

FTA range from 175% to 220% based on the estimation method using the most reliable estimation techniques. In addition, we found that the difference in benefits of joining FTAs on trade flows between rich (bigger) and poor (smaller) countries is rather small or insignificant. Instead, both countries benefit significantly from joining a FTA. Our results were robust to different testing procedures.

References

Anderson, J. E., and van Wincoop, E., 2003. "Gravity with gravitas: a solution to the border puzzle," *American Economic Review*, 93 (1), 170.192.

Anderson, T.W., and Cheng Hsiao, 1982, "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, 18, 47-82.

Angrist, J. D. and J. Pischke, 2009. Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press.

Arellano, B. and S. Bond, 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 277-297.

Baier S. L., J. H. Bergstrand, 2007. "Do Free Trade Agreements Actually Increase Members' International Trade?" *Journal of International Economics*, 71, 72-95.

Buthe, T. and H. Milner, 2008. "The Politics of Foreign Direct Investment Into Developing Countries: Increasing FDI through Policy Commitment via Trade Agreements and Investment Treaties", *American Journal of Political Science*, 52(4), 741-762

Egger, Peter, 2000. A note on the proper econometric specification of the gravity equation. *Economics Letters*, 66, 25.31.

Egger, H., P. Egger and D. Greenaway, 2008. "The Trade Structure Effects of Endogenous Regional Trade Agreements," *Journal of International Economics*, 74, 278-298.

Feenstra, Robert, 2004. Advanced International Trade: Theory and Evidence. Princeton University Press, Princeton, NJ.

Jang, H. 2010, 23 Things They Don't Tell You About Capitalism, Allen Lane Publishing Co.

Levi Yeyati, E., E. Stein, and C. Daude, 2003. "Regional Integration and the Location of FDI," Inter-American Development Bank (IAD), Working Paper.

Stiglitz, J. And A. Charlton, 2006, Fair Trade for All, Oxford University Press

Wooldridge, J.M., 2010. Econometric Analysis of Cross-Section and Panel Data. MIT Press, Cambridge, MA.

Appendix

Appendix Table 1 | Table 1. Definition of Variables and Descriptive Statistics

Name of variables	Description	# obs	Mean	Std.	Min	Max
rltf	logged real trade flow (export) from the exporting country to the importing country	70613	6.302	5.076	0.000	19.162
fta	free trade agreement, 1 = established between two countries at year t, 0 = otherwise	82080	0.023	0.151	0	1
big	1 if real GDP of the exporting country > real GDP of the importing country, 0 = otherwise	82080	0.498	0.500	0	1
rich	1 if per capita real GDP of the ith country > per capita real GDP of the jth country, 0 = otherwise	82080	0.498	0.500	0	1
lrgdpexp	logged real GDP of the exporting country	78375	16.844	2.175	11.707	22.919
lrgdpimp	logged real GDP of the exporting country	78375	16.844	2.175	11.707	22.919
lgdpexp	logged nominal GDP of the exporting country	78185	16.255	2.221	10.316	23.007
lgdpimp	logged nominal GDP of the importing country	78185	16.255	2.221	10.316	23.007
adj	1 if two countries are adjacent, 0 = otherwise	82080	0.025	0.156	0	1
lang	1 if two countries have common languages, 0 = otherwise	82080	0.059	0.236	0	1
ldist	logged distance between two countries	82080	8.259	0.756	4.841	9.424
fta_i	# of FTAs signed at year t of the exporting country	82080	2.231	4.491	0	26
fta_j	# of FTAs signed at year t of the importing country	82080	2.231	4.491	0	26

Appendix Table 2 | Probit Estimation Results

	Dependent Variable: FTA
fta_i	0.150***
	(32.92)
fta_j	0.146***
	(32.31)
ldist	-1.085***
	(-36.76)
adj	0.089
	(1.20)
lang	1.064***
	(17.38)
lrgdpexp	0.003
	(0.25)
lrgdpimp	0.022
	(1.57)
Time FE	Yes
N	36563

Note: t statistics in parentheses,
* p<.10, ** p<.05, *** p<.01

CHAPTER 11

Floating Exchange Rates and Macroeconomic Independence

by
*Yoonbai Kim**
(University of Kentucky)
Lian An
(University of North Florida)
Yu You
(University of Kentucky)

Abstract

In this paper, we investigate the nature of macroeconomic interdependence and the empirical validity of the insulation property of floating exchange rates in Japan vis-à-vis the U.S. We employ the method of cointegration and error correction modeling due to Johansen to test various hypotheses related to international transmission and movement of interest rates and goods prices. When monetary independence is measured as long-term freedom to manage the domestic interest rate, it seems to have increased during the more recent period of floating exchange rates. However, there is little evidence that capital controls before 1980 contributed to it in Japan.

* Corresponding Author. Department of Economics, University of Kentucky, Lexington, KY 40506. PH. (859) 257-2838, FAX (859) 323-1920, E-mail ykim01@uky.edu

1. Introduction

The concept of monetary independence (MI) under floating exchange rates is model-specific and depends on the relative size of the economy, the existence of nontraded goods, the substitutability of goods and assets, and other factors. Nonetheless, the notion that floating rates promote MI is prevalent among economists and policy makers. As the argument goes, the autonomy of monetary policy is enhanced under floating rates as the central bank is relieved of the obligation to maintain the fixed exchange rate. The economy is also better insulated from disturbances originating from abroad as necessary changes in the real exchange rate are made through changes in the nominal exchange rate instead of the prices of domestic goods. Thus, the breakup of the Bretton Woods system and the introduction of the modern float in the early 1970s have often been viewed as attempts to pursue monetary policies independently of external influences.

MI usually refers to independent management of interest rates and inflation rate without being forced to match foreign rates. The two parity conditions, purchasing power parity (PPP) and uncovered interest parity (UIP), provide a convenient explanation. In a world of perfectly substitutable goods and assets, domestic inflation and interest rate may not deviate from foreign counterparts under fixed exchange rates, leading to a complete loss of monetary control in a small economy. Exchange rate changes put a wedge between domestic and foreign inflation rates or interest rates and therefore deviations of domestic rates from foreign rates become feasible.

Whether floating exchange rates provide insulation from external shocks and/or enhance independent pursuit of monetary policy is an empirical issue. The vector autoregressive (VAR) model is a popular choice among researchers given the uncertainty regarding the 'right' model of international transmission. Of the early contributions using the VAR model, Genberg et al. (1987), Lastrapes and Koray (1990), and Hutchison and Walsh (1992) address the issue of international interdependence. All use variance decomposition as a measure of interdependence. Genberg et al. (1987) find strong influence of foreign (U.S. and Germany) on Swiss output, price, and interest rate. Their results suggest that, compared with the fixed rate period, the Swiss economy is hardly more independent from foreign disturbances during the floating rate period. Lastrapes and Koray (1990) find that the experience is more country-specific: the United Kingdom has been successful, in the short run, in using flexible exchange rates to isolate the domestic economy from external (U.S.) shocks while the German economy seems to become more dependent upon U.S. shocks. Hutchison and Walsh (1992) use a structural VAR model in which the small-open-economy assumption is used to identify structural shocks. In their variance decomposition, the portion of innovations in Japanese output that is explained by foreign shocks is much larger under floating rates than under fixed rates. They attribute the result to drastic changes in the nature of shocks affecting the economy instead of the change in the exchange rate system. In particular, the variance of domestic supply shocks seems to have become much smaller in the floating rate period. They rely on the impulse responses to investigate international transmission. According to their results, floating rates were successful in mitigating the effects of both external and domestic shocks on domestic output. Broda (2001) shows that, in the developing world, flexible regimes can insulate the economy better from real shocks. Edwards and Yeyati (2005) find that flexible rates function as shock absorbers.

Another strand of empirical study of MI is focused on the linkages between the domestic interest and the external (base) interest rate. Clarida, Gali, and Gertler (1999) show that while the US, Germany, and Japan follow their own shocks, even large countries participating in the EMS significantly curtail the extent that they follow their own shocks and instead follow the German interest rate. Kim (2001) reports that US interest rates have an impact on output in the other six G-7 countries.

Schambaugh (2004) show that floating countries appear to respond less, and with more delay, to movements in the base rate. He also argues that the gap between pegs and floats becomes clear when actual pegs and floats are carefully distinguished. According to di Giovanni and Schambaugh (2004), high foreign interest rates have a contractionary effect on the domestic economy and this effect is centered on countries with fixed exchange rates.[1] Similarly, Miniane and Rogers (2007) show that pegged open-financial-market countries follow base shocks more closely and Bluedorn and Bowler (2008) find that open capital-market pegs may respond even more strongly to exogenous shocks to the base interest rate. Borenstein, et al (2001) report that, in interest rate movement, Hong Kong reacts more than Singapore while Mexico and Argentine are somewhat similar.[2]

In contrast, Frankel (1999) shows evidence in which floats may follow the base interest rate even more than pegs while Frankel, Schmukler, and Serven (2004) find that results show limited autonomy for all but the largest floating countries. Along the same line, Forssback and Oxelheim (2006) find that in EU nations in the 1979-2000 period, there is no clear evidence of monetary policy constraint differing across exchange rate regimes. Bordo and MacDonald (1997) also find that countries on the gold standard had some flexibility in their monetary policy. Jensen (2008) shows that the Dutch had some limited autonomy despite pegging to Germany due to its tight target zone against the DM. A target zone can provide some autonomy.

Comparison of the fixed exchange rate period of the Bretton Woods system with the floating rate period is difficult because the two periods are different in more than just exchange rate arrangement. One of the most fundamental changes relevant for our study is increased capital mobility through reduction or elimination of the capital controls that had been put in place in the 1960s and 1970s. By mid 1980s, most industrial countries had eliminated restriction on capital inflow or outflow. As well-known from the Mundell-Fleming model, increased capital mobility further reduces the ability of the monetary authority to control money supply under fixed exchange rates. However, increased capital

---

1 Obstfeld, Schambaugh, and Taylor (2005) show that while the trilemma does not bind entirely, there is still a significant difference between the loss of monetary policy autonomy under pegs as compared to under floats.

2 Another important stylized fact in international finance is fear of floating: countries are unable to pursue an independent monetary policy due to factors like lack of credibility, exchange-rate pass-through, and foreign-currency liabilities. While formally or legally floating, they may peg their currencies and de facto “importing” the monetary policy of major-currency countries.

mobility also tends to magnify the reactions of the exchange rate to monetary or nonmonetary disturbances from home or abroad. Given the multitude of channels of international transmission occasioned by exchange rate changes, whether countries will be more independent under floating rates than under fixed rates is uncertain. Moreover, whether exchange rate changes absorb the burden of adjustment on the interest rate and prices is an empirical issue.

Most studies of comparative economic performance under different exchange rate regimes fail to take into account the changing degree of capital mobility and attribute differences in the behavior of output and prices to the change in the exchange rate system.[3] Given the fundamental importance of capital mobility in the determination of macroeconomic outcome, the approach seems unwarranted. The present study pays an explicit attention to the degree of capital mobility in a small model of international transmission in industrial countries. The purpose of this paper is to formulate a framework to evaluate the performance of floating exchange rates using recently developed methods of cointegration and error correction. The focus is on whether and how international transmission is different under different exchange rate systems and different levels of capital mobility. We consider movements of goods prices and interest rates to obtain a picture of interdependence between countries. With the model, we also investigate whether Japan has been able to achieve macroeconomic independence using systematic deviation of domestic interest rates from the foreign rate through capital restrictions. The next section illustrates the Johansen test of cointegration and error correction and how the test is applied in the present context. Empirical results are reported in Section 3. Section 4 concludes with some discussion.

3 In a recent contribution, Aizenman et al (2010) simultaneously investigate the three legs of trilemma – MI, exchange rate stability, and capital mobility.

2. Methodology

In this study, we develop a VAR model of two countries. Instead of giving an exclusive emphasis on the examination of variance decomposition and impulse responses, we employ the analysis of cointegration and error correction models and their implications. The methodology is due to Johansen (1988 and 1991) and Johansen and Juselius (1990). For a comprehensive discussion of the method, see Johansen (1995) and Juselius (2006). In the following, we describe the model and test procedures.

