From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2023

Essays on Regional Inequality and Mobility

Jin, Jonas Yang

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01s7526g678

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Essays on Regional Inequality and Mobility

Jonas Jin

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY
THE DEPARTMENT OF
Economics
Adviser: Richard Rogerson

May 2023

© Copyright by Jonas Jin, 2023. All rights reserved.

Abstract

This dissertation studies sources and consequences of regional inequality and mobility.

Chapter 1 uses individual-level longitudinal survey data from the Indonesian Family Life Survey to produce estimates of mobility in Indonesia. I find that intergenerational income persistence is low relative to other developed countries. I explore mechanisms of mobility such as education, location, and sector of work. Contrary to previous evidence in the United States, I find no evidence of neighborhood effects between urban and rural areas.

Chapter 2, joint with Yu (Alan) Yang and Xiaoyang Ye, studies the effects of region-based education and migration policies in China that produce regional inequities in quality of and access to opportunity. We build and calibrate a structural spatial overlapping generations model that captures China's institutional setting. We then use the calibrated model to quantify the effects of changing these policies on intergenerational mobility, welfare, education, and the income distribution, with particular interest in outcomes for children born at the bottom of the parental income distribution. We find that lifting migration restrictions and equalizing public spending levels increases intergenerational mobility and improves outcomes for the poorest born in poor provinces at the expense of children born in rich provinces. Modifying college admissions to a merit-based or equity-based system has modest aggregate effects, with positive (and negative) effects concentrated among children born to high-income parents.

Chapter 3, joint with Janet Currie and Molly Schnell, uses quarterly county-level data from 2006–2014 to examine the direction of causality in the relationship between per capita opioid prescription rates and employment-to-population ratios in the United States. We address concerns of simultaneity between economic conditions and opioid prescription rates using instrumental variables. We instrument for opioid prescription rates of the working-age population using opioid prescription rates of the elderly, and we instrument for employment-to-population ratios using a shift-share instrument. Our findings suggest that after control-

ling for province-level fixed effects, there is no simple causal relationship between economic conditions and opioid abuse, and thus improving economic conditions is unlikely on its own to curb the opioid epidemic.

Acknowledgements

This work would not have been possible without the assistance and encouragement of numerous faculty, colleagues, friends, and family.

I am especially grateful to Richard Rogerson, whose excellent and prompt research advice supported me through the last seven years. Not only was his guidance invaluable to my progress, but his kindness and patience enabled me to prioritize my well-being during even the most stressful periods of graduate school. There were so many instances when I felt stuck or lost that were resolved after a 30-minute meeting with Richard,

My work and my ability as a researcher also benefited greatly from Leah Boustan's and Thomas Fujiwara' insightful comments. Their advice always provided new angles and questions to explore, which was particularly helpful in framing my research to appeal across different fields. I also thank Karen McGuinness, who provided numerous opportunities for me to develop the skills, experience, and flexibility necessary to develop into a better instructor.

I will fondly recall the memories with my friends who made these past years such an incredible experience—while graduate school was certainly not easy by any means, I feel extremely fortunate to have been able to enjoy life alongside Molly Chaney, Dan Kozuch, Michael Palmer, Sean King, Lena Basta, and the rest of the Squa(d) almost from day one. Whether it was watching sports or Bachelor, eating insane amounts of sushi, spilling wine at formal events, spending nights at DBar, or discussing plans to purchase a group hot tub during dinner, our experiences together greatly enriched my life at Princeton.

I thank Tyler Maxey, Eric Qian, Robert Wagner, Sebastian Guarda, and other members of the A95 office community for creating such a lively, productive, and unconditionally supportive work environment. Because of you all, my ability at darts has never been better, and I'll never forget to say "bless you" when someone sneezes.

Anyone who has met me probably knows that intramural sports were instrumental to my life at Princeton. Playing with the Soggy Seals, #trusttheprocess, Moneyball, Volleylobbers,

and other creatively-named teams has provided me with many free T-shirts and even more incredible memories on and off the pitch, court, or diamond: winning (and losing) championships, navigating the bumpy landscape of Poe Field, playing against club soccer, going on an adventure to Ithaca, finally lifting the summer softball trophy, enduring two hospital trips (sorry Rocky and Aaron!)[1], eating delicious croissants from The Little Chef, and just hanging out over lunch. A huge shout out to Rocky Diegmiller, Jordan Dull, Valentin Skoutnev, Jack Guenther, Sam Day-Weiss, Alan Kaplan, Nick Cho, Alex Johnson, Francisco Cabezon, Yannick Eatmon, Eric Blow, Theo Gibbs, Holly Burrows, Eve Beaury, Kyle Luther, Sam Saskin, Michael Soskind, Eric O'Neill, Tom Underwood, Denis Shiskin, and countless others who are much better athletes than myself and propelled me to a wildly successful intramural career.

Finally, thank you to my family for their continued support not only throughout graduate school, but throughout life. To my sister, Shan, for being a role model growing up and for locking me in the basement until I learned to enjoy math from a young age. To Malcolm, Kaylee, Iris, and Lily, who understand nothing of my day-to-day responsibilities beyond my role as the fun uncle. And to my parents, Daqing and Guizhen, for enabling me to focus on my studies and pursue my dreams, emphasizing self-care, and keeping me grounded even when things weren't going well. I cannot overstate how fortunate and thankful I am to have parents who selflessly provided me with all the care and resources that they possibly could; my accomplishments are as much theirs as they are mine.

[1] If I had a nickel...

To my parents, whose hard work and sacrifice inspired my research and blessed me with the opportunities of a lifetime.

Contents

Abstract . . . 3

Acknowledgements . . . 5

List of Tables . . . 11

List of Figures . . . 14

1 Intergenerational Poverty Traps: Evidence from Indonesia **16**

1.1 Introduction . . . 16

1.2 Background . . . 18

1.3 Data . . . 19

1.3.1 Indonesian Family Life Survey . . . 19

1.3.2 Sample Selection . . . 21

1.3.3 Income . . . 22

1.4 Results . . . 25

1.4.1 Economy-Wide Intergenerational Mobility . . . 25

1.4.2 Urban and Rural Intergenerational Mobility . . . 32

1.4.3 Nonlinearities in Relative Mobility . . . 45

1.5 Robustness Checks . . . 47

1.6 Conclusion . . . 52

Appendix **61**

1.A.1 Discussion of Empirics (adapted from Mazumder (2005)) . . . 62

1.A.2 Rank-Rank Mobility Estimates using Within-Region Income Distributions . . . 63
1.A.3 Full Survey Results: Gender Differences . . . 63
1.A.4 Contribution of Agricultural Sector to IGE . . . 65

2 Intergenerational Mobility Effects of Region-Based Education and Migration Policies in China 67

2.1 Introduction . . . 67
2.2 Background . . . 73
2.3 Model . . . 75
2.A.1 Adult Consumption and Investment Decisions . . . 76
2.A.2 Children's Education . . . 78
2.A.3 Migration Decision and Migration Costs . . . 79
2.A.4 Housing Prices . . . 81
2.A.5 Steady State Equilibrium . . . 81
2.4 Model Calibration . . . 86
2.A.1 Data Sources . . . 86
2.A.2 Aggregating the Data to Two Regions . . . 87
2.A.3 Calibration Procedure . . . 87
2.5 Policy Evaluation . . . 91
2.A.1 Steady State Effects . . . 94
2.A.2 Distributional Effects and Mechanisms . . . 95
2.6 Discussion and Conclusion . . . 110

Appendix 113

2.A.1 Rich vs. Poor Provinces . . . 113
2.A.2 College Seat Allocation . . . 114

2.A.3 Steady-State Solution Process . . . 115
2.A.4 Solving for the Transition Path . . . 115
2.A.5 Calibration: Non-Targeted Moments . . . 116
2.A.6 Combining Hukou and Public Spending Policies . . . 117
2.A.7 Scaling College Allocation Counterfactuals . . . 120

3 U.S. Employment and Opioids: Is There a Connection? 122

3.1 Introduction . . . 122
3.2 Background . . . 124
3.3 Data . . . 125
3.3.1 Summary Statistics . . . 127
3.4 Methods . . . 131
3.5 Results . . . 134
3.6 Robustness . . . 137
3.7 Discussion and Conclusions . . . 140

Appendix 144

Bibliography 146

List of Tables

1.1 Summary Statistics: IFLS 21
1.2 Summary Statistics: Sample 24
1.3 Estimates of Intergenerational Elasticity by Country 26
1.4 Estimates of Relative Mobility 29
1.5 IGE with Controls 30
1.6 Determinants of Child Observables 31
1.7 Urban and Rural Estimates of Relative Mobility 33
1.8 IGE with Controls: Urban vs. Rural 34
1.9 Determinants of Child Observables: Urban vs. Rural Children 54
1.10 Mincer Regressions for Urban and Rural 55
1.11 Expected Daily Childhood Income for Given Parent Income ($US PPP) . 55
1.12 Selection into Migration 56
1.13 Measures of School Quality: Rural vs. Urban 56
1.14 Nonlinearities in IGE 57
1.15 Access to and Usage of Financial Institutions 57
1.16 IGE: Nonlinearities, Urban. vs. Rural 58
1.17 IGE vs. Rank-Rank 59
1.18 Rank-Rank: Ranks within Distribution 59
1.19 IGE with Unadjusted Income 59
1.20 IGE with Parent and Child Age Controls 60
1.21 Agricultural vs. Non-Agricultural 60

1.22 IGE with Controls, Adjusting for Financial Crisis 60
1.A.1 Gini Estimates by Country 61
1.A.2 Urban and Rural Estimates of Rank-Rank Mobility: Within-Sample . . . 63
1.A.3 Full Sample Estimates of Relative Mobility 64
1.A.4 Full Sample IGE with Controls 64
1.A.5 Full Sample Urban and Rural Estimates of Relative Mobility 65

2.1 Results of Province Aggregation 87
2.2 Externally-Calibrated Parameter Values 89
2.3 Internally-Calibrated Parameters and Model Fit 91
2.4 Steady State Differences across Policies 94
2.5 Hukou: Effects on Child's Expected Welfare 97
2.6 Hukou: Effects on Migration % 98
2.7 Hukou: Effects on Human Capital, College Attendance %, and Log(Income) 99
2.8 Hukou: Differences Between Winners vs. Losers from Policy 100
2.9 Public Spending: Effects on Child's Expected Welfare 101
2.10 Public Spending: Effects on Children's Migration % 102
2.11 Public Spending: Effects on Children's Human Capital, College Attendance %, and Log(Income) 103
2.12 Public Spending: Differences Between Winners vs. Losers from Policy . . 104
2.13 Merit: Effects on Child's Expected Welfare 105
2.14 Merit: Effects on Children's Human Capital, College Attendance %, and Log(Income) 106
2.15 Merit: Differences Between Winners vs. Losers from Policy 107
2.16 Equity: Effects on Child's Expected Welfare 108
2.17 Equity: Effects on Children's Human Capital, College Attendance %, and Log(Income) 109

2.18 Equity: Differences Between Winners vs. Losers from Policy 110
2.A.1 Rich vs. Poor Provinces 113
2.A.2 Quartile-Quartile Transition Matrix 117
2.A.3 Steady State Differences across Policies 118
2.A.4 Hukou + Public Spending: Effects on Child's Expected Welfare 119
2.A.5 Hukou + Public Spending: Effects on Children's Human Capital, College Attendance %, and Log(Income) 119
2.A.6 Hukou + Public Spending: Differences Between Winners vs. Losers from Policy 120
2.A.7 Steady State Differences across College Allocation Policies (Scaled) 121
2.A.8 Merit: Effects on Child's Expected Welfare (Scaled) 121
2.A.9 Equity: Effects on Child's Expected Welfare (Scaled) 121

3.1 Mean Employment to Population and Prescriptions Rates 128
3.2 Effects of Opioid Prescriptions on Employment: without County Fixed Effects 136
3.3 Effect of Opioid Prescriptions on Employment: with County Fixed Effects 137
3.4 Effect of Employment on Opioid Prescriptions: without County Fixed Effects 138
3.5 Effect of Employment on Opioid Prescriptions: with County Fixed Effects 139
3.6 Effect of Opioid Prescriptions on Employment: controlling for County Fixed Effects and % Insured 141
3.7 Effect of Employment on Opioid Prescriptions: controlling for County Fixed Effects and % Insured 142
3.A.1 Age Groupings in Prescription and Employment Data 144
3.A.2 Mapping from NAICS Codes to SIC Codes 145
3.A.3 Comparison of QWI and QCEW OLS Estimates for Working Age Individuals 145

List of Figures

1.1 Indonesia vs. Comparison Countries 20
1.2 Labor Force Participation and Earnings by Gender and Age 22
1.3 Lifecycle Bias, Indonesia 29
1.4 HS and College Graduation Rates by Parent Income 36
1.5 Mobility by Urban Status 38
1.6 Mobility with Education Controls 39
1.7 Urban/Rural Children by Income Quartile, Top- and Bottom-Quartile Parents 40
1.8 Education 40
1.9 Rural Stayers vs. Migrants vs. Urban 42
1.10 Rural-Urban Migration with Controls 43
1.11 Childhood Exposure Effects 44
1.12 Absolute Mobility: Log Income vs. Percentiles 48
1.13 Absolute Mobility: Unadjusted Income 49
1.14 Absolute Mobility: Age Controls 50
1.15 Absolute Mobility: Non-Agricultural vs. Agricultural 51

2.1 Provincial K - 12 Per-Student Public Spending vs. GDP per Capita 74
2.2 4-year College Enrollment vs. GDP per Capita 75
2.3 Timeline of Model Events and Decisions 82
2.4 Steady State Regional Quartile Compositions 97
2.A.1 Transition Paths: Location Size 117

2.A.2 Transition Paths: Income 118

3.1 Opioid Prescriptions per Capita Among Individuals Aged 20-64 129

3.2 Employment-to-Population Ratios Among Individuals Aged 19-64 130

3.3 QWI vs. QCEW 131

3.4 Log(Employment to Population) vs. Log(Prescriptions per Capita) 132

Chapter 1

Intergenerational Poverty Traps: Evidence from Indonesia

1.1 Introduction

While a large and long-established literature has sought to understand intergenerational mobility and its mechanisms, studies have largely focused on developed countries.[1] In developing countries, the high prevalence of poverty makes this an especially compelling topic, as low intergenerational mobility suggests children born into poverty may be stuck in "intergenerational poverty traps." It is unclear how well findings generalize from developed to developing countries, as many characteristics of low-income countries could potentially make the levels and mechanisms of persistence particularly interesting. Developing countries typically have a larger informal sector not encompassed by administrative tax data, which are often used for studies in developed countries. Moreover, low-income countries often have relatively large agricultural sectors and rural populations often consisting of farmers who live near subsistence consumption with limited saving and investment opportunities.

In this paper, I use longitudinal panel data from the Indonesian Family Life Survey (IFLS) from 1993 - 2014 to examine intergenerational income mobility in Indonesia. The paper contributes to the literature in two main ways. First, I build on literature on intergenerational mobility by producing estimates of mobility in a large developing country. I adopt definitions of relative and absolute mobility from previous work in order to compare mobility levels in

[1] See Corak (2013) for a relatively recent list of references, with recent influential work in the U.S. by Chetty et al. (2014b)

Indonesia with those in developed countries. Second, I study mechanisms by exploring the factors that affect the relationships between parent's and child's income such as education, location, and sector of work. While I am unable to unveil causal relationships, correlations can still help us understand the channels through which upward mobility occurs, allowing us to better inform policy targeted at promoting upward mobility.

The degree of intergenerational income mobility and the channels of transmission have been long-standing questions of economic interest. Early models of persistence can be traced to Galton (1877). Becker and Tomes (1986) establishes a robust model that incorporates many facets of the intergenerational tranmission of earnings, assets, and consumption, such as investment in human capital, bequests, and credit constraints. Since, then a number of empirical studies have produced estimates across various developed countries, such as the United States (Mazumder (2005); Corak (2006); Chetty et al. (2014b)), Canada (Corak and Heisz (1999); Grawe (2004); Chen et al. (2016)), and Nordic countries such as Sweden, Denmark, Finland, and Norway (Bjorklund and Jantti (2000); Jantti et al. (2006)), among others.

While the concept has been well-studied in developed countries, evidence on developing countries lags behind. Given previous theory and findings, it is reasonable to believe that differently-structured economies and a greater proportion of individuals living in or near poverty in developing countries could cause a large discrepancy in intergenerational mobility. Part of the issue is perhaps due to the lack of longitudinal data sets; studies in developed countries typically use administrative tax income data, which make linking income of parents and children (dependents) relatively simple. However, similar data in developing countries is often either unavailable or unrepresentative of the economy due to large informal sectors. Early studies calculated educational mobility rather than income mobility (Behrman et al. (2001)). Later studies, largely focused in South America, impute income data based on father's education and occupation using a two-step process (Ferreira and Veloso (2006); Nunez and Miranda (2010); Azevedo and Bouillon (2010)) or have small sample sizes (Ng (2007)), both of which undesirable in obtaining a precise measurement of intergenerational income persistence. Recent work by Asher et al. (2018) on India and Alesina et al. (2020) on Africa has introduced new evidence and methods specific to mobility in developing countries. I believe that my work contributes to the literature by presenting evidence regarding the degree and mechanisms of intergenerational income mobility in a large Asian country during a period of rapid development and using self-reported, individual-level longitudinal data rather than imputed data to measure income.

My paper also complements a related body of literature concerning the agricultural productivity gap, misallocation of labor, and rural-to-urban migration in developing countries (Gollin et al. (2014); Restuccia and Rogerson (2017); Hamory et al. (2021); Pulido et al. (2018); Alvarez (2020)). Though I do not directly measure productivity gaps between agricultural and non-agricultural workers, my results suggest that benefits to rural-urban migration may depend on circumstances during childhood such as parent income.

My analysis yields four main results. First, the overall intergenerational income elasticity in Indonesia is .259, which ranks low compared to both developed and other developing countries. Parent's income has no statistically significant effect on child's income after controlling for child's education and sector, suggesting that the channel of persistence works through the indirect effect of parent's income on child's education and sector rather than a direct relationship between parent's and child's income. Second, children growing up in urban locations exhibit higher levels of income persistence than rural locations under some specifications of relative mobility, though differences are not significant in the baseline regression. Third, further splitting the sample of rural children into "rural-to-urban migrants" and "always rural" individuals reveals that a large portion of upward mobility for rural children comes from migration into an urban area. Finally, despite the existence of significant neighborhood effects in previous literature and differences in urban and rural school quality, there are no neighborhood effects from growing up in urban areas in Indonesia.

The paper proceeds as follows. Section 2 discusses relevant economic conditions and trends in Indonesia during the time period. Section 3 describes the empirical framework and data, including explanation of certain data selection decisions. Section 4 reveals the main results on intergenerational mobility. Section 5 presents a series of robustness checks. Section 6 concludes.

1.2 Background

Indonesia possesses a number of features that make it an ideal setting for the study. It is a developing country with a population of 270 million and an economy comparable to other developing countries: a large agricultural sector, a low-skilled but increasingly-educated workforce, and a somewhat high, but decreasing, rate of extreme poverty. Considered one of the five "Tiger Cub Economies"[2], Indonesia has recently sustained strong and relatively stable growth. From 1980 - 1997, the country's GDP grew at an annual rate of 6.6% (World

[2]The "Tiger Cub Economies" refer to the five countries (Indonesia, Philippines, Malaysia, Vietnam, and Thailand) that followed the rapid export-driven growth of the "Asian Tigers" that preceded them.

Bank). The country fell into a financial crisis from 1997 - 2000, but has since recovered to previous levels of growth.[3]. The growth has drawn large inflows of investment and capital into the Indonesian economy, which makes Indonesia a compelling environment by allowing me to study patterns of mobility in a country on a stable path of development.

Figure 1.1 displays the trajectories of key economic metrics in Indonesia alongside the other Tiger Cub economies (excluding Malaysia and Singapore, which developed earlier) as well as India. Indonesia has sustained steadily decreasing rates of extreme poverty, agricultural work, and population living in rural areas, and a steadily increasing GDP per capita and employment-to-population ratio. Indonesia is similar to these other countries by these benchmarks. While this does not imply that findings on mobility in Indonesia should be applied everywhere, the similarity of economic conditions suggests that evidence from Indonesia may provide useful insight into factors affecting mobility in economies at similar stages of development.

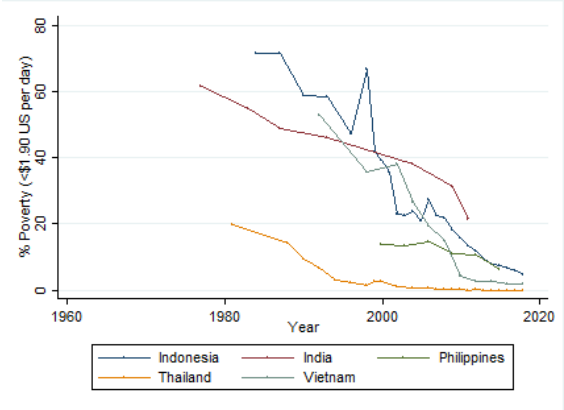
1.3 Data

1.3.1 Indonesian Family Life Survey

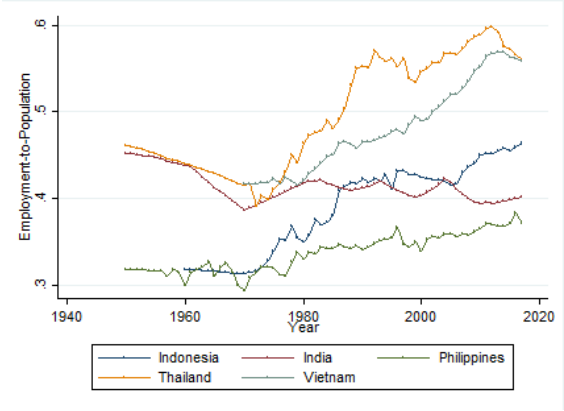
Data comes from the Indonesian Family Life Survey, a survey conducted by the RAND Corporation for the purpose of studying processes of economic development. The IFLS is an on-going longitudinal survey containing individual- and household-level data for 30,000 individuals in 13 of the 27 provinces of Indonesia, representative of 83% of the Indonesian population. The survey so far has occurred in five waves: IFLS1–IFLS5 (conducted in 1993, 1997, 2000, 2007, and 2014, respectively). The IFLS panel features a number of qualities that make it ideal for this study. It allows the user to match and track parents and children with individual-level characteristics such as income, education, migration, sector of work, and various other socioeconomic factors, as well as household-level data on location (urban or rural), income, access to health services, and more. Having individual-level survey data is particularly valuable in my work; many longitudinal surveys in developing countries keep only household-level data and do not track household members who leave. In addition, recent work shows that survey data fares better than administrative data in minimizing measurement error, especially at the bottom of the distribution distribution due to higher non-response rates (Abowd and Stinson (2013); Hokayem et al. (2015)). Moreover, the IFLS

[3]I address the effect of the financial crisis on intergenerational mobility estimates in more detail in the Robustness section.

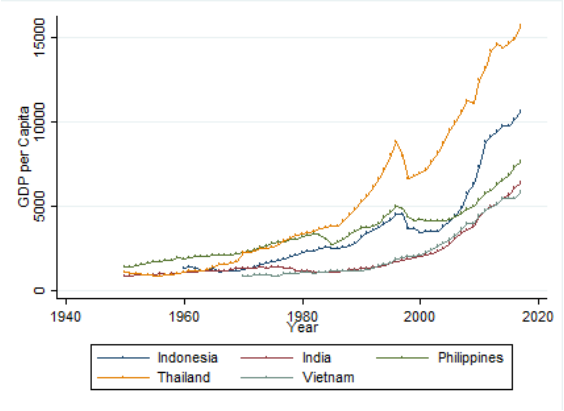
Figure 1.1: Indonesia vs. Comparison Countries



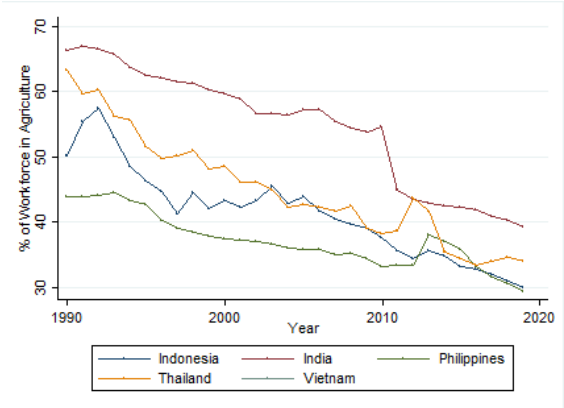
(a) Poverty rate (%)



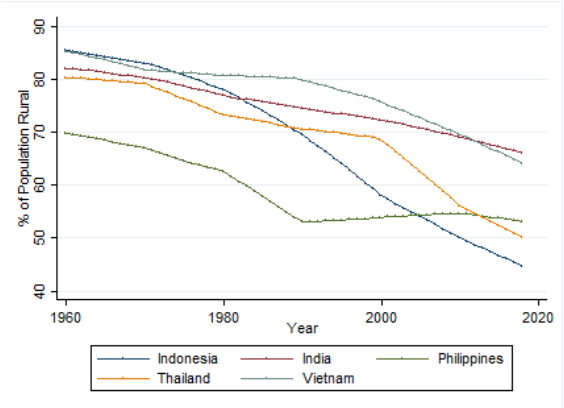
(b) Employment-to-Population Ratio



(c) GDP per Capita (2011 $US PPP)



(d) Agricultural Sector Workers (%)



(e) Rural population (%)

Table 1.1: Summary Statistics: IFLS

	Survey				
	IFLS1	IFLS2	IFLS3	IFLS4	IFLS5
Income ($US PPP)*	3,458	3,352	2,455	3,391	4,640
Age	33.82	33.48	33.13	34.57	35.99
Male (%)	48	48	48	49	50
Urban (%)	52	50	51	55	60
Agricultural (%)	-	12	14	13	13
Non-Agricultural (%)	-	88	86	87	87
Rural (%)	48	50	49	45	40
Agricultural (%)	-	59	62	60	56
Non-Agricultural (%)	-	41	38	40	44
Worker (%)	63	55	60	54	59
Years of Education	7.25	7.81	8.28	8.92	9.06
HS Graduate (%)	27	31	34	39	39
College Graduate (%)	10	11	12	13	13
N	20,163	25,015	33,269	42,321	47,974

*Income conditional on working and reporting income. Those without income omitted from calculation, not included as zeros.

has maintained a high re-contact rate; of the 7,224 households (over 22,000 individuals) originally interviewed in IFLS1, 6,341 (87.8%) of them were either interviewed in all 5 waves or died. The low attrition rate lowers the risk of selection bias.

Table 1.1 describes all individuals between the ages of 15 and 65 in each wave of the IFLS survey. A few interesting trends emerge from the cross-sections. As Indonesia develops over the course of the survey, workers become more educated, more likely to live in urban locations, and more likely to work in the non-agricultural sector. Given the country's growth during this period, these factors are likely also important inputs to upward mobility. The survey years also encompass the Asian financial crisis in 1997, which caused rampant inflation and political fallout in Indonesia and explains the decline in real incomes between IFLS1-3. While this does affect my measurement of permanent income, my analysis adjusts for this in Section 5.