The VAR model consists of 5 variables: foreign and domestic prices and interest rates and the exchange rate. A 5-dimensional vector autoregressive model can be written as

$$X_t = A_1 X_{t-1} + \cdots + A_k X_{t-k} + \mu + \delta D_t + \varepsilon \qquad (1)$$

where $X' = (p, i, e, p^*, i^*)$. p, i , and e refer to price, interest rate and the exchange rate, respectively. ('*' denotes foreign variables.) D includes exogenous variables such as the change in real price of oil and seasonal dummies. Eq. (1) can be reformulated in the error-correction form:

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k-1} + \Pi X_{t-1} + \mu + \delta D_t + \varepsilon \qquad (2)$$

In this form, changes in the variables are related to past changes of themselves and lagged levels. The coefficients matrices Γ_i capture short-run interaction among variables while Π portrays long-run relationships. Variables are said to be cointegrated if they are individually nonstationary but some linear combination of them is stationary. In the Johansen method, the hypothesis of cointegration is formulated as a reduced rank of the Π-matrix:

$$\Pi = \alpha\beta' \qquad (3)$$

where α and β are $n \times r$ matrices of full rank, where n is the number of variables in the system and r is the rank of Π. If $0 < r < n$, variables are cointegrated and the r vectors of β are called cointegrating vectors. Cointegration implies that $\beta' X_t$ is stationary while X_t is nonstationary. $\beta' X_t$ represents equilibrium errors or magnitudes of disequilibrium. The r vectors of matrix α measure how the equilibrium errors are cleared by changes in the corresponding endogenous variables. Incorporating the decomposition of the long-run relationship, Eq. (3), in Eq. (2) and rewriting the latter, we obtain

$$\begin{aligned}\Delta p_t &= \gamma_{11}(L)\Delta p_{t-1} + \gamma_{12}(L)\Delta i_{t-1} + \gamma_{13}(L)\Delta e_{t-1} + \gamma_{14}(L)\Delta p^*_{t-1} + \gamma_{15}(L)\Delta i^*_{t-1} \\ &+ \alpha_{11}(\beta_1 X_{t-1}) + \cdots + \alpha_{1r}(\beta_r X_{t-1}) + \mu_1 + \delta_1 D_t + \varepsilon_{1t}\end{aligned} \tag{4a}$$

$$\begin{aligned}\Delta i_t &= \gamma_{21}(L)\Delta p_{t-1} + \gamma_{22}(L)\Delta i_{t-1} + \gamma_{23}(L)\Delta e_{t-1} + \gamma_{24}(L)\Delta p^*_{t-1} + \gamma_{25}(L)\Delta i^*_{t-1} \\ &+ \alpha_{21}(\beta_1 X_{t-1}) + \cdots + \alpha_{2r}(\beta_r X_{t-1}) + \mu_2 + \delta_2 D_t + \varepsilon_{2t}\end{aligned} \tag{4b}$$

$$\begin{aligned}\Delta e_t &= \gamma_{31}(L)\Delta p_{t-1} + \gamma_{32}(L)\Delta i_{t-1} + \gamma_{33}(L)\Delta e_{t-1} + \gamma_{34}(L)\Delta p^*_{t-1} + \gamma_{35}(L)\Delta i^*_{t-1} \\ &+ \alpha_{31}(\beta_1 X_{t-1}) + \cdots + \alpha_{3r}(\beta_r X_{t-1}) + \mu_3 + \delta_3 D_t + \varepsilon_{3t}\end{aligned} \tag{4c}$$

$$\begin{aligned}\Delta p^*_t &= \gamma_{41}(L)\Delta p_{t-1} + \gamma_{42}(L)\Delta i_{t-1} + \gamma_{43}(L)\Delta e_{t-1} + \gamma_{44}(L)\Delta p^*_{t-1} + \gamma_{45}(L)\Delta i^*_{t-1} \\ &+ \alpha_{41}(\beta_1 X_{t-1}) + \cdots + \alpha_{4r}(\beta_r X_{t-1}) + \mu_4 + \delta_4 D_t + \varepsilon_{4t}\end{aligned} \tag{4d}$$

$$\begin{aligned}\Delta i^*_t &= \gamma_{51}(L)\Delta p_{t-1} + \gamma_{52}(L)\Delta i_{t-1} + \gamma_{53}(L)\Delta e_{t-1} + \gamma_{54}(L)\Delta p^*_{t-1} + \gamma_{55}(L)\Delta i^*_{t-1} \\ &+ \alpha_{51}(\beta_1' X_{t-1}) + \cdots + \alpha_{5r}(\beta_r' X_{t-1}) + \mu_5 + \delta_5 D_t + \varepsilon_{5t}\end{aligned} \tag{4e}$$

where $\gamma_{ij}(L)$ a polynomial in the lag operator, L. Using the above setup, we can test various hypotheses regarding international transmission and interdependence using the concepts of cointegration and error correction.

2. 1. Long-run interdependence

A1. Cointegration of p and p^*: International transmission of inflation and comovement of prices implies that domestic and foreign prices are cointegrated. If a country is able to achieve independent inflation different from foreign inflation, prices will not be cointegrated.[4]

4 Noncointegration may occur for other reasons such as the existence of nontraded goods or productivity shocks that change relative prices.

A2. Cointegration of p, p^* and e: In the absence of real shocks, the real exchange rate would be stationary. This implies that (p, p^*, e) are cointegrated with cointegrating vector (1, -1, -1). In the case of fixed exchange rates, purchasing power parity (PPP) implies that (p, p^*) are cointegrated with cointegrating vector (1, -1). Due to the presence of nontraded goods, measurement errors and other reasons, we also test for PPP as a hypothesis that (p, p^*, e) are cointegrated with no restriction on the cointegrating vector. (Froot and Rogoff, 1995.)

A3. Cointegration of i and i^*: Advocates of floating exchange rates maintain independent management of domestic interest rates as a major advantage. Persistent deviation of the domestic interest rate from the foreign rate would imply noncointegration between them.

A4. Cointegration of i, i^* and e: In a world of uninhibited capital mobility and perfect substitutability of assets, uncovered interest parity (UIP) holds and (i, i^*, x) are cointegrated with cointegrating vector (1, -1, -1) where x is the expected rate of depreciation of the domestic currency. If the condition holds and x depends on the current level of the exchange rate (for instance, as in regressive expectations), the exchange rate serves to put a wedge between i and i^*. Other proxies for x are can be considered. Johansen and Juselius (1992), for instance, consider that the expected rate of depreciation is proportional to the relative price, $p - p^*$.

The role of flexible exchange rates as an instrument to increase macroeconomic independence would imply that domestic and foreign interest rates are cointegrated if the exchange rate is added in the relation but they are not if the exchange rate is omitted.[5]

A5. Implications of capital mobility

Restrictions on capital flows provide a wedge between the domestic and foreign interest rates with or without the exchange rate. Thus, they are likely to be non-cointegrated. In this sense, capital controls provide an additional instrument for macroeconomic policy and enhance monetary independence.

---

5 Noncointegration of (i, i^*, e) may occur or UIP fails if domestic and foreign assets are imperfect substitutes even if capital mobility is perfect in the sense that desired portfolio can be achieved with no friction.

2. 2. The small open economy assumption

Most previous studies assume a small open economy (SOE) model. For instance, Genberg et al. (1988) assume that Switzerland is a small economy taking U.S. or German variables as exogenously given. Hutchison and Walsh (1992), in a model for the Japanese economy, assume that domestic shocks have no long-run effects on U.S. income. Given the status of the U.S. in the world economy, those assumptions seem harmless. However, neither paper provided statistical justification for the assumption.

The SOE assumption can be rephrased to mean Granger exogeneity of foreign variables. In our model, it means that the foreign variables can be explained by their own past values and not by any domestic variables including the exchange rate. Thus it involves two hypotheses.

B1. $\alpha = 0$ in the equations for foreign variables

In this case, error correction terms, $\beta' X_{t-1}$, are absent in the equations for foreign variables. Those variables are termed "weakly exogenous" to the parameters of long run relationships. If the hypothesis is not rejected, the cointegrating vectors can be obtained from the equations for the domestic variables only. This implies that adjustment to long-run equilibria will be made by domestic variables and not by foreign variables. Another important implication of the result is that common trends in domestic variables are in fact due to foreign shocks and permanent shocks are generated from abroad. Domestic shocks will be responsible for merely transitory deviations from such trends.

B2. Changes in domestic variables are absent in the equations for foreign variables

If the equations for the foreign variables do not contain domestic variables as well as the error correction terms, those foreign variables are not Granger-caused by domestic variables. In this case, foreign variables are not affected by domestic variables in the short run or in the long run. If only B2 is accepted and B1 is not, foreign variables are unaffected by domestic variables only in the long run while there can be short-run interactions between them. For the purpose of identifying international transmission and macroeconomic interdependence in the long run, we only test for weak-exogeneity of foreign variables.

2. 3. Insulation property of floating exchange rates

The coefficients α can be considered adjustment to long-run equilibrium. The insulation property of floating exchange rates implies that, under floating rates, the α's in the price and interest rate equations will be insignificant while those in the exchange rate equation will be significant as the burden of external adjustment is taken by exchange rate changes instead of changes in output or prices. For instance, independence of inflation in the presence of perfectly substitutable goods (i.e., PPP) would imply that $\alpha_{11} = \alpha_{21} = 0$ and nonzero (presumably positive) α_{31} for equilibrium error from the cointegrating vector $\beta_1' X_{t-1} = (p_{t-1} - p^*_{t-1} - e_{t-1})$. The latter is also called departure from PPP.[6]

3. Empirical Results

We employ the consumer price index (CPI) and short-term interest rate as the price index and interest rate. The exchange rate is period average. The monthly data of Japan covers from 1957:1 through 2110:12. The U.S. is taken as the foreign country. The model also includes seasonal dummies and the current and lagged changes in the price of oil relative to U.S. CPI as exogenous variables. All data are obtained from *International Financial Statistics* (CD-ROM). The lag length is set at 3, which produces white noise errors. Reducing it often creates anomalous errors while increasing it rapidly depletes the degrees of freedom.

We divide the data into 4 periods. The first period begins at 1957:1 and ends at 1971:7. This period is characterized by global fixed

6 The presence of the *changes* in foreign variables in the domestic variables equations suggests that domestic variables will still be affected by foreign variables even if the error correction terms are insignificant. Nonetheless, I focus on the error correction term to investigate the insulation property because the coefficients on the first difference terms capture short-run interactions that might reflect the influences of other factors such as omitted variables. Without error correction, variables will not show comovement.

exchange rates under the Bretton Woods system and widespread capital controls. The end of the period roughly corresponds to the time when the global exchange rate system was in disarray and many currencies including the Japanese yen started floating. For period I, we estimate the 4-variable model without the exchange rate. The second period starts at 1971:8 and ends at 1980:12. This is a period with the floating exchange rate against the US dollar. However, extensive capital controls were retained until the end of 1980. The separation of the data for the floating exchange rate period is arbitrary. The two periods are distinguishable in some respects, however. First, period II is characterized as one of numerous real shocks. To name important ones, there were a couple of major oil price shocks in the 1970s and divergent fiscal policies among the U.S., Japan, and Germany. Central bank intervention was more sporadic in period II. It is only in period III that industrial countries including the U.S. started actively intervening in the foreign exchange market in a concerted manner with some target zones in mind. (Funabashi 1988) Also, many industrial countries including Japan had embarked on a series of financial liberalization during period II. See Ito (1983) and Frankel (1984) for the Japanese experience. As a result, the degree of capital mobility throughout period III can be considered much higher than that of period II.

Post-Period II is characterized by floating exchange rates and the absence of capital controls. Given the imbalance of the sample size among the sub-periods, we divide the last period into two. Thus, Period III covers from 1981:1 through 1990:12. Period IV is from 1991:1 to the end of 2010. One of the most contrasting differences between the two periods would be the long-term recession after the bursting of equity and real estate bubbles in the early 1990s.

We present a VAR model of 5 variables. The variables are ordered (RF, PF, ER, RD, PD). In period I, ER is omitted. The foreign variables are placed ahead of the domestic variables given the dominant position of the U.S. economy especially in monetary/financial matters. For each set of domestic and foreign variables, the interest rate is ordered before the price level on the assumption that the former reflects the policy decision of the government as well as endogenous reactions to other

| Table 1 | Variance Decompositions

	ε_{RF}	ε_{PF}	ε_{ER}	ε_{RD}	ε_{PD}
Period I					
RF	97.5 / 77.6	0.3 / 3.8	na	2.0 / 14.6	0.2 / 4.0
PF	1.5 / 8.8	86.4 / 56.2	na	0.6 / 4.2	11.4 / 30.9
RD	18.5 / 23.2	0.2 / 2.9	na	78.7 / 48.9	2.7 / 25.0
PD	0.6 / 7.0	11.5 / 10.3	na	0.2 / 2.1	87.7 / 80.6
Period II					
RF	62.8 / 38.2	3.5 / 5.6	29.6 / 24.6	3.6 / 17.9	0.4 / 13.7
PF	41.4 / 12.3	45.5 / 34.9	7.2 / 47.4	1.5 / 3.6	4.4 / 1.8
ER	10.4 / 24.0	1.2 / 0.8	84.4 / 55.3	3.3 / 13.5	0.7 / 6.3
RD	49.8 / 31.8	2.2 / 5.5	3.0 / 18.6	43.0 / 26.7	2.0 / 17.4
PD	22.3 / 30.9	6.5 / 9.2	0.7 / 21.4	9.8 / 13.4	60.7 / 25.0
Period III					
RF	91.2 / 75.2	1.0 / 1.2	3.5 / 12.5	3.7 / 10.6	0.6 / 0.5
PF	37.3 / 24.2	54.9 / 25.3	6.7 / 44.7	0.4 / 3.3	0.8 / 2.5
ER	11.5 / 26.4	0.5 / 1.0	85.4 / 55.4	1.4 / 14.8	1.3 / 2.4
RD	16.3 / 36.6	0.8 / 2.6	1.2 / 10.9	80.4 / 49.5	1.3 / 0.4
PD	11.1 / 34.2	23.4 / 16.9	6.2 / 26.8	3.7 / 11.9	55.6 / 10.2
Period IV					
RF	91.3 / 57.5	0.3 / 2.4	7.0 / 28.7	1.3 / 4.7	0.1 / 6.7
PF	1.1 / 6.6	87.4 / 62.9	0.7 / 8.5	0.5 / 1.2	10.5 / 20.8
ER	2.4 / 6.0	0.6 / 2.1	91.7 / 77.7	0.3 / 5.8	5.1 / 8.4
RD	21.4 / 24.2	0.1 / 3.4	2.0 / 3.3	72.7 / 45.3	3.7 / 23.8
PD	0.8 / 8.2	12.4 / 10.8	1.2 / 3.1	0.5 / 4.7	85.1 / 73.2

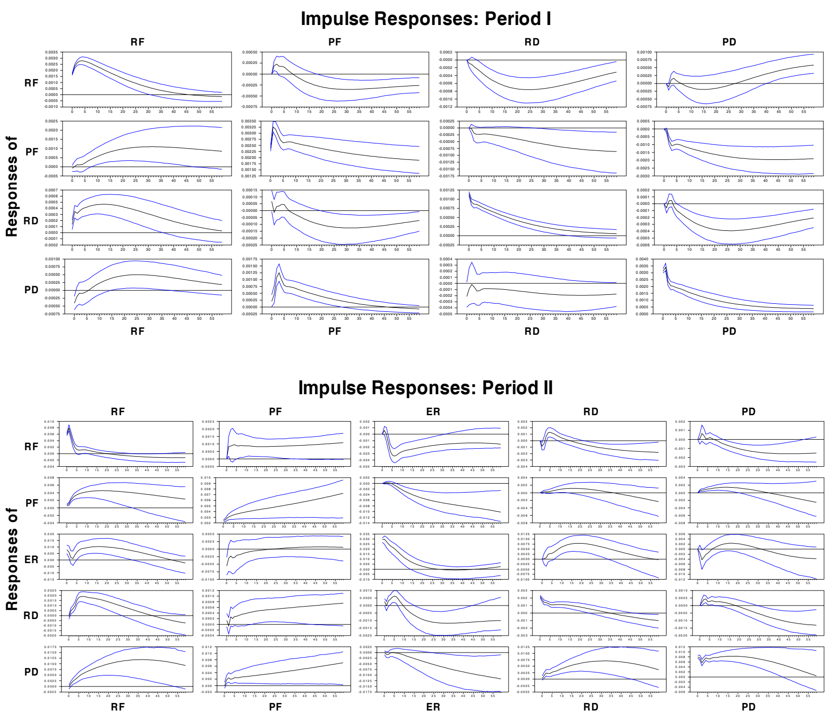
Note: For each cell, the first (second) entry is variance decomposition at 12 (60)-month horizon.

variables. Finally, the exchange rate is placed before the other domestic variables.

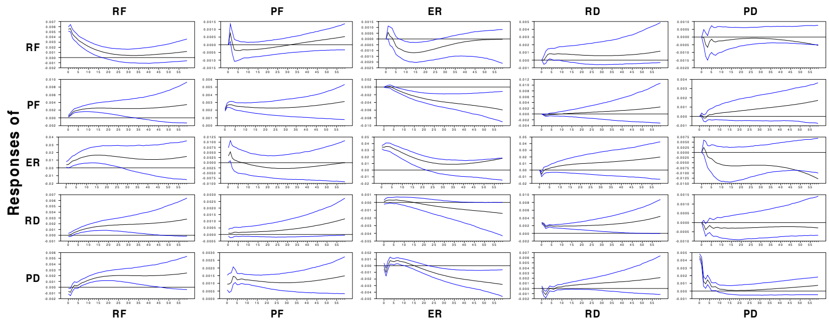
Table 1 presents variance decompositions.12 and 60-month forecast horizons are chosen as short-run and long-run effects. Our main focus is the determination of the domestic interest rate, especially the influence of the foreign interest rate. The role of foreign interest rate shocks seems fairly even throughout the whole period except period II, explaining about 20 percent in the short run and rising to nearly 30 percent in the medium run. Interestingly, it is highest during period II. Alternatively, the domestic interest rate is mainly explained by own shocks, reflecting a fair amount of insulation from foreign influence in interest-rate setting. Here again, the role of own shocks is lowest in period II. As the exchange rate was floating with capital controls still in place, period II is supposed to show highest monetary independence.

Figure 1 shows impulse responses. We focus on the mutual interaction between the two interest rates. The panels in the first column show the responses to an increase in RF. It affects the domestic interest rate immediately and induces highly correlated movements between the two rates. The domestic interest rate appears to be more persistent than

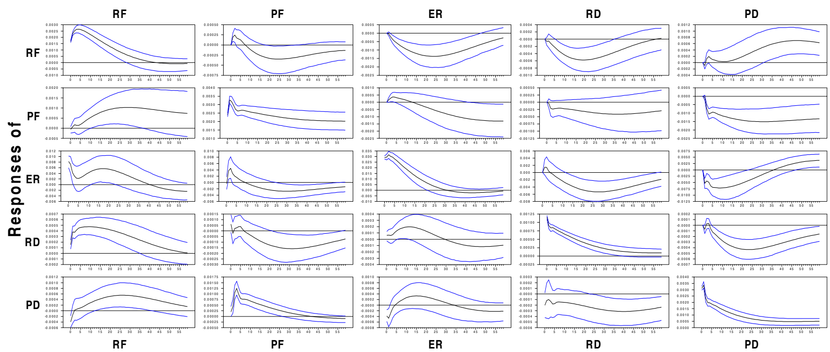
Figure 1 Impulse Responses



Impulse Responses: Period III



Impulse Responses: Period IV



the original foreign interest rate shock. Period III is the only exception, in which the domestic interest rate moves much more gradually and persistently over time.

In the case of shocks to the domestic interest rate, the foreign rate moves little as in periods II and III or declines significantly as in period I and IV, generating a negative correlation between the two rates. It suggests that for the U.S., Japan is still a small country vis-à-vis the U.S. while its interest rate movement may induce offsetting change by other countries represented by the U.S. In response to an increase in the foreign interest rate, the domestic currency depreciates in all three periods of floating exchange rates. On the other hand, a domestic interest rate shock depreciates the yen as in periods II and III or appreciates it as in period IV.

| Table 2 | Unit-root Tests

Unit root tests for levels						
Variable	Method	All periods	I	II	III	IV
ER	DF	-2.47	-2.67	-2.32	-1.62	-1.49
	PP	-2.63	-3.19	-2.05	-2.28	-1.87
	KPSS	0.67**	0.36**	0.17*	0.21*	0.27**
RD	DF	-3.33+	-2.63	-2.81	0.39	-4.59**
	PP	-2.65	-1.88	-1.54	0.23	-3.44*
	KPSS	0.92**	0.24**	0.18*	0.29**	0.66**
RF	DF	-1.79	-3.39*	-0.99	-2.48	-2.9
	PP	-2.17	-3.08	-1.5	-2.12	-1.74
	KPSS	1.51**	0.19*	0.25**	0.32**	0.19*
PD	DF	0.51	-3.85*	-1.18	-1.31	-2.81
	PP	1.24	-2.93	-0.54	-2.49	-2.95
	KPSS	2.26**	0.3**	0.39**	0.25**	0.55**
PF	DF	-0.27	0.86	-1.69	-1.28	-2.52
	PP	0.47	2.15	-1.11	-2.96	-2.87
	KPSS	1.51**	0.59**	0.19*	0.2*	0.11

Unit root tests for differenced data						
ER	DF	-10.28**	-6.12**	-3.41+	-3.88*	-7.1**
	PP	-23.88**	-14.23**	-9.97**	-10.19**	-14.65**
	KPSS	0.05	NA	0.08	0.16*	0.09
RD	DF	-6.65**	-3.99*	-2.42	-3.43+	-4.47**
	PP	-25.55**	-13.03**	-10.6**	-11.37**	-17.31**
	KPSS	0.05	NA	0.11	0.17*	0.2*
RF	DF	-11.5**	-5.95**	-5.59**	-5.5**	-3.3+
	PP	-18.2**	-7.91**	-6.64**	-9.96**	-9.39**
	KPSS	0.02	NA	0.07	0.05	0.1
PD	DF	-6.16**	-5.91**	-2.93	-3.97*	-5.44**
	PP	-22.22**	-14.44**	-8.83**	-11.38**	-13.27**
	KPSS	0.46**	NA	0.13+	0.19*	0.11
PF	DF	-5.19**	-4.02**	-2.33	-4.15**	-7.26**
	PP	-12.95**	-11.68**	-6.89**	-5.77**	-9.05**
	KPSS	0.93**	NA	0.14+	0.23**	0.04

Note: '+', '*', and "**' denote significance at the 10, 5 and 1% level, respectively. All tests choose 6 lags with interc ept and trend. DF, PP, and KPSS stand for (augmented) Dickey-Fuller, Phillips-Perron and Kwiatkowski-Phillips-Schmidt-Shin (1922) tests.

In principle, Japan has had the ability to retain monetary independence in all 4 periods. Trilemma suggests that capital controls during periods I and II or floating exchange rates during periods II, III, and IV allow the monetary authority to control its own interest rate independently of external influence. The results do not support such prediction. In fact, both impulse responses and variance decomposition show that the domestic interest is most strongly affected by the foreign interest during period II. Nonetheless, Japan seems to have retained MI to a substantial degree, especially in the more recent period of floating rates after 1980.

Table 2 presents the unit root tests. We employ 3 popular tests: Augmented Dickey-Fuller (DF), Phillips-Perron (PP), and Kwiatkowski-Phillips-Schmidt-Shin (KPSS). Unit-root tests are applied both for the levels and the first differences. Note that the null hypothesis posits the series as a nonstationary (unit-root) process under the DF and PP tests and stationary under the KPSS test. In levels, DF and PP fail to reject the null hypothesis of nonstationarity while KPSS strongly rejects the null of stationarity in most series for most periods although the three tests do not always agree as in the case of RD in period IV. According to the bottom half of the table, first differencing renders most series to become stationary.

Choosing the number of cointegration vectors is tricky and subject to discretion. We use the trace statistic and the maximum eigenvalue statistic to find the approximate rank. To determine the cointegration rank, we also examine other properties of the time series such as stationarity of variables, whether a variable is absent in all cointegrating vectors (long-run exclusion), and whether a variable does not respond to equilibrium errors (weak exogeneity).[7] Table 3 summarizes the trace and the maximum eigenvalue statistics. Table 4 reports the test results of each of the three restrictions at the chosen number of cointegrating vectors. For the sake of convenience, we report just the result of the test

7 Whether a variable is stationary or nonstationary is tested in the multivariate context. In the Johansen procedure, the inclusion of a stationary variable increases the rank of cointegration by one. Tests of cointegration vectors can then be used to determine whether there is a cointegrating vector that consists of the variable only, which establishes the stationarity of the variable.

at the 5 percent significance level. These tests on the time series properties can be exploited to formulate cointegrating vectors as shown in next 2 tables.

| Table 3 | The Johansen Tests of Cointegration

		L-max	Trace	Trace[a]
	r			
Period I				
	0	47.8*	90.0*	73.73*
	1	24.8*	42.2*	na
	2	8.9	17.4	na
	3	8.5	8.5	na
Period II				
	0	57.3*	153.4*	86.96*
	1	42.8*	96.1*	35.61*
	2	39.0*	53.3*	12.43*
	3	8.7	14.3	10.95
	4	5.6	5.6	na
Period III				
	0	53.0*	100.6*	79.7*
	1	23.0	47.6	17.0
	2	12.5	24.6	9.8
	3	8.1	12.1	1.2
	4	4.0	4.0	na
Period IV				
	0	70.7*	145.6*	134.0*
	1	35.9*	74.9*	37.7
	2	22.0*	39.0*	9.8
	3	11.8	17.0	3.9
	4	5.2	5.2	0.8

Note: L-max and Trace are the maximum eigenvalue statistic and the trace statistics, respectively. (Trace[a] is the trace statistic with the Bartlett corrections.) The former is the likelihood ratio test for the hypothesis that there are r cointegrating vectors as opposed to $r+1$. The latter is the likelihood ratio test of there being at most r cointegrating vectors in a set of n variables. The critical values of the maximum eigenvalue statistic at the 5-percent significance level, for r = 0, 1, …, are 33.46, 27.07, 20.97, 14.07, and 3.76. The corresponding critical values for the trace statistic are 76.81, 53.94, 35.07, 20.16, and 9.14. In a system of 4 variables, the relevant critical values begin with the second number in each statistic. (Osterwald-Lenum, 1992) '*' denotes significance at the 5 percent level.

❙ Table 4 ❙ Long-Run Exclusion, Stationarity and Weak Exogeneity

	*p**	*p*	*i**	*i*	*e*
Period I	*r*=2				
Long-run exclusion	o	o	x	+	
Stationarity	o	o	o	x	
Weak exogeneity	x	x	o	x	
Period II	*r*=3				
Long-run exclusion	x	x	x	x	x
Stationarity	+	x	+	x	x
Weak exogeneity	x	x	x	x	o
Period III	*r*=1				
Long-run exclusion	x	x	o	x	x
Stationarity	x	x	o	x	x
Weak exogeneity	x	o	o	o	o
Period IV	r=3				
Long-run exclusion	x	o	o	x	x
Stationarity	o	o	o	o	o
Weak exogeneity	x	x	x	x	x

Note: Long-run exclusion tests the hypothesis that a variable is absent in all cointegrating vectors. Stationarity tests whether a variable is stationary by itself. Weak exogeneity tests the hypothesis that all error correction terms are jointly insignificant in the equation. All three are distributed as χ^2 with varying degrees of freedom. Reported are the results of the tests of hypotheses at the 5 percent significance level. "o," "+" and "x" denote that the hypothesis is 'not rejected (at 10%),' 'rejected (at 10%)' and 'rejected (at 5%),' respectively.

| Table 5 | Tests of Hypotheses

	I	II	III	IV
Number of cointegrating vectors	2	3	1	3
Hypothesis				
A1. $p\text{-}p^*$	0.17	0.01	0.19	0.01
A2. $p\text{-}a_1p^*$	0.08	0.01	0.49	0.42
A3. $p\text{-}p^*\text{-}e$	na	0.00	0.45	0.88
A4. $p\text{-}a_1p^*\text{-}a_2e$	na	0.15	0.29	0.46
A5. $i\text{-}i^*$	0.00	0.06	0.51	0.01
A6. $i\text{-}a_1i^*$	0.72	0.05	0.24	0.89
A7. $i\text{-}i^*\text{-}a_1(p\text{-}p^*)$	0.09	0.46	0.32	0.18
A8. $e\text{-}a_1(i\text{-}i^*)$	na	0.15	0.65	0.62
A9. $p\text{-}a_1i$	0.03	0.03	0.45	0.70
A10. $p^*\text{-}a_1i^*$	0.65	0.10	0.67	0.89

When there are more than one cointegrating vectors, only the cointegrating space is determined while the set of independent cointegrating vectors comprising the space is indeterminate. In choosing the model, we test for various combinations of variables as suggested by hypotheses of long-run independence, the small-open-economy assumption or the insulation property of floating exchange rates. The significance levels (p-value) of the restrictions are reported in Table 5. In the table, group A tests the stationarity of various combinations of variables.[8] It includes, in hypotheses A1 through A8, comovement of foreign and domestic prices and interest rates with or without the exchange rate. The hypothesis A9 (A10) is the combination of the domestic (foreign) variables, which may be taken as the money market condition in the home (foreign) country.[9]

8 Note that these tests, as in typical unit root tests, have nonstationarity as the null hypothesis. Thus the rejection of the null establishes stationarity.