1.3.2 Sample Selection

I take a number of steps to clean the raw survey data to obtain a sample for which permanent income can be measured with reasonable accuracy. First, because women's labor force participation in developing countries involves heavy selection, I restrict the sample to fathers

and sons[4]. Second, all individuals in the sample must belong to one of the following four groups in at least one survey:

1. Their primary activity the past week was working,
2. They worked, tried to work, or helped earn income for at least 1 hour the past week,
3. They have a job or business but were temporarily not working the past week, or
4. They worked a family-owned (farm or non-farm) business during the past week.

These criteria establish a consistent definition of employment using IFLS data, following Pulido et al. (2018). Finally, I exclude individuals who never report a numerical value for income between the ages of 25 - 55 in any survey. Labor force participation and income appears to stabilize across that age range for males (Figure 1.2), allowing me to reduce selection and lifecycle bias in my measurement of an individual's earnings.[5] After trimming the dataset, sons are matched to the male household head. The resulting data set includes 4,125 father-son pairs.

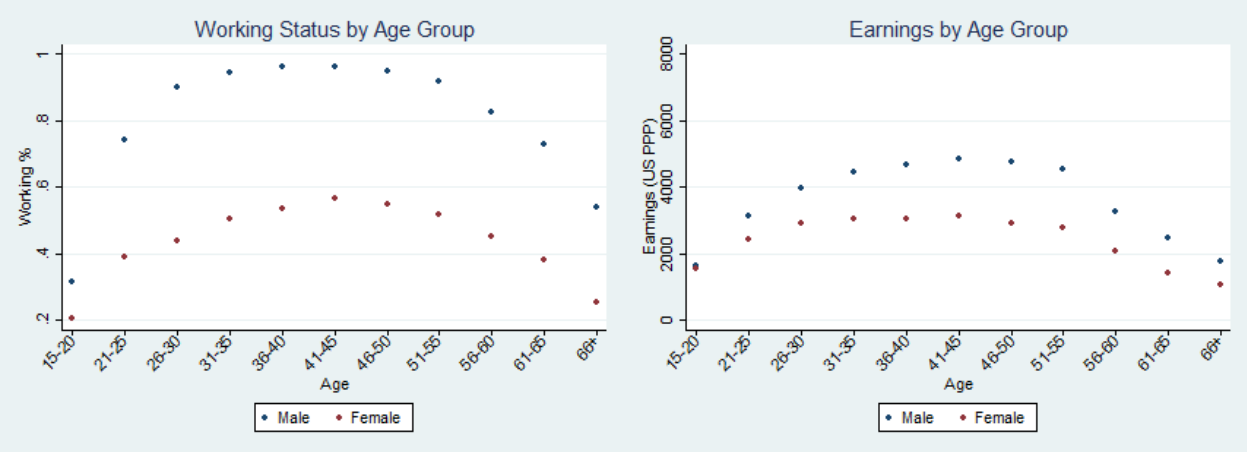


Figure 1.2: Labor Force Participation and Earnings by Gender and Age

1.3.3 Income

In this section, I describe my measurement of permanent income.[6] Since I am interested in the relationship between father's and son's earnings, I measure income on the individual

[4]I repeat the analysis including daughters in the Appendix

[5]While previous studies in developed countries use a much smaller age range, I am unable to do so in my study without significantly limiting the sample size.

[6]There are many well-documented difficulties in using survey data from developing countries to measure permanent income; I address these in Section 4.1.

level. Each individual's income is calculated by taking the average individual employment income across all years in which the individual reports working and earning a numerical value for income while between the ages of 25 - 55. The IFLS includes data on up to 15 years of income: 1988 - 2000, 2007, and 2014.[7]. I incorporate as many income observations as possible to accurately measure a person's yearly earnings potential across his prime working-aged career.[8]

Data on individual-level employment income exists in up to two sources within each survey, Book K and Book 3. In Book K, only the household head reports the income of each individual on the household roster. In Book 3, each individual over the age of 15 is interviewed separately and reports his or her own income. Because self-reported income is likely more accurate than household head-reported income, I prioritize income reported in Book 3; if income from Book 3 is unavailable, I use income from Book K. Any surveys in which an individual has no recorded income (in both Books 3 and K) are considered missing and not included when calculating average income. [9] While mobility of those with missing observations of income is theoretically interesting, I omit them in my calculations. To avoid single-year outliers introducing excessive variation into individual's permanent income, I winsorize incomes at the 2.5th and 97.5th percentiles according to the income distribution within each respective survey year.

The IFLS also includes additional data on education, work sector, migration history, and health which can be matched on the individual level, and location (urban vs. rural) which can be matched on the household level. Incomes may differ mechanically across urban and rural locations due to different prices; higher incomes in urban areas may not necessarily reflect a higher standard of living, but rather higher prices for expenditures such as housing. In my analysis, I adjust rural incomes upward using a 1981 urban-rural price index adjustment for Indonesia from Ravallion et al. (1989).[10]

Table 1.2 describes the fathers and sons in the sample. In this table, I classify an individual as "rural" ("urban") if he lives in a rural (urban) location every year in which he is surveyed. An individual is "urban during childhood" if he lives in an urban location

[7]The IFLS collected data about work history in 2007 and 2014 as well, but only asked about income in years 1988 - 2000.

[8]Bowles (1972) and Solon (1989, 1992) show that single-year transitory fluctuations in incomes will cause attenuation and thus underestimates in intergenerational income persistence.

[9]In most cases, a person did not report income if they did not work or were not surveyed in that year. However, there are a few who do work, yet do not report income. There are also some who work and report income equal to zero, but not missing. The latter are included while the former are not.

[10]Results using unadjusted incomes are similar (albeit with mechanically smaller urban-rural gaps) and reported in Section 5.

Table 1.2: Summary Statistics: Sample

	Sons	Fathers
Average Income ($US PPP)	4,914	5,229
	(4,616)	(6,071)
Median Income ($US PPP)	3,576	3,391
Urban (%)	60	55
Agricultural (%)	7	15
Non-Agricultural (%)	93	85
Rural (%)	40	45
Agricultural (%)	65	91
Non-Agricultural (%)	35	9
Urban during Childhood (%)	48	-
Years of Education	10.67	6.87
	(4.43)	(4.55)
HS Graduate (%)	54	20
College Graduate (%)	19	7
Average Age	30.1	44.8
	(3.32)	(6.52)
# Income Observations	2.13	9.76

Standard deviations in parentheses.

before turning 18, when he presumably lives with his parents and does not have a choice over location. Since intergenerational mobility concerns the relationship between childhood circumstances (i.e. initial conditions that are exogenous to the child's own decision-making) and future outcomes, "rural (urban) during childhood" is used as the criterion for an rural (urban) individual for the remainder of the paper. While it is still of interest to study differential intergenerational income patterns in urban and rural areas separately, I address this in my analysis by considering migration to an urban area during adulthood as a possible mechanism of mobility. I classify an individual as "agricultural" ("non-agricultural") if he works in the agricultural (non-agricultural) sector for his main job every year in which he works. Unlike rural or urban status, working in the agricultural sector occurs only during adulthood, so there is no way to isolate childhood conditions in this case.

The differences between generations in Table 1.2 mirror the trends in Table 1.1: sons are more likely to live in urban areas and work in non-agricultural jobs than fathers, reflecting the increase in rural-urban migration in Indonesia and many other developing countries. Likewise, education is much higher among sons than fathers: the rate of high school and college graduation for sons is almost three times higher than that of fathers. These two trends show the structural transformation of Indonesia's economy toward higher-skilled, non

manual labor-intensive work. Given this information, it may be surprising that the average income of sons in the sample is lower than that of fathers, but this likely reflects the younger age at which sons in the sample work.[11] The difference in the number of income observations between fathers and sons can be explained by the fact that individual-level income data is available from years 1988 - 2000, 2007, and 2014, and sons are likely to have fewer years as a working-aged adult in the survey.

1.4 Results

1.4.1 Economy-Wide Intergenerational Mobility

In this section, I report measures of intergenerational mobility for the entire economy and incorporate evidence regarding the channels of transmission. I follow much of the existing literature in using the intergenerational income elasticity (IGE) as my main measure of intergenerational mobility, which is the estimated coefficient of a regression of the log of child's income on the log of parent's income. That is, IGE is the estimated value of β in the following OLS regression:

$$y_{i,c} = \alpha + \beta y_{i,p} + \varepsilon_i \tag{1.1}$$

where $y_{i,c}$ is the natural log of child's income, $y_{i,p}$ is the natural log of parent's income, and i indexes a parent-child observation. Running this regression for the full sample yields an estimate of $\beta = .259$ (s.e. $= .031$). This estimate is relatively small compared to both estimates for the US, which are around .5 in most studies, and estimates for other developing countries. Table 1.3 lists estimates of intergenerational elasticity in other countries (Corak (2013)). At .259, Indonesia ranks among higher-mobility countries in terms of IGE. Perhaps most notably, it ranks highest out of all non-OECD countries on the list; though developing countries tend to have low mobility, Indonesia appears to be somewhat of an exception. Whether this results from the two-stage approach used to estimate income in other studies of developing countries, lower income persistence in Indonesia, or another source (such as measurement error and attenuation bias) is unclear. This estimate runs somewhat contrary to previous findings in that poorer countries tend to have higher inequality (World Bank), and inequality tends to have a positive relationship with IGE (and thus a negative relationship

[11] Since age changes across the sample and I average income over many years, I do not control for age in my baseline analysis. Section 5 includes regressions controlling for *average* age.

with mobility).[12] The Gini coefficient in my sample is .448, which ranks similarly to many other developing countries according to World Bank estimates.

Table 1.3: Estimates of Intergenerational Elasticity by Country

Country	Estimate
Peru	.67
China	.60
Brazil	.58
Chile	.52
United Kingdom	.50
Italy	.50
Argentina	.49
United States	.47
Switzerland	.46
Pakistan	.46
Singapore	.44
France	.41
Spain	.40
Japan	.34
Germany	.32
New Zealand	.29
Sweden	.27
Australia	.26
Canada	.19
Finland	.18
Norway	.17
Denmark	.15

Source: Corak (2013)

Previous work shows that the IGE is highly sensitive to a number of empirical issues, such as lifecycle bias, including zeros, and differences in the variances of parent's and child's income distributions. Chetty et al. (2014b) as well as other studies in recent literature argue that using a regression of child income *rank* on parent income *rank* is less susceptible to some of these issues. That is, rather than using equation (1), relative mobility is estimated by the coefficient δ in the regression:

$$p_{i,c} = \gamma + \delta p_{i,p} + \varepsilon_{2i} \tag{1.2}$$

[12]The relationship between inequality and IGE is well-documented through studies of the Great Gatsby Curve, which plots a country's IGE against its Gini coefficient. A table with World Bank estimates of the Gini coefficient is included in the Appendix.

where $p_{i,p}$ is parent's percentile in the parent's income distribution and $p_{i,c}$ is child's percentile in the child's income distribution. This regression yields a coefficient of $\delta = .268$ (s.e. $= .028$), still a low estimate relative to other countries both rich and poor. While IGE and rank mobility are conceptually different, they are both useful measures of mobility (Mazumder (2001)). Since IGE provides a longer history of literature with comparable estimates, I focus my attention on IGE for the remainder of the paper.[13]

As noted above, one important concern is the possibility that low estimates of the IGE are explained by measurement error in parent income causing attenuation in relative mobility estimates. Following Friedman (1957), most analyses of current income employ a standard statistical framework to account for transitory fluctuations in income:

$$y_{i,p,s} = y_{i,p} + e_{i,p,s} \tag{3}$$

$$y_{i,c,t} = y_{i,c} + e_{i,c,t} \tag{4}$$

$$e_{i,p,s} = \rho e_{i,p,s-1} + \xi_{i,s} \tag{5}$$

where $y_{i,p,s}$ denotes parent's earnings in year s and $y_{i,c,t}$ denotes child's earnings in year t. Both consist of a permanent income component ($y_{i,p}$ and $y_{i,c}$ for parent and child, respectively) and a transitory shock to earnings ($e_{i,p,s}$ and $e_{i,c,t}$). Equation (5) imposes a first-order autoregressive structure on the parent's transitory shock. Obtaining a consistent estimate for β in Equation (1) requires an accurate measure of parent's permanent income $y_{i,p}$. In practice, researchers approximate $y_{i,p}$ by averaging income over a large number of years to minimize these shocks, though the estimate of β will still be attenuated by a factor that decreases with the number of years in the average (Mazumder (2005); see Appendix for further discussion). Therefore, using one or few observations to noisily measure parent's permanent income may be problematic.

Previous studies of intergenerational mobility have addressed this issue in various ways. One useful exercise is to find the coefficient of educational mobility by regressing child's education on parent's education under the assumption that education is a less noisy estimate of productivity. Hertz et al. (2008) produces estimates for 42 countries between 1994 and 2004 and finds that the unweighted average coefficients across developing countries in Latin America (.79), Africa (.80), and select countries in Asia (.69) were higher than those in Nordic countries (.34) and the United States (.54). Repeating this exercise for Indonesia yields

[13] I address rank-rank and its relationship with IGE in more detail in Section 5. IGE and rank mobility produce roughly similar results in Indonesia.

an estimate of .493, which falls more in line with estimates from developed rather than developing countries. Another common method using education to address measurement error is to instrument for parents' earnings. Solon (1992) argues that an IV estimate using father's years of education to instrument for father's log earnings serves as an upper bound for the true IGE; the estimate from my sample is .512 (se = .049).

Another method excludes observations from individuals with fewer than two years of income in the sample, so as to decrease the effect of attenuation (Mazumder (2005)). This method yields an estimate of .233 (se = .027), providing some evidence that low IGE estimates in my study are not being driven by attenuation. Finally, given the possibility of greater measurement error at the low end of the income distribution as well as outliers and influential points at each end of the parent income distribution, I run the regression on the sample of parents between the 10th and 90th percentiles of parent income, which yields an estimate of .270 (s.e. = .021). These estimates are still relatively low and more consistent with developed countries. The estimates remaining generally consistent provides evidence that low IGE estimates in my study are not being driven by attenuation, and instead that sons in Indonesia indeed exhibit relatively low income persistence.

Another concern regarding IGE measurement is lifecycle bias, which affects both the individual's observed income and the variance of the transitory component of father's income may change over the life cycle. Individual earnings typically flatten out around ages 30 - 50, and Baker and Solon (2003) and Mazumder (2001) find that the variance of income deviations from the mean follows a U-shape with age, a trend that appears to hold in the IFLS data (Figure 1.3). The result is that measures of IGE will also depend on the age distribution of working fathers in the sample. Typical methods of addressing this issue to minimize the effects of transitory shocks include incorporating more years of father's earnings and centering the averages of children's and parents' incomes around the age of 40, when observed incomes are closest to permanent incomes. In my study, it is difficult to deal with lifecycle bias using any of these methods without severely limiting the size of the dataset. However, Figure 1.3 shows that the parental age distribution lines up rather well with ages at which transitory shocks have the smallest variance.

Table 1.4 below summarizes our measures of relative mobility for the sample of fathers and sons.

Channels of Transmission

An important goal of studying intergenerational mobility is to uncover the channels through which persistence occurs. Starting with the seminal work of Becker and Tomes (1986), one

Figure 1.3: Lifecycle Bias, Indonesia

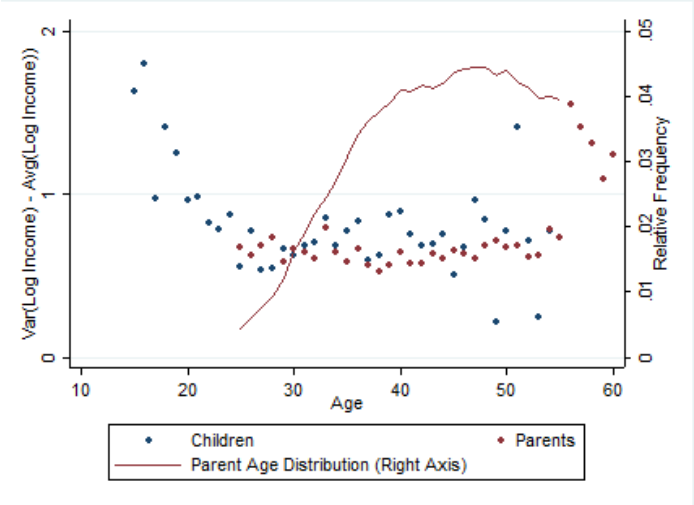


Table 1.4: Estimates of Relative Mobility

Measure	Value	N
IGE	.259***	4,125
	(.031)	
IGE: $10^{th} - 90^{th}$	.306***	3,303
	(.043)	
IGE: 2 or more observations	.233***	1,818
	(.027)	
Rank-rank	.268***	4,125
	(.028)	
Rank-rank: $10^{th} - 90^{th}$	.270***	3,303
	(.021)	
IGE: IV	.512***	3,932
	(.049)	
Education	.493***	3,893
	(.012)	

Standard errors in parentheses. *=.05, **=.01, *** = .001.

focus in the literature has been to assess the relative importance of genetic and economic channels of transmission. Genetic transmission refers to parents with high innate ability tending to have children with high innate ability; if ability is correlated with income, then genetic transmission will cause income to be correlated across generations. Economic transmission reflects the possibility that holding innate ability constant, higher parental income yields beneficial effects on children's income through investment in children. Becker and Tomes

(1986) and a large subsequent literature emphasize the investment channel; in a world with credit constraints, investment in a child's education will be affected by parental resources. Other potential economic channels exist as well: if personal networks affect economic opportunities, then children with higher income parents might receive superior opportunities via their parents' connections. If parents are viewed as role models, then parental income might influence a child's perception of available opportunities.

Absent an exogenous measure of innate ability, it is not possible to distinguish between genetic and economic channels; children of high income parents may obtain more education, but this could be driven by high innate ability or high income. But the data do allow me to shed light on possible channels of transmission by studying how the effects of parental variables are changed when I include controls for certain child outcomes. I focus on two characteristics, years of education and sector of employment (agricultural or non-agricultural). It is well-established in many settings that both education and sector/occupation have significant effects on income, and this is true in Indonesia as well. Results from this analysis

Table 1.5: IGE with Controls

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Parent Income)	0.259***	0.191***	0.224***	0.165**	0.110*	0.151**	0.115
	(0.031)	(0.040)	(0.045)	(0.054)	(0.053)	(0.058)	(0.061)
Father's Education		0.035***		0.036***	0.010	0.030***	0.007
		(0.007)		(0.009)	(0.009)	(0.009)	(0.008)
Father Agricultural			-0.308***	-0.228***	-0.181**	0.005	-0.021
			(0.067)	(0.066)	(0.067)	(0.090)	(0.089)
Son's Education					0.068***		0.062***
					(0.008)		(0.009)
Son Agricultural						-0.514***	-0.353**
						(0.103)	(0.109)
R^2	0.028	0.038	0.044	0.052	0.075	0.071	0.093
N	4125	3932	2754	2741	2714	2071	2070

Standard errors in parentheses. *=.05, **=.01, *** = .001.

are presented in Table 1.5. Columns (1)-(4) show how child's income relates to parental characteristics. The first column simply reports the IGE coefficient. Columns (2) - (4) add controls for father's education, father's agricultural status, and both, respectively. Father's income and education are positively correlated with child's income, and father's agricultural status is negatively correlated with child's income. One should exercise caution in interpreting the coefficients, as the result could be mechanically driven by large measurement error

in father's income relative to education.[14] But interestingly, column (4) shows that all three variables have a statistically significant relationship with child's income.

Columns (5) - (7) in Table 1.5 further condition the regression on education and agricultural status of the child, altering the effects of parental variables. Columns (5) and (6) show separately how controlling for son's education and agricultural status affect the estimates on parental variables. When only including child's agricultural status, parental education remains significant, and when only including child's education, parental agricultural status remains significant. However, when including both in column (7), both father's education and sector cease to be significant. Taken together, these results suggest that the intergenerational transmission channel operates through education and sector of work. That is, to the extent that father's income (and other observables) has causal effects on child's income, these effects manifest largely through the intermediate channels of educational attainment and agricultural status, and not directly on income.

Table 1.6: Determinants of Child Observables

Variable	(1)	(2)	(3)	(4)
Father's Education	0.493***	0.407***	0.444***	0.388***
	(0.012)	(0.015)	(0.016)	(0.018)
Log(Parent Income)		0.772***		0.624***
		(0.073)		(0.089)
Father Agricultural			-1.126***	-0.752***
			(0.164)	(0.169)
R^2	0.255	0.281	0.277	0.292
N	3893	3893	2714	2714

(a) Child's Education

Variable	(1)	(2)	(3)	(4)
Father Agricultural	0.459***	0.435***	0.407***	0.401***
	(0.014)	(0.018)	(0.017)	(0.019)
Log(Parent Income)		-0.023**		-0.008
		(0.007)		(0.008)
Father's Education			-0.013***	-0.012***
			(0.002)	(0.002)
N	2079	2079	2071	2071

(b) Child Agricultural (Average Marginal Effects)

Standard errors in parentheses. *=.05, **=.01, *** = .001.

[14] Consider an extreme case in which father's (true) income and father's education are perfectly correlated, but father's income is measured with error. When including both in a regression, education will appear to be more predictive simply due to less noisy measurement.

To better understand these intermediate channels, I next present evidence on the relationship between child's education and work sector and father's income, education, and work sector. Table 1.6a reports coefficients of an OLS regression of child's education on parent observables, while 1.6b reports average marginal effects from a logistic regression of a dummy for child's agricultural status on the same explanatory variables. Table 1.6a shows that child's education is statistically significantly related to father's income, education, and work sector, with father's education providing the most explanatory power. The effects are economically significant; an additional year of education for the father increases the son's education by an expected .39 years. Though the effect of father's income on son's education is statistically significant, the magnitude of the effect is rather small: a 10% increase in father's income is associated with an increase in son's education by just .06 years.

Table 1.6b shows that father's work sector and education are also strongly associated with children's work sector. Again, the estimated effects are economically significant: the father working in the agricultural sector increases the child's probability of agricultural work by 40%[15]. Interestingly, the effect of father's income is not significant after controlling for education. This suggests that child agricultural status is largely inherited from father's agricultural status and is not attributed to a financial decision. In fact, 95% (470/495) of children who work in the agricultural sector also had parents in the agricultural sector.

1.4.2 Urban and Rural Intergenerational Mobility

In this section, I present estimates of relative mobility for urban and rural subsamples of children (β_u and β_r, respectively), investigate the channels of transmission for each, and discuss potential reasons for differences in relative mobility between the groups. Then, I discuss differences in *absolute* mobility and highlight the correlates of upward mobility for each group.

Relative Mobility

While IGE for the full sample suggests a moderately strong relative mobility compared to other countries, this does not guarantee that different groups in Indonesia exhibit the same degree of mobility. Table 1.7 reports relative mobility estimates for urban and rural children

[15]Recall that for an individual to be coded as "agricultural," he must work in the agricultural sector for *all* income observations.

using specifications from Section 4.1. [16] Non-upper bound IGE estimates for both urban and rural children remain low relative to other country-wide estimates. The baseline elasticity is .256, and estimates range from .240 - .305 for urban children, while rural IGE estimates range from .162 - .284. Some specifications suggest that urban and rural estimates of relative mobility are significantly different at the 5% level, which I discuss in a later section.

Table 1.7: Urban and Rural Estimates of Relative Mobility

Measure	Urban	Rural	Difference
IGE	.256**	.162**	.094
	(.041)	(.057)	(.070)
IGE: $10^{th} - 90^{th}$	.240**	.284***	-.044
	(.065)	(.085)	(.107)
IGE: 2 or more observations	.305**	.162***	.142**
	(.046)	(.038)	(.060)
Rank-rank	.290**	.218***	.072*
	(.028)	(.028)	(.034)
Rank-rank: $10^{th} - 90^{th}$	.276**	.268***	.008
	(.036)	(.036)	(.051)
IGE: IV	.627**	.516***	.111†
	(.089)	(.167)	-
Education	.468**	.442***	.026
	(.019)	(.023)	(.029)

Standard errors in parentheses. *=.05, **=.01, *** = .001.

† The limiting distribution of this comparison is $\chi^2(1)$, and the χ^2-statistic is .47 with a corresponding p-value of .492.

Channels of Transmission

Tables 1.8 and 1.9 reproduce Tables 1.5 and 1.6, respectively, using subsamples of urban and rural children. In Table 1.10, the coefficients on father's income, education, and sector are no longer significant for both urban and rural children after adding child education and sector controls, similarly to results in Table 1.5. The most noticeable difference from results for the full sample is that child's agricultural status is insignificant for urban children, though it is still significant and associated with a 33.4% decrease in income for rural children. This is perhaps unsurprising given the low prevalence of agricultural work for urban children. This suggests that fathers affect children's incomes through intermediate channels, though they

[16] For the rank-rank regression, urban and rural income ranks are given by the *national* parent and child income distributions, not separate income distributions for urban and rural parents and children. Results for relative mobility are similar in both cases and are provided in the Appendix.

differ across groups: education for urban children and education and agricultural status for rural children.

Table 1.8: IGE with Controls: Urban vs. Rural

Panel A: Urban

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Parent Income)	0.256***	0.140**	0.253***	0.143*	0.095	0.122*	0.075
	(0.041)	(0.049)	(0.053)	(0.056)	(0.055)	(0.060)	(0.060)
Father's Education		0.045***		0.042***	0.009	0.039**	0.009
		(0.009)		(0.011)	(0.011)	(0.012)	(0.011)
Father Agricultural			-0.151	-0.066	0.050	0.098	0.069
			(0.099)	(0.103)	(0.106)	(0.124)	(0.122)
Son's Education					0.082***		0.080***
					(0.011)		(0.013)
Son Agricultural						-0.444**	-0.109
						(0.155)	(0.168)
R^2	0.023	0.040	0.027	0.041	0.080	0.039	0.072
N	1419	1389	1036	1035	1025	909	908

Panel B: Rural

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Parent Income)	0.162**	0.118	0.145*	0.103	0.068	0.077	0.049
	(0.057)	(0.066)	(0.073)	(0.085)	(0.083)	(0.072)	(0.075)
Father's Education		0.037*		0.046**	0.024	0.045**	0.022
		(0.014)		(0.017)	(0.018)	(0.015)	(0.015)
Father Agricultural			-0.148	-0.094	-0.084	-0.007	-0.049
			(0.153)	(0.151)	(0.151)	(0.158)	(0.157)
Son's Education					0.058***		0.068***
					(0.014)		(0.015)
Son Agricultural						-0.481**	-0.334*
						(0.149)	(0.153)
R^2	0.009	0.016	0.010	0.019	0.034	0.054	0.081
N	1553	1514	1022	1011	1008	646	646

Standard errors in parentheses. *=.05, **=.01, *** = .001.

Table 1.9 presents evidence on the intermediate channels. As before, a one-year increase in father's education is associated with roughly a .45-year increase in rural and urban child's education, whereas the effect of parent's income is statistically significant but small. In addition, child's agricultural status largely reflects father's agricultural status and not income, though the correlation is much stronger in rural areas: rural children born to fathers in the agricultural sector are about 64% more likely to work in the agricultural sector themselves, whereas urban children are only 14% more likely.

Differences in Urban and Rural Relative Mobility

Differences in relative mobility estimates for urban and rural children are not statistically significant for every specification in Table 1.9. However, there is some suggestive evidence that relative mobility is lower in urban than rural areas. In this section, I discuss reasons as to why urban and rural mobility could potentially differ.

First, the difference could be mechanical. β can be written as:

$$\beta = \frac{cov(y_{i,p}, y_{i,c})}{var(y_{i,p})}$$

and we can rewrite the ratio of IGEs for urban and rural children as:

$$\frac{\beta_u}{\beta_r} = \left(\frac{cov(y_{i,p,u}, y_{i,c,u})}{var(y_{i,p,u})}\right) \Big/ \left(\frac{cov(y_{i,p,r}, y_{i,p,r})}{var(y_{i,p,r})}\right) = \left(\frac{cov(y_{i,p,u}, y_{i,c,u})}{cov(y_{i,p,r}, y_{i,p,r})}\right)\left(\frac{var(y_{i,p,r})}{var(y_{i,p,u})}\right)$$

The difference in elasticities could thus be generated by attenuation from either noisy realizations of income (e.g. large transitory shocks) or errors in measurement that increase variance in incomes. IGE being measured as lower in rural areas could simply be driven by higher attenuation if rural incomes are measured with more error than urban incomes (for example, if rural locations have higher rates of informal work and are more subject to transient shocks such as weather). However, the ratio of rural-urban variances of parent income in the sample is $\frac{var_{i,p,r}}{var_{i,p,u}} = \frac{1.028}{.744} = 1.382$, far lower than the urban-rural IGE ratio ($\frac{\beta_u}{\beta_r} = \frac{.256}{.162} = 1.580$), suggesting that differences in variance alone cannot fully explain the difference in IGE. In addition, insofar as rural incomes are subject to higher measurement error, methods to reduce the effects of measurement error should increase rural estimates of relative mobility *more* than urban estimates, decreasing the disparity between them. This is not the case–using IV, rank-rank, or more observations of income makes differences in relative mobility larger and significant in some specifications. While the estimates themselves could potentially be attenuated, the relative comparison between the groups should not have the same issue.

Thus, one must look beyond mechanical sources in studying disparities in relative mobility. Given that father's income appears to affect child's income through intermediate channels, it is important to consider each channel individually: how father's income corresponds to children's characteristics, and separately, how children's characteristics correspond to children's income. Table 1.9, column (2) examines the first part of the channel by regress-

ing child's education and agricultural status on parent's income. Tables 8a and 8c show that a 10% increase in parent's income is associated with an increase in child's education by .184 and .121 years for urban and rural children, respectively, a difference which is significant at the 1% level.

Looking at the relationship between parent's income and child's level of schooling attained provides a more descriptive understanding of the relationship. Figure 1.4 shows that urban children are more likely than rural children to graduate high school at each parent income percentile. Moreover, the relationship between income and rate of college graduation appears to be stronger for urban children: while children of below-median parents graduate college at extremely low rates, higher parent incomes above the median increase the college graduation rate at a significantly higher rate for urban children than rural children. Whether these features reflect disparities in physical or financial access to education, genetic transmission of higher ability, or another factor, the relationship between parent's income and child's education appears to be stronger for children growing up in urban areas.

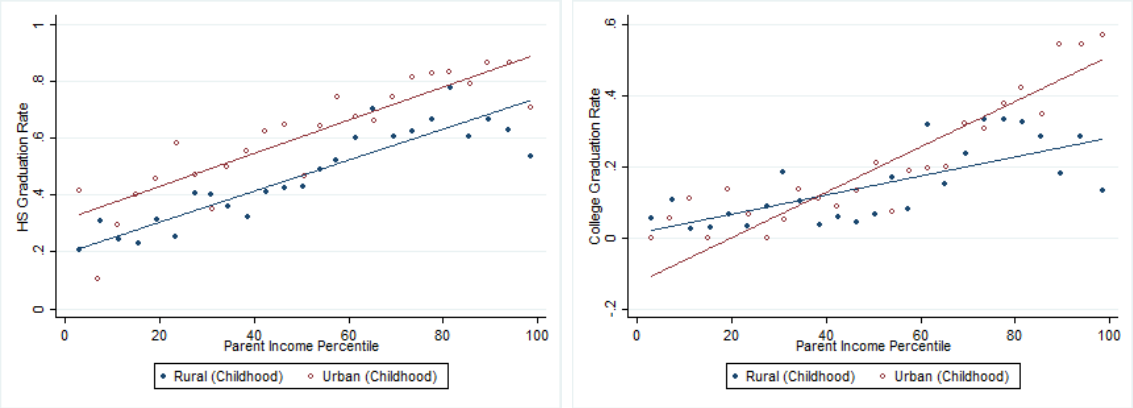


Figure 1.4: HS and College Graduation Rates by Parent Income

Column (7) of Table 1.8 shows how intermediate channels affect relative mobility for each subset of children. I begin with the contribution of education to differences in relative mobility. From the results in this regression, it is unclear whether education has different effects for urban and rural children. However, Mincer regression results in Table 1.10 show that urban children receive a significantly higher return on an additional year of education than rural children, largely driven by the difference in work sector composition for children growing up in each location–children working in the agricultural sector (and more likely to grow up and live in rural areas) receive a lower return to education due to the low-skilled nature of agricultural work. Combining results from Tables 1.7—1.9, the relationships between (i) parent's income and child's education and (ii) child's education and child's income

appear to be stronger for urban than rural children, contributing to a higher IGE for urban children.

Absolute Mobility

Absolute mobility concerns differences in child income *conditional* on parent income rather than correlations between parent and child incomes. In particular, absolute mobility as a concept helps define a metric for *upward* mobility: fixing a level (or small range) of parent income, what characteristics are associated with higher child incomes, and how do these characteristics differ across groups? While absolute mobility for the full sample provides no additional information to relative mobility, differences in absolute mobility between urban and rural children allows for level comparisons. I choose to report estimates of average child income at two benchmarks: \$1.90 US PPP per day, which is the World Bank definition of the international poverty line; and \$10 US PPP per day, which has been widely used as the lower threshold for the "middle class" in developing countries (World Bank; Birdsall, 2010; Birdsall, Lustig, and Meyer 2013).[17]

Table 1.11 reports average daily incomes for children born at the poverty and middle-class thresholds. In the full sample, children born at the poverty line are expected to earn approximately \$4.40 per day. Children born at the middle class threshold are expected to earn about \$6.65 per day. Discrepancies in absolute mobility become clear when comparing across urban and rural groups. [18] While rural and urban children born to parents in poverty are expected to earn roughly the same, higher IGE for urban children implies that a difference emerges at higher incomes. On average, urban children born to middle class parents earn about \$7.16/day while rural children earn \$5.82, and the difference increases to about 35 log points for the richest parents. [19]

The positive gap between urban and rural children contrasts Chetty et al. (2014b), who study absolute mobility gaps at the 25^{th} percentile of parent income and find that children from urban areas exhibit *lower* absolute upward mobility (41.7) than those from rural areas

[17] Providing one benchmark and the appropriate estimate of relative mobility is sufficient to determine the average income levels at all benchmarks since they are taken from the regression line, but I report both benchmarks simply as reference levels of interest.

[18] It is important for this section to recall that rural incomes are scaled upward according to the 1981 price index proposed by Ravallion et al. (1989). My reported income differences thus reflect differences in *adjusted* incomes, whereas differences using raw incomes would be greater.

[19] The fact that children on average earn less than their middle class parents seems surprising, but this is because children in the sample are on average younger than parents, and a middle-class parent earns an above-average income, so that their children earning less represents regression to the mean.

(45.8) in the US. [20] The result is perhaps surprising, but emphasizes that countries at different stages of development undoubtedly have different economic conditions associated with urban and rural locations, so that outcomes in a developed country cannot be generalized to developing countries.

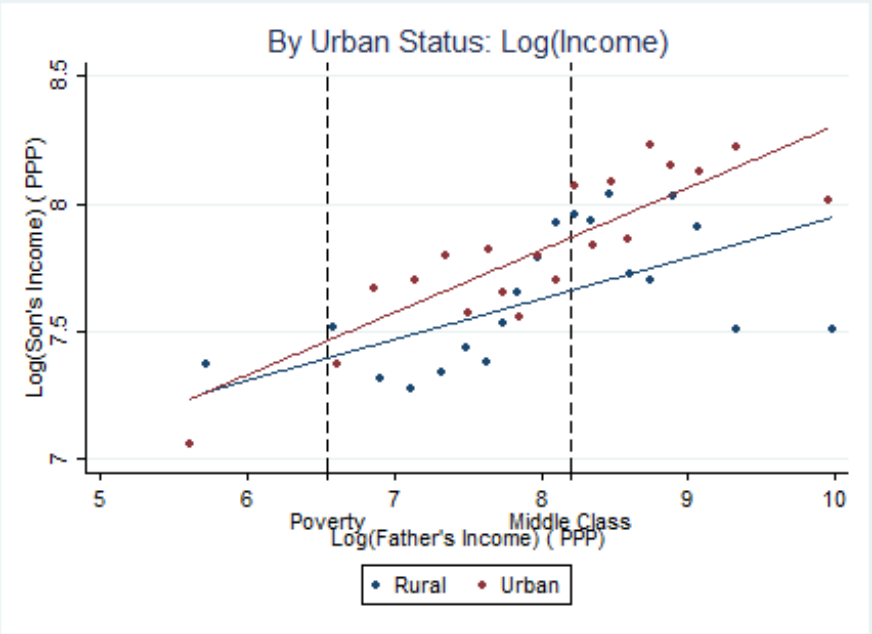


Figure 1.5: Mobility by Urban Status

There are many potential theories justifying higher absolute mobility for urban children. Individuals living in urban areas could be selected on the basis of education, productivity, or other individual qualities. In addition, growing up in an urban area may provide children access to higher-quality public services and social networks as a child and employment networks as an adult (a *network* or *local area* effect). I analyze these by incorporating data on education (for selection effects) and rural-urban migration (for network/area effects).

Education

Education affects human capital and income and thus plays an indisputable role in upward mobility. Corak (2013) finds that countries with higher returns to schooling also have higher IGEs. This further entrenches the importance of education: as education becomes more valuable, people at the bottom of the distribution with less education suffer relatively more, leading to worse mobility. In a developing country where the poorest individuals may lack

[20] The analogous percentile estimates in my sample are 46.81 for urban children and 45.38 for rural children.

quality (or quantity) of schooling through lack of financial access or otherwise, this could drive large wedges between the rich and poor. As previously reported in Table 1.10, Mincer regressions also suggest differential returns to education for children in urban and rural areas, further driving mobility differences.

How much of the urban-rural absolute mobility gap can be explained by differences in education alone? The gap ranges up to 35 log points at the top of the parental income distribution (Figure 1.5), but controlling for education causes this gap to shrink significantly down to 6 log points for the richest parents, a decrease of 83% (Figure 1.6). Table 1.9, column (2) suggests that the education of children in urban areas appears more responsive to changes in parent's income, so it is unsurprising that controlling for education lowers income more for urban children relative to rural children born to rich parents.

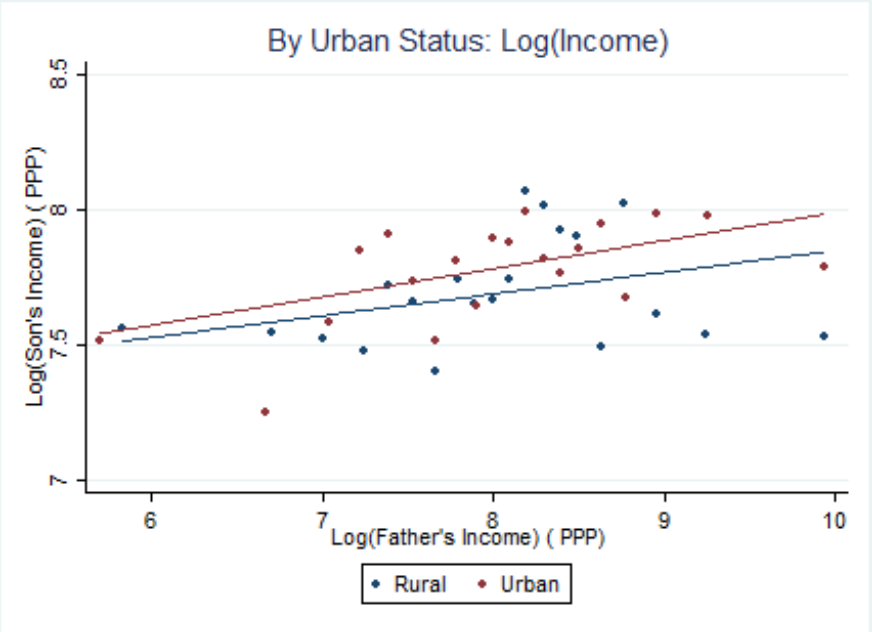


Figure 1.6: Mobility with Education Controls

The results indicate that children at different points in the parental income distribution potentially benefit differently from the mechanisms affecting upward mobility. To study this, I first bin parents and children into quartiles. Then, fixing children born to parents in the top and bottom quartiles (the least and most "fortunate" children), I observe changes in average characteristics conditional on each child income quartile to distinguish characteristics which are associated with upward mobility. The purpose of this exercise is not to identify causal relationships about the mechanisms of upward mobility, but rather to study the characteristics of individuals who demonstrate upward mobility conditional on their childhood

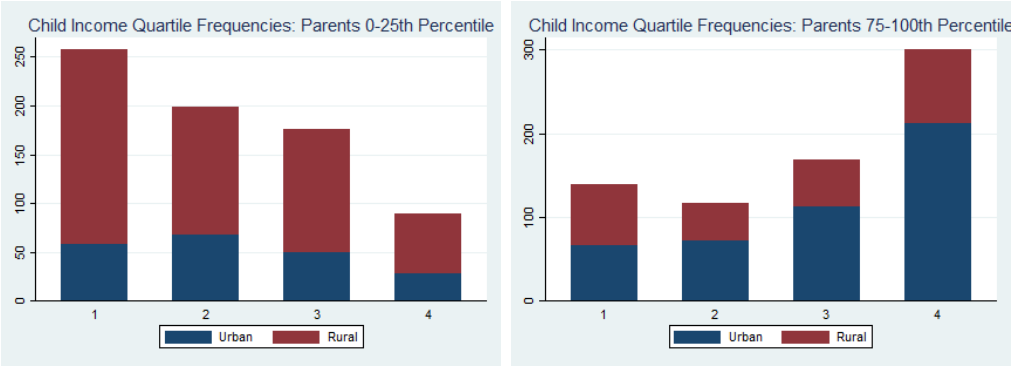
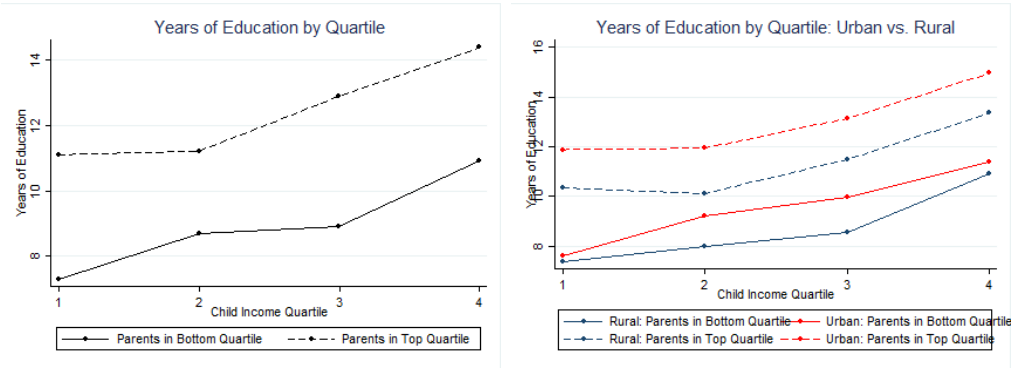


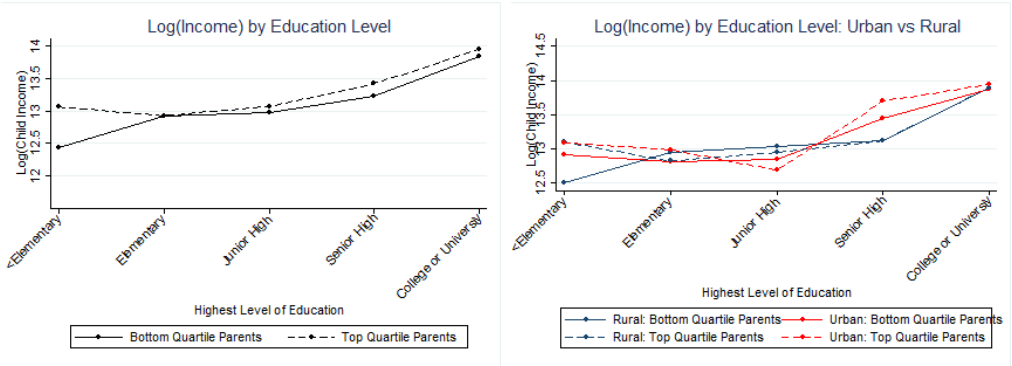
Figure 1.7: Urban/Rural Children by Income Quartile, Top- and Bottom-Quartile Parents

circumstances. To better understand the composition of each quartile, Figure 1.7 shows the rural-urban and income quartile distribution of children within the top and bottom parent quartiles, respectively. Most children in the bottom parent quartile grow up in rural areas, while most of the children in the richest parent quartile grow up urban.

Figure 1.8: Education



(a) Log(Income) by Parent Income Quartile



(b) Income by Education

Figure 1.8a shows average years of education conditional on child outcomes. Education is clearly associated with upward mobility for all children, with urban children on average receiving more education than rural children in the same quartile. The urban-rural education gap in education is larger for children of the richest than the poorest parents: urban children receive 1.5 - 2 more years of education than rural children within each child quartile. The gap in education between the rich and poor is rather striking–children with poor parents on average receive less education than children with rich parents, regardless of the child's own economic success. In fact, even the most successful children earn less than a high school degree on average. Figure 1.8b shows average log incomes conditional on child education level. Conditional on education level, parent incomes do not appear to have a strong effect on children's incomes. Combining the two graphs suggests that successful children born to the poorest parents obtain their success through some other means in addition to education.

Rural-to-Urban Migration

The availability of high-paying jobs and amenities in urban locations has led to an explosion of rural-to-urban migration in Indonesia as well as other similar developing countries. A large body of literature studying spatial sorting of labor finds substantial cross-sectional productivity gaps across sectors and locations (Miguel and Hamory (2009), Alvarez (2020)), which are consistent with large benefits to migration that are left unexploited in developing countries due to migration costs (Bryan et al. (2014), Pulido et al. (2018), Bryan and Morten (2019)). However, recent work has questioned the magnitude of the urban-to-rural (or non-agricultural-to-agricultural) gap after adjusting for selection (Hamory et al. (2021), Alvarez (2020)). Moreover, other papers have found that selective migration accounts for little of the skill difference between less dense and denser areas and suggest a larger role for "sorting at birth," or sorting among the parent generation (Mion and Naticchioni (2009), Combes et al. (2012)).

In this section, I exploit rural-urban migration both as a method to identify neighborhood effects during childhood and as a potential channel for upward mobility. I first split the subsample of rural children into two groups. I define an *"always rural"* individual as a child who stays in a rural location throughout the survey, both during childhood and adulthood. I define a *rural-to-urban migrant* as a child who spends his childhood in a rural location and is observed in an urban location at least once after turning 18. An important characteristic of these migrants is that they have decided to move as adults, not as children when they

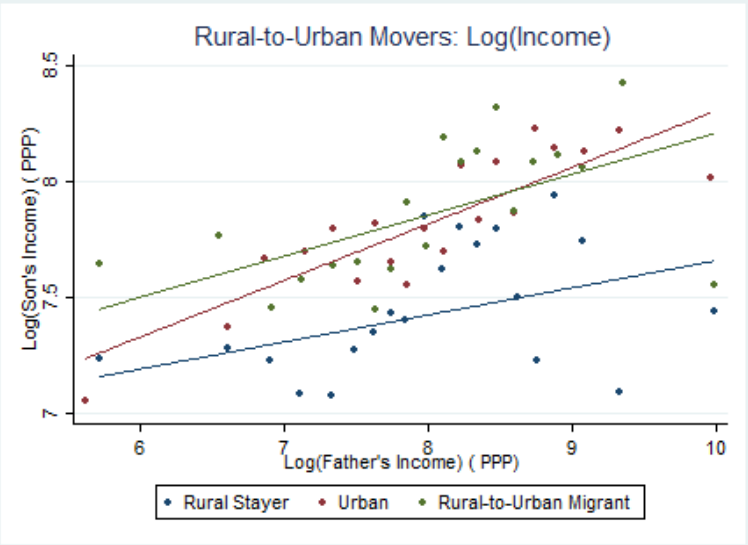


Figure 1.9: Rural Stayers vs. Migrants vs. Urban

presumably lack control over their location.[21] 45.4% of rural children in the sample are rural-to-urban migrants.

Figure 1.9 shows binned scatterplots for rural stayers, urban children, and rural-to-urban migrants. Rural-to-urban migrants appear to have extremely similar absolute mobility levels to urban children. This suggests a minimal influence of childhood area effects: migrants, despite growing up in rural areas, on average fare roughly just as well as children growing up in urban areas at each level of parent income. In addition, rural nonmigrants suffer a wide disparity in absolute mobility compared to urban children, up to 65 log points at the top of the parent income distribution. This represents a vast difference from the 35-log point gap between urban and rural children without separating for migrants in the previous analysis, suggesting that a significant portion of upward mobility can be attributed to rural children migrating to urban areas to earn higher incomes rather than rural children capitalizing on opportunities available in rural areas.

However, this does not imply that any child in a rural area would experience upward mobility simply by moving to an urban area. Rural-urban migration inevitably encounters the issue of selection bias—individuals consciously decide to migrate based on perceptions of the risks and benefits. Table 1.12 shows that rural-to-urban migrants are likely to have richer, more educated, and non-agricultural parents, and are more educated and substantially less

[21] While migration in the opposite direction could also be used to help identify effects of growing up in each location, the number of children who undergo this move is too small to include in any analysis.

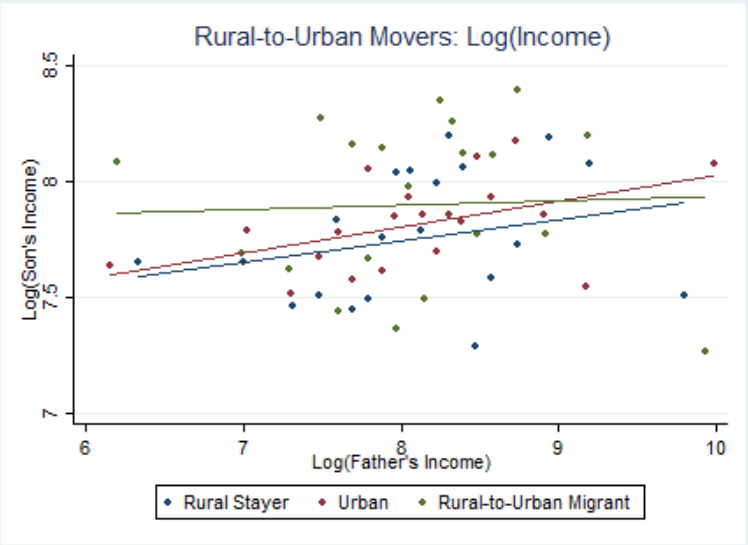


Figure 1.10: Rural-Urban Migration with Controls

likely to work in the agricultural sector themselves. However, they are not as advantaged as children growing up in urban areas on these metrics.

Figure 1.10 builds on Figure 1.9 by controlling for child's education and work sector. Migrants born to poor parents seem to fare better than other children with similar economic backgrounds. There are a few potential interpretations. Part of this could be a high return to migration. If one believes that controlling for parent's and children's sectors and education levels captures all effects of selection, then this implies that rural children of the poor can reap large benefits simply by migrating to an urban area. Previous research finds evidence consistent with this fact; Pulido et al. (2018)find that there are large benefits to migration that are not being realized due to the existence of substantial barriers beyond selection on observables. However, this explanation seems unlikely: children growing up *and* working in urban areas also fall below these migrants. Alternatively, one could speculate that migrants possess some intangible or unobserved trait that helps them succeed; for example, for a poor child to be able to migrate in the first place, he must be sufficiently hard-working and risk-taking so as to earn a living in an unfamiliar area.

Both Figures 1.9 and 1.10 support the notion that there is no benefit to growing up in a urban relative to a rural area beyond a higher likelihood of working in an urban area during adulthood. A large portion of the difference between urban children and rural non-migrants in Figure 1.9 can be explained by selection on education and sector in Figure 1.10. In addition, rural-to-urban migrants fare better than both across most of the income distribution. This result runs contrary to Chetty et al. (2014b); Chetty and Hendren (2018) and similar

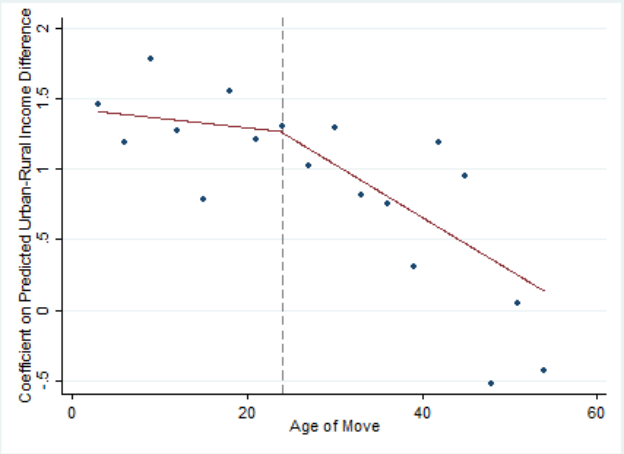


Figure 1.11: Childhood Exposure Effects

literature documenting the existence of childhood neighborhood effects in the US. These papers typically find that children who move to better neighborhoods earlier during childhood receive greater benefits. One method exploits variation in the age of migration and finds that children converge to outcomes of permanent residents of the destination area at 4.4% per year of exposure before age 23 (Figure IV of Chetty et al. (2014b)). In Figure 1.11, I reproduce the exercise using my data and find that for every three years earlier that an individual moves from a rural to an urban area before turning 24, he converges to urban area incomes at a rate of 2% (or just .66% per year). [22] Moreover, the slope sharply declines after the age of 24. This suggests that neighborhood effects of urban areas in Indonesia are much smaller than those in the United States—the age of migration during childhood matters little beyond allowing the individual to earn a higher income in an urban area for a longer period of time.

The minimal impact of neighborhood effects is somewhat surprising given previous literature. Access to key public goods and quality of amenities are consistently lower in rural than urban areas in developing countries (Gollin et al. (2014)). In particular, a World Bank (2013) report finds that in developing countries, rural areas are inferior to urban areas on almost all measures of health and human capital, and the greatest disparities are found in the poorest countries. Chetty and Hendren (2018) link public services to mobility, noting that higher absolute upward mobility is associated with higher school quality among other characteristics such as lower inequality and better family structure. While I cannot directly link individuals in the sample to the quality of their schooling during childhood, the IFLS

[22] I cannot replicate the method exactly due to sample size restrictions, so I bin age of migration into three-year bins and take average income instead of income at age 24. My exercise is an approximation.

includes community-level data on schools that allows me to include school quality metrics across urban and rural areas. Table 1.13 shows that urban teachers are typically higher-paid, spend more hours teaching, and more likely to have graduated from college, university, or a teaching college. However, urban areas also tend to have larger student-teacher ratios. It is unclear whether these effects offset or simply have no impact; in either case, these differences in education quality during childhood do not appear to manifest in adulthood outcomes.