9 The conventional money market equilibrium condition, $m - p = -a_1 i + a_2 y + \varepsilon$, where ε denotes monetary shocks, can be rewritten as $p - a_1 i = m + a_2 y + \varepsilon$. Thus the linear combination of the price and the interest rate represents shocks to the money

As explained above, if the hypothesis that all error correction terms are absent in a particular variable, the variable is termed weakly exogenous. More importantly, it is responsible for one of stochastic trends in the data. The hypothesis of weak exogeneity is of interest for the test of the SOE assumption. Also, independent movements of the price or the interest rate due to capital restrictions or other reasons imply weak exogeneity.

At the 5 percent significance level, the L-max and trace tests agree with the number of cointegrating vectors, except period IV. However, the trace statistics with the Bartlett corrections often suggest different (smaller) ranks. We determine that there are 2, 3, 2, and 2 cointegration vectors in the 4 periods.

In period I, there seem two cointegrating vectors. RF appears weakly exogenous and the cointegration relationships are mainly about interest rates. (Table 4) Among other things, the null hypothesis that the interest rate differential is nonstationary is strongly rejected. This suggests that the foreign and domestic interest rates moved very closely in the long run. In period II, we find 3 cointegration vectors. The small-sample test with Bartlett corrections also agrees with it. No variable is excluded in the long-run relationships. Interestingly, the exchange rate is the only one that appears weakly exogenous. This suggests that, during this period, the exchange rate did not move in the directions to remove disequilibrium errors in goods or financial markets. Table 5 shows strong evidence of goods market integration as well as financial market integration. Both PPP and IRP hold with strong statistical significance. As in period I, the null hypothesis that the interest rate differential is nonstationary is strongly rejected. The results indicate that neither the float nor the capital controls allowed independent movement of the domestic interest rate in Japan during this period.

In period III, all three tests point to one cointegration vector. All variables other than the foreign price level appear to be weakly exogenous. This suggests that most variables may not be responding to long-run adjustment. Moreover, the foreign interest rate may not be in

market equilibrium condition that have to be cleared by changes in the price or interest rate.

the long-run relationship with any other variables. (Table 4) Indeed, as Table 5 indicates, none of the hypothesis is significantly rejected which implies that the interest rates, the price levels, and the exchange rate are not related in a manner that can be identified as predicted in theory. According to the definition of monetary independence, period III seems to have the highest degree of independence.

In period IV, the rank of cointegration seems to be 3. No variable seems to be weakly exogenous and PD and RF may not be present in the cointegration relationships. In Table 5, we also see both the relative price (*p-p**) and the interest rate differential (*i-i**) are stationary with the p-value of less than 1 percent. These suggest that the Japanese economy was very closely integrated with the U.S. in both goods and financial markets. According to our metric, monetary independence is not likely to be high, in comparison to other periods.

Moving average representation

A VAR model (of order 2 for simplicity)

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \alpha\beta' X_{t-1} + \varepsilon_t$$

has a corresponding moving-average representation which is given by $\Delta X_t \equiv (1-L)X_t = C(L)\varepsilon_t$ where $C(L)$ can be expanded as $C(L) = C(1) + C^*(L)(1-L)$.[10] Then, the representation of X_t is given by

$$X_t = \frac{C(1)}{1-L}\varepsilon_t + C^*(L)\varepsilon_t = C\sum_{i=1}^{t}\varepsilon_{t-i} + C^*(L)\varepsilon_t$$

where $C = C(1)$ is the long-run impact matrix that shows how each variable is influenced by the cumulated disturbances. It has a rank less than p and can be found from the cointegration parameters α and β as follows: $C = \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1}\alpha'_\perp = \tilde{\beta}_\perp \alpha'_\perp$ where $\tilde{\beta}_\perp = \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1}$. Based on the above, the $p \times (p-r)$ matrix $\tilde{\beta}_\perp$ can be given an interpretation

10 See Engle and Granger (1987) and Juselius (2006, pp. 256-258).

Table 6 The Moving Average Representation

A. Period I (r = 2)

	ε_{PF}	ε_{RF}	ε_{PD}	ε_{RD}	ε_{ER}
$\alpha'_{\perp,1}$	0.703 (1.67)	1.000 (na)	-0.207 (-1.68)	0.000 (na)	
$\alpha'_{\perp,2}$	-0.244 (-1.92)	0.000 (na)	0.084 (2.255)	1.000 (na)	
	$\tilde{\beta}'_{\perp,1}$	$\tilde{\beta}'_{\perp,2}$			
PF	-2.334 (-1.83)	-11.462 (-2.99)			
RF	0.298 (2.15)	1.177 (2.84)			
PD	-6.458 (-2.75)	-12.69 (-1.81)			
RD	0.273 (2.59)	0.742 (2.35)			

B. Period II (r = 3)

	ε_{PF}	ε_{RF}	ε_{PD}	ε_{RD}	ε_{ER}
$\alpha'_{\perp,1}$	0.952 (0.82)	0.594 (1.03)	0.000 (na)	1.000 (na)	0.337 (1.09)
$\alpha'_{\perp,2}$	-0.23 (-1.07)	0.110 (0.11)	1.000 (na)	0.000 (na)	-0.707 (-1.27)
	$\tilde{\beta}'_{\perp,1}$	$\tilde{\beta}'_{\perp,2}$			
PF	3.797 (0.66)	0.701 (0.29)			
RF	-0.481 (-0.81)	-0.351 (-1.39)			
PD	10.450 (1.07)	4.625 (1.12)			
RD	-0.222 (-0.39)	-0.250 (-1.05)			
ER	-0.497 (-0.12)	-1.352 (-0.76)			

C. Period III (r = 1)

	ε_{PF}	ε_{RF}	ε_{PD}	ε_{RD}	ε_{ER}
$\alpha'_{\perp,1}$	1.000 (na)	0.000 (na)	0.000 (na)	0.000 (na)	0.451 (0.83)
$\alpha'_{\perp,2}$	0.000 (na)	0.000 (na)	0.000 (na)	1.000 (na)	-0.126 (-0.71)
$\alpha'_{\perp,3}$	0.000 (na)	0.000 (na)	1.000 (na)	0.000 (na)	0.030 (0.25)
$\alpha'_{\perp,4}$	0.000 (na)	1.000 (na)	0.000 (na)	0.000 (na)	-0.181 (-0.66)
	$\tilde{\beta}'_{\perp,1}$	$\tilde{\beta}'_{\perp,2}$	$\tilde{\beta}'_{\perp,3}$	$\tilde{\beta}'_{\perp,4}$	
PF	1.065 (0.48)	0.485 (0.32)	1.933 (1.08)	5.345 (1.40)	
RF	0.239 (1.17)	0.048 (0.35)	-0.435 (-2.68)	0.076 (0.22)	
PD	0.416 (0.43)	0.201 (0.31)	1.613 (2.10)	2.291 (1.39)	
RD	0.300 (1.74)	0.865 (7.42)	0.107 (0.78)	0.066 (0.23)	
ER	2.805 (0.81)	-0.221 (-0.09)	1.080 (0.39)	-0.536 (-0.09)	

D. Period IV (r = 3)

	ε_{PF}	ε_{RF}	ε_{PD}	ε_{RD}	ε_{ER}
$\alpha'_{\perp,1}$	-0.075 (-0.38)	1.000 (na)	-0.076 (-0.23)	0.000 (na)	-0.065 (-1.81)
$\alpha'_{\perp,2}$	0.481 (1.43)	0.000 (na)	1.028 (1.82)	1.000 (na)	0.082 (1.34)
	$\tilde{\beta}'_{\perp,1}$	$\tilde{\beta}'_{\perp,2}$			
PF	12.717 (0.90)	0.953 (0.24)			
RF	-0.809 (-0.66)	0.211 (0.61)			
PD	-0.444 (-0.19)	0.961 (1.53)			
RD	0.248 (0.60)	-0.092 (-0.78)			
ER	-10.142 (-0.93)	-1.359 (-0.45)			

as the loadings to the $p-r$ common stochastic trends $\alpha_{\perp}' \sum_{i=1}^{t} \varepsilon_i$.[11] This above shows that the cointegrated VAR model can be represented as common stochastic trends (created by the sum of past residuals each of which has a permanent effect on X_t) plus stationary disturbances.

Therefore, if X_t is integrated of order one and cointegrated, C has reduced rank and hence only $p-r=n$ elements of $C\sum \varepsilon_{t-i}$ have independent effect on X_t. See Kasa (1992) and Juselius (2006) for examples. The loadings $\tilde{\beta}_{\perp}$ show how the variables react to the common trends.

Table 6 reports the estimates of $\alpha_{\perp}'$ and $\tilde{\beta}_{\perp}$for the 4 periods. In period 1, there 2 common trends (CT). CT(1) is dominated by RF and CT(2) by RD. Both price levels are adjusting. With either common trend, both interest rates react strongly and almost by equal amounts, especially with CT(1). This confirms that the domestic interest rate tended to move together with the foreign rate. Interestingly, with CT(2) which is largely driven by shocks on the domestic interest rate, the foreign (U.S.) rate moves more. It appears the shocks on the domestic interest rate may have been the result of shocks that affect not only Japan but other countries in the world.

For period II, 2 common trends − CT(1) and CT(2) − affect the interest rates in the same direction and by almost equal amounts. This is hardly a sign of long-run monetary independence. In period III, there are 4 common trends, each having quite different long-run consequences on the interest rates. CT(1), driven mainly by shocks on the foreign interest rate, affects the domestic interest rate more than the foreign rate. Since there are three other common shocks that affect the interest rates, the relationship between the two interest rates is more difficult to read. In period IV, two common trends are associated with the two interest rates as in period I. However, each trend moves the two interest rates in opposite directions. Moreover, the domestic rate moves much less than the foreign rate.

11 For given $p \times r$ matrices α and β, the orthogonal complements $\alpha_{\perp}$ and $\beta_{\perp}$ of full rank with dimension $p \times (p-r)$ so that $\alpha'\alpha_{\perp}=0$ and $\beta'\beta_{\perp}=0$ and rank $(\alpha,\alpha_{\perp})=p$ and rank$(\beta,\beta_{\perp})=p$.

4. Conclusion

Conventional wisdom holds that floating exchange rates are more conductive to macroeconomic independence. The notion became more forceful as fixed exchange rates are increasingly difficult to defend in a world of high capital mobility. (Obstfeld and Rogoff, 1995 and Eichengreen, 1994) Has macroeconomic independence increased under floating exchange rates? Does the imposition of controls on capital mobility promote monetary independence? This paper uses the method of VAR, cointegration and error correction to investigate international interdependence and the role of the exchange rate during the Bretton Woods period and the modern float for Japan. Moving average representation of the VAR model is applied first time in the analysis of macroeconomic interdependence. Given the drastic changes in the structure of the economy and the exchange rate regime, we divide the whole sample into 4 sub-periods. Presumption is that MI is higher in the float period. In addition, capital controls provide addition ammunition to the monetary authority and thus MI would be lower in the more recent periods when when capital controls have been liberalized.

In a nutshell, we find evidence that floating exchange rates promote MI. Thus, Japan seems to have enjoyed much higher level of MI during periods III and IV than in period I. However, our study shows that capital controls may not necessarily promote MI as indicated by the finding that period II – the period of float and capital controls – has the lowest degree of MI.

Greater mobility of capital has become a fact of life for most countries while independent management of monetary policy remains the cherished goal for policy makers. Our limited evidence suggests that fixed exchange rates would be more likely to be incompatible with the goal of MI. It would be difficult to maintain unrestricted management of interest-rate policy. The experience of Japan shows that capital controls would not be the answer at least in the long run.

References

Aizenman, J., M. Chinn, and H. Ito (2010) “The Emerging Global Financial Architecture: Tracing and Evaluating the New Patterns of the Trilemma's Configurations," *Journal of International Money and Finance* 29 (June), 615-641.

Bluedorn, J.C. and C. Bowdler (2010), “The Empirics of International Monetary Transmission: Identification and the Impossible Trinity,” *Journal of Money, Credit and Banking* 42 (June), 679-713.

Broda, C. (2001), “Coping with Terms-of-Trade Shocks: Pegs versus Floats,” *American Economic Review* 91 (2), 376-380.

Burdekin, Richard C. K., “International Transmission of U.S. Macroeconomic Policy and the Inflation Record of Western Europe,” *Journal of International Money and Finance* 8 (September 1989), 401-424.

Calvo, G. and C. Reinhart (2002), “Fear of Floating,” *Quarterly Journal of Economics* 117 (2), 379-408.

Chin, M. and H. Ito (2006), “What Matters for Financial Development,” *Journal of Development Economics* 81, 163-192.

di Giovanni, J. and J.C. Schambaugh (2008), “The Impact of Foreign Interest Rates on the Economy: The Role of the Exchange Rate Regime,” *Journal of International Economics* 74, 341-361.

Dornbusch, Rudiger, “Flexible Exchange Rates and Interdependence,” *IMF Staff Papers* 30 (March 1983), 3-30.

Dornbusch, Rudiger, and Alberto Giovannini, “Monetary Policy in the Open Economy,” in B. M. Friedman and F. H. Hahn, eds., *Handbook of Monetary Economics, Vol. II* (North Holland: Amsterdam, 1990)

Edwards, S. and E.L. Yeyati (2005), “Flexible Exchange Rates as Shock Absorbers,” *European Economic Review* 49, 2079-2105.

Eichengreen, Barry, *International Monetary Arrangements for the 21st Century* (Brookings Institution, 1994).

Flood, Robert P., and Andrew K. Rose, “Fixing Exchange Rates: A Virtual Quest for Fundamentals,” *Journal of Monetary Economics* 36 (August 1995), 3-37.

Frankel, Jeffrey A., *The Yen/Dollar Agreement: Liberalizing Japanese Capital Markets* (Washington, D. C.: Institute for International Economics, 1984).