My result also relates to the literature documenting large rural and urban (or agricultural and non-agricultural sector) productivity gaps and the subsequent puzzle of low rural-urban migration in developing countries.[23] While Gollin et al. (2014) document large gaps in productivity between sectors even after controlling for hours worked and human capital, Hamory et al. (2021) and Alvarez (2020) find that these gaps can be largely explained by sorting on individual-level observables. My findings do not contradict either conclusion, but refine them by suggesting that productivity gaps may differ depending on economic conditions during childhood and potentially unobservable individual traits.

1.4.3 Nonlinearities in Relative Mobility

Previous literature on intergenerational mobility includes a significant subliterature regarding potential nonlinearities in the relationship between parent and child log earnings. The seminal work of Becker and Tomes (1979, 1986) introduces a model in which credit constraints impede child investment, causing the relationship to be concave if poor households are borrowing-constrained. Follow-up work posits an alternative hypothesis of convexity driven by educational policies and institutions designed to provide equal opportunities to human capital formation, which would be particularly beneficial to children of poorer parents (Han and Mulligan (2001), Grawe and Mulligan (2002)). Empirical studies have provided varying evidence as to whether these relationships exist in the data: Couch and Lillard (2004) and Mazumder (2005) find evidence of credit constraints and a concave relationship in the US, but Mulligan (1999) finds that estimates of persistence are the same for both the financially constrained and unconstrained using PSID data. Bratsberg et al. (2007) uncover a convex relationship in Nordic countries, and Corak and Heisz (1999) find the same in Canada. The nature of the relationship is particularly interesting in a poor country context, where financial institutions are less developed and accessible, potentially causing credit constraints to bind for a relatively large share of the population.

[23] I focus on childhood location rather than sector because intergenerational mobility concerns the effect of childhood conditions on adulthood outcomes, and working in the agricultural sector occurs during adulthood. Analysis of agricultural productivity gaps conditional on parent income is included in Section 5.

The true functional form of the relationship between parent and child log income remains unclear, or at least varies across different countries. Previous studies in developed countries address this question using different methods. Bratsberg et al. (2007) report earnings elasticities at the 10^{th}, 50^{th}, and 90^{th} percentiles by adding higher-order polynomials in addition to the linear model for IGE. Mazumder (2005) compares elasticities in net worth for varying thresholds of poor vs. rich households in the U.S. and find that credit constraints might play a role in decreasing mobility for the poor. It is undoubtedly a worthwhile study in a developing country context, especially given the relatively large proportion of poor families and the potential prevalence of credit constraints. Table 1.14 replicates these methods to explore the possible existence of nonlinearities in the father-son log income relationship in Indonesia.

Using a fourth-order polynomial provides no evidence of credit constraints; in fact, relative mobility appears to be *greater* for children of the poorest parents. IGE is .181 at the 10th percentile, which is lower than the general population estimate. However, IGE for children of below-median parents is somewhat higher than children of above-median parents, suggesting that children of poorer parents may indeed face credit constraints. Data on knowledge of and access to financial institutions is only available in 2007 and 2014 waves of the survey when most children in the survey are already income-earning adults, so I am unable to match most individuals to their parent's access to credit during childhood. However, Table 1.15 reports mean rates of access to borrowing (formal or informal), access to banking, successful and rejected borrowing (conditional on attempting to borrow), and amounts borrowed in 2007 and 2014 for each child income quartile. It appears that richer children have slightly better access to formal banking and higher credit, and insofar as access to financial institutions along the relative income distribution has not changed between generations, this suggests that credit constraints could lower mobility by hindering investments in businesses or human capital. However, it is difficult to draw conclusions without sufficient data on parental borrowing during childhood, and I leave further exploration to future work.

Nonlinearities: Urban and Rural

Given the potential existence of nonlinearities and the fact that rural children are more likely to be born to poorer parents, it is reasonable to believe that the two groups could face different severities of credit constraints as well. However, sample sizes become somewhat restrictive and standard errors become large, leading to many insignificant estimates in Table 1.16. While rural children between the $10^{th} - 50^{th}$ parents percentiles seem to exhibit low relative mobility relative to the full rural sample, including the $0^{th} - 10^{th}$ percentiles

nullifies this result. Given the imprecision of estimates, it is difficult to identify nonlinearities or differences in nonlinearities between urban and rural children, and one should take caution in interpreting these results.

1.5 Robustness Checks

In this section, I perform a series of robustness checks. First, I compare estimates of relative mobility and differences in absolute mobility using percentiles instead of log income.

The statistical relationship between IGE and rank-rank relative mobility is as follows:

$$\text{IGE} = \rho_{y_c,y_p} \frac{SD(y_c)}{SD(y_p)}$$

$$\text{Rank-Rank Slope} = \rho_{p_c,p_p}$$

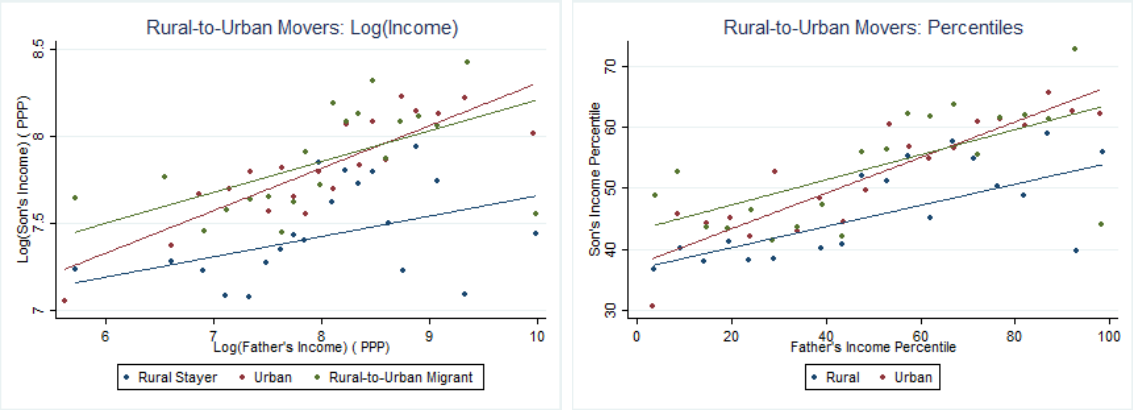
where ρ denotes the correlation between variables in the subscript and SD() denotes standard deviation. ρ_{y_c,y_p} and ρ_{p_c,p_p} are closely related scale-invariant measures of the association between children's and parent's income. Thus, IGE differs from the rank-rank slope in that IGE depends on changes in the variances of income across generations—intuitively, higher variance in children's incomes increases the effect of a given change in parents' incomes when the variance of children's incomes is higher than that of parents' incomes. Given that variance of father's income is greater than variance of son's income in my sample (Table 1.2), it is not surprising that the rank-rank estimate is slightly higher.

Table 1.17, Panels A and B produce analogous estimates from Columns (1) and (7) of Tables 1.5 and 1.8, respectively, using income percentiles instead of log income. Results are generally quite similar: the regression without controls suggests that for an increase in father's income by 10 percentiles, child's expected rank increases by 2.68 percentiles for the full sample, 2.18 percentiles for rural children, and 2.90 percentiles for urban children (Column 1). However, one notable difference is that father's income percentile remains significant even after controlling for son's education and work sector in all three samples (Column 7). The fact that parent's income percentile remains significant could reflect the decrease in variance in incomes between generations: while a change in parent's income rank significantly affects child's income *relative* to other children in the income distribution, a change in parent's log income does not significantly affect child's log income.

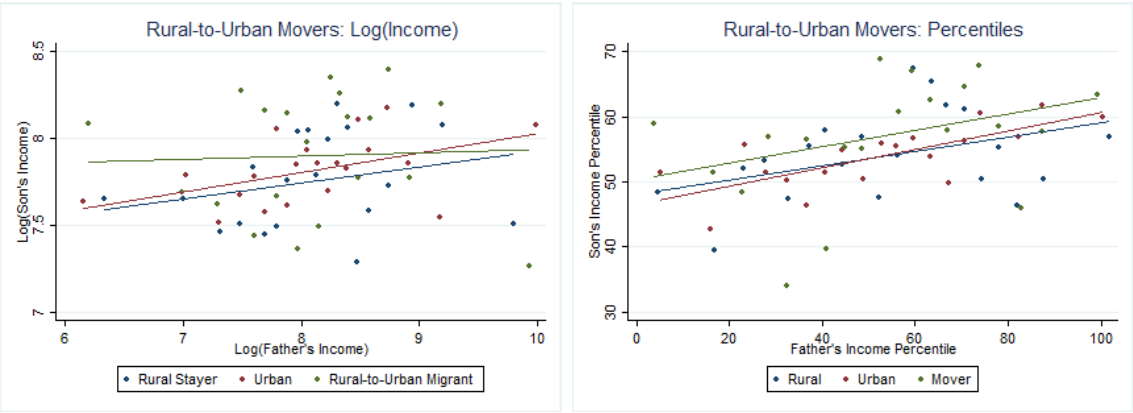
An alternative method to calculate income rank in measuring rank-rank relative mobility is to calculate parent and child income ranks based on ranks *within* a relevant income

distribution (for example, comparing rural children only to other rural children and not the entire income distribution).[24] Table 1.18 provides estimates for urban and rural subsamples using the local income distribution, and results are likewise robust to this change.

Figure 1.12: Absolute Mobility: Log Income vs. Percentiles



(a) No Controls

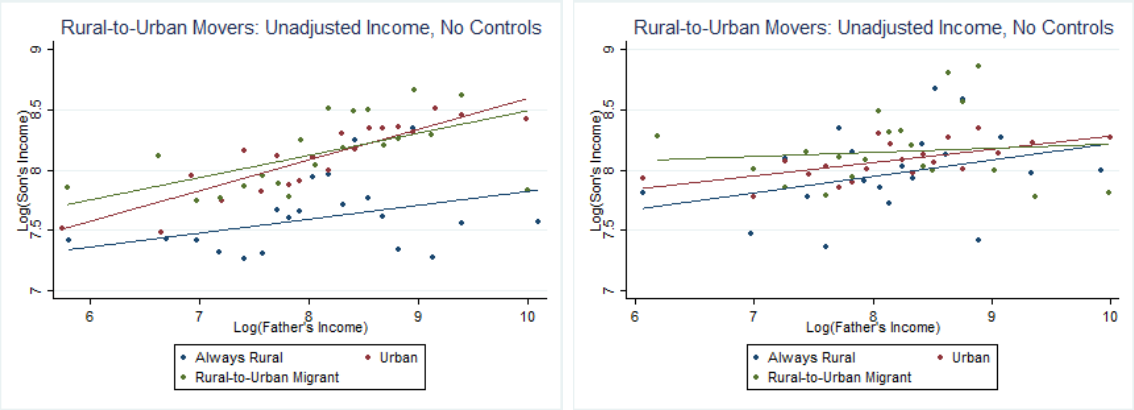


(b) Education and Sector Controls

Income percentiles also produce similar absolute mobility results. Figure 1.12 shows that while absolute mobility gaps are expressed in terms of percentiles instead of log points, the general patterns in comparing rural children, urban children, and rural-to-urban movers remain roughly the same. If anything, the disparity between migrants and children who stay urban or rural becomes more clear, especially after controlling for observables: a gap of roughly 4 percentiles persists across the entire parent income distribution, again suggesting that there may be unobserved attribute of migrants beyond education, sector, and parent's income that contributes to better outcomes.

[24]Previous work (Chetty et al. (2014b)) does this but finds similar estimates to using the national distribution, so they omit it from the main analysis.

Figure 1.13: Absolute Mobility: Unadjusted Income



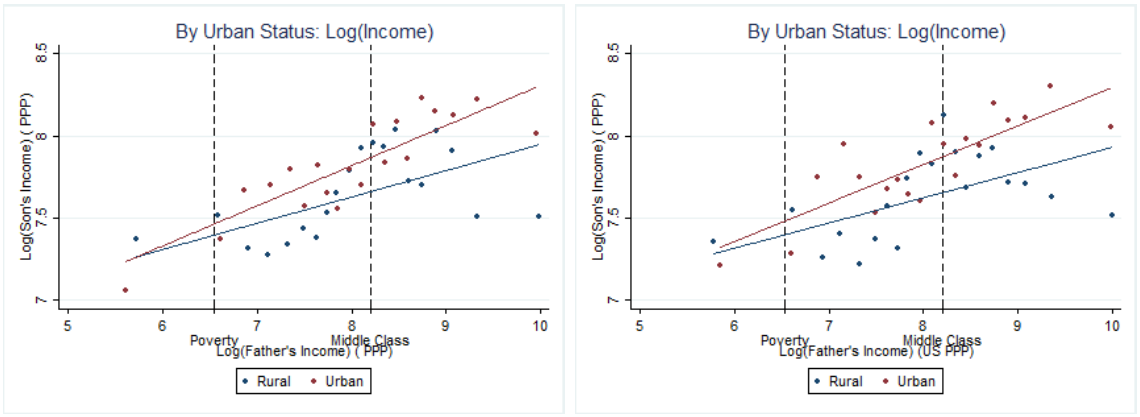
Recall that incomes earned in rural areas are adjusted upward using Indonesia's 1981 rural-to-urban price adjustment index by Ravallion et al. (1989) to account for income differences that reflect differences in quality of life and not simply higher prices. Repeating my analysis in Table 1.20 and Figure 1.13 shows that relative mobility results are robust to this change, while the gap in absolute mobility for children of the richest parents mechanically increases from 65 to 77 log points.

Another issue in my data is lifecycle variation in income: namely, differences in the variances of transitory earnings at different points in the lifecycle. Since the earnings profile tends to flatten around age 40, realizations of income at ages that deviate too far from this age are noisy measures of an individual's true earnings potential. Specifically, Grawe (2006) finds that earnings persistence estimates decrease by 50% when fathers are observed at age 55 rather than age 40, and estimates increase with the age at which sons' incomes are measured. [25] Since my measure of income averages over multiple observations, I cannot control for age in my regressions. However, I can repeat the analysis controlling for father's and son's *average* age across income measurements. Table 1.20 reproduces columns (1) and (7) from Tables 1.5 and 1.8 show that relative mobility results are largely robust to including age controls, and reproducing Figures 1.5 and 1.9 with unadjusted incomes in Figure 1.14 shows that absolute mobility results are barely affected as well.

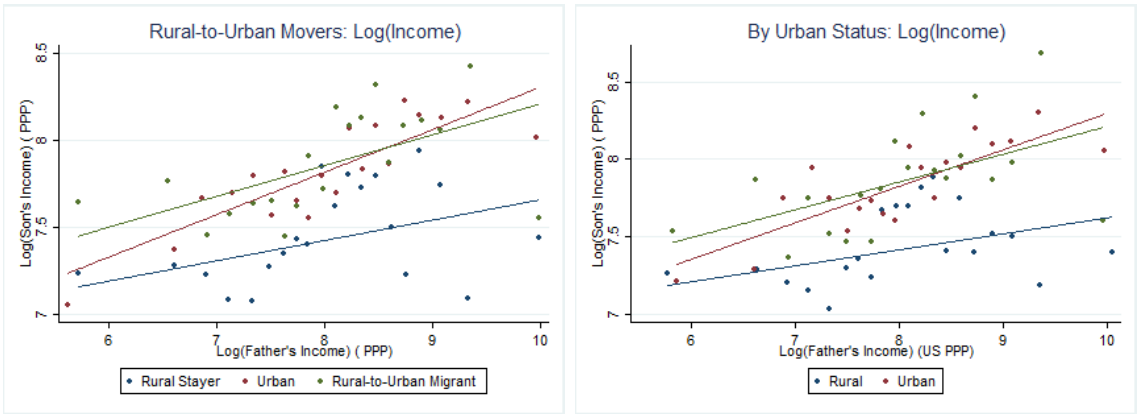
Next, I repeat the analysis on groups defined by sector of work (agricultural or non-agricultural) rather than location. Table 1.21 replicates Columns (1) and (5) from Table 1.5 and reveals that that IGE for agricultural workers is higher than that for non-agricultural

[25] Haider and Solon (2006) also demonstrate that IV estimates are likely downward-biased if children's incomes are measured under age 40 and fathers' incomes over age 40. My main results are not IV estimates, so I omit them from my analysis.

Figure 1.14: Absolute Mobility: Age Controls



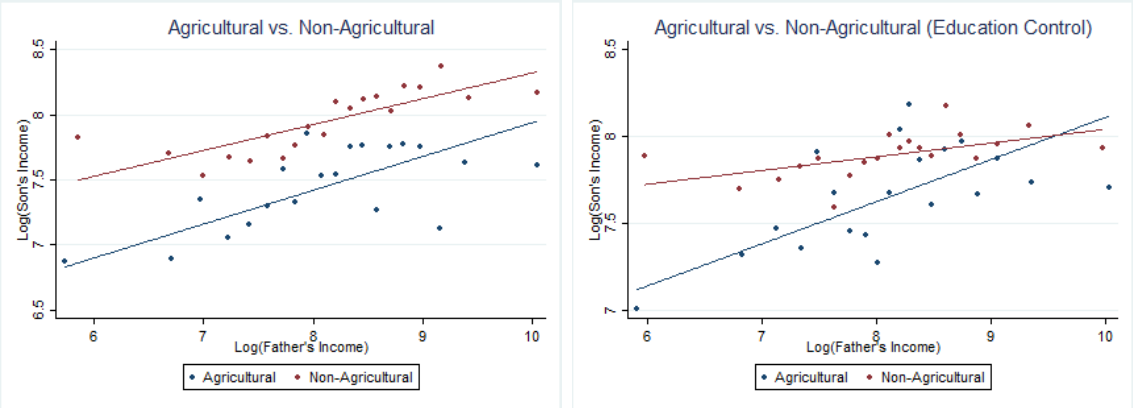
(a) Rural vs. Urban Children



(b) Rural-to-Urban Migration

workers, though the difference is not statistically significant. Given that agricultural work strongly correlates with rural children, this result is somewhat surprising in that it reflects the reverse of the observed relationship between urban and rural relative mobility. If the difference does indeed exist, it is not immediately clear what causes it. Another peculiar result is that parent's income is still significant for agricultural workers even after controlling for father's and son's education and father's sector. This could potentially explained by the physical nature of agricultural work, which would diminish the effect of education conditional on parent's income and make income highly dependent on the capacity of an individual to produce output. Then, if children produce on land inherited from their parents, then parent's income could have a strong effect as a proxy for the scale of the child's agricultural operation. This explanation is pure speculation, but since working (and staying) in the agricultural sector does not condition on childhood conditions, it is beyond the scope of a study on intergenerational mobility and is left for future work. Figure 1.15 shows that

Figure 1.15: Absolute Mobility: Non-Agricultural vs. Agricultural



differences in non-agricultural and agricultural absolute mobility are slightly larger than urban and rural differences, though the pattern is consistent with differences between always urban and always rural children in Figure 1.9. Controlling for education mostly shrinks the gap at the top of the distribution just as it does for urban and rural children.

Finally, I adjust my income calculations to account for incomes earned between 1997 - 2000, which are likely lower in response to the Indonesian financial crisis in August 1997. In 1997, Indonesia fell into a financial crisis after the rupiah's value was allowed to float. The rupiah to dollar exchange rate preceding the crisis was 2,600 rupiah to $1 US. By mid-1998, this had ballooned to over 14,000:1. The economy had largely recovered by 2000 and has since sustained 4-6% growth during the 2000s. The financial crisis introduces a particular challenge in measuring income in my sample as incomes measured between 1997 - 2000 are likely depressed due to the financial crisis. Since each individual earns income in different years, those with incomes measured in those years would then have mechanically lower incomes that would bias my results. To address this, I repeat the analysis using two methods. First, I add year dummies for both fathers and sons (separately) to the specification that are equal to 1 if the individual's average income includes income from that year, and 0 if not. Second, I calculate an alternative measure of income by dropping income between years 1997 - 2000. Table 1.22 below reproduces columns (1) and (7) of Tables 1.5 and 1.8 with these two alternative specifications along with the baseline regression for comparison. Results are largely robust to these alternative specifications. In particular, the relative comparisons of relative mobility for urban and rural subsamples remain the same.

1.6 Conclusion

In this paper, I use representative panel survey data from the IFLS to analyze intergenerational mobility in Indonesia and the factors that affect it. My results show that parent-child income persistence is relatively low compared to other countries, both developing and developed. I find suggestive evidence that income persistence is not driven through a direct mechanism from parent's income to children's income, but rather through its effects on intermediate channels such as children's education and work sector. While the difference in IGE for urban and rural children is not statistically significant, alternative specifications suggest that rural children could exhibit less persistent income mobility than urban children. This appears to be driven by a stronger relationship for urban children between parent's income and child's education, and subsequently, child's education and child's income.

While rural children exhibit better *relative* mobility than urban children, urban children are still expected to earn higher incomes conditional on parent income across most of the parent income distribution. This is particularly true for urban children raised by rich parents, as the gap in expected income grows to 35 log points for children of the richest parents. Most of this can be explained by selection, as controlling for education shrinks this gap to 6 log points. Separating rural children into rural-to-urban migrants and always rural individuals reveals additional findings about the effect of childhood circumstances. First, about half of upward mobility for rural children can be associated with migration to urban areas as opposed to earning higher incomes in rural areas. Second, children who migrate earn incomes comparable to children growing up in rural areas. After controlling for these observables, migrants still compare favorably to children in either location type, especially at the bottom of the income distribution.

Contrary to previous findings in developed countries and indirectly related work in developing countries, I find no neighborhood effects for children growing up in either location. Most of the difference between urban and rural children can be explained by selection, and migrants growing up in rural areas and working as adults in urban areas fare just as well (if not better) than children growing up *and* working in urban areas. This occurs despite apparent differences in metrics of public school quality.

While results appear to be robust to adjustments, such as those to address measurement error and attenuation, the data is still limited by a somewhat small sample size compared to previous studies in developed countries. In particular, more precise estimates of IGE for rural and urban children could reveal a statistically significant difference that could help us understand the true nature of intergenerational mobility in these areas. Future work

incorporating upcoming waves of the IFLS could improve precision as well as shed light on whether patterns of persistence change over time.

Despite the fact that Indonesia is a developing country with a large poor population, children appear to exhibit strong relative mobility and are unlikely to fall into intergenerational poverty traps. Given that these findings somewhat diverge from previous estimates of relative mobility in developing countries, it is certainly possible that Indonesia is an outlier among them. Moreover, the channels diverge from mechanisms previously established in developed countries. These discrepancies demonstrate that intergenerational mobility in a developing country as a complex and interesting question, potentially unique to conditions within each country. The dearth of literature hopefully encourages future work in the area, which can help reveal areas of need for policy intervention targeting childhood circumstances and engender economic opportunities for the world's most vulnerable children.

Table 1.9: Determinants of Child Observables: Urban vs. Rural Children

Variable	(1)	(2)	(3)	(4)	(5)
Father's Education	0.468***		0.397***	0.447***	0.399***
	(0.019)		(0.024)	(0.022)	(0.027)
Log(Parent Income)		1.841***	0.760***		0.562***
		(0.130)	(0.143)		(0.165)
Father Agricultural				-1.686***	-1.441***
				(0.357)	(0.366)
R^2	0.265	0.141	0.283	0.316	0.325
N	1375	1400	1375	1025	1025

(a) Child's Education (Urban)

Variable	(1)	(2)	(3)	(4)	(5)
Father Agricultural	0.181***		0.167***	0.139***	0.139***
	(0.009)		(0.011)	(0.010)	(0.011)
Log(Parent Income)		-0.046***	-0.015		0.001
		(0.009)	(0.008)		(0.008)
Father's Education				-0.009***	-0.009***
				(0.002)	(0.002)
N	910	1219	910	909	909

(b) Child Agricultural (Average Marginal Effects, Urban)

Variable	(1)	(2)	(3)	(4)	(5)
Father's Education	0.442***		0.381***	0.408***	0.367***
	(0.023)		(0.025)	(0.029)	(0.031)
Log(Parent Income)		1.205***	0.660***		0.581***
		(0.107)	(0.111)		(0.135)
Father Agricultural				-0.632	-0.221
				(0.324)	(0.337)
R^2	0.168	0.081	0.190	0.149	0.166
N	1508	1547	1508	1008	1008

(c) Child's Education (Rural)

Variable	(1)	(2)	(3)	(4)	(5)
Father Agricultural	0.688***		0.661***	0.649***	0.639***
	(0.072)		(0.077)	(0.072)	(0.075)
Log(Parent Income)		-0.108***	-0.030		-0.014
		(0.015)	(0.018)		(0.019)
Father's Education				-0.018***	-0.017***
				(0.005)	(0.005)
N	652	1006	652	646	646

(d) Child Agricultural (Average Marginal Effects, Rural)

Standard errors in parentheses. *=.05, **=.01, *** = .001.

Table 1.10: Mincer Regressions for Urban and Rural

Variable	Urban	Rural	Difference
Years of Education	0.067***	0.054***	0.013*
	(0.004)	(0.005)	(0.007)
Age	0.286***	0.265***	0.021
	(0.018)	(0.025)	(0.030)
Age Squared	-0.004***	-0.003***	-.000
	(0.000)	(0.000)	(0.000)
N	5,375	3,762	
R^2	0.186	0.158	

(a) Mincer Regressions for Urban and Rural

Variable	Non-Agricultural	Agricultural	Difference
Years of Education	0.065***	0.043***	0.022**
	(0.004)	(0.008)	(0.009)
Age	0.329***	0.208***	0.121***
	(0.019)	(0.029)	(0.034)
Age Squared	-0.004***	-0.003***	-.002**
	(0.000)	(0.000)	(0.001)
N	6,282	1,795	
R^2	0.202	0.096	

(b) Mincer Regressions for Non-Agricultural and Agricultural

Standard errors in parentheses. *=.05, **=.01, *** = .001. Results come from repeated cross-section instead of panel data to avoid issues with ages changing across the panel. Individuals could appear more than once if they work multiple times during the survey.

Table 1.11: Expected Daily Childhood Income for Given Parent Income ($US PPP)

	Parent Income	
	Poverty ($1.90/day)	Middle Class ($10/day)
Full Sample*	4.40	6.65
Rural	4.45	5.82
Urban	4.76	7.16

*Values for the full sample are only mechanically related to relative mobility but are provided for reference.