----- "Ambiguous Policy Multipliers in Theory and in Empirical Models," in R. C. Bryant, et al., eds., *Empirical Macroeconomics for Interdependent Economies* (Brookings Institution, 1988), 17-26.

Frankel, J.A., S.L. Schmukler, and L. Serven (2004), "Global Transmission of Interest Rates: Monetary Independence and Currency Regimes," *Journal of International Money and Finance* 23, 701-733.

Froot, Kenneth A., and Kenneth Rogoff, "Perspectives on PPP and Long-Run Real Exchange Rates," in G. M. Grossman and K. Rogoff, eds., *Handbook of International Economics*, Vol. III (North-Holland, 1995)

Funabashi, Yoichi, *Managing the Dollar: From the Plaza to the Louvre* (Washington, D. C.: Institute for International Economics, 1988).

Genberg, Hans, Michael K. Salemi, and Alexander Swoboda, "The Relative Importance of Foreign and Domestic Disturbances for Aggregate Fluctuations in the Open Economy," *Journal of Monetary Economics* 19 (1987), 45-67.

Giavazzi, Francesco, Stefano Micossi, and Marcus Miller, *The European Monetary System* (Cambridge University Press, 1988).

Girton, Lance, and Don Roper, "A Monetary Model of Exchange Market Pressure Applied to the Postwar Canadian Experience," *American Economic Review* 67 (September 1977), 537-548.

Hausmann, R., U. Panizza, and E. Stein (2001), "Why do countries float the way they float," *Journal of Development Economics* 66 (2), 378-414.

Hutchison, Michael, and Carl E. Walsh, "Empirical Evidence on the Insulation Properties of Fixed and Flexible Exchange Rates," *Journal of International Economics* 32 (May 1992), 241-263.

Ito, Takatoshi (1983), "Capital Controls and Covered Interest Parity," NBER Working Paper No. 1187.

Johansen, S., "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control* 12 (1988), 231-254.

_____, "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica* 59 (1991), 1551-1580.

_____, *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models* (Oxford Univ. Press, 1995).

_____ and K. Juselius, "The Full Information Maximum Likelihood Procedure for Inference on Cointegration," *Oxford Bulletin of Economics and Statistics* 52 (1990), 169-210.

_____ and _____, "Testing Structural Hypothesis in a Multivariate Cointegration Analysis of the

PPP and UIP for UK," *Journal of Econometrics* 53 (1992), 211-244.

Juselius, K. and S. Johansen (1992), "Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and the UIP for UK," *Journal of Econometrics* 53, 211-244.

Juselius, K. (2006), *The Cointegrated VAR Model: Econometric Methodology and Macroeconomic Applications*. Oxford: Oxford University Press.

Kaminsky, G. and S. Schmukler (2002), "Short and Long-run Integration: Do Capital Controls Matter?" World Bank Working Paper No. 2912, 125-178.

Kasa, Kenneth (1992), "Common Stochastic Trends in International Stock Markets," *Journal of Monetary Economics* 29, 95-124.

Kim, S. (2001), "International Transmission of U.S. Monetary Policy Shocks: Evidence from VAR's," *Journal of Monetary Economics* 48, 339-372.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt & Y. Shin(1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?", *Journal of Econometrics* 54, pp 159-178.

Lastrapes, William D., and Faik Koray, "International Transmission of Aggregate Shocks under Fixed and Flexible Exchange Rate Regimes: United Kingdom, France, and Germany, 1959 to 1985," *Journal of International Money and Finance 9* (December 1990), 402-423.

McKinnon, Ronald I., "The Exchange Rate and Macroeconomic Policy: Changing Postwar Perceptions," *Journal of Economic Literature* 19 (June 1981), 531-557.

Meese, Richard A., and Kenneth Rogoff, "Empirical Exchange Rate Models of the Seventies," *Journal of International Economics* 14 (February 1983), 3-24.

Obstfeld, Maurice, and Kenneth Rogoff, "The Mirage of Fixed Exchange Rates," *Journal of Economic Perspectives* 9 (Fall 1995), 73-96.

Obstfeld, M., J.C. Schambaugh and A.M. Taylor (2004), "Monetary Sovereignty, Exchange Rates, and Capital Controls: the Trilemma in the Interwar Period," *IMF Staff Papers* 71 (Special issue), 75-108.

Obstfeld, M., J.C. Schambaugh and A.M. Taylor (2005), "The Trilemma in History: Tradeoffs among Exchange Rates, Monetary Policies, and Capital Mobility," *Review of Economics and Statistics* 87 (3), 423-438.

Osterwald-Lenum, Michael, "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics," *Oxford Bulletin of Economics and Statistics*," Oxford Bulletin of Economics and Statistics 54 (May 1992), 461-472.

Pesaran, M.H. and R. Smith (1995), "Estimating Long-run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics* 68 (1), 79-113.

Rnagvid, Jesper (2001), "Increasing Convergence among European Stock Markets? A Recursive

Common Stochastic Trends Analysis," *Economics Letters* 71, 383-389.

Rose, A.K. (1996), "Explaining Exchange Rate Volatility: An Empirical Analysis of the Holy Trinity of Monetary Independence, Fixed Exchange Rates, and Capital Mobility," *Journal of International Money and Finance* 15, 925-945.

Schambaugh, J.C. (2004), "The Effect of Fixed Exchange Rates on Monetary Policy," *Quarterly Journal of Economics* 119, 301-352.

CHAPTER 12

Uncovered Interest Parity Puzzle: Asymmetric Responses*

by
Byung-Joo Lee
(University of Notre Dame)

Abstract

This paper estimates UIP slope parameter using a large number of cross-country bilateral exchange rates. The exchange rates analyzed here include a broad spectrum of developed and developing countries. Based on the empirical evidence, short-term(one month) UIP holds well and UIP puzzle is largely observed in the key currencies.

We introduce the *key currency bias* to explain the empirical failure of UIP. UIP fails more often when a key currency is involved in the bilateral exchange rate relationship especially when the key currency offers higher return on capital than when only non-key currencies are involved. This paper presents an empirical evidence for a state-dependent asymmetric response in exchange rate changes depending on the direction of the forward premium.

* This paper is greatly benefitted by numerous comments from my colleague, Nelson Mark.
Department of Economics Notre Dame, IN 46556 U.S.A. 574-631-6837, blee@nd.edu

1. Introduction

Exchange rate between two national currencies is determined by the economic funda-mentals of the countries involved, and its dynamics are heavily influenced by the macroeco-nomic policies of each country. One important potential factor determining the exchange rate is the uncovered interest parity (UIP). The UIP theory asserts forward market ef-ficiency and states that a country's currency is expected to depreciate against a foreign currency when its interest rate is higher than the foreign country's, due to international capital arbitrage. However, as is well documented, numerous empirical tests fail to support the UIP theory, thus producing the so-called forward market anomaly. Froot and Thaler (1990) report average slope estimates of -0.88 using a survey of 75 published estimates (Froot, 1990). Among others, Backus, Gregory and Telmer (1993), Froot and Frankel (1989), and McCallum (1994) all report negative relations on the UIP condition using the currencies of major developed countries. Eichenbaum and Evans (1995), Mark and Moh (2004) report that UIP is sensitive to the monetary policies. Chinn and Meredith (2004), Alexius (2001), and Mehl and Cappielo (2009) have some evidence of long-term UIP. Lee (2011) examines the cross-sectional UIP, and found positive and significant UIP using panel regression model.

This paper investigates the asymmetric UIP using a large number of bilateral UIP relationships. Using monthly time-series data, the bilateral exchange rates of one country against all other countries are calculated, thus producing a large number of bilateral ex-change rates at each time period.[1] Bansal and Dahlquist (2000) examined the weekly data for 28 countries and concluded that there may exist a non-linear asymmetric relationship in UIP for positive and negative forward premiums. They found that the violation of the UIP is not pervasive and the puzzle is largely confined to the high-income countries, and in particular to when U.S. interest rates are higher than foreign rates. Therefore, there may exist an asymmetric relationship between forward premium and exchange rate changes depending on the sign of the forward premium.

1 For 37 currencies, there are 666 bilateral cross-country exchange rates.

UIP asymmetry has previously been in-vestigated by Wu and Zhang (1996) and by Bansal and Dahlquist (2000), but this paper uses a broader spectrum of currencies to investigate UIP asymmetry.

With a few exceptions, most of the existing studies have focussed on exchange rates of major developed countries. Flood and Rose (2001) and Bansal and Dahlquist (2000) expanded their samples to include several important developing countries. However, even when the sample is expanded to include a broader spectrum of countries, tests of the UIP hypothesis have focused mainly on exchange rates with the U.S. dollar. Mark and Wu(1998) considered the cross-country rates for UIP hypothesis, but only in a few cases such as against the Mark or the Yen.

The next section briefly summarizes the UIP theory and introduces this paper's econo-metric model. Section 3 explains the data set and presents empirical results for bilateral UIP estimates. It also investigates the asymmetry of UIP estimates. Section 4 summarizes the main findings of the paper and outlines a plan for further investigation of the UIP puzzle.

2. The forward premium puzzle

Consider the following UIP relationship in natural log form.

$$E_t(s_{t+k}) - s_t = f_{t,k} - s_t = i_t - i_t^* \qquad (1)$$

where f_{t+k} is the k-period forward rate, s_t is the spot rate at time t, and both are in natural logs expressed as the domestic currency price of one unit of the foreign currency. Increase of the spot (forward) rate refers to the depreciation of the domestic currency. i_t and i_t^* are domestic and foreign k-period maturity risk-free bond yields expressed in respective currency terms. Under forward market efficiency and rational expectations hypothesis, UIP states that the expected exchange rate from t to $(t+k)$ equals interest rate differential at time t. UIP relationship does not hold as in Equation (1) if there is a capital control between two countries. See Ito and Chinn (2007) for UIP decomposition.[2] Since $E_t(s_{t+k})$ is unobservable at time *t*, assuming rational expectations for the future spot rate, the econometric model to test the UIP hypothesis uses *ex post* realized spot rate s_{t+k} for $E_t(s_{t+k})$. The econometric model is:

$$s_{t+k} - s_t = \beta_0 + \beta_1(f_{t,k} - s_t) = \varepsilon_{t+k} \tag{2}$$

UIP theory tests forward market efficiency if the joint hypothesis of $\beta_0 = 0$ and $\beta_1 = 1$ holds, i.e., the forward rate is an unbiased predictor of future spot rate. Typical UIP investigations have focused on the time-series estimate of slope parameter β_1 considering β_0 to be the constant risk premium. Overwhelming majority of empirical studies have found that the slope estimates are negative and often statistically significant, let alone being the unity predicted by the UIP. This anomaly has provoked numerous attempts to examine different sample periods with different exchange rates. Few of these investigations have found evidence supporting the UIP theory. This paper investigates that UIP relationship of Equation (1) can also be a cross-sectional property in Section 4.

The negative slope estimate is the evidence of bias of forward rate for the future spot rate. There are several alternative explanations for the negative slope estimates. Fama (1984) first introduced the risk premium, defined as $rp = f_{t,k} - E_t(s_{t+k})$, to explain the negative relationship between the exchange rate and the forward premium. Engel (1996) presents an excellent survey on the forward discount anomaly, focusing on the risk premium explanation. However, if the risk premium hypotheses holds for negative slope estimates, then the risk premium is negatively correlated with the expected depreciation and the variance of

---

2 Consider the following identity,

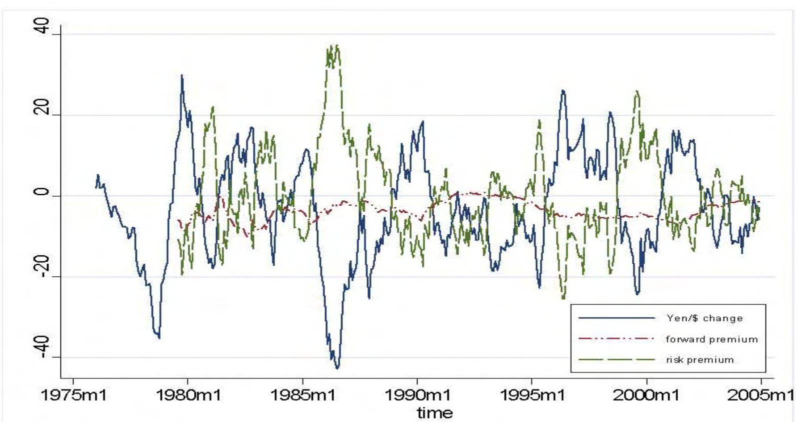
$$i_t - i_t^* = [i_t - i_t^* - (f_{t,k} - s_t)] + (f_{t,k} - E_t(s_{t+k})) + (E_t(s_{t+k}) - s_t)$$

Equation (1) holds if the .rst two terms of the identity are equal to zero. The .rst term is the covered interest parity, and the second term is for the equality of forward rate and the expected exchnage rate. I appreciate the anonymous referee to point out this decomposition.

the risk premium should be greater than that of expected exchange rate depreciation.[3] McCallum (1994) reports that the average of the slope estimates is -4, which is typical of many other studies. This estimate implies that the standard deviation of risk premium is five times larger than that of the forward discount. The surprisingly large standard deviation of the risk premium is not well supported empirically. Figure 1 is time-series plot of one year percentage change of Japanese Yen against U.S. Dollar, one year forward premium and *ex post* (estimated) risk premium for the sample period. This is a typical time-series plot of exchange rate changes, forward premium and estimated risk premium of other developed countries. It is clear that risk premium and exchange rate changes are negatively correlated, with correlation coefficient being -0.88, but the risk premium does not appear to be significantly more volatile than the expected change in exchange rates.