Table 1.12: Selection into Migration

	Rural (Stayer)	Rural-to-Urban Migrant	**Difference**	Urban
Parent Income ($US PPP)	3289	3705	**416***	6077
			(226)	
Years of Education (Parent)	4.96	6.06	**1.09****	8.27
			(.204)	
Years of Education	8.78	10.65	**1.87****	11.82
			(.214)	
Agricultural (Parent)	.897	.729	**-.168****	.140
			(.023)	
Agricultural	.628	.137	**-.492****	.075
			(.027)	

Table 1.13: Measures of School Quality: Rural vs. Urban

		Year									
		1993		**1997**		**2000**		**2007**		**2014**	
School Level	**Variable**	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
Elementary	Student-Teacher Ratio	36.9	41.0	25.3	23.6	26.7	8.4	12.8	19.9	15.0	20.3
	Teacher Salary ($US PPP)	3748	4200	4852	4768	5394	5488	5667	6915	6237	7506
	Hours Teaching	33.9	31.4	29.3	27.3	31.8	30.9	30.2	31.1	29.2	31.9
	Need More Classrooms (%)			52.1	52.2	47.3	47.9	63.4	56.1	48.3	60.3
	College-Educated Teacher (%)*	12.7	29.4	44.8	64.1	42.5	47.5	63.7	78.6	87.8	95.1
Junior High	Student-Teacher Ratio	21.9	32.9	16.3	18.2	16.1	17	13.3	14.9	12.1	16.2
	Teacher Salary ($US PPP)	2479	3253	2746	3650	2965	3965	7243	6963	6684	7402
	Hours Teaching	22.5	24.9	20.9	22	21.7	23.6	25.4	28.8	26.4	32.2
	Need More Classrooms (%)			73.2	66.6	71.7	57.8	68.5	60.1	75.5	69.1
	College-Educated Teacher (%)*	82.6	95.2	89.5	97.4	85.9	92.0	86.5	93.7	97.1	98.9
Senior High	Student-Teacher Ratio	26.7	32.4	13.7	15.4	12.3	16.2	10.4	10.9	11.1	13.9
	Teacher Salary ($US PPP)	2631	3304	3037	3848	3266	3880	3178	5595	3335	7930
	Hours Teaching	21.6	27.8	19.7	21.5	21.8	22.7	25.5	30	27.6	33
	Need More Classrooms (%)			73.3	70.9	69.8	64.8	74.6	66.9	74.9	72.6
	College-Educated Teacher (%)*	98.1	98.7	98.6	99.7	92.6	96.6	93.5	97.7	98.1	98.8

*Includes graduating from a teaching college

Table 1.14: Nonlinearities in IGE

Percentile Range	Estimate	Percentile	Estimate
≤ 50	.266***	10	.181
	(.072)	25	.233
>50	.195**	50	.284
	(.072)	75	.287
10 —50	.326***	90	.234
	(.097)		
50 —90	.249		
	(.131)		
10 —90	.306***		
	(.043)		
10 —100	.276***		
	(.034)		
0 —90	.272**		
	(.038)		

* = .05, ** = .01, *** = .001. Standard errors in parentheses. Percentile range estimates come from IGE regressions using parents in the corresponding percentile bracket (Mazumder (2005)). Percentile point estimates use a fourth-order polynomial in Log(Father's Income) to measure IGE at different points along the parent income distribution (Bratsberg et al. (2007)).

Table 1.15: Access to and Usage of Financial Institutions

Child Income Quartile	Access to Borrowing	Access to Formal Banking	Borrowed Succesfully	Rejected from Borrowing	Amount Borrowed (2007) (US PPP)	Amount Borrowed (2014) (US PPP)
1	.96	.82	.98	.13	1568	1605
2	.96	.85	.97	.10	2392	2876
3	.97	.87	.96	.10	3234	3400
4	.99	.93	.95	.09	7759	9994

Table 1.16: IGE: Nonlinearities, Urban. vs. Rural

Parent Percentile(s)	Urban	Rural	Difference	Percentile	Urban	Rural
≤ 50	.176*	.146	.031	10	.264	.061
	(.087)	(.097)	(.131)	25	.258	.224
>50	.260**	-.099	.359	50	.246	.279
	(.099)	(.205)	(.228)	75	.252	.139
10—50	.082	.333*	-.251	90	.285	-.164
	(.126)	(.143)	(.191)			
50—90	.185	.022	.163			
	(.190)	(.282)	(.340)			
10—90	.240***	.284***	-.044			
	(.065)	(.085)	(.107)			
10—100	.263***	.219**	.044			
	(.048)	(.076)	(.090)			
0—90	.239***	.186**	.053			
	(.052)	(.02)	(.081)			

* = .05, ** = .01, *** = .001. Standard errors in parentheses when applicable. Urban and rural percentiles based on economy-wide, not subsample, percentiles. The first four estimates run IGE regressions on subsamples of data. All others use a fourth-order polynomial in Log(Father's Income) to measure IGE at different points along the parent income distribution.

Table 1.17: IGE vs. Rank-Rank

Panel A: No Controls

	Full Sample: IGE	Full Sample: RR	Rural: IGE	Rural: RR	Urban: IGE	Urban: RR
Log(Parent Income)	0.259***		0.162**		0.256***	
	(0.031)		(0.057)		(0.041)	
Parent Income Percentile		0.268***		0.218***		0.290***
		(0.015)		(0.028)		(0.028)
R^2	0.028	0.072	0.009	0.041	0.023	0.069
N	4125	4125	1553	1553	1419	1419

Panel B: Full Controls

	Full Sample: IGE	Full Sample: RR	Rural: IGE	Rural: RR	Urban: IGE	Urban: RR
Log(Parent Income)	0.115		0.049		0.075	
	(0.061)		(0.075)		(0.060)	
Parent Income Percentile		0.128***		0.097*		0.127**
		(0.026)		(0.045)		(0.041)
Father's Education	0.007	0.259	0.022	0.353	0.009	0.316
	(0.008)	(0.160)	(0.015)	(0.306)	(0.011)	(0.239)
Father Agricultural	-0.021	0.802	-0.049	1.977	0.069	-0.948
	(0.089)	(1.659)	(0.157)	(3.140)	(0.122)	(3.392)
Son's Education	0.062***	1.792***	0.068***	1.822***	0.080***	2.125***
	(0.009)	(0.160)	(0.015)	(0.261)	(0.013)	(0.264)
Son Agricultural	-0.353**	-8.238***	-0.334*	-10.593***	-0.109	-3.477
	(0.109)	(1.855)	(0.153)	(2.497)	(0.168)	(4.154)
R^2	0.093	0.187	0.081	0.177	0.072	0.177
N	2070	2070	646	646	908	908

Standard errors in parentheses. IGE uses log income; RR measures the rank-rank slope

Table 1.18: Rank-Rank: Ranks within Distribution

	Rural: Baseline	Rural: Within	Urban: Baseline	Urban: Within
Parent Income Percentile	0.218***	0.199***	0.290***	0.276***
	(0.028)	(0.025)	(0.028)	(0.025)
R^2	0.041	0.040	0.069	0.076
N	1553	1553	1419	1419

Standard errors in parentheses. "Within" uses income distribution *within* rural and urban subsamples.

Table 1.19: IGE with Unadjusted Income

	Full Sample	Rural	Urban	Full Sample	Rural	Urban
Log(Parent Income)	0.267***	0.161**	0.236***	0.098	0.053	0.083
	(0.030)	(0.054)	(0.044)	(0.055)	(0.075)	(0.063)
Edu/Sector Controls	No	No	No	Yes	Yes	Yes
R^2	0.033	0.009	0.022	0.111	0.097	0.082
N	4125	1553	1419	2070	646	908

Standard errors in parentheses.

Table 1.20: IGE with Parent and Child Age Controls

	Full Sample	Rural	Urban	Full Sample	Rural	Urban
Log(Parent Income)	0.254***	0.168**	0.259***	0.092	0.057	0.059
	(0.031)	(0.061)	(0.043)	(0.060)	(0.079)	(0.060)
Edu/Sector Controls	No	No	No	Yes	Yes	Yes
R^2	0.036	0.016	0.030	0.106	0.094	0.080
N	4113	1550	1416	2063	643	906

Standard errors in parentheses.

Table 1.21: Agricultural vs. Non-Agricultural

	Non-Agricultural	Agricultural	Difference	Non-Agricultural	Agricultural
Log(Parent Income)	0.196***	0.280***	0.084	0.032	0.274***
	(0.030)	(0.060)	(0.095)	(0.044)	(0.073)
Edu/Father's Sector Controls	No	No	No	Yes	Yes
R^2	0.017	0.031		0.064	0.043
N	2372	682		1581	489

Table 1.22: IGE with Controls, Adjusting for Financial Crisis

Full Sample

	Baseline	Yr Dummies	Ex. 1997 - 2000	Baseline	Yr Dummies	Ex. 1997 - 2000
Log(Parent Income)	0.259***	0.279***	0.190***	0.115	0.109	0.024
	(0.031)	(0.031)	(0.026)	(0.061)	(0.059)	(0.042)
Year Dummies	No	Yes	No	No	Yes	No
Edu/Sector Controls	No	No	No	Yes	Yes	Yes
R^2	0.028	0.066	0.017	0.093	0.137	0.086
N	4125	4125	3950	2070	2070	2014

Urban

	Baseline	Yr Dummies	Ex. 1997 - 2000	Baseline	Yr Dummies	Ex. 1997 - 2000
Log(Parent Income)	0.256***	0.288***	0.210***	0.075	0.102	0.038
	(0.041)	(0.045)	(0.042)	(0.060)	(0.064)	(0.057)
Year Dummies	No	Yes	No	No	Yes	No
Edu/Sector Controls	No	No	No	Yes	Yes	Yes
R^2	0.023	0.055	0.018	0.072	0.101	0.068
N	1419	1419	1390	908	908	897

Rural

	Baseline	Yr Dummies	Ex. 1997 - 2000	Baseline	Yr Dummies	Ex. 1997 - 2000
Log(Parent Income)	0.162**	0.162**	0.125*	0.049	0.024	0.044
	(0.057)	(0.060)	(0.053)	(0.075)	(0.075)	(0.067)
Year Dummies	No	Yes	No	No	Yes	No
Edu/Sector Controls	No	No	No	Yes	Yes	Yes
R^2	0.009	0.042	0.006	0.081	0.155	0.074
N	1553	1553	1533	646	646	640

Standard errors in parentheses.

Appendix

Table 1.A.1: Gini Estimates by Country

Country	Year	Gini Estimate
Ukraine	2018	26.1
Iceland	2015	26.8
Norway	2017	27.0
Belgium	2017	27.4
Finland	2017	27.4
Netherlands	2017	28.5
Denmark	2017	28.7
Sweden	2017	28.8
Poland	2017	29.7
France	2017	31.6
Germany	2016	31.9
Switzerland	2017	32.7
Portugal	2017	33.8
Canada	2013	33.8
Greece	2017	34.4
Australia	2014	34.4
Spain	2017	34.7
United Kingdom	2016	34.8
Luxembourg	2017	34.9
Italy	2017	35.9
Romania	2017	36.0
Serbia	2017	36.2
Russian Federation	2018	37.5
United States	2016	41.4
Turkey	2018	41.9
Mexico	2018	45.4
Myanmar	2017	30.7
Egypt, Arab Rep.	2017	31.5
Korea, Rep.	2012	31.6
Bangladesh	2016	32.4
Japan	2013	32.9
Vietnam	2018	35.7
Lao PDR	2012	36.4
Thailand	2018	36.4
India	2011	37.8
China	2016	38.5
Burundi	2013	38.6
Indonesia	2018	39.0
Uruguay	2018	39.7
Sri Lanka	2016	39.8
Senegal	2011	40.3
Iran, Islamic Rep.	2017	40.8
Malaysia	2015	41.0
Argentina	2018	41.4
Madagascar	2012	42.6
Peru	2018	42.8
Uganda	2016	42.8
Philippines	2015	44.4
Chile	2017	44.4
Venezuela, RB	2006	46.9
Colombia	2018	50.4
Brazil	2018	53.9

1.A.1 Discussion of Empirics (adapted from Mazumder (2005))

Recall equations (3) - (5), which define the framework for transitory fluctuations in income (Friedman (1957)):

$$y_{i,p,s} = y_{i,p} + e_{i,p,s} \tag{3}$$

$$y_{i,c,t} = y_{i,c} + e_{i,c,t} \tag{4}$$

$$e_{i,p,s} = \rho e_{i,p,s-1} + \xi_{i,s} \tag{5}$$

as well as equation (1), the regression for IGE:

$$y_{i,c} = \alpha + \beta y_{i,p} + \varepsilon_i \tag{1}$$

where $y_{i,p,s}$ denotes parent's earnings in year s and $y_{i,c,t}$ denotes child's earnings in year t. Both consist of a permanent income component ($y_{i,p}$ and $y_{i,c}$ for parent and child, respectively) and a transitory shock to earnings ($e_{i,p,s}$ and $e_{i,c,t}$). Obtaining a consistent estimate for β in Equation (1) relies on obtaining an accurate measure of parent's permanent income $y_{i,p}$. In practice, researchers approximate $y_{i,p}$ by averaging income over a number of years T to minimize these shocks:

$$\text{plim}\hat{\beta} = \beta\lambda_T$$

where

$$\lambda_T = \frac{\sigma^2_{y_p}}{\sigma^2_{y_p} + \frac{1}{T}\alpha\sigma^2_{e_p}}$$

and

$$\alpha = 1 + 2\rho\frac{T - \frac{1-\rho^T}{1-\rho}}{T(1-\rho)} \tag{6}$$

Thus, the estimate is attenuated by a factor of λ_T, also called the "reliability ratio." When $\rho = 0$ (or no serial correlation in the error term), increasing T decreases the influence of the variance of the transitory error term. However, in the presence of serial correlation,

λ_T becomes much more complicated in response to changes in T. Mazumder (2005) simulates attenuation coefficients for various parameterizations of ρ and finds that for fixed ρ, increasing T decreases the amount of attenuation. He finds that for $\rho = .5$ and a reasonable parameterizations of variances, a 5-year average of fathers' earnings would cause esatimtes of β to be downwardly biased by approximately 30%. However, he also notes that the equations may not provide a sufficient framework for the earnings process. For more details, see Mazumder (2005).

1.A.2 Rank-Rank Mobility Estimates using Within-Region Income Distributions

In this subsection, we repeat the rank-rank regressions defining income percentile ranks within rural and urban subpopulations instead of the national income distribution. Results are similar in absolute and relative magnitude to the rank-rank results presented in Table 1.8: mobility is higher for children growing up in rural areas.

Table 1.A.2: Urban and Rural Estimates of Rank-Rank Mobility: Within-Sample

Measure	Urban	Rural	Difference
Rank-rank	.276***	.199***	.077*
	(.025)	(.025)	(.036)
Rank-rank: $10^{th} - 90^{th}$	.245***	.256***	-.011
	(.035)	(.034)	(.049)
IGE: IV	.571***	.512***	.049†
	(.052)	(.069)	-

Standard errors in parentheses. *=.05, **=.01, *** = .001.
† The limiting distribution of this comparison is $\chi^2(1)$, and the χ^2-statistic is .45 with a corresponding p-value of .501.

1.A.3 Full Survey Results: Gender Differences

In this section, I repeat some of the baseline analysis for the full survey, including daughters in addition to sons. Below, Tables 1.A.3—1.A.5 reproduce Tables 1.4, 1.5, and 1.7, respectively.

A few patterns emerge from comparing Tables 1.5 and 1.A.4: coefficients on parent income and father's education are higher for the full sample than for the subsample of sons in columns (1) and (7), respectively. This suggests that daughter's incomes are more dependent on father's income and education than son's income, driving the coefficients upward; in fact, a regression of daughter's log income on father's log income yields an IGE of .354.

Table 1.A.3: Full Sample Estimates of Relative Mobility

Measure	Value	N
IGE	.289***	7,974
	(.021)	
IGE: $10^{th} - 90^{th}$	.389***	6,379
	(.033)	
IGE: 2 or more observations	.265***	3,262
	(.019)	
Rank-rank	.272***	7,974
	(.011)	
Rank-rank: $10^{th} - 90^{th}$	.271***	6,379
	(.015)	
IGE: IV	.622***	7,576
	(.039)	
Education	.521***	7,504
	(.008)	

Standard errors in parentheses. *=.05, **=.01, *** = .001.

Table 1.A.4: Full Sample IGE with Controls

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Parent Income)	0.295***	0.169***	0.218***	0.133***	0.077**	0.128***	0.078*
	(0.020)	(0.023)	(0.026)	(0.030)	(0.029)	(0.032)	(0.033)
Father's Education		0.055***		0.051***	0.022***	0.050***	0.022***
		(0.005)		(0.006)	(0.006)	(0.006)	(0.006)
Father Agricultural			-0.490***	-0.375***	-0.323***	-0.209**	-0.229***
			(0.052)	(0.052)	(0.052)	(0.068)	(0.067)
Child's Education					0.072***		0.072***
					(0.006)		(0.007)
Child Agricultural						-0.580***	-0.390***
						(0.100)	(0.101)
R^2	0.033	0.047	0.052	0.065	0.089	0.087	0.108
N	7974	7576	5314	5287	5235	4192	4187

Standard errors in parentheses. *=.05, **=.01, *** = .001.

Table 1.A.5 reveals a similar pattern, particularly for children growing up in urban areas. The IGE for women growing up in urban areas is .345, relative to .169 for women growing up in rural areas. The corresponding female urban and rural rank-rank mobility estimates are .309 and .194, respectively. Moreover, the gaps between urban and rural mobility estimates using the full sample are statistically significant, which is not a feature when using sons alone. This suggests that in particular, women growing up in urban areas exhibit higher intergenerational income persistence and lower mobility.

Table 1.A.5: Full Sample Urban and Rural Estimates of Relative Mobility

Measure	Urban	Rural	Difference
IGE	.345**	.169**	.176**
	(.034)	(.040)	(.053)
IGE: $10^{th} - 90^{th}$	.424**	.323***	.101
	(.059)	(.066)	(.088)
IGE: 2 or more observations	.298**	.191***	.142**
	(.035)	(.033)	(.060)
Rank-rank	.309**	.194***	.114***
	(.021)	(.020)	(.029)
Rank-rank: $10^{th} - 90^{th}$	.308**	.241***	.067
	(.027)	(.027)	(.038)
IGE: IV	.745***	.590***	.155†
	(.076)	(.100)	-
Education	.466**	.464***	.002
	(.014)	(.016)	(.021)

Standard errors in parentheses. *=.05, **=.01, *** = .001.

† The limiting distribution of this comparison is $\chi^2(1)$, and the χ^2-statistic is 1.53 with a corresponding p-value of .217.

There could be various explanations for this finding. In developing countries, low female labor force participation is often driven by economic and social factors such as educational attainment and social norms. Daughters especially in higher-income households likely enjoy substantially better educational opportunities and may face less social pressure with high-achieving parents who serve as role models, which would significantly increase their probability of working. We leave further exploration of the mechanisms to future work.

1.A.4 Contribution of Agricultural Sector to IGE

Suppose the "true" models of income for urban children ($y_{u,c}$) and rural children ($y_{r,c}$), respectively, are:

$$y_{u,c} = a_0 + a_1 edu_c + \varepsilon_1 \qquad \text{(Urban)}$$

$$y_{r,c} = b_0 + b_1 edu_c + b_2 ag_c + \varepsilon_2 \qquad \text{(Rural)}$$

where edu_c is child's education, ag_c is child's agricultural status, and $y_{u,c}$ and $y_{r,c}$ are urban and rural child's incomes, respectively. The variables are chosen based on the significant

variables in each respective regression from Table 1.8 Panels A and B, Column (7). [26] I then run the IGE regressions:

$$y_{u,c} = c_0 + c_1 y_{u,p} + u_1 \qquad \text{(Urban)}$$
$$y_{r,c} = d_0 + d_1 y_{r,p} + u_2 \qquad \text{(Rural)}$$

where $y_{u,p}$ and $y_{r,p}$ denote urban and rural parent's incomes, respectively. The IGE coefficients c_1 and d_1 are then given by:

$$\begin{aligned} c_1 &= \frac{cov(y_{u,c}, y_{u,p})}{var(y_{u,p})} \qquad \text{(Urban)} \\ &= \frac{a_1 cov(edu_c, y_{u,p}) + cov(\varepsilon_1, y_{u,p})}{var(y_{u,p})} \\ d_1 &= \frac{cov(y_{r,c}, y_{r,p})}{var(y_{r,p})} \qquad \text{(Rural)} \\ &= \frac{b_1 cov(edu_c, y_{r,p}) + b_2 cov(ag_c, y_{r,p}) + cov(\varepsilon_2, y_{u,p})}{var(y_{r,p})} \end{aligned}$$

where cov(.,.) and var(.) are the covariance and variance functions, respectively.

[26] While these are not the true models of income, using this framework can help us understand how each component contributes to IGE.

Chapter 2

Intergenerational Mobility Effects of Region-Based Education and Migration Policies in China

2.1 Introduction

Understanding the mechanisms of intergenerational income mobility and inequality has long been a question of interest for economists. Previous literature has found childhood location to be a key factor, having substantial impacts on future social mobility and income. (Benabou (1993), Fernandez and Rogerson (1996, 1998), Chetty et al. (2014b), Eckert and Kleineberg (2021)). There are many channels underlying this relationship, with location-based government policies often serving as a primary factor. Many countries employ policies that generate dispersion in regional opportunity, which can have drastic consequences for inequality and intergenerational mobility. While most previous work studying this connection has focused on developed countries, higher incidence of poverty in developing countries makes it a compelling topic in low-income contexts, as low mobility implies that children born in poverty tend to stay in poverty as adults. It is thus especially important to evaluate the role and impact of regional policies on promoting social mobility and reducing inequality in developing countries.

China represents an ideal economic and institutional setting to study this question. Its rapid growth over the past 40 years has given rise to prosperous cities and lifted many out of poverty, but has also been accompanied by substantial increases in regional inequality (Kanbur and Zhang (2005), Xu (2011), Zhang and Zou (2012)). Three region-based policies

have potentially contributed to this outcome. First, public K-12 education spending is decentralized to local governments, with the amount of spending in each province coming from local government revenue and community contributions. Second, the central government (specifically, the Ministry of Education) dictates the number of students from each province who are admitted to public colleges in China. Third, China's *Hukou*, or residential permit, system requires individuals to obtain a local provincial Hukou in order to access key public resources such as health care, public schooling, subsidized housing markets, social pension, and other important provisions; individuals can and often do migrate without a Hukou, but cannot access these resources.

Each of the policies affects the opportunities available to children based on birth location. Children born in richer regions receive higher levels of per-student spending on K-12 public education and receive larger allocations of seats at public colleges relative to the number of applicants in the province. The Hukou system limits migration, especially from rural to urban areas, by imposing substantial moving costs for potential migrants. Children born in low-income areas with lower-quality public and private resources thus not only face inequities in the quality of opportunities available to them at birth, but also cannot access opportunities in better-resourced regions through parental migration. These policies generate intergenerational persistence in education and location, leading to intergenerational persistence in income.

To study the impacts of these policies, we build a two-province heterogeneous-agent model of overlapping generations that captures key features of the Chinese institutional context and use it to evaluate the effects of the country's public education and migration policies on aggregate outcomes such as per capita income, inequality, intergenerational mobility, and welfare, as well as individual outcomes such as education, individual earnings, and migration. To do this, we first aggregate individual- and province-level data on college admissions, income, and migration across rich and poor provinces to generate an "aggregate rich" and an "aggregate poor" province. Second, we calibrate the model's steady state to match key features of the aggregated provinces. Finally, we use the calibrated model to carry out several counterfactual experiments to understand how China's institutional setting impacts the outcomes of interest.

We conduct four policy experiments, which we call "Hukou," "Public Spending," "Merit," and "Equity." The Hukou policy lifts Hukou migration costs, removing a longstanding barrier to regional migration in China. This allows individuals to move freely from the poor to the rich province with full access to public resources and housing markets, improving efficiency

in sorting across locations. The Public Spending counterfactual equalizes per-student public spending across provinces, eliminating quality differences in childhood education.

Equity and Merit modify provincial college seat allocations, holding total seats fixed. In Equity, each province sends an equal proportion of children to each tier of college, improving college opportunities for children in the poor province. In Merit, students with the highest level of pre-college human capital are assigned seats at the top college with seats available regardless of birth region, effectively removing the provincial aspect of college allocations.

The policies interact, with education policies determine the level of opportunity within regions, and the Hukou policy determines who can take advantage of these opportunities. Together, they contribute to a widely-held notion that rural, low-income families have few options for upward mobility, leading to criticism and social unrest. While the Hukou system has been somewhat relaxed in recent years, significant obstacles remain: in June 2022, Shanghai eased Hukou requirements by waiving social pension payments for graduates of the world's top 50 universities.[1] Still, the situation leaves scope for further adjustments, motivating the analysis using our model.

We estimate the policy effects on long-run, steady state outcomes as well as short-term effects on a fixed population (namely, the baseline steady state). The former exercise examines aggregate effects on the entire population while the latter allows us to identify effects on interesting subgroups of the population (e.g. regional populations or income quartiles) without encountering selection issues regarding differences in regional compositions of human capital or income between counterfactual steady states.

For the remainder of the paper, we express welfare changes in terms of the compensating consumption equivalent as a percentage of baseline consumption.[2] We find that the Hukou and Public Spending counterfactual policies have significant effects on spatial distributions of income and education, intergenerational mobility, and welfare in the steady state. We observe that lifting Hukou migration restrictions decreases steady state inequality by 3.7% and intergenerational elasticity by 4.2%, increases population size in the rich province, and produces overall welfare gains of 8%. We also find that short-run impacts on children vary across regions and parental income quartiles. The policy increases average welfare by 14.1% for children born in the poor province and decreases welfare by 2.2% in the rich province, with children born to higher-income parents in the rich region particularly worse off. Even though the policy lifts migration costs, the primary short-term beneficiaries in the short term are not migrants themselves, but rather individuals who do *not* migrate: welfare in

[1] Chen (2022)

[2] We describe this computation in more detail in the Policy Evaluation section.

poor regions increases from the option value of unrestricted migration and lower housing prices after net outflow of migration.

Equalizing regional per-student public spending decreases steady state intergenerational persistence by 6.7%, produces overall welfare gains of 3.7%, and increases population size in the poor province. As expected, the policy benefits almost all children born in the poor province and harms almost all children born in the rich province, increasing welfare in the poor province by 14.5% and decreasing welfare in the rich province by 12.9%. We find that welfare gains in the poor province are driven by improved human capital accumulation and less migration, while the only beneficiaries of the policy in the rich province are children whose parents do not have a Hukou and thus cannot access the rich province's public schooling.

Changing college allocations has modest effects on steady state inequality, intergenerational elasticity, population size, and welfare, but has noticeable short-term impacts on the distribution of college admissions for children born across quartiles of parent income. Under the Merit counterfactual, we reallocate seats from the poor to the rich province, decreasing (increasing) welfare in the poor (rich) province by 4.2% (9.4%) through lower (higher) educational attainment and human capital. Under the Equity counterfactual, we reallocate seats from the rich to the poor province, increasing welfare for those born in poor provinces by 3.7% by providing better access to education. Interestingly, many children in the rich province are also better off from this policy as a smaller college allocation incentivizes greater parental investment into children's human capital by raising standards for college admission from the rich province. For both of these allocations, effects are most concentrated among children of higher-income parents who are most likely to have educational attainment outcomes improved or worsened by the policy.

Previous Literature

Our work follows several strands of related literature. First, we build on an extensive body of work using overlapping-generations models to study intergenerational outcomes and the income distribution, beginning with the seminal works of Becker and Tomes (1979, 1986) and Loury (1981). Later work augments these models with spatial equilibria and various public school finance policies. Durlauf (1996) develops a theory of income inequality driven by endogenous selection into neighborhoods. Fernandez and Rogerson (1996, 1998) and Bénabou (2002) add public education expenditure policy to this theory and find that redistributive policies can improve outcomes for poor communities as well as steady-state income and welfare for the overall population. Lee and Seshadri (2019) build a model with multiple stages of human capital investment across the parent's and child's life cycle, identifying the

importance of early childhood investments (when parents are most likely to face borrowing constraints) in the calibrated model. Their results suggest that education subsidies provided during early childhood have a significant effect on income persistence. In a similar vein, Eckert and Kleineberg (2021) assess the spatial general equilibrium effects of equating school funding, finding that positive effects on children's education in low-skilled families are attenuated by residential choices of those families toward locations with lower education quality in general equilibrium. Zheng and Graham (2022) find that the use of property taxes to fund public education engenders inequality and low intergenerational mobility, as low-income households cannot afford to live in neighborhoods with high housing prices and their corresponding high-quality schools. Kotera and Seshadri (2017) evaluate different education funding program options across districts within states and conclude that funding programs with more equal distributions of public school spending result in higher rank-rank intergenerational mobility. Notably, most of this literature centers around the United States and other developed countries, with much less work on developing countries. Our work adapts various features of previous models into our own, with China representing a unique institutional and economic setting to study these types of policies.

Second, we contribute to the literature estimating economic effects of policies specific to China, with the college admissions mechanism and Hukou policy attracting significant attention. Yang (2020) and Guo et al. (2018) estimate effects of merit-based rather than provincial quota-based college admissions. While they report dissenting conclusions on the change in implied regional shares of the student body and income, both acknowledge the role of college admissions in perpetuating regional inequality.

Previous studies of Hukou policy have focused on cross-sectional effects on migration, urbanization, labor force composition, or regional income disparities (Whalley and Zhang (2007); Bosker et al. (2012); Dreger and Zhang (2017); Wang et al. (2020); Dulleck et al. (2020)). Our overlapping generations model brings a fresh perspective to examine the intergenerational impacts of these policies, exploring an additional channel through which the Hukou effects occur. Sieg et al. (2020) uses a model to estimate the effect of the Hukou on migration and human capital development in a way similar to ours, but our framework includes the college admissions mechanism and studies a broader set of policies and outcomes. To our knowledge, our paper builds the richest model of China's institutions that can quantify both short- and long-term individual and economy-wide policy effects of college admissions and Hukou policies on inequality and intergenerational mobility.

Third, we build on previous literature concerning intergenerational mobility in developing countries. The question of intergenerational mobility has attracted significant attention, but

has largely focused on developed countries (Chetty et al. (2014b,a), Solon (1999), Corak (2013), Patrizio (2007); see Black and Devereux (2011) for a relatively recent list of references). It is unclear whether these findings hold in developing countries, where different existing income distributions and institutions could change the mechanisms underlying income persistence. Hnatkovska et al. (2013) and Asher et al. (2018) studies the effect of the caste system on mobility in India. Alesina et al. (2020) examines intergenerational mobility in educational mobility across 27 countries in Africa, revealing that historical and geographical features had a substantial impact on mobility. Fan et al. (2021) and Yan and Deng (2022) estimate the intergenerational elasticity in China and obtain a value of .41, which we use as a target in our model. Our work builds a structural framework to better understand the institutions and other channels underlying intergenerational mobility in China.

Finally, we contribute to literature studying the effects of college admissions policies on income, welfare, and intergenerational mobility. One strand of this literature builds similar overlapping-generations models to assess policy counterfactuals (Abbott et al. (2019), Capelle (2020), Herskovic and Ramos (2019)). Abbott et al. (2019) conduct policy experiments varying the size and nature of the federal grant program and loan limits, finding increases in GDP and welfare attributed to status-quo federal aid along with greater gains from potential aid expansion. Capelle (2020) similarly experiments with financial aid and government transfers to colleges to study changes in welfare, inequality, and intergenerational mobility. Herskovic and Ramos (2019) evaluate Affirmative Action college admissions policies in Brazil, finding that these policies increase aggregate output and welfare while reducing intergenerational elasticity. This literature also includes studies of centralized college admissions systems, which are common in both developing and developed countries. These studies are largely empirical, employing changes in college admissions mechanisms or specific features of the existing mechanisms such as test score cutoffs to identify treatment effects. Otero et al. (2021) finds that affirmative action in the centralized admissions system in Brazil increased income for targeted students and decreased income for non-targeted students in almost a one-to-one manner, increasing equity without affecting overall efficiency. Mari et al. (2020) detect persistent effects on enrollment and elite formation from introducing a purely meritocratic admissions system in Japan due to displacement of rural applicants by urban applicants in higher education. We augment previous models with important characteristics of the Chinese institutional setting; namely, we integrate location-based education spending, college admissions, and Hukou policies to produce a unique framework for studying effects of college admissions policies.

Our paper proceeds as follows. Section 2 describes the institutional setting in China. Section 3 details the model. Section 4 outlines the data sets used and the model calibration process. Section 5 describes results of the model. Section 6 discusses the significance and potential policy implications of our results and concludes.

2.2 Background

China's economy and social structures have evolved significantly over the past 40 years, largely guided by the central government's planning and policy decisions. While the country has developed into an economic powerhouse during this period, there have been fears that growth has favored rich urban areas and left poor rural areas behind. Aforementioned regional education and migration policies exacerbate the issue by producing inequities in both levels of and access to opportunities and resources within a region, becoming a source of social discontent within the country.

Location-based education policies influence the level of opportunity within a region. Decentralized education spending causes K-12 public education funding in rich areas to be higher than that in poor areas. Tsang (1996) provides a detailed account of China's decentralization of public education finance. In 1985, China reassigned responsibility for financing of secondary education from the central government to lower-level governments.[3] As a result, poorer areas tend to have poorly-funded public schools. Figure 2.1 plots logged provincial per-student K-12 public education spending against logged provincial GDP per capita, revealing a clear positive relationship. The government has the capacity to mitigate gaps in public education quality by providing categorical grants and education subsidies to poorer areas. However, efforts to reduce these inequities have been insufficient to eliminate persistent gaps in educational quality.[4]

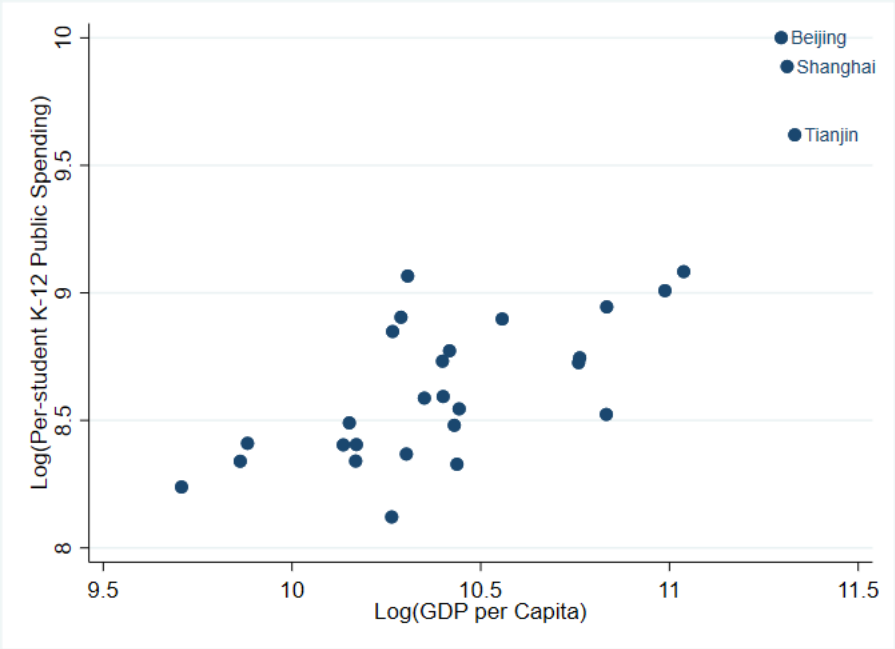
Moreover, the Ministry of Education allocates seats for admission to public colleges at the provincial level. Students are assigned to schools based on their preferences and performance on the *Gaokao*, or National College Entrance Exam (NCEE). The test varies by province, and students with the highest scores on the NCEE within their respective province receive seats at their top-choice schools until the school's provincial quota is filled.[5] Richer provinces receive more seats proportional to the number of applicants at higher-quality colleges, particularly

[3]Financing for *primary* education was the responsibility of the provincial government.

[4]Using 1989 data for 29 regions, Tsang (1996) calculates the correlation between per capita gross regional product with budgeted regional per-student primary and secondary education expenditures to be .9.

[5]The process of college seat allocation is detailed in the Appendix, but the mechanism itself is not important to our paper.

Figure 2.1: Provincial K - 12 Per-Student Public Spending vs. GDP per Capita



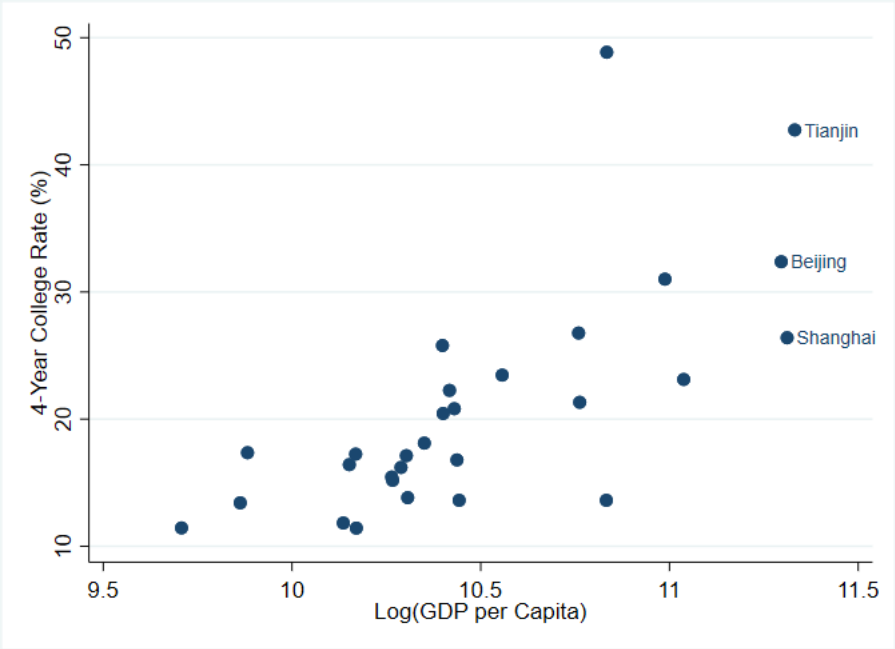
4-year and elite colleges. Figure 2.2 plots the 4-year college attendance rate (relative to the provincial age-18 population) against logged provincial GDP per capita, again showing that richer provinces send a higher percentage of their students to 4-year colleges. This is especially pertinent in China, where a cultural emphasis on higher education as a signal of both productivity and social status produces a strong correlation between an individual's schooling and his or her future success.

Previous literature (Yang (2020); Loyalka et al. (2017)) studying the efficiency of the allocation mechanism (in terms of allocating students with the highest human capital to the best schools) finds mixed results. Our model aligns with Yang (2020) in that eliminating provincial quotas provides more seats to richer regions; in other words, the "marginal student" (the lowest-skilled student admitted to a college) from a rich province has higher human capital than the marginal student from a poor province.[6]

China's *Hukou*, or residential permit, policy institutes barriers to migration into richer provinces, particularly from rural to urban regions. The Hukou system was originally instituted in 1958 to maintain a large rural workforce and curb urban population growth. Though the system is not as strict as it once was, it still continues to restrict rural-to-urban migration

[6]We note that development of pre-college human capital includes the effects of childhood conditions on cognitive development, which could explain the gap between the rich province's and the poor province's marginal student. This creates a challenge in defining the "fair" allocation of seats across provinces; for this reason, we describe the mechanism in terms of efficiency instead of fairness.

Figure 2.2: 4-year College Enrollment vs. GDP per Capita



today: individuals must possess a local Hukou to access the province's college allocation, K-12 public schooling, social insurance, subsidized housing markets, healthcare, and other key public provisions. Obtaining a Hukou in a desirable location is difficult and costly, often requiring some combination of employer sponsorship, a degree from a high-ranked college, a government job, a "special talent," and contribution into the region's social pension system for a number of years. This restricts access to better public resources and educational opportunities for current and potential migrants and their children. The issue is prevalent for a large portion of China's population: rural migrant workers comprise 36% of the labor force in China, with 60% of them long-distance migrants (China Labour Bulletin (2022)).

2.3 Model

In this section, we develop our spatial, heterogeneous-agent model with overlapping generations that characterizes salient features of the Chinese institutional setting. We use the model to study effects of Chinese policies on inequality, intergenerational mobility, and welfare.

The basic setup of the model comes from the standard Becker and Tomes (1979) overlapping generations framework. The economy is populated by a unit mass of dynastic households. Each household consists of two individuals, an adult and a child. Individuals live for

two periods, one period as a child and one period as an adult, with each adult having one child.

We augment the model with features that encapsulate China's major regional policies. Our model has two representative provinces (a "rich," or high-productivity , and a "poor," or low-productivity, province) that differ in effective wages, housing costs, social security benefits, public education spending levels, and schooling allocations. First, we model Hukou restrictions by imposing migration costs on poor-to-rich migrants who did not attend the elite college. Individuals are permitted to migrate across provinces without a Hukou, but they receive lower wages, pay higher housing costs, and cannot access local social security benefits, college allocations, or public schooling for their children.[7]

Second, we model educational policies. The rich province receives higher public education spending, resulting in higher childhood education quality in the rich province. We model regional college quotas by distinguishing three levels of educational attainment (high school, four-year college, and tier-1, or elite, college), where the two colleges admit different shares of children from each province. Each college's provincial quota combined with the human capital distribution in each province implies a human capital threshold for each province and college, where students are admitted to a college if their pre-college human capital exceeds the respective threshold for their households' Hukou region.

In the following subsections, we detail how our model emulates the institutional setting and describe each step in individual decision-making.

2.A.1 Adult Consumption and Investment Decisions

At the beginning of each period, each dynasty is characterized by five state variables: its province ($r \in \{\ell, h\}$, corresponding to the low-productivity and high-productivity province, respectively), the effective human capital of the parent (H_2), an idiosyncratic income shock (ξ_y), an idiosyncratic preference shock for location (ξ_r), and Hukou status $m \in \{0, 1\}$, where $m = 0$ denotes that the parent possesses the local Hukou. Migration decisions for the child occur at the end of each period immediately before the child becomes an adult, so r designates the region where the adult works and the household resides in the current period.

Adults earn labor income y given by:

[7] In reality, children of non-Hukou migrants either stay in their hometown with extended family ("left-behind" children) or migrate with their parents. In the former case, previous literature has shown that that left-behind children suffer worse human capital outcomes (Zhang et al. (2014)). In the latter case, children typically cannot attend local public schools and instead attend private migrant schools, which are expensive and lower quality than the public schools in the area. We abstract away from the difference and assume that the two options are effectively the same in terms of human capital accumulation.

$$y = \bar{w}\alpha_r H_2 \xi_{y,r} \tag{2.1}$$

where $\bar{w}$ is the wage per unit of human capital, α_r represents a provincial wage premium (normalized to 1 for $r = \ell$), and H_2 is the adult's human capital after college. The idiosyncratic income shock $\xi_{y,r}$ is realized before the current period but after the migration decision in the previous period, and is lognormally distributed with mean 0 and standard deviation $\sigma_{y,r}$. It is i.i.d both across dynasties and time.

Income is taxed at a provincial rate of τ_r to fund provincial spending on primary and secondary education. Each household purchases one unit of housing at province-specific housing price P_r. Individuals maximize utility $u(c, \xi_r)$ by allocating remaining income between consumption c and private investment in child's human capital x. Government social insurance guarantees that adults holding a local Hukou for province r receive at least $\bar{y}_r$ in income after housing expenditure. They cannot save or borrow, so the dynasty's only mechanism to transfer resources to future periods is investment in the human capital of the child. Adult holders of a local Hukou ($m = 0$) thus face the following budget constraint:

$$c + x = \max\{(1 - \tau_r)y - P_r, \bar{y}_r\} \tag{2.2}$$

All households residing in the poor province hold the local Hukou ($m = 0$). A household without a Hukou ($m = 1$) in the rich province faces several costs:

(i) It incurs a "migrant investment tax" $\tau_{m,x}$, which represents the cost of educating their child at low-quality, private migrant schools in the rich province.[8]

(ii) It suffers a pecuniary "migrant wage tax" $\tau_{m,y}$, which captures the lower pay relative to Hukou holders that migrants tend to receive, conditional on human capital.

(iii) It pays a "housing tax" $\tau_{m,P}$, which represents their lack of access to subsidized local housing markets.

(iv) It does not receive the rich region's social insurance $\bar{y}_h$, instead receiving $\bar{y}_\ell$.

(v) The household's child competes with children from the poor province for college seats.

[8] Equivalently, the cost of educating a "left-behind" child.

This specification captures the costs of migration for individuals who move from poor to rich regions without a Hukou. We can re-express their consumption decision as simply the rich Hukou holder's problem with the following modifications. A household in the rich region with Hukou status $m = 1$ earns income y given by $(1^{'})$:

$$y = (1 - \tau_{m,y})\bar{w}\alpha_r H_2 \xi_{y,r} \tag{1'}$$

and faces the alternative budget constraint $(2^{'})$ as adults:

$$c + (1 + \tau_{m,x})x = max\{(1 - \tau_h)(1 - \tau_{m,y})y - (1 + \tau_{m,P})P_h, \bar{y}_\ell\} \tag{2'}$$

The adult faces these costs for one generation, and the child receives the rich province Hukou ($m = 0$) if he remains there in the next generation.

2.A.2 Children's Education

For the remainder of the paper, we denote the next generation with $'$ when the notation is unambiguous. Children are born in parent's province of work r, and their human capital accumulation occurs in two stages. In the first stage, they attend K-12 schooling in the region where the household resides. Their human capital from schooling depends on (i) parent's human capital H_2, (ii) private investment x, (iii) region-specific public funding $\bar{x}_r$, and (iv) human capital shock $\xi_H^{'}$, occurring after the private investment decision. Pre-college human capital $H_1^{'}$ is given by:

$$H_1^{'} = H_2^{\nu}(\omega x^{\rho} + (1 - \omega)\bar{x}_r^{\rho})^{\frac{\mu}{\rho}}\xi_H^{'} \tag{2.3}$$

where $\nu \in [0, 1]$, $\omega \in [0, 1]$, $\mu \in [0, 1]$, and $\rho \in [0, 1]$ are parameters.[9] ν, μ, and ρ are elasticities: ν is the elasticity of child's human capital with respect to parent's human capital; ρ is the elasticity of substitution between public and private education spending; and μ is the elasticity of child's human capital with respect to private and public composite investment. The constant ω is the weight on private relative to public education spending. The shock $\xi_H^{'}$ is iid across children and generations and is lognormally distributed with mean 0 and standard deviation σ_H. It is realized after the private human capital investment decision.

In the second stage, children potentially accumulate additional human capital depending upon college attendance. After completing K - 12 education, children are allocated to the

[9] Kotera and Seshadri (2017) use a similar formulation for the US.

four-year college ($e^{'}=2$) or the elite college ($e^{'}=3$) based on pre-college human capital $H_1^{'}$ and each college's provincial quota. Each quota implies a province-specific human capital cutoff $\bar{H}_{r,e}$, where children are admitted to college e if $H_1^{'}$ exceeds $\bar{H}_{r,e}$ for their birth region r. Children who do not attend either college finish with a high school education ($e^{'}=1$). That is,

$$e^{'}=\begin{cases}1, & \text{if } H_1^{'}\in[0,\bar{H}_{r,2})\\ 2, & \text{if } H_1^{'}\in[\bar{H}_{r,2},\bar{H}_{r,3})\\ 3, & \text{if } H_1^{'}\in[\bar{H}_{r,3},\bar{H})\end{cases} \tag{2.4}$$

Children born to households in the rich province without the local Hukou cannot access local public schools or the rich province's college allocation; they accumulate human capital according to ($3^{'}$):

$$H_1^{'}=H_2^{\nu}(\omega x^{\rho}+(1-\omega)\bar{x}_{\ell}^{\rho})^{\frac{\mu}{\rho}}\xi_H^{'} \tag{$3^{'}$}$$

and compete with children in the poor province, receiving education according to the following rule ($4^{'}$):

$$e^{'}=\begin{cases}1, & \text{if } H_1^{'}\in[0,\bar{H}_{\ell,2})\\ 2, & \text{if } H_1^{'}\in[\bar{H}_{\ell,2},\bar{H}_{\ell,3})\\ 3, & \text{if } H_1^{'}\in[\bar{H}_{\ell,3},\bar{H})\end{cases} \tag{$4^{'}$}$$

College is assumed to be costless, and the return to college is increasing in quality so that children will attend the highest-quality school where they are admitted. Children complete college with effective human capital $H_2^{'}=\alpha_{e'}H_1^{'}$.

2.A.3 Migration Decision and Migration Costs

After college, children realize an idiosyncratic location preference shock $\xi_r^{'}\sim N(\mu_{M,r},\sigma_M^2)$, which represents their preference to remain in the birth location. If the child chooses *not* to migrate, they realize $\xi_r^{'}$ in the next period. Then, children in both provinces make a migration decision to maximize expected continuation utility.

Children who relocate from the rich region to the poor region receive a Hukou and migrate costlessly ($m' = 0$). However, children who migrate from the poor to the rich province receive a Hukou only if they attended the elite college. Otherwise, while they live and work in the rich region, they face Hukou costs ($m' = 1$). The asymmetry between poor-to-rich and rich-to-poor migration is intentional: it is difficult obtain a Hukou in rich cities, where higher wages and better quality of public resources drive competition for living space, but not in poorer areas.

After the migration decision, the child realizes a lognormal wage shock $\xi_{y,r'} \sim logN(0, \sigma_{y,r'})$, becomes an adult, and has his own child. The next state begins with state variables $r^{'}$, $H_2^{'}$, $\xi_{y,r'}$, $\xi_{r'}$, and $m^{'}$.

Figure 2.3 below is a timeline of the events in the model.

2.A.4 Housing Prices

Regional housing prices are given by $P_r = A_r N_r^{\sigma_P}$, where A_r is a province-specific constant that captures housing demand, N_r is population size of province r, and σ_P is the inverse price elasticity of housing supply.

2.A.5 Steady State Equilibrium

Our baseline results study both differences in steady state equilibria across counterfactuals and transition paths between steady states. In this subsection I define a steady state equilibrium.[10] There are several values that households take parametrically: regional housing prices P_r, regional tax rates τ_r, regional public spending levels $\bar{x}_r$, and college admission thresholds $\bar{H}_{r,e}$. These values are constant in equilibrium so I suppress them as arguments.

Let $V_r(H, \xi_y, \xi_r, m)$ represent the value function for an adult living and working in region r with human capital H, income shock ξ_y, home-province preference shock ξ_r, and Hukou status m. Adults allocate income less taxes and housing costs into current-period consumption and investment to maximize dynastic utility. Investment increases future utility by improving children's human capital and subsequent education and migration outcomes. These tradeoffs are made explicit in the Bellman equations that follow. The Bellman equation for a household in the poor province is given by:

[10] We discuss transition dynamics in a later section.

Figure 2.3: Timeline of Model Events and Decisions

$$
\begin{aligned}
V_\ell(H_2, \xi_{y,\ell}, \xi_\ell, 0) &= \max_{c,x} \; u(c, \xi_\ell) + \beta \; max \; E \begin{cases} V_\ell(H_2^{'}, \xi_{y,\ell}^{'}, \xi_\ell^{'}, 0), \\ \mathbb{1}[e^{'} \leq 2] V_h(H^{'}, \xi_{y,h}^{'}, 0, 1), \\ \mathbb{1}[e^{'} = 3] V_h(H^{'}, \xi_{y,h}^{'}, 0, 0) \end{cases} \\
\text{s.t. } c + x &= max((1-\tau_\ell)\bar{w} H_2 \xi_{y,\ell} - P_\ell, \bar{y}_\ell) \\
H_1^{'} &= H_2^{\nu}(\omega x^{\rho} + (1-\omega)\bar{x}_\ell^{\rho})^{\frac{\mu}{\rho}} \xi_H^{'} \\
e^{'} &= \begin{cases} 1, & \text{if } H_1^{'} \in [0, \bar{H}_{\ell,1}) \\ 2, & \text{if } H_1^{'} \in [\bar{H}_{\ell,1}, \bar{H}_{\ell,2}) \\ 3, & \text{if } H_1^{'} \in [\bar{H}_{\ell,2}, \bar{H}) \end{cases} \\
H_2^{'} &= \alpha_{e'} H_1^{'} \\
r' &= \underset{r^* \in \{\ell, h\}}{argmax} E(V_{r^*}(H_2^{'}, \xi_{y,r^*}^{'}, \mathbb{1}[r^* = \ell]\xi_\ell^{'}, \mathbb{1}[r^* = h]\mathbb{1}[e^{'} < 3]) \\
log\xi_{y,r'}^{'} &\sim N(0, \sigma_{y,r'}^2) \\
log\xi_H^{'} &\sim N(0, \sigma_H^2) \\
\xi_\ell^{'} &\sim N(\mu_{M,\ell}, \sigma_M^2)
\end{aligned}
$$

Note that because the children are born in the poor province, they realize preference shock $\xi_\ell^{'}$ as an adult if they decide to not migrate, and they face one period of migration costs if they do not obtain an elite college education.