Rogoff (1980) argues that in small samples exchange rates may have fat tails, and that the convergence to normal distribution is slow. Davis, Miller and Prodan (2009) proposes an asymmetric response of UIP using risk premium in the UIP model. Mark and Wu (1998) show that

Figure 1 % change of spot rate, forward premium and risk premium



3 Considering the short-horizon nature of this paper, the *expected* exchange rate change might will be zero, and this condition might always hold.

the risk premium explanation is not consistent with the intertemporal asset pricing model and that the empirical data provide a weak support for the noise-trader model. Coakley and Feurtes (2001) use the exchange rate over-shooting argument as a novel solution to explain the forward premium anomaly.

Next section introduces data and starts with the time-series UIP estimation as a base model to confirm results from previous literature.

3. Empirical analysis

3.1. Data description

Data consist of the currencies of 36 countries and the Euro, totaling 37 currencies [4] [5] The exchange rate data comes from the IMF's International Financial Statistics (IFS). The exchange rates are the monthly rate of the national currency per U.S. Dollar from January 1975 to December 2004, total 360 monthly observations for each country. Euro country local currency exchange rates end at December 1998 and Euro rates start from January 1999 to the end of sample period, December 2004. Therefore, there is no arbitrage oppor-tunities between Euro countries starting January 1999. International currency tradings are mostly conducted through major trading currencies such as Dollar, Euro, Yen and Pound. Many other currency exchanges are conducted indirectly through those major currencies. Therefore, bilateral exchange rates are calculated as the relative rates through U.S. Dollar exchange rates. For example, the bilateral rate between South Korea and Hong

---

4 Countries included in our study are in alphabetic order: Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, China, Denmark, Finland, France, Germany, Greece, Hong Kong, India, Indonesia, Ireland, Italy, Japan, Korea, Malaysia, Mexico, Netherlands, New Zealand, Norway, Peru, Philippines, Russia, Singapore, Spain, Sweden, Switzerland, Thailand, U.K., U.S., Venezuela, Euro.

5 Among 37 national currencies, 21 (including Euro) are classifed as the developed economy currencies and 16 are currencies from the emerging and developing economies. Develpoed countries are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, U.K., U.S., and Euro.

Kong is calculated as relative ratio of South Korean Won per U.S. Dollar to Hong Kong Dollar per U.S. Dollar. In addition, many developing countries have varying degrees of capital con-trols for foreign exchange trading, and it incurs heavy trading costs. Therefore, the official exchange rates may not be a perfect measure for international arbitrage of UIP condition. However, since the main purpose of this paper is to examine the UIP puzzle (negative slope estimate of Eq. 2) in a cross-sectional context, we follow the standard literature in using official exchange rates for UIP investigation. Since forward exchange rates are not widely available for many developing countries, interest rate differentials are used to measure the forward premium. We use four different maturities of interest rate: one month, three month, six month, and one year rates. Interest rate data come from the Datastream, which provides a wealth of detailed information on various interest rates.[6] Euro-currency rates are used for most of the developed countries whenever they are available. When Euro-currency rates are not available, the equivalent interbank rate is used. For some developing countries the interbank rates are used first, when they are available. When they are not available bank deposit rates are used.[7] The interest rate data starts from January 1975 for most of the developed countries but there are several developing countries whose data do not start until mid or late 1990s.[8] We will start with the standard UIP analysis based on U.S. dollar exchange rates to confirm previous findings in the literature.

3.2. UIP with U.S. dollar rate

We will start with the conventional UIP tests using country-by-country exchange rates per U.S. dollar. The baseline econometric model

6 Datastream provides three different kinds of interest rates, bid rate, offer rate and middle rate whenever they are available. We use the middle rate for our analysis.

7 Bank deposit rates are not equivalent to the interbank rates, but deposit rates are the only available interest rates for some developing countries. They are: Argentina, Brazil, Chile, China (one-month rate), Greece (one-year rate), India, Indonesia, Korea, Malaysia, and Venezuela.

8 Details about the interest rate data is available upon request.

is Equation (2).

$$s_{t+k} - s_t = \beta_0 + \beta_1(i_t - i_t^*) = \varepsilon_{t+k} \tag{3}$$

The next two tables report UIP slope estimates for the each country's exchange rate per U.S. dollar using monthly observations for each different maturities, one-, three-, six-, and 12-months. Each country has different start and end dates for different interest maturities depending on data availability. The available monthly observations start from January 1975 and ends at December 2004. Since this equation involves k period forward observations, error terms are subject to the serial correlation of $MA(k\text{-}1)$ process. To correct the serial correlation on ε_{t+k}, this equation is estimated using the Newey-West procedure to calculate the serial correlation robust standard errors. Following standard

Table 1 UIP Slope Estimates for Developed Countries: U.S. Dollar rate

	b1m	se(1m)	b3m	se(3m)	b6m	se(6m)	b1y	se(1y)
Australia	-1.268	0.882	-1.042	0.788	**-1.326**	0.585	**-1.380**	0.539
Austria	-0.661	1.463	-0.510	1.315	-0.539	1.042	-0.611	0.957
Belgium	-0.111	0.777	0.040	1.028	-0.283	0.555	-0.524	0.557
Canada	**-1.403**	0.495	**-0.917**	0.367	*-0.660*	0.363	-0.615	0.455
Denmark	-0.614	0.692	-0.713	0.798	-0.910	0.934	-0.865	0.994
Finland	1.366	1.193	1.462	1.150	1.426	1.134	1.087	1.038
France	-0.158	0.835	0.076	0.644	0.120	0.679	0.215	0.712
Germany	-0.549	0.922	-0.470	0.738	-0.544	0.629	-0.326	0.577
Greece	-0.969	1.058	*-0.295*	0.162	**-0.690**	0.279	**-1.290**	0.171
Ireland	1.169	0.962	0.398	0.895	0.142	1.073	**-5.410**	1.798
Italy	0.514	0.659	**1.311**	0.665	**1.725**	0.601	**1.901**	0.536
Japan	**-2.834**	0.871	**-3.007**	0.667	**-2.933**	0.599	**-2.729**	0.538
Netherlands	**-1.774**	0.797	-1.246	0.765	-1.119	0.702	-0.738	0.637
New Zealand	*-1.523*	0.808	*-1.186*	0.640	**-1.406**	0.485	**-1.406**	0.566
Norway	0.256	0.987	-0.262	0.838	-0.619	0.677	-0.689	0.641
Spain	0.964	1.183	1.246	1.064	1.005	0.990	0.925	1.127
Sweden	-1.587	1.497	*-2.224*	1.131	*-2.406*	1.273	**-2.764**	1.010
Switzerland	-1.328	0.812	-1.086	0.675	*-1.025*	0.566	**-0.954**	0.474
UK	**-1.594**	0.748	-1.270	0.775	-1.135	0.759	-0.799	0.683
Euro	**-6.443**	2.295	**-6.465**	1.732	**-6.556**	0.890	**-6.615**	0.709

Note: Bold numbers are 5% significant and italics are 10% significant.
Standard errors are Newey-West serial correlation robust errors.

classification of countries, Table 1 and 2 report slope estimates and standard errors for developed countries and developing countries, respectively.

As we can see from these tables, many developed countries have statistically significant negative slope estimates. Japan, Canada, and the U.K. all have statistically significant negative estimates. The Euro has strong negative slope estimates, but since the Euro data starts from January 1999, its sample point consists of at most 5 year's monthly observations. Italy is a lone exception with statistically significant positive estimates for three, six and one year UIP. Finland and Spain also have positive estimates for all maturities, but these are not statistically significant. These estimates are generally in line with the findings from previous research for developed countries. For developing countries, only a few slope estimates are statistically significant. Russia and Peru have statistically significant positive estimates while Chile has statistically significant negative estimates for one and three month exchange rate changes.

Table 2 UIP Slope Estimates for Developing Countries: U.S. Dollar rate

	b1m	se(1m)	b3m	se(3m)	b6m	se(6m)	b1y	se(1y)
Argentina	0.171	0.567	-0.194	0.344	.	.	-0.036	0.478
Brazil	-0.079	0.132	.	.	.	.	.	.
Chile	*-2.990*	1.647	*-2.528*	1.472	.	.	.	.
China	3.136	3.074	2.980	2.344	2.436	1.720	1.095	0.838
Hong Kong	-0.034	0.077	-0.037	0.052	0.005	0.030	0.028	0.016
India	0.365	0.962	-0.233	1.419	-0.799	1.682	-0.138	1.242
Indonesia	-0.291	1.796	-1.227	1.234	**-1.967**	0.539	**-1.758**	0.498
Korea	.	.	0.066	0.653	-0.546	0.615	-0.040	0.550
Malaysia	0.227	0.677	0.069	0.553	0.005	0.481	-0.031	0.389
Mexico	-0.156	0.750	-0.111	0.218	0.034	0.227	-0.005	0.180
Peru	.	.	.	.	**1.242**	0.452	**0.783**	0.283
Philippines	0.046	0.376	-0.237	0.440	-0.452	0.453	-0.691	0.448
Russia	**0.669**	0.209	**0.521**	0.216	.	.	.	.
Singapore	-1.407	1.134	*-1.347*	0.697	*-0.887*	0.513	-0.816	0.621
Thailand	0.802	1.780	0.260	1.316	.	.	-0.153	0.947
Venezuela	0.758	1.044	1.066	0.704	.	.	.	.

Note: Bold numbers are 5% significant and italics are 10% significant.

| Table 3 | Rejection of UIP test for each currency: U.S. Dollar rate

Maturity	Developed countries	Developing countries	All countries
b1m	9/20 (0.45)	5/14 (0.36)	14/34 (0.37)
b3m	12/20 (0.60)	7/14 (0.50)	19/34 (0.56)
b6m	14/20 (0.70)	7/10 (0.70)	21/30 (0.70)
b1y	14/20 (0.70)	7/12 (0.58)	21/32 (0.66)

Note: Fractions are in the parenthesis

We tested the UIP hypothesis of $H_0 : \beta_1 = 1$, and rejected the null hypothesis for 9, 12, 14 and 14 out of 20 developed countries, respectively for one-, three-, six- and twelve-month changes.[9] Test results are summarized in Table 3. The UIP hypothesis is rejected slightly more often for developed countries than developing countries. Even if we did not reject the null hypothesis for 11 out of 20 developed countries for one month exchange rate changes, this is more likely due to the large standard errors of the estimates rather than the estimates being close to one. Similar conclusions hold for all other monthly changes. These results mostly agrees to the previous literature. Table 4 is a mean and median of bilateral slope estimates. Since China has fixed its exchange rates for a long period of time and Russia does not have a credible official exchange market, these two countries are excluded from the summary statistics.

It is very difficult to find any clear pattern in these figures, but the slope estimates for developed countries (either including or excluding

| Table 4 | Summary of all slope estimates

	All countries		Developed countries		Developed excluding Euro		Developing countries	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
b1m	-0.661	-0.224	-0.927	-0.815	-0.637	-0.661	-0.216	0.006
b3m	-0.644	-0.279	-0.808	-0.611	-0.510	-0.510	-0.371	-0.153
b6m	-0.728	-0.619	-0.887	-0.675	-0.588	-0.660	-0.374	-0.452
b1y	-0.853	-0.615	-1.179	-0.769	-0.893	-0.738	-0.260	-0.040

9 Rejection for one month UIP: Australia, Canada, Denmark, Japan, Netherlands, New Zealand, Switzer-land, U.K., and Euro.

Euro) tend to be more negative than those of developing countries. The mean slope estimates are generally more negative than those of the median, which suggests that there are more extreme negative estimates than positive ones. Since the Euro has a relatively short sample period, summary statistics are presented with and without the Euro for fair comparison. As with the previous literature, this paper also found numerous negative slope estimates for US dollar-based time-series UIP.

3.3. Key currency bias: UIP for bilateral exchange rates

Next, we estimate the UIP equation for all *bilateral exchange rates* for 37 currencies.[10] Since China and Russia are excluded from our analysis, the maximum number of slope estimates is 595 for each interest maturity with complete data for all countries, but due to the missing data problem, we only have 519 estimates for one and three month, and 414 and 476 for six month and one year UIP estimates. Table 5 gives the summary statistics of UIP regressions for all countries, grouped into developed and developing countries. The means of slope estimates have relatively large standard errors, indicating that the slope estimates are widely dispersed for each bilateral UIP relationship. This result is similar to an individual slope estimate that has relatively large standard error. This is in line with the Baillie and Bollerslev (2000)'s argument that the forward premiums are highly persistent and slope estimates are widely spread. The fourth column of Table 5 shows the acceptance percentage of the UIP theory ($H_0 : \beta_1 = 1$) at 5% significance for two-sided test. There are several interesting results coming out of these statistics.

First, unlike with numerous previous empirical results, the UIP slope estimates are far less negative. The acceptance rate here is much better than that of U.S. Dollar based UIP from Table 3. For all country estimates, the acceptance rate for the UIP theory is close to 78% for one month forward premium. This proportion gradually decreases as forward maturity moves to the longer periods but even for one year

10 Local (*country$_i$*) and foreign (*country$_j$*) designation is interchangeable. Therefore, U.S.-Japan UIP produces exactly the same slope estimates for Japan-U.S. UIP.

| Table 5 | Bilateral UIP slope estimates

Estimates	Mean	Standard Error	Accept UIP(%)	Median	95% Confidence Interval for Median		Sample Size
All Countries							
b1m	-0.050	0.078	77.84	0.078	-0.019	0.242	519
b3m	-0.256	0.078	61.66	-0.044	-0.137	0.038	519
b6m	-0.315	0.092	50.97	-0.118	-0.234	0.030	414
b1y	-0.299	0.085	49.58	-0.131	-0.231	-0.005	476
Between Developed Countries							
b1m	-0.158	0.144	74.13	0.210	-0.111	0.402	201
b3m	-0.143	0.140	58.71	0.144	-0.021	0.210	201
b6m	-0.138	0.140	51.00	0.106	-0.091	0.217	200
b1y	-0.452	0.144	45.50	-0.067	-0.284	0.059	200
Between Developed and Developing Countries							
b1m	-0.065	0.104	82.14	0.041	0.034	0.271	252
b3m	-0.397	0.110	66.27	-0.117	-0.268	-0.001	252
b6m	-0.466	0.139	53.64	-0.312	-0.504	-0.079	179
b1y	-0.167	0.126	55.20	-0.138	-0.313	0.019	221
Between Developing Countries							
b1m	0.340	0.172	72.73	0.000	-0.118	0.320	66
b3m	-0.061	0.137	53.03	-0.161	-0.368	0.074	66
b6m	-0.550	0.178	37.14	-0.403	-0.872	0.007	35
b1y	-0.273	0.118	41.82	-0.107	-0.524	0.154	55

maturity the acceptance rate is close to 50%. A similar pattern holds for different groups of countries. This result shows that the UIP theory holds best for short maturity forward premium and the relationship becomes weaker as maturity becomes longer. It is interesting to compare this result with the claim by Chinn and Meredith (2004) and Alexius (2001) that the UIP relationship holds better with a longer-term horizon (5 to 10 years) than with a shorter one. However, the maximum time horizon focused on in this analysis is one year, and within this time frame short-term UIP holds better than longer-term. Therefore, this result is not necessarily inconsistent with Chinn and Meredith (2004) and with Alexius (2001). Chaboud and Wright (2003) focus on the short-term daily UIP using high-frequency 5 minute exchange data, and claim that UIP holds over a very short period. Our empirical result is generally favorable to the short-term UIP theory. However, one needs to be careful when interpreting this result. Similar to the U.S. results, the generous acceptance of UIP theory is mainly due to the large standard errors of slope estimates rather than the closeness of the slope estimates to one. Secondly, the strongest UIP relationship holds between developed and developing countries, while the weakest relationship is between developing countries. Since the developing countries have less sophisticated financial markets, and their currencies are not easily convertible for international arbitrage, it is reasonable to expect the UIP theory to hold less well between these countries. The acceptance rate is weaker for between-developed countries exchange rates than that of between developed and developing countries. This may be partially due to the *key currency bias* which we discuss below. Lastly, the second and the third panels of countries have medians generally greater than their means, implying that the empirical distribution of slope estimates are negatively skewed. Both groups involve deveolped countries for UIP estimation. The skewed empirical distribution suggests that the medians are better representations of the group statistics than the means. Between developed countries, the medians have similar patterns as the means as they tend to move away from one as interest horizon becomes longer. For other groups of countries, there is no clear trends in medians as with the means, but the medians generally decrease as interest horizon becomes longer. For developed countries, all three maturities have positive medians of slope estimates which are different from other studies.

The next question is to understand why there have been so many puzzling empirical results for UIP theory. We offer a couple of explanations for the apparent miserable failure of the UIP theory. First, most of the existing studies have focussed on a small number of bilateral exchange rates, mostly among developed countries, and on the U.S. dollar exchange rate in particular. This may be the root of some puzzling findings. Although Froot (1990) reports average estimates of -0.88 from the survey of 75 published estimates, most of the estimates come from the studies of a small number of mostly developed countries on U.S. dolar exchange rate. By comparing Table 4 and 5, the UIP slope estimates of U.S. Dollar rates are more negative than those of bilateral UIP for all groups of countries. In Table 3 and 5, the UIP hypothesis is rejected more often for U.S. Dollar rates than for bilateral exchange

Table 6 Regression of UIP slope on key currencies

	b1m		b3m		b6m		b1y	
U.S.	**-0.651**	0.162	0.414	0.294	-0.444	0.238	*-0.593*	0.047
	(0.325)	(0.348)	(0.325)	(0.351)	(0.360)	(0.389)	(0.346)	(0.375)
Key currency		**-1.048**		**-0.912**		**-0.905**		**-0.832**
		(0.188)		(0.190)		(0.217)		(0.206)
constant	-0.009	**0.225**	**-0.230**	-0.026	**-0.284**	-0.060	**-0.260**	-0.068
	(0.081)	(0.089)	(0.081)	(0.090)	(0.095)	(0.108)	(0.088)	(0.099)

rates. However, U.S. Dollar effect is completely mitigated when we expand the set of currencies to the major key currencies. We define the key currencies as those of the countries with a high level of economic power in international economy, together with the Euro.

Table 6 reports two sets of OLS regressions of bilaterla UIP slope estimates for each interest maturity. The first set is the regression of the slope estimates on U.S. dummy with intercept, while the second regression set includes key currency dummy together with U.S. dummy. U.S. dummy has negative estimates for all interest maturities, but only the one month and one year estimates are statistically significantly negative at 5% and 10%, respectively. Key currency dummy is statistically significantly negative at 5% for all maturities, and the impact in absolute terms is the strongest at one month and gradually decreasing, but remains strongly negative. This result suggests that whenever the UIP theory is tested with key currencies, the slope estimates are significantly negative.[11]

In fact, most of the existing literature focuses on empirical tests of the UIP theory involving one or more key currencies, and the results are not surprising given our empirical findings. We call this phenomenon the *key currency bias*. The concept is similar to the home bias in the equity portfolio holdings first observed by French and Poterba (1991) with regard to the fact that the U.S. and Japan have surprisingly large shares of domestic equity holdings despite the existence of more profitable opportunities internationally. This phenomenon is not limited

11 Key currency dummy equals to one for bilateral UIP both for between key currecnies and for key currency and other currencies.

to U.S. and Japan but occurs in many OECD countries. Further analysis of the key currency bias is presented in the next section.

A second possible explanation for the UIP puzzle could be a statistical anomaly due to the small number of slope estimates, mostly for developed countries, used in most studies. This explanation is the cross-sectional equivalent to the peso problem. The UIP theory may hold well for a large number of cross-country bilateral exchange rates, but most studies have focussed mainly on a small number of developed countries. For empirical testing of the UIP theory a few exceptions like Bansal and Dahlquist (2000) and Flood and Rose (2002) have expanded the sample set of countries to include the developing countries. However, even their expanded sets of countries included at most 28 and 23 countries respectively and both studies focussed only on U.S. dollar exchange rates. They had less than 30 estimates of UIP slope parameters. This paper uses a far greater number of estimates to study the statistical properties of the UIP slope parameter. We may have only 37 exchange rates, slightly more than those of Bansal and Dahlquist (2000) and Flood and Rose (2002), but since we are considering bilateral exchange rates, we can have over 600 cross-country bilateral UIP slope estimates to study. Even if we exclude China and Russia for the statistical analysis, we have a much greater number of slope estimates on which to base statistical inferences regarding the UIP slope parameter. Based on the statistical properties of large number of bilateral slope estimates, we can conclude that UIP theory appears to hold well in general, but less well (the relationship is statistically significantly negative) when whenever the U.S. dollar is involved, and among developed countries. However, the dollar effect vanishes when we expand the set of currencies to include the key currencies. Key currencies appear to be the main cause for negative UIP slope estimates. The next section provides empirical evidence for state-dependent asymmetric key currency bias in UIP.

3.4. Asymmetric Response of UIP

If key currencies have a statistically significant negative effect on the UIP slope, then the next step is to investigate the UIP slope estimates more closely to understand key currency bias. Key currency bias is similar to home equity bias in the sense that although UIP theory calls for the expected appreciation of the local currency when interest rate for key currency countries is higher than the domestic interest rate, the key currency is preferred to the local currency whenever the key currency offers higher interest rate. This implies that there may be an asymmetric response of capital movement depending on the sign of the interest rate differentials. UIP theory is an arbitrage condition for equalizing the return on capital between two countries. However, key currency bias suggests that there is no arbitrage when a key currency provides higher return on capital.

Eichenbaum and Evans (1995) found that a contractionary shock due to U.S. monetary policy leads to persistent, significant appreciation in U.S. nominal and real exchange rates, significant deviations from the UIP theory. Wu and Zhang (1996) examined the yen and deutschemark and found that the slope estimates are asymmetric in the direction of forward premium. Bansal and Dahlquist (2000) using weekly data for 28 countries found that the negative slope estimates are more pronounced when U.S. interest rates are higher than that of other countries. Therefore, there may exist an asymmetric relationship between forward premium and exchange rate changes depending on the sign of the forward premium. This section tests the asymmetric response of key currencies. Bansal and Dahlquist (2000) used the state-dependent econometric model as:

$$s_{t+k} - s_t = \beta_0 + \beta_i^+ x_t^+ + \beta_1^- x_t^- + \varepsilon_{t+k} \qquad (4)$$

where x_t^+ t and x_t^- represent positive and negative forward premium. This model assumes only the slope difference of the state-dependency of the UIP theory. This paper relaxes their model to allow not only the slope but also the intercept of the UIP model to be state-dependent. The estimation model is as follows.

$$\begin{aligned} s_{t+k} - s_t &= \beta_0^+ + \beta_1^+ x_t^+ + \varepsilon_{t+k}^+ \quad for \;\; x_t \geq 0 \qquad (5) \\ s_{t+k} - s_t &= \beta_0^- + \beta_1^- x_t^- + \varepsilon_{t+k}^- \quad for \;\; x_t < 0 \end{aligned}$$

Intercepts and slopes are estimated separately depending on the positive or negative for-ward premium.[12] This model is more general than using dummy variable to distinguish x_t^+ and x_t^- in a sense that error terms can have different variances in each state. Ta-ble 7 is F-test result for each bilateral exchange rate pair to test the null hypothesis of no state-dependence: $H_0 : \beta_0^+ = \beta_0^+$ and $\beta_1^+ = \beta_1^-$ with 5% significance. *F*-test is con-ducted only when both states have minimum of 20 observations to estimate parameters accurately. Therefore, when one country has higher or lower interest rates than the other country throughout the entire sample period, only one set of parameters is estimated and no *F*- test is performed. Figure 2 is an example of U.K.-Japan one month interest rates for the sample period. Since U.K. has always higher interest rate than Japan for entire sample period, we only estimate (β_0^+, β_1^+) for U.K.-Japan UIP.

The null hypothesis of symmetric UIP is rejected more often when key currencies are involved and especially when U.S. Dollar is involved. As interest maturity becomes longer, rejection is more frequent than shorter maturity. It is clear that UIP theory is heavily dependent on the direction of forward premium, and the dependency becomes much more apparent for key currencies including U.S. Dollar. This is yet more evidence of key currency bias.

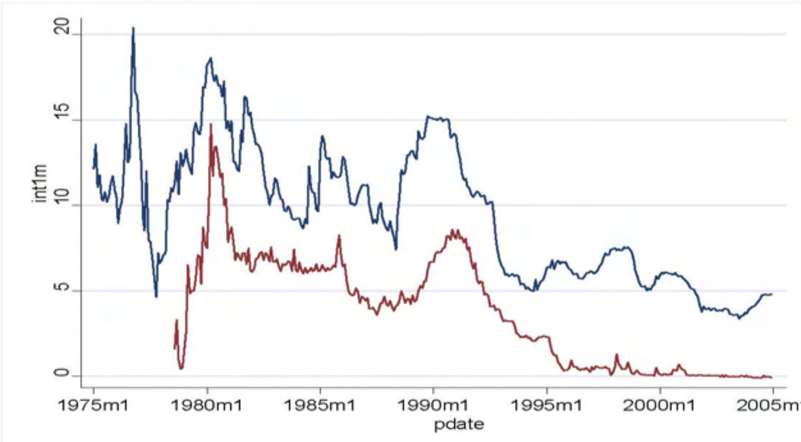
Next, we move to investigate the statistical properties of state-dependent slope esti-mates. For state-dependent UIP estimation, we consider two different situations. First, we estimate the case when U.S. interest rates are higher than the rest of all other sam-ple countries. This is the same state-dependent model as Bansal and Dahlquist (2000). Second, we also consider the case when interest rates from the key currency countries are higher than the rest of other sample countries.[13] Table 8 is the summary statistics for the slope estimates of bilateral cross-country UIP equation.[14]

12 Since x_t^+ and x_t^- could alternate in a consecutive time period, time series observations are not in sequence. However, proper lag structures are all maintained.

13 Among key currencies, the ordering is as following: U.S., Euro, Japan, U.K., and Germany. For example, between Euro and Japanese Yen, $b1m_K^-$ is a slope estimate when Japanese Yen Eurorate is lower than that of Euro.

14 Intercepts are also estimated differently for positive and negative forward premium,

| Figure 2 | One month Eurorate for British Pound and Japanese Yen



| Table 7 | Asymmetric UIP relationship: F-test

	All Countries			Key Currencies			U.S.		
Maturity	n	reject	%	n	reject	%	n	reject	%
One month	196	17	8.7	59	9	15.3	19	6	31.6
Three month	186	59	31.7	56	24	42.9	19	13	68.4
Six month	148	84	56.8	51	30	58.8	18	14	77.8
One year	150	102	68.0	47	35	74.5	18	16	88.9

The first panel reports state-dependent slope estimates for each interest maturity for all countries.[15] The means and medians of slope estimates for positive forward premium are more positive than those of negative forward premium. In fact, for six month and one year maturity for all country group (the first panel), the median estimates are statistically significantly negative. The next two panels report asymmetric slope estimates of state-dependent forward premium against U.S. rates and key currency rates, respectively.[16] There is a strong evidence in the U.S. results that the exchange rates respond differently on positive or

but they are not reported here.