The Bellman equation for a household with a Hukou in the rich province is given by:

$$
\begin{aligned}
V_h(H_2, \xi_{y,h}, \xi_h, 0) &= \max_{c,x} \; u(c, \xi_h) + \beta \; max \; E \begin{cases} V_\ell(H_2^{'}, \xi_{y,\ell}^{'}, 0, 0), \\ V_h(H_2^{'}, \xi_{y,h}^{'}, \xi_h^{'}, 0) \end{cases} \\
\text{s.t. } c + x &= max((1-\tau_h)\alpha_h \bar{w} H_2 \xi_{y,h} - P_h, \bar{y}_h) \\
H_1^{'} &= H_2^{\nu}(\omega x^{\rho} + (1-\omega)\bar{x}_h^{\rho})^{\frac{\mu}{\rho}} \xi_H^{'} \\
e^{'} &= \begin{cases} 1, & \text{if } H_1^{'} \in [0, \bar{H}_{h,2}) \\ 2, & \text{if } H_1^{'} \in [\bar{H}_{h,2}, \bar{H}_{h,3}) \\ 3, & \text{if } H_1^{'} \in [\bar{H}_{h,3}, \bar{H}) \end{cases} \\
H_2^{'} &= H_1^{'} \alpha_{e'} \\
r' &= \underset{r^* \in \{\ell, h\}}{argmax} E(V_{r^*}(H_2^{'}, e^{'}, \xi_{y,r^*}^{'}, 0, 0)) \\
log\xi_{y,r'}^{'} &\sim N(0, \sigma_{y,r'}^2) \\
log\xi_H^{'} &\sim N(0, \sigma_H^2) \\
\xi_h^{'} &\sim N(\mu_{M,h}, \sigma_M^2)
\end{aligned}
$$

Finally, the Bellman equation for a household without a Hukou in the rich province is given by:

$$
\begin{aligned}
V_h(H_2, \xi_{y,h}, 0, 1) &= \max_{c,x} \; u(c,0) + \beta \; max \; E \begin{cases} V_\ell(H_2^{'}, \xi_{y,\ell}^{'}, 0, 0), \\ V_h(H_2^{'}, \xi_{y,h}^{'}, \xi_h^{'}, 0) \end{cases} \\
\text{s.t. } c + (1+\tau_{m,x})x &= max((1-\tau_h)(1-\tau_{m,y})\alpha_h \bar{w} H_2 \xi_{y,h} - (1+\tau_{m,P})P_h, \bar{y}_\ell) \\
H_1^{'} &= H_2^{\nu}(\omega x^{\rho} + (1-\omega)\bar{x}_\ell^{\rho})^{\frac{\mu}{\rho}} \xi_H^{'} \\
e^{'} &= \begin{cases} 1, & \text{if } H_1^{'} \in [0, \bar{H}_{\ell,2}) \\ 2, & \text{if } H_1^{'} \in [\bar{H}_{\ell,2}, \bar{H}_{\ell,3}) \\ 3, & \text{if } H_1^{'} \in [\bar{H}_{\ell,3}, \bar{H}) \end{cases} \\
H_2^{'} &= H_1^{'} \alpha_{e'} \\
r' &= \underset{r^* \in \{\ell,h\}}{argmax} E(V_{r^*}(H_2^{'}, \xi_{y,r^*}^{'}, 0, 0)) \\
log\xi_{y,r'}^{'} &\sim N(0, \sigma_{y,r'}^2) \\
log\xi_H^{'} &\sim N(0, \sigma_H^2) \\
\xi_h^{'} &\sim N(\mu_{M,h}, \sigma_M^2)
\end{aligned}
$$

Note the differences in costs of investment ($\tau_{m,x}$), income ($\tau_{m,y}$), housing ($\tau_{m,P}$), human capital accumulation ($\bar{x}_\ell$ instead of $\bar{x}_h$), and thresholds for college attendance ($\bar{H}_{\ell,e}$ instead of $\bar{H}_{h,e}$) from the previous specification for a rich province Hukou holder.

We proceed to define the steady state. Let $\mu_r(H_2, \xi_{y,r}, \xi_r, m)$ represent the measure over effective human capital H_2, income shocks human capital $\xi_{y,r}$, location preference shock ξ_r, and Hukou status m in province r. Let $T(\mu_\ell, \mu_h) = (\mu_\ell^{'}, \mu_h^{'})$ be the mapping from the measures in period t to the measures in period $t+1$. We define a steady state equilibrium as follows: given housing prices $\{P_\ell^*, P_h^*\}$, regional human capital thresholds $\{H_{\ell,2}^*, H_{\ell,3}^*, H_{h,2}^*, H_{h,3}^*\}$, and public spending levels $\{\bar{x}_\ell, \bar{x}_h\}$, a steady state in our model is a pair (μ_ℓ^*, μ_h^*) such that:

1. Households maximize utility.
2. (μ_ℓ^*, μ_h^*) is a fixed point of the mapping T, i.e. $T(\mu_\ell^*, \mu_h^*) = (\mu_\ell^*, \mu_h^*)$.
3. Housing and labor markets clear.

2.4 Model Calibration

In this section, we discuss our model calibration. We first describe our sources of data. We then outline the steps we take to aggregate province-level data into moments that discipline the representative poor province and rich province in our model. Finally, we detail the calibration process, in which we choose parameters to calibrate the steady state of the model to match important features of China such as educational attainment, intergenerational mobility, and the income distribution. For the remainder of the paper, we refer to the model calibrated to current Chinese data the "baseline" model.[11]

2.A.1 Data Sources

We combine data from multiple sources to calculate key moments that guide our calibration. We use administrative province-level college enrollment data in 2011, which informs the number of applicants and admitted students for each tier of college (tier 1, tier 2, 4-year, Project 211, Project 985, and Peking or Tsinghua University) from each province.[12] We combine this with provincial age-7 population counts from the 2000 Chinese Census (as these children applied for college in 2011) to calculate admissions rates for each province and tier of schooling. Our primary source for other province-level data is the National Bureau of Statistics of China (NBS), which provides information on regional GDP, population size, government social insurance (or *dibao*), and enrollment and per-student public spending by K-12 schooling level (primary, middle, and high school).

Individual- and family-level data comes from the China Family Panel Survey (CFPS), a household survey with waves every two years from 2010-2018. It features data on location, Hukou status, family- and individual-level income, education, and other characteristics such as spending on education and housing. We use this data to estimate Hukou cost parameters and migration flows. Survey questions asking directly about migration appear to severely undercount the number of migrants based on the number of people changing location from their birthplace, so we approximate migration flows by comparing an individual's province at age 12 to their province as an adult. We are particularly interested in whether an individual moves from a poor province during childhood to a rich province during adulthood or vice-versa.

[11] A detailed explanation of the steady-state solution process can be found in the Appendix.

[12] Project 211 (less prestigious) and Project 985 (more prestigious) are two projects created by the Ministry of Education to raise education and research quality and expand college access. The Project 211 tier has 116 schools, while Project 985 includes 39 of those 116. Peking and Tsinghua University are the two most prestigious colleges in China.

2.A.2 Aggregating the Data to Two Regions

First, we must categorize provinces as low-income or high-income to generate the representative regions in the model. We use 40,000 yuan/year in 2011 GDP per capita as an income cutoff to distinguish between "poor" and "rich" provinces in the data, and we aggregate the provinces within each group into the representative poor and rich province, respectively. The resulting categorization is consistent with that used in previous literature (Whalley and Zhang (2007)). The calibration targets for the two provinces in our model are based on population-weighted data from provinces within each group.so

Table 2.1 summarizes the results of this exercise. Our aggregate poor province contains 63.9% of the population, has an average income of 28,588 yuan, and spends 4,693 yuan per student on public education. 16.2% of children with a poor-province Hukou attend a four-year college, and 3.9% of these children attend an elite college. Our aggregate rich province holds 36.1% of the population, has an average income of 56,319 yuan, and spends 7,439 yuan per student on funding public education. 24.8% of children with a rich-province Hukou attend a four-year college, and 6.7% of these children attend an elite college. The list of provinces in each aggregate province can be found in the Appendix.

Table 2.1: Results of Province Aggregation

	Population (%)	GDP per capita (RMB)	K-12 spending (RMB/student)	4-year (%)	Tier 1 (%)
Poor	63.9	28,588	4,693	16.2	3.9
Rich	36.1	56,319	7,439	24.8	6.7

2.A.3 Calibration Procedure

The only unspecified functional form is the period utility function; we assume that $u(c, \xi_r)$ is given by $log(c + \xi_r)$. The model has 31 parameters to be calibrated: discount factor β; human capital accumulation parameters $\nu, \omega, \rho, \delta$, and σ_H^2; income parameters $\bar{w}, \alpha_h, \sigma_{y,\ell}^2$, and $\sigma_{y,h}^2$; returns to education α_1, α_2, and α_3; housing parameters A_ℓ, A_h, and ϕ; public social insurance levels $\bar{y}_\ell$ and $\bar{y}_h$; public investment levels $\bar{x}_\ell$ and $\bar{x}_h$ and tax rates τ_ℓ and τ_h; migration cost parameters $\tau_{m,x}, \tau_{m,p}$, and $\tau_{m,y}$; college thresholds $\bar{H}_{2,\ell}, \bar{H}_{3,\ell}, \bar{H}_{2,h}$, and $\bar{H}_{3,h}$; and preference shock means $\mu_{M,\ell}, \mu_{M,h}$ and variance σ_M^2. We proceed to calibrate the model in two steps. First, we set 17 parameters outside of the model using estimates from data and previous literature. Then, we jointly identify the remaining 14 parameters by targeting

14 moments in the steady-state equilibrium of our model to 14 equivalent moments in the data. We describe each step in the following subsections.

External Calibration

Table 2.2 lists externally-calibrated parameters. Many of these estimates are based on previous literature using similar models or producing values relevant to our model. The discount factor β is calibrated at .55 based on an annual discount rate of .98 for a period length of 30 years. The weight on private investment $\omega = .324$, return to total investment $\mu = .334$, and elasticity of substitution between public and private investment $\rho = .589$ come from Kotera and Seshadri (2017), who incorporate public and private investment in a similar specification for human capital accumulation. Education premia relative to high school for the four-year college ($\alpha_2 = 1.44$) and elite college ($\alpha_3 = 2.02$) are derived from the return to an average 4-year college in Zhong (2011) and the return to elite education in Jia and Li (2021), respectively. The estimate of inverse price elasticity of housing supply $\sigma_P = 2$ comes from Chow and Niu (2015). The income penalty for migrants without a Hukou is set at $\tau_{m,y} = .08$, which is consistent with recent work estimating the urban Hukou wage premium at 6-10%. (Boffy-Ramirez and Moon (2018), Dreger and Zhang (2017)).

The remaining parameters in Table 2.2 come from data from the CFPS and the National Bureau of Statistics of China (NBS). Provincial education spending tax rates τ_ℓ and τ_h are calculated from public education expenditure as a share of provincial GDP within each aggregate province. Social insurance payments and per-student public spending levels for the representative poor and rich provinces are informed by NBS data on social security and K - 12 public spending. We calculate $\bar{y}_\ell$ and $\bar{y}_h$ by aggregating over province-level data on unemployment insurance expenses and the number of beneficiaries. Similarly, we compute per-student public spending levels by combining NBS provincial data on schooling expenditures with the number of K-12 students. We identify the non-Hukou migrant expenditure taxes on private investment ($\tau_{m,x}$) and housing ($\tau_{m,p}$) using CFPS data with the following regression:

$$log(spend_{f,t}) = \alpha_0 + \alpha_1 log(y_f) + \gamma_r + \tau_{m,t} richprov_f(1 - hukou_f) + \varepsilon_f$$

where $\tau_{m,t}$ is the coefficient of interest. $t = x$ or $t = p$ denotes education or housing, respectively, $spend_{f,t}$ is family spending, y_f is family income, $richprov_f$ is an indicator for whether the individual lives in a rich province, and $hukou_f$ is an indicator for whether the

Table 2.2: Externally-Calibrated Parameter Values

Parameter	Value	Source	Description
β	0.55	Standard	Discount factor
ω	0.37	Kotera and Seshadri (2017)	Weight on private investment
ρ	0.589	Kotera and Seshadri (2017)	Elast. of subs. b/t investment types
μ	0.36	Kotera and Seshadri (2017)	Return to total investment
α_1	1	Normalization	Return to high school education
α_2	1.44	Zhong (2011)	Return to four-year college
α_3	2.02	Jia and Li (2021)	Return to elite college
σ_P	2	Chow and Niu (2015)	Inverse price elasticity of housing supply
$\tau_{m,y}$	.08	Boffy-Ramirez and Moon (2018), Dreger and Zhang (2017)	Income penalty w/o Hukou
$\bar{y}_\ell$	1,716	NBS	Province ℓ social insurance
$\bar{y}_h$	3,456	NBS	Province h social insurance
$\bar{x}_\ell$	4,693	NBS	Province ℓ public investment
$\bar{x}_h$	7,439	NBS	Province h public investment
τ_ℓ	.023	NBS	Province ℓ public education tax
τ_h	.014	NBS	Province h public education tax
$\tau_{m,x}$	.55	CFPS	Investment cost w/o Hukou
$\tau_{m,p}$	.40	CFPS	Migrant housing cost w/o Hukou

family's household head holds a local Hukou. Since spending on housing and education are family-level decisions, we use f to denote the use of family-level variables. Our regressions yield $\tau_{m,x} = .55$ and $\tau_{m,p} = .4$, implying that migrant families living in a rich province without a local Hukou spend 55% more on education and 40% more on housing than families with one, controlling for income.

Internal Calibration

We calibrate the remaining 14 parameters using a Nelder-Mead simplex algorithm to choose 14 parameters to minimize a loss function based on the weighted sum of squared errors

between 14 model-based moments and analogous moments in the data. While all internally-calibrated parameters are jointly identified by the moments of the model, each of these parameters is intuitively directly related to one particular moment. Table 2.3 lists the internally-calibrated parameters alongside their associated moments in the data and the model. In this section, we explain each parameter-moment relationship.

We use provincial enrollment rates in the administrative data to construct the four human capital thresholds $H_{2,\ell}$, $H_{3,\ell}$, $H_{2,h}$, and $H_{3,h}$, choosing thresholds to send the correct proportions of students from each province to the four-year and elite college in the baseline steady state.

Housing constants A_ℓ and A_h are pinned down by mean provincial housing expenditures as a share of earnings in the NBS. We find that households in poor provinces spend 12.5% on housing while households in rich provinces spend 10.5%. We incorporate these expenditure shares with targeted average earnings in each province to calculate housing prices P_r, which we combine with targeted provincial population sizes N_r to back out housing constants $A_\ell = \frac{P_\ell}{(N_\ell)^{\sigma_P}}$ and $A_h = \frac{P_h}{(N_r)^{\sigma_P}}$.

$\bar{w}$, α_r, $\sigma^2_{y,\ell}$, and $\sigma^2_{y,h}$ relate to the income distribution. Wage per effective human capital $\bar{w}$ is chosen to match average earnings in the poor province, while $\alpha_r > 1$ sets the rich-province wage premium to match average earnings in the rich province. The variance of the regional income shocks $\sigma^2_{y,r}$ are informed by the standard deviation of the regional log earnings distributions.

An individual's migration decision depends on his preference across locations, which depends on the realization of location preference shock $\xi_r \sim N(\mu_{M,r}, \sigma^2_M)$. We calibrate the mean of preference shock $\mu_{M,ell}$ to match the proportion of the population living in poor provinces in China. Increasing $\mu_{M,\ell}$ implies an increased preference for children born in the poor region to stay and work close to their family, which increases steady-state poor province population size. Preference shock variance σ^2_M determines the likelihood that an individual receives a shock large enough to migrate. Thus, σ^2_M is informed by matching excess migration flows; in our model with two provinces, this amounts to migration volume from the rich province to the poor province. [13]

The final parameters in Table 2.3 are used to match measures of social and intergenerational mobility. Since human capital shock ξ_H affects pre-college human capital, it greatly influences children's capacity to receive a college education, especially for children of low-

[13] We recognize a caveat in interpreting our model: empirical poor-to-rich migration volume exceeds rich-to-poor migration volume in the CFPS. This suggests that provincial population sizes have not yet reached steady state in China, whereas our calibration treats the economy as if they have. However, analysis in our model requires a definitive target for the steady state, which is difficult to project in the future.

income parents who receive relatively little private investment in human capital. We thus calibrate σ_H^2 to match the proportion of children from the bottom parent income quintile attending at least a 4-year college, which serves as a measure of access to opportunity. ν is the elasticity of children's human capital with respect to parent's human capital absent investment. The parameter characterizes the hereditary aspect of human capital and subsequently income transmission. We estimate the intergenerational elasticity (IGE) in the model by taking the slope coefficient in a regression of child's log earnings on parent's log earnings in the steady state, and we choose ν to match it to the estimate for China produced in Yan and Deng (2022) (IGE = .410).

Table 2.3: Internally-Calibrated Parameters and Model Fit

Parameter	Description	Value	Moment (Source)	Value	
				Target	**Model**
$\bar{H}_{2,\ell}$	Four-year college threshold, province ℓ	176	% province ℓ attending 4-year college (admin data + Census)	16.2	16.2
$\bar{H}_{3,\ell}$	Elite college threshold, province ℓ	299	% province ℓ attending elite college (admin data + Census)	3.9	3.9
$\bar{H}_{2,h}$	Four-year college threshold, province h	217	% province h attending 4-year college (admin data)	24.8	24.7
$\bar{H}_{3,h}$	Elite college threshold, province h	402	% province h attending elite college (admin data)	6.7	6.7
$\mu_{M,\ell}$	Mean of poor province pref shock ξ_ℓ	181	Location size in poor province (%) (NBS)	63.9	64.1
A_ℓ	Housing constant, province ℓ	8,773	Housing expenditure share, province ℓ (%) (NBS)	12.5	13.0
A_h	Housing constant, province h	45,592	Housing expenditure share, province h (%) (NBS)	10.5	10.3
$\bar{w}$	Wage per effective human capital	188	Average yearly earnings in poor province (NBS)	28,588	27,543
α_h	Rich-province premium in earnings	1.033	Average yearly earnings in in rich province (NBS)	56,318	57,809
$\sigma_{y,\ell}^2$	Variance of income shock ξ_y, poor province	1.03	Std. dev. of log earnings, poor province (CFPS)	1.16	1.12
$\sigma_{y,h}^2$	Variance of income shock ξ_y, rich province	.28	Std. dev. of log earnings, rich province (CFPS)	1.07	1.06
σ_H^2	Variance of human capital shock ξ_H	.58	% from first quintile attending college from poor province (CFPS)	6.1	7.0
σ_M^2	Variance of preference shocks $\xi_{M,r}$	636	Migration volume from rich to poor (%) (CFPS)	4.9	6.4
ν	Elasticity of child's HC wrt parent's HC	.411	Intergenerational elasticity (Fan et al. (2021))	.410	.400

Our calibrated model matches key characteristics of income and regional mobility, educational attainment, and the income distribution well. We also verify the performance of our model in matching non-targeted moments in the Appendix.

2.5 Policy Evaluation

We study the effects of policy changes on aggregate outcomes such as intergenerational mobility, earnings, inequality, and welfare. Moreover, we are interested in how the effects of policy are distributed across the parent income distribution—in particular, whether policy can provide better opportunities or outcomes for vulnerable subsets of the population, such as poor households with little capacity to invest in children's human capital and without the option to freely migrate.

The current policies in China either determine the quality of opportunity available within a region or access to opportunity available elsewhere. Thus, our policy experiments modify specific features of the Chinese institutional context along one of these dimensions. We evaluate four counterfactual policies, which we call Hukou, Public Spending, Merit, and Equity. We choose these policies because each one's implementation can be justified through equity or efficiency concerns. The Hukou counterfactual lifts Hukou restrictions, eliminating a longstanding policy that continues to distort human capital investment and migration decisions. In the model, we remove the five migration costs for poor-to-rich migrants without an elite education, allowing all individuals to efficiently sort across locations. The Public Spending counterfactual equates per-student public education spending across the two regions (as if the central government equally allocated all educational spending), holding the national public spending-earnings ratio fixed. This eliminates differences in human capital accumulation caused by differences in public school quality and removes one major source of regional inequality during childhood, promoting more equitable outcomes.

The latter two policies directly affect the allocation of college seats between the two regions. In the Merit counterfactual, we lift the provincial element of college seat allocation by equating human capital thresholds across provinces. This allocates children with the highest levels of pre-college human capital to the best college seat available regardless of location, improving the aggregate efficiency of the allocation in terms of total earnings. Each college's human capital thresholds are higher in the rich province than the poor province at baseline, implying that the Merit counterfactual allocates more seats to the rich province and is thus a *regressive* intervention. In the Equity counterfactual, we set provincial quotas such that equal proportions of children from both regions attend each tier of college. This effectively mitigates equity concerns over the "regional fairness" of the college policy by providing children born in each region equal opportunity to attend college. Note that a higher proportion of children from the rich province attend both levels of college in the baseline model, making the Equity counterfactual a *progressive* intervention by improving access to college for children in the poor province.

Each counterfactual creates incentives to live in a region and could thus change equilibrium location sizes. This induces corresponding changes in regional quotas, per-student public spending levels, and housing prices. In changing regional quotas, we hold each college's *total* capacity constant from baseline. The Hukou and Public Spending college allocations adjust to location size changes by reallocating provincial seats at each college equal to the baseline overall attendance rate at that college multiplied by the net migration flow relative

to baseline in each respective region.[14] In the Equity and Merit counterfactuals, the new college allocation rules dictate how to set the quotas. We adjust provincial per-student public education spending by fixing from baseline the ratio of spending to average earnings in each province (except the Public Spending counterfactual, in which both provinces receive the same ratio of spending to average earnings across the country as in baseline). Finally, housing prices adjust to population sizes according to the previously-defined relationship $P_r = A_r N_r^{\sigma_P}$.

Steady state results allow for comparisons in aggregate outcomes such as total income and welfare, population size, intergenerational mobility, and inequality. However, comparing steady state outcomes for different subgroups becomes difficult to interpret because the steady-state compositions of subgroups could differ across counterfactuals. For this reason, it is more natural to study policy effects on a *fixed* population. We accomplish this by solving for the transition path between baseline and counterfactual steady states.[15] We then apply the decision rules from the first period along the transition path after the policy has been enacted to the baseline steady state and simulate one generation for each policy, comparing each outcome to another generation of the baseline steady state. These results allow us to better study policy effects effects on specific fixed individuals or groups relative to baseline.

As we are interested in examining regional effects of policy on the poor, we study outcomes for children born in different regions and to different quartiles of the parental income distribution. These exercises help us understand how each policy affects the regional allocation of opportunities and how effectively the policy targets children whose parents lack sufficient private resources. In particular, we are able to better comprehend the *mechanisms* underlying changes in mobility and welfare driven by policy in isolation–by comparing differences in *individual* outcomes, we can discern both *which* individuals are affected and *why* they are affected.

In what follows, we define household welfare as the household's dynastic utility. In tables that include comparisons of average welfare between baseline and counterfactual populations or subgroups, we report differences in welfare as the compensating equivalent consumption for each counterfactual as a percentage of baseline consumption in each group. In other words, we calculate the percentage increase in consumption from baseline for all individuals within each group that is required to match mean welfare in the counterfactual for that group.

[14] For example, if 25% of the population attends the 4-year college in the baseline steady state and a province grows (shrinks) by 1,000 individuals in the Hukou or Public Spending counterfactual, that province also gains (loses) 250 seats at the four-year college.

[15] Details for the solution process of solving for the transition path can be found in the Appendix.

2.A.1 Steady State Effects

We begin by focusing on differences in aggregate outcomes across counterfactual steady-state equilibria listed in Table 2.4. We are interested in changes in income, spatial population distribution, intergenerational mobility (measured by the intergenerational elasticity), inequality (measured by the Gini), and welfare.

Table 2.4: Steady State Differences across Policies

	Baseline	Hukou	Public Spending	Merit	Equity
Average Income	37,988	36,966	38,369	37,955	37,491
Population size (%), province ℓ	63.9	59.0	67.3	62.8	64.3
IGE	0.401	0.384	0.374	0.411	0.390
Gini	0.564	0.543	0.563	0.565	0.557
Welfare	21.743	21.820	21.779	21.748	21.766
Δ Welfare in % Consumption (CE %)	-	8.0	3.7	.4	1.3

The Hukou counterfactual lifts migration costs in moving from the poor to the rich province without an elite education, effectively treating poor-to-rich migrants the same as residents of the rich province. This has important implications for decision-making, particularly in the human capital investment and migration decisions. We find that lifting Hukou restrictions also has relatively significant effects in the steady state. Removing migration costs encourages individuals to move from the poor to the rich region, increasing the rich province's population size by about 5% of the population. Intergenerational persistence decreases by .017 (about 4.2% or .11 expected generations required to regress to the mean), driven by improved labor mobility and access to better K-12 education for the children of migrants from the poor province. The Gini coefficient decreases by about 3.7%. Interestingly, while average national earnings *decreases* by about 2.7%, welfare increases by about 8.0% of baseline consumption.

In the Public Spending counterfactual, we equalize per-student public spending levels $\bar{x}_\ell$ and $\bar{x}_h$, holding fixed the total spending. This is equivalent to the government collecting all education tax revenue and allocating it equally across the population rather than each province distributing its own revenue. This eliminates one substantial, policy-relevant source of inequity in human capital development for children born in rich and poor provinces by improving the opportunities available in the poor province. Equalizing per-student public spending across regions has sizable effects on steady state outcomes. The average income increases by about 1%, driven by providing higher-quality education to a larger population in the poor province. Equating public spending makes province ℓ a more attractive loca-

tion, resulting in a larger steady-state population. Higher public spending in the poor region reduces the correlation between parent's income and public schooling quality, reducing intergenerational income persistence by 6.7% from baseline or .18 expected generations to regress to the mean. Overall inequality remains roughly the same. Average welfare increases by 3.7% of average baseline consumption.

The latter two policies, Merit and Equity, affect regional opportunity by reallocating college seats across provinces. The Merit counterfactual maximizes efficiency in allocating individuals to college by sending the highest-human capital individuals to the best college seat available regardless of birth province, maximizing the expected returns to college in terms of total earnings. The Equity counterfactual equalizes the proportion of applicants from each province who attend each college. We again note that relative to the baseline model, the Merit counterfactual reallocates seats from the poor province to the rich province, and vice versa in the Equity counterfactual.

In the Merit counterfactual, individuals no longer face higher human capital thresholds in the rich province, attracting individuals to the rich province. Moreover, it increases the association between parent's income and children's education by providing children in the rich province with more college seats, resulting in a 2.5% increase in the IGE. The Equity counterfactual generates changes in the opposite direction: more college seats makes the poor province slightly more attractive and decreases the association between income and college attendance, decreasing IGE by 2.7%. Neither policy substantially changes aggregate inequality, welfare, or income.

These findings suggest that despite public controversy regarding the fairness of the college allocation mechanism in China, revising the mechanism with an efficient or equitable allocation does not appear to produce large aggregate effects, with modest effects on intergenerational persistence. The reasoning for this likely stems from the fact that in 2011, only about 20% of age-18 children across the country attended a four-year college, so that reallocating college seats does not shift the education outcomes of many individuals relative to the general population. However, China's recent educational expansion beyond 2011 gives scope for these policies to have greater effects with more college seats available.

2.A.2 Distributional Effects and Mechanisms

We are also interested in studying heterogeneity in household effects across the income distribution as well as the mechanisms of these effects. To study this, we enact each policy in the baseline steady state and compute the transition path between baseline and counterfactual

steady states. This allows us to fix one generation of households (namely, the baseline steady state) and directly compare outcomes from enacting the policy for each individual household (or groups of households, e.g. birth locations and/or parent income quartiles) between one additional generation of baseline policy and the first period of the transition path after the policy.[16].

In order to isolate the effects of policy in this exercise, we fix the child's idiosyncratic human capital shock $\xi_H^{'}$ and location preference shock $\xi_r^{'}$. Policy welfare effects are calculated using *expected* welfare with respect to location preference shock ξ_r and income shock $\xi_{y,r}$, as these shocks affect welfare and decision making in the *subsequent* period.