15 Superscript + and - refesr to the positive and negative forward premium, respectively.

16 $b1m_{US}^{-}$ is the slope estimate of one month UIP when local (domestic) has higher interest rate than U.S. (foreign), a positive forward premium, and $b1m_{US}^{-}$ is when local country has lower interest rate than U.S., a negative forward premium.

negative forward premium for U.S. Both means and medians of slope esti-mates for positive forward premium are all positive and *do not* reject the UIP theory. The means of one and three month slope estimates are statistically significantly positive. For negative forward premium, the slope estimates are all negative and statistically significant at 5%. The medians of negative premium are all statistically significant. When the sample is expanded to include other key currencies besides the U.S. dollar, the asymmetric UIP response still remains strong. Since key currency results include a greater number of esti-mates than the U.S. dollar results, median confidence intervals become much tighter than those relating to the dollar. While slope estimates of positive premium

Table 8 Asymmetric UIP slope estimates

Estimate	Mean	Standard Error	Median	95% interval	confidence for median	Sample size
			All Countries			
$b1m^+$	-0.027	0.170	0.090	-0.049	0.219	375
$b3m^+$	0.022	0.192	0.001	-0.111	0.142	363
$b6m^+$	0.094	0.168	0.159	0.023	0.357	302
$b1y^+$	0.005	0.132	0.021	-0.131	0.143	340
$b1m^-$	-0.309	0.185	0.134	-0.028	0.327	340
$b3m^-$	-0.580	0.189	-0.130	-0.343	0.027	342
$b6m^-$	-0.551	0.194	-0.279	-0.488	-0.089	258
$b1y^-$	-0.440	0.159	-0.387	-0.612	-0.197	285
			U.S.			
$b1m^+_{us}$	0.844	0.412	0.278	-0.146	1.401	30
$b3m^+_{us}$	0.926	0.292	0.631	-0.065	1.125	30
$b6m^+_{us}$	0.417	0.570	0.381	-0.264	1.678	28
$b1y^+_{us}$	0.436	0.510	0.297	-0.039	1.645	30
$b1m^-_{us}$	-4.378	0.782	-3.694	-6.816	-1.779	21
$b3m^-_{us}$	-4.679	1.056	-4.277	-5.740	-1.593	21
$b6m^-_{us}$	-3.731	0.830	-2.447	-5.423	-0.879	19
$b1y^-_{us}$	-2.889	0.549	-2.952	-3.506	-1.041	19
			Key Currency			
$b1m^+_k$	-0.065	0.338	0.129	-0.159	0.298	129
$b3m^+_k$	0.070	0.386	0.050	-0.219	0.403	128
$b6m^+_k$	-0.241	0.341	0.104	-0.175	0.366	115
$b1y^+_k$	-0.251	0.241	0.036	-0.196	0.321	122
$b1m^-_k$	-2.477	0.520	-2.049	-3.463	-0.915	71
$b3m^-_k$	-2.864	0.675	-1.863	-3.859	-1.035	69
$b6m^-_k$	-1.817	0.536	-0.836	-2.047	-0.195	60
$b1y^-_k$	-1.911	0.407	-1.202	-2.386	-0.477	59

| Table 9 | Paired t test for the means of + and - forward premium

Maturity	b^+	b^-	Difference	Std. Error	t-stat	p-value	n
			All Countries				
$b1m$	-0.137	-0.453	0.316	0.402	0.786	0.433	196
$b3m$	0.333	-0.716	1.049	0.526	1.993	0.048	186
$b6m$	0.560	-0.505	1.065	0.418	2.550	0.012	148
$b1y$	0.436	-0.302	0.738	0.333	2.220	0.028	150
			U.S.				
$b1m$	1.370	-4.619	5.989	1.021	5.866	0.000	19
$b3m$	1.443	-4.974	6.418	1.252	5.125	0.000	19
$b6m$	0.734	-3.879	4.614	1.131	4.078	0.001	18
$b1y$	1.159	-2.997	4.156	0.975	4.265	0.001	18
			Key Currencies				
$b1m$	-0.280	-2.442	2.162	0.851	2.542	0.014	59
$b3m$	0.705	-2.703	3.408	1.345	2.535	0.014	56
$b6m$	0.304	-1.487	1.791	0.877	2.041	0.047	51
$b1y$	0.463	-1.609	2.072	0.613	3.380	0.002	47

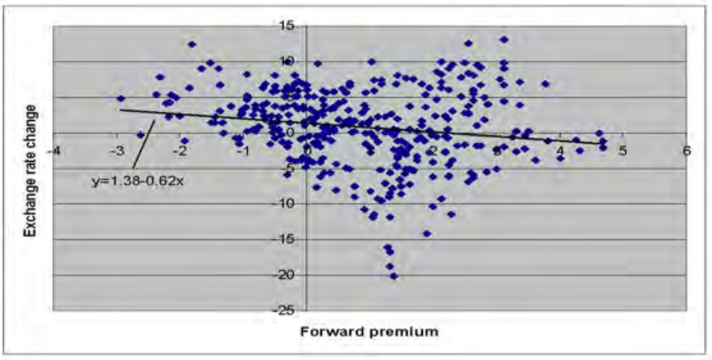
still reject the UIP theory, the 95% confidence interval includes positive values. Those of negative premium are statistically significantly negative. The differences of means for positive and negative forward premiums are statistically significant except for one month for all country group. There is no statistical evidence of asymmetric UIP of one month for all country group. As interest rate maturity becomes longer, b^+ and b^- are statistically different for all groups of countries. Table 9 is a paired t-test results for different interest maturity and country group.[17]

The following two figures are scatter plots of exchange rate changes and forward premium between local currency and one of the key currencies. Figure 3 is a scatter plot of one year percentage changes of Canadian Dollar-U.S. Dollar exchange rate and one year for-ward premium, and Figure 4 is a scatter plot of one year change of Norwegian Krone-U.K. Pound and one year forward premium between two countries. UIP regression line using all observations is plotted in thick line and two separate regressions for positive and negative forward premium are plotted in thin (negative premium) and dotted (positive premium) lines. Scatter plots clearly show that there is a negative

17 Paired t -tets is conducted when UIP regression has slope estimates for both states.

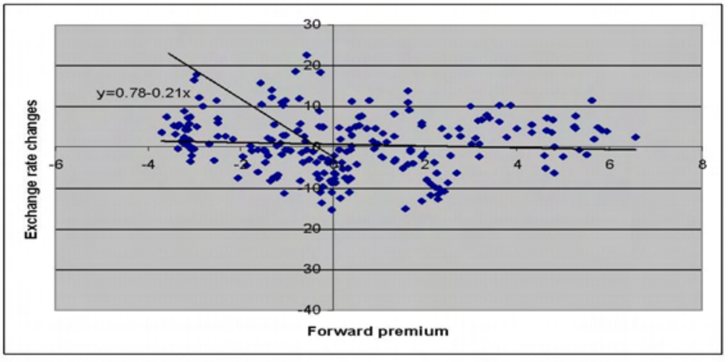
relationship for negative forward premium while positive relationship for positive forward premium. F-statistic to test the equivalence of two equations are 9.88 (p-value: 0.000) and 11.88 (p-value: 0.000), respectively for Canadian Dollar and Norwegian Krone. With regard to the UIP puzzle, this is yet more evidence of negative slope estimates if we do not consider the asymmetric state-dependent nature of UIP.

| Figure 3 | Canadian Dollar-U.S. Dollar 1 year change &1 year forward permium.



Note: There are 348 observations to produce y_t =1.38(0.73)-0.62(0.46) x_t . Standard errors in the parenthesis. For 248 positive premium observations, the estimated UIP equation is y_t =-0.79(1.38)+0.47(0.68) x_t , and for 100 negative premium observation, it is y_t =2.16(1.89)-1.08(1.00) x_t .

| Figure 4 | Norwegian Krone-U.K. Pound 1 year change & 1 year forward premium.



Note: There are 216 observations to produce y_t =0.78(1.38)-0.21(0.43) x_t . Standard errors in the parenthesis. For 111 premium observations, the estimated UIP equation is y_t =-248(2.09)+1.03(0.49) x_t , and for 105 negative premium observations, it is y_t =-1.92(3.35)-2.13(1.12) x_t .

From these two tables and two figures, it is very clear that the exchange rates respond very differently depending on the state of forward premium. Asymmetric responses are prevalent for the key currency UIP condition and become more pronounced for the U.S. dollar. This is more evidence for key currency bias. UIP asserts international capital arbitrage between two countries through exchange rate adjustment. However, even taking into account the transactions costs, UIP does not hold when key currencies are involved. The reason for this is a subject for future research. One possible conjecture for key currency bias is as following. UIP is an arbitrage condition by adjusting exchange rate for capital movement. Ultimately, there will be no capital movement in equilibrium because of the UIP. Underlying the UIP theory is an implicit assumption that when capital moves from the domestic to a foreign country to seek a higher return, after the return is realized, capital will return to the domestic country. However, on the basis of the empirical evidence, when key currency offers higher return, capital flows into the key currency with no immediate intention to repatriate into the local currency. Capital may ultimately be converted to the local currency, but not within the one-year time frame analyzed in this study. In this sense, longer-term UIP proposed by Chinn and Meredith (2004), Alexius (2001), and Mehl and Cappielo (2009) may be valid.

4. Conclusion

This paper investigated empirical evidence relating to the UIP puzzle. We showed that there is no evidence of UIP puzzle in the cross-sectional UIP. Cross-sectional UIP slope estimates are statistically positive for all interest rate maturities, and the relationship becomes stronger as interest rate maturity becomes longer. This is the first paper to investigate the statistical property of cross-sectional UIP slope estimates.

Time-series UIP seems to hold well among developed-developing country exchange rate pair. UIP puzzle is largely confined to the key currencies and is more prevalent when a key currency offers higher return on capital. For the cross-country bilateral UIP test, we accepted the null hypothesis of UIP over 77% for one month maturity for all countries, and the acceptance rate gradually decreases as the maturity becomes longer, up to one year. For country groups of developed and developing countries, one month UIP hypothesis is accepted at least 72 % of the time. This result sharply contradicts most of the existing literature on the failure of the UIP theory. However, no previous study has produced a large number of slope estimates to draw meaningful statistical conclusions about the slope parameter. Most of existing literature presents only a few slope estimates, mostly among developed country-pair UIP relationships.

This paper also presented statistical evidence for asymmetric response of the UIP relationship. Bansal and Dahlquist (2000) only conjectured the possibility of the state-dependent asymmetric UIP; they did not provide the statistical evidence for state-dependence. This paper used bilateral exchange rates and produced a large number of slope estimates to draw valid statistical inferences about the UIP theory. Negative forward premium of key currencies was found to be the main reason for negative slope estimates. The UIP theory is not rejected for positive forward premium for the U.S.

References

Alexius, Annika (2001), "Uncovered interest parity revisited," *Review of International Economics*, vol. 9, 505-517.

Backus, David, Allan Gregory and Chris Telmer (1993), "Accounting for forward rates in markets for foreign currency," *Journal of Finance*, vol. 48, 1887-1908.

Baillie, Richard T. and Tim Bollerslev (2000), "The forward premium anomaly is not as bad as you think," *Journal of International Money and Finance*, vol. 19, 471-488.

Baillie, Richard T., Cecen and Han (2000), "High frequency Deutsche mark-U.S. dol-lar returns: FIGARCH representations and non linearities," *Multinational Finance Journal*, vol. 4, 247-267.

Bansal, Ravi and Magnus Dahlquist (2000), "The forward premium puzzle: Different tales from developed and emerging economies," *Journal of International Economics*, vol. 51, 115-144.

Chaboud, Alain P. and Jonathan H. Wright (2003), "Uncovered interest parity: It works, but not for long," working paper, Board of Governors of the Federal reserve System.

Chinn, Menzie D. and Guy Meredith (2004), "Monetary policy and long-horizon uncovered interest parity," IMF Staff Papers, vol. 51, No.3.

Coakley, Jerry and Ana-Maria Feurtes (2001), "Exchange rate overshooting and the forward premium puzzle," working paper, University of Essex.

Eichenbaum, Martin and Charles L. Evans (1995), "Some empirical evidence on the effects of shocks to monetary policy on exchange rates," *Quarterly Journal of Economics*, vol. 110, 975-1009.

Engel, Charles (1996), "The forward discount anomaly and the risk premium: A survey of recent evidence," *Journal of Empirical Finance*, vol. 3, 123-192.

Fama, Eugene (1984), "Forward and spot exchange rates," Journal of Monetary Economics, vol. 14, 319-338.

Fama, Eugene F. and James D. MacBeth (1973), "Risk, return and equilibrium: Empirical test," *Journal of Political Economy*, vol. 81, 607-636.

Flood, Robert P. and Andrew K. Rose (1996), "Fixes: Of the forward discount puzzle," *Review of Economics and Statistics*, vol. 78, 748-752.

Flood, Robert P. and Andrew K. Rose (2002), "Uncovered interest parity in crisis: The interest rate defense in the 1990s," IMF Staff Papers, vol. 49, No. 2.

French, K. and J. Poterba (1991), "Investor diversification and international equity markets," *American Economic Review*, vol. 81, 222-226.

Froot, Kenneth (1990), "Short rates and expected asset returns," NBER working paper No. 3247.

Froot, Kenneth and J. Frankel (1989), "Forward discount bias: Is it an exchange risk premium*?" Quarterly Journal of Economics*, vol. 104, 139-161.

Froot, K. and Thaler, R. (1990), "Anomalies: foreign exchange," *Journal of Economic Perspective*, vol. 4, 197-92.

Krasker, William (1980), "The .peso problem. in testing the efficiency of forward exchange market," *Journal of Monetary Economics*, vol. 6, 269-276.

Lee, Byung-Joo (2011), "Uncovered interest parity: cross-sectional evidence," Review of International Economics, vol. 19, No.2, 219-231.

Mark, Nelson C. and Young-Kyu Moh (2004), "O¢ cial intervention and occasional violations of uncovered interest parity in Dollar-DM market," working paper, University of Notre Dame.

[22] Mark, Nelson, C. and Yangru Wu (1998), "Rethinking deviations from uncovered interest parity: The role of covariance risk and noise," *The Economic Journal*, vol. 108, 1686-1706.

McCallum, Bennett T. (1994), "A reconsideration of the uncovered interest parity relationship," *Journal of Monetary Economics*, vol. 33, 105-132.

Reinhart, Carmen M. and Rogoff, Kenneth S. (2004), "The modern history of exchange rate arrangements: A reinterpretation," *Quarterly Journal of Economics*, vol. 119, 1-48.

Rogoff, Kenneth (1980), "Tests of martingale model for foreign exchange futures market," in *Essays on expectations and exchange rate volatility*. Ph.D. Dissertation, Massa-chusetts Institute of Technology.