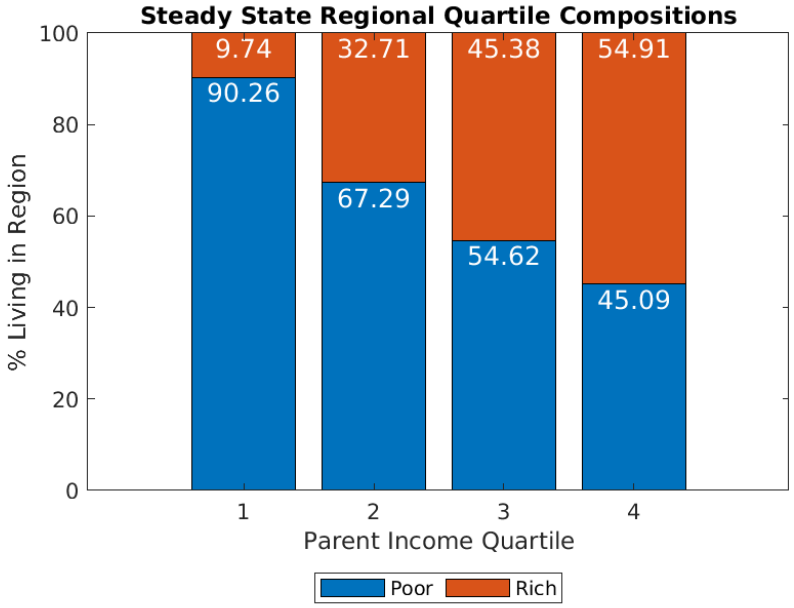
In the following subsections, we examine effects from each policy by first studying effects on average welfare, human capital, college attendance, and income for children born to each quartile of the parental income distribution in each province. Then, we directly study the mechanisms of improved welfare outcomes by comparing the "winners" and "losers" of each policy: our heterogeneous agent model allows us to compare children who are better off in expectation under the policy relative to baseline ("winners") to those who are worse off ("losers"). In other words, we attribute gains in expected welfare to differences in characteristics between the winners and losers from the policy. We examine four quantities that could change expected welfare: (i) higher counterfactual human capital (and thus income), (ii) higher counterfactual education, (iii) differential migration, and (iv) elite migration (defined as migration to the rich region with an elite college education) for children born in the poor province, and parent's Hukou status for children born in the rich province.

Note that in the Hukou and Public Spending counterfactuals, since the policies apply at the beginning of the period when population sizes and regional college seat allocations are fixed and the college allocation mechanism has not changed (as it does in Equity and Merit), provincial college attendance rates should not change for children born in both provinces. However, we can still observe whether college seats are distributed differently across quartiles *within* each province due to policy.

Figure 2.4 provides a regional decomposition of parent income quartiles in the baseline steady state, which helps to contextualize the scale of effects on parent income quartiles within a region and better understand the overall impacts of policy.

[16]Details regarding the computation of transition paths can be found in the Appendix

Figure 2.4: Steady State Regional Quartile Compositions



Hukou

Table 2.5: Hukou: Effects on Child's Expected Welfare

	Province ℓ		Province h	
	CE%	% Better Off	CE%	% Better Off
Overall	14.1	84.5	-2.2	31.5
Q1	16.3	89.3	10.0	69.8
Q2	14.4	87.0	0.5	45.5
Q3	13.3	84.6	-3.0	33.0
Q4	10.5	71.1	-5.3	14.8

Table 2.5 lists the welfare effects of lifting the Hukou policy on children born in different parent income quartiles in each region. We again report welfare effects as the consumption equivalent relative to baseline consumption within the subgroup. Most individuals in the poor province and the poorest quartile in the rich province are better off at the cost of the top two quartiles in the rich province, with children born in the poor province benefiting in the range of 10.5 to 16.3% with the greatest gains at the bottom of the parental income distribution.

Table 2.6: Hukou: Effects on Migration %

	Province ℓ		Province h	
	Baseline	Hukou	Baseline	Hukou
Overall	5.7	33.2	10.7	45.3
Q1	3.1	27.3	26.9	72.7
Q2	4.2	29.3	17.9	62.9
Q3	5.0	33.0	11.5	51.8
Q4	13.8	50.9	2.9	24.6

Next, we study the mechanisms of these changes. Table 2.6 shows that eliminating Hukou costs induces substantial changes in migration flows in both directions; poor-to-rich migration increases from 5.7 to 33.2% while rich-to-poor migration increases from 10.7 to 45.3%. The fact that poor-to-rich migration increases and drives welfare gains is unsurprising, as the primary channel affected by lifting the Hukou policy is migration, and individuals can now migrate and realize the benefits of living in the richer region. The Hukou policy also allows for efficient sorting, encouraging certain individuals to move to the poor region for lower housing prices without fear of facing migration costs to move back in a later generation. This option value lessens the welfare gap between rich and poor provinces, leading to higher rich-to-poor migration flows as well.

Next, we examine effects on average human capital, college attendance, and income across quartiles in Table 2.7. Human capital and college attendance outcomes do not substantially change except perhaps slightly for children in the bottom quartile in the rich province, driven by migrants without a Hukou newly accessing the rich province's higher social insurance and public schooling. Similarly, college attendance rates do not change substantially for any quartile. Incomes in the poor province increase in the range of 2 to 4%, while incomes in the rich province decrease between 1 to 6%. These results suggest that welfare gains and losses are largely driven by migration and not other outcomes.

Table 2.7: Hukou: Effects on Human Capital, College Attendance %, and Log(Income)

	Province ℓ		Province h	
	Baseline	Hukou	Baseline	Hukou
Human Capital				
Overall	106.2	105.9	174.5	174.8
Q1	81.5	81.4	72.0	74.1
Q2	96.0	95.9	113.1	113.5
Q3	111.3	111.1	147.0	146.9
Q4	165.0	164.2	253.1	253.3
College Attendance %				
Overall	16.1	16.3	24.0	24.6
Q1	7.8	8.0	2.5	2.5
Q2	11.8	12.0	7.5	7.8
Q3	17.4	17.6	14.7	15.4
Q4	37.6	37.8	45.3	46.1
Log(Income)				
Overall	9.72	9.74	10.30	10.27
Q1	9.46	9.48	9.38	9.37
Q2	9.62	9.64	9.86	9.84
Q3	9.80	9.83	10.16	10.10
Q4	10.29	10.32	10.85	10.83

Finally, we study differences in observables between winners and losers from the Hukou policy. Each cell in Table 2.8 represents the proportion of individuals realizing each outcome for the winners ("W") and losers ("L") from the policy. For example, in the "Higher HC" column under province ℓ, we calculate the percentage of children who have higher human capital in the counterfactual among those who are better off (winners) and compare it to the percentage of children who have higher human capital in the counterfactual among those who are worse off (losers). Results from this table help us understand what characteristics drive welfare gains compared to welfare losses.

Table 2.8: Hukou: Differences Between Winners vs. Losers from Policy

Province ℓ	Higher HC			Higher Edu			Migration			Elite Migration		
	W	L	Diff	W	L	Diff	W	L	Diff	W	L	Diff
Overall	0.0	0.0	0.0	0.1	0.0	0.1	27.2	99.0	-71.8	0.1	56.9	-56.8
Q1	0.0	0.2	-0.2	0.1	0.0	0.1	22.8	100.0	-77.2	0.0	24.7	-24.7
Q2	0.0	0.0	0.0	0.1	0.0	0.1	24.4	99.1	-74.8	0.1	41.9	-41.8
Q3	0.0	0.0	0.0	0.1	0.0	0.1	27.4	98.8	-71.4	0.1	56.1	-56.0
Q4	0.0	0.0	0.0	0.1	0.0	0.1	41.3	98.5	-57.1	0.1	87.4	-87.3

Province h	Higher HC			Higher Edu			Migration			Non-Hukou Parent		
	W	L	Diff	W	L	Diff	W	L	Diff	W	L	Diff
Overall	12.9	3.9	9.0	1.9	0.0	1.9	86.0	30.1	55.8	12.6	0.2	12.3
Q1	29.0	1.1	27.8	0.1	0.0	0.1	77.8	57.5	20.3	29.0	1.1	27.8
Q2	8.3	0.1	8.1	0.5	0.0	0.5	88.3	44.8	43.5	8.3	0.1	8.1
Q3	7.9	0.2	7.7	1.7	0.0	1.7	89.5	37.9	51.6	7.9	0.2	7.7
Q4	14.5	8.0	6.5	8.5	0.0	8.5	83.0	17.6	65.4	12.3	0.3	12.0

We find that better short-term outcomes for children born in the poor province are not driven by increases in human capital accumulation or educational attainment, and they are much less likely to migrate to the rich province than those who are worse off. This happens for two main reasons. First, lifting the Hukou policy on adds the option value of freely migrating to province h for later generations of residents in province ℓ, indirectly increasing welfare from living in the poor province. Second, children who migrate to the rich province, including those with an elite education, under the baseline and Hukou policy settings are relatively worse off without migration restrictions because the influx of migrants requires them to pay higher housing prices in the next period. This explains why migrants lose from a policy that lifts migration costs–migration is always the best option for these individuals under both baseline and counterfactual, but it does not provide the same return as before.

Winners born in the rich province typically migrate to the poor province for lower housing costs, especially for children born to higher-income parents. Human capital outcomes improve as well; observing the similarity between the "Higher HC" and "Non-Hukou Parent columns," we find that children who enjoy better human capital outcomes are almost exactly the ones whose parents receive a Hukou due to the policy, with the only difference in

the highest parent income quartile. These children and their parents benefit from no longer facing Hukou costs in earnings, education, housing, and investment.

Taken together, results suggest that lifting Hukou restrictions produces positive welfare effects through two main mechanisms. First, it improves human capital outcomes for children whose parents previously faced Hukou restrictions. Second, it lessens the welfare gap between the rich and poor province by allowing individuals to efficiently sort without facing migration costs in a later period. Both of these channels emphasize the welfare improvement from better access to the rich province.

Public Spending

Table 2.9 decomposes the welfare effects by parent income quartile and birth location in the Public Spending counterfactual. As expected, all quartiles in the poor province are better off, gaining 8 to 17.4% in welfare and 14.5% overall. Lower-income parents in the poor province gain the most from the policy because they are relatively more dependent on public rather than private investment in children's human capital accumulation. The opposite holds in the rich province: most children are worse off, losing 9.3 to 14.5% of income in consumption equivalent and 12.9% overall. The bottom quartile of the rich province appears to suffer the lowest welfare loss on average, but this is due to this group largely consisting of non-Hukou holders; we expand on this in a later table.

Table 2.9: Public Spending: Effects on Child's Expected Welfare

	Province ℓ		Province h	
	CE%	% Better Off	CE%	% Better Off
Overall	14.5	96.4	-12.9	6.2
Q1	17.4	99.0	-9.3	18.2
Q2	15.9	98.9	-14.5	8.1
Q3	13.4	97.3	-14.3	5.5
Q4	8.0	86.1	-11.5	3.4

We next examine effects on migration in Table 2.10. Equalization in public spending unsurprisingly makes the poor province more attractive, inducing migration into the region particularly by children born to lower-income parents in the rich province. The justification for the small increase in migration rates in the opposite direction is less obvious; however, equating per-student public spending across regions also incentivizes poor-to-rich migration

by lessening the cost of moving from without receiving a Hukou, as children of migrants no longer face a difference in schooling quality upon migration.

Table 2.10: Public Spending: Effects on Children's Migration %

	Province ℓ		Province h	
	Baseline	Public Spending	Baseline	Public Spending
Overall	5.7	6.7	9.8	17.9
Q1	3.4	4.3	25.5	40.7
Q2	4.1	5.2	16.4	29.5
Q3	5.1	6.3	10.1	19.4
Q4	13.5	14.4	2.7	5.6

We report effects on children's human capital, college attendance, and logged income in Table 2.11. As expected, average child human capital increases for children across all quartiles in the poor province and decreases across all quartiles in the rich province, with larger percentage gains and losses in lower quartiles of the poor and rich province, respectively. This is met with slight increases in college attendance (about 3 - 5% of baseline) in the bottom two quartiles and slight decreases in the upper two quartiles for children born in the poor province, as children with poorer parents benefit relatively more from increases in public relative to private investment. However, this comes at a cost–children of poorer parents in the rich province suffer the largest relative decreases in human capital and college-going, with attendance increasing for the top quartile. The richest individuals can most readily compensate losses in public funding with private investment, allowing them to maintain their children's status at the top of the human capital and education distribution. These factors lead to the largest income gains for children of the poorest in the poor province and the largest income reductions for the children of the poorest in the rich province.

Table 2.11: Public Spending: Effects on Children's Human Capital, College Attendance %, and Log(Income)

	Province ℓ		Province h	
	Baseline	Public Spending	Baseline	Public Spending
Human Capital				
Overall	105.5	110.2	173.2	164.9
Q1	80.8	85.8	73.3	68.7
Q2	95.1	100.3	112.0	103.8
Q3	111.0	115.6	143.5	134.5
Q4	164.1	167.8	252.9	244.6
College Attendance %				
Overall	16.1	16.1	24.0	23.8
Q1	7.8	8.2	2.5	1.9
Q2	11.8	12.2	7.5	6.9
Q3	17.4	17.2	14.7	14.0
Q4	37.6	36.5	45.3	46.0
Log(Income)				
Overall	9.71	9.76	10.30	10.24
Q1	9.46	9.52	9.41	9.35
Q2	9.62	9.67	9.86	9.78
Q3	9.80	9.84	10.14	10.07
Q4	10.26	10.28	10.85	10.81

Table 2.12 reveals the mechanisms of welfare gains and losses. We first note that from Table 2.9, almost all children born in the poor province are better while almost all children born in the rich province are worse off from the policy, as would be expected from reallocating public schooling funds from rich to poor regions. Welfare losers born in the poor province are migrants who receive an elite education and obtain a Hukou for the rich province in both the baseline and the counterfactual, whose children then do not receive higher quality schooling as before. The children born in the rich province who gain from the policy consists of those born to migrants without a Hukou who do not have access to the rich province's public education or those who migrate to the poor province under both the baseline and policy counterfactual settings (or both). As with the Hukou policy, the children with better human capital outcomes in the rich province are mostly the same as the children of parents without a Hukou. However, under this policy, rather than children of migrants benefiting

from access to public schools in the rich region, they benefit from improved schooling quality in the poor region.

Table 2.12: Public Spending: Differences Between Winners vs. Losers from Policy

Province ℓ	Higher HC			Higher Edu			Migration			Elite Migration		
	W	L	Diff	W	L	Diff	W	L	Diff	W	L	Diff
Overall	99.9	97.2	2.7	0.4	0.0	0.4	3.5	92.1	-88.7	1.4	71.7	-70.3
Q1	99.8	91.3	8.5	0.5	0.0	0.5	3.6	97.5	-93.9	1.2	6.3	-5.1
Q2	100.0	95.8	4.2	0.5	0.0	0.5	4.1	94.7	-90.6	2.1	31.6	-29.4
Q3	100.0	99.0	1.0	0.3	0.0	0.3	3.6	97.9	-94.4	1.6	72.0	-70.4
Q4	100.0	97.5	2.5	0.0	0.0	0.0	1.8	89.9	-88.1	0.2	82.9	-82.8

Province h	Higher HC			Higher Edu			Migration			Non-Hukou Parent		
	W	L	Diff	W	L	Diff	W	L	Diff	W	L	Diff
Overall	62.2	0.9	61.3	8.8	0.0	8.8	47.4	16.7	30.6	62.2	0.9	61.3
Q1	91.6	7.3	84.2	0.5	0.0	0.5	48.0	39.7	8.4	91.6	7.3	84.2
Q2	69.6	0.3	69.3	0.5	0.0	0.5	45.0	29.0	16.0	69.6	0.3	69.3
Q3	60.0	0.3	59.7	0.5	0.0	0.5	53.0	18.4	34.6	60.0	0.3	59.7
Q4	26.9	0.7	26.2	33.0	0.0	33.0	43.7	4.5	39.2	26.9	0.7	26.2

In summary, the Public Spending counterfactual accomplishes similar welfare outcomes to the Hukou counterfactual, particularly improving outcomes for low-income populations in the poor province. However, whereas the Hukou counterfactual lessens the poor-rich gap by allowing access to the rich province's resources, the Public Spending counterfactual accomplishes its outcome by improving resources in the poor province, lifting individuals who are born or attend public school in the region.

Merit and Equity

The Merit and Equity policies may impact outcomes within subgroups if college seats are differentially allocated across the groups or they generate differential incentives for child investment. The effect of these policies on child investment is unclear–on one hand, lower human capital thresholds in a region implies that parents need not invest as much to guarantee the same probability of admission (prior to the human capital shock) as in baseline. On

the other hand, further marginal investments into children could increase admission probability and produce much better future outcomes by attaining a higher tier of education.

Table 2.13: Merit: Effects on Child's Expected Welfare

Merit	Province ℓ		Province h	
	CE%	% Better Off	CE%	% Better Off
Overall	-4.2	31.4	9.4	94.9
Q1	-1.7	46.1	1.8	85.3
Q2	-3.4	33.4	5.1	90.3
Q3	-5.2	20.3	8.5	95.3
Q4	-9.2	12.5	14.5	99.0

We begin by analyzing the welfare effects of the Merit counterfactual in Table 2.13. In the baseline counterfactual, the marginal student from the rich province has higher human capital, meaning that lifting provincial quotas with this policy reallocates seats from the poor province to the rich province. On average, children born in the poor province are worse off by 4.2% and children born in the rich province are better off by 9.4%, where losses and gains are concentrated at the top end of the parental income distribution in both the poor and rich provinces, respectively.

Examining human capital, college attendance, and earnings outcomes in Table 2.14 helps to explain these changes in welfare. The policy slightly reduces human capital and greatly reduces college attendance for children in the poor province, with a corresponding 2% overall decrease in income. In the rich province, average human capital stays roughly the same; combined with increases in college attendance, this drives a 4% increase in earnings. The largest percentage point losses and gains in earnings and college attendance are again concentrated among children of high-income parents in both regions. This trend is explained by the fact that children of high-income parents are most likely to be on the margin of college attendance and thus have college attainment affected by the reallocation.

Table 2.14: Merit: Effects on Children's Human Capital, College Attendance %, and Log(Income)

	Province ℓ		Province h	
Human Capital	Baseline	Merit	Baseline	Merit
Overall	106.3	105.9	174.5	175.2
Q1	81.5	81.4	71.8	71.9
Q2	96.1	95.9	113.0	113.2
Q3	111.4	110.7	146.9	147.6
Q4	165.2	164.0	253.1	254.3
College Attendance (%)				
Overall	16.3	11.9	24.9	32.1
Q1	7.6	5.0	1.7	2.6
Q2	12.3	8.6	8.3	13.0
Q3	17.3	12.2	16.6	24.0
Q4	38.3	30.0	45.9	55.6
Log(Income)	9.46	9.45	9.38	9.38
Overall	9.72	9.70	10.30	10.34
Q1	9.46	9.45	9.38	9.38
Q2	9.62	9.60	9.86	9.88
Q3	9.80	9.77	10.16	10.20
Q4	10.29	10.24	10.85	10.91

Table 2.15 shows that welfare improvements in the poor province through the Merit reallocation are not driven by improved human capital accumulation or educational outcomes, but rather by migration to the rich province with a Hukou, particularly for children of the richest parents. The subsequent generation of these households then access a better college allocation relative to the baseline allocation in the rich province. Winners born in the rich province are more likely than losers to have higher human capital and better educational outcomes, migrate toward the worse college allocation much less often, and tend to have parents who hold the local Hukou so that they are able to access the rich province's allocation upon birth.

Table 2.15: Merit: Differences Between Winners vs. Losers from Policy

Province ℓ	Higher HC			Higher Edu			Migration			Elite Migration		
	W	L	Diff	W	L	Diff	W	L	Diff	W	L	Diff
Overall	0.0	0.0	0.0	0.0	0.0	0.0	8.6	3.7	4.9	6.6	0.0	6.6
Q1	0.0	0.0	0.0	0.0	0.0	0.0	3.8	3.0	0.8	1.3	0.0	1.3
Q2	0.0	0.0	0.0	0.0	0.0	0.0	5.0	3.0	1.9	3.3	0.0	3.3
Q3	0.0	0.0	0.0	0.0	0.0	0.0	9.1	3.7	5.4	8.0	0.0	8.0
Q4	0.3	0.2	0.1	0.0	0.0	0.0	57.0	5.2	51.8	56.3	0.0	56.3

Province h	Higher HC			Higher Edu			Migration			Non-Hukou Parent		
	W	L	Diff	W	L	Diff	W	L	Diff	W	L	Diff
Overall	55.0	13.4	41.6	11.3	0.0	11.3	4.8	93.8	-89.1	3.2	10.2	-7.0
Q1	7.0	1.6	5.4	1.3	0.0	1.3	12.2	95.7	-83.5	21.6	16.8	4.8
Q2	20.9	5.7	15.2	6.1	0.0	6.1	7.1	98.5	-91.4	3.5	3.0	0.6
Q3	53.6	24.7	28.9	10.1	0.0	10.1	5.5	93.7	-88.3	1.9	8.9	-6.9
Q4	82.4	45.6	36.8	16.7	0.0	16.7	1.8	61.8	-59.9	1.0	39.7	-38.7

Given that reallocating college seats to the rich province tends to increase welfare in the rich province and decrease welfare across all quartiles of the poor province, one might surmise that the Equity counterfactual would have the reverse effects. Table 2.16 shows that this holds true in the poor province as children improve welfare by 3.7% on average, with effects concentrated at the top of the income distribution. However, while overall welfare does slightly decrease in the rich province, this is driven by a decrease in average welfare only for children of the highest-income parents; in fact, most children born in all subgroups are better off under the Equity allocation than baseline, a perhaps surprising result.

Table 2.16: Equity: Effects on Child's Expected Welfare

Equity	Province ℓ		Province h	
	CE%	% Better Off	CE%	% Better Off
Overall	3.7	79.8	-0.3	79.9
Q1	2.1	72.4	3.7	95.1
Q2	3.2	80.0	3.8	94.7
Q3	4.4	87.8	1.7	89.8
Q4	6.6	84.6	-4.4	59.9

Table 2.17 sheds light on the effects on human capital, college attendance, and earnings. With lower human capital thresholds in the poor province, parental incentive to invest decreases, resulting in lower human capital but increases college attendance across children in the poor province. These offset to generate only small income gains overall (1%, up to 3% in the top quartile). Higher human capital thresholds in the rich province induce greater investment and higher human capital outcomes for children in the rich province. This offsets the decrease in college attendance so that incomes only decrease by 1% (4% in the top quartile). Again, effects are magnified in the top quartiles, where children are most likely to be at the margin of college attendance and thus affected by the policy.

Table 2.17: Equity: Effects on Children's Human Capital, College Attendance %, and Log(Income)

	Province ℓ		Province h	
Human Capital	Baseline	Merit	Baseline	Merit
Overall	106.3	105.5	174.5	176.4
Q1	81.5	80.5	71.8	72.8
Q2	96.1	95.3	113.0	115.4
Q3	111.4	110.8	146.9	149.2
Q4	165.2	164.6	253.1	254.5
College Attendance (%)				
Overall	16.3	19.3	24.9	19.1
Q1	7.6	9.3	1.7	1.6
Q2	12.3	15.1	8.3	5.4
Q3	17.3	21.0	16.6	11.6
Q4	38.3	43.4	45.9	36.8
Log(Income)				
Overall	9.72	9.73	10.30	10.29
Q1	9.46	9.45	9.38	9.39
Q2	9.62	9.62	9.86	9.87
Q3	9.80	9.81	10.16	10.16
Q4	10.29	10.32	10.85	10.81

Finally, we study the characteristics of winners and losers from the Equity policy in Table 2.18. Welfare improvements for children born in province ℓ are driven by higher college attendance from the improved allocation, while welfare losses are driven by migrants who attend the elite college and then face the rich province's lower college allocation in future generations. As before, education and migration differentials are concentrated in the top quartile. In the rich province, higher welfare indeed appears to be driven by a greater proportion of improved human capital outcomes relative to losers as well as migration to the poor province or children of non-Hukou parents, who now access the improved college quota in province ℓ.

Table 2.18: Equity: Differences Between Winners vs. Losers from Policy

Province ℓ	Higher HC			Higher Edu			Migration			Elite Migration		
	W	L	Diff	W	L	Diff	W	L	Diff	W	L	Diff
Overall	0.0	0.0	0.0	5.1	0.0	5.1	2.0	23.1	-21.1	1.3	19.3	-18.0
Q1	0.0	0.0	0.0	2.8	0.0	2.8	1.6	7.5	-5.9	0.4	4.1	-3.7
Q2	0.0	0.0	0.0	4.2	0.0	4.2	1.5	15.7	-14.2	0.8	11.8	-11.0
Q3	0.0	0.0	0.0	5.6	0.0	5.6	1.7	34.1	-32.4	1.4	27.9	-26.5
Q4	0.0	0.0	0.0	9.4	0.0	9.4	3.7	82.9	-79.3	3.4	80.0	-76.7

Province h	Higher HC			Higher Edu			Migration			Non-Hukou Parent		
	W	L	Diff	W	L	Diff	W	L	Diff	W	L	Diff
Overall	92.4	80.4	12.0	0.2	0.0	0.2	12.3	1.2	11.1	5.9	2.9	3.0
Q1	50.1	6.6	43.5	0.2	0.0	0.2	26.1	23.0	3.2	80.3	17.9	62.4
Q2	97.5	73.4	24.1	0.1	0.0	0.1	17.2	3.6	13.6	25.7	2.2	23.4
Q3	98.6	88.0	10.6	0.2	0.0	0.2	11.3	2.6	8.7	10.1	1.4	8.7
Q4	92.1	81.0	11.1	0.2	0.0	0.2	4.9	0.2	4.7	1.8	1.2	0.6

One curious result is that both counterfactual allocations alter incentives such that the rich province invests more into children and the poor province invests less. This suggests that parents in the high-income province are able to more readily increase private investment in children's human capital in response to both upward or downward movements in human capital thresholds. We also observe that both allocations have the greatest effects on children of high-income parents in both regions; the largest impacts are concentrated among children who are on the margin of college attendance, and thus positive (or negative) effects from changing college allocation mechanisms born to the lowest-income parents are limited.

2.6 Discussion and Conclusion

China's regional education and migration policies have generated lasting unrest over fairness in treatment by the government. In this paper, we explore the effects of these policies on intergenerational mobility, inequality, welfare, and the income distribution, with particular attention on the effects of policy on different subgroups. Our counterfactual policy exer-

cises each address one dimension of regional inequity, affecting either the quality of existing opportunities in the poor province or access to better opportunities in the rich province.

The Hukou and Public Spending policies generate positive aggregate and distributional effects for children born to the poorest parents in the low-income province, decreasing intergenerational persistence and increasing overall welfare by 8.0% and 3.7%, respectively. Notably, they achieve this with different end results for the spatial distribution of individuals. Eliminating Hukou migration costs allows individuals to freely access opportunities in the rich province by migrating from the poor province, inducing net migration into the rich province. Equalizing provincial per-student public spending improves existing K-12 education quality in the poor province, attracting individuals to the poor province. We find that lifting Hukou migration costs and equating public spending across provinces reduces intergenerational income persistence, but at the cost of worsening outcomes in the rich province.

One common thread between these policies is that both improve the value of living in the poor province. Lifting Hukou restrictions decreases the welfare gap between the two locations through the increased option value of free migration to the rich province as well as decreased housing prices in the poor province and increased housing prices in the rich province; children who migrate with or without a Hukou under both the baseline and Hukou settings actually tend to be worse off from the policy due to higher housing prices. Setting public schooling quality equal increases the value of growing up in the poor relative to the rich province. However, the distribution of these effects differs–the Hukou policy improves welfare for most individuals (children who do not obtain an elite education and non-migrants) in the poor province and some (children of migrant parents without a Hukou) in the rich province. The Public Spending policy improves welfare for almost all individuals in the poor province and very few (children of migrants without a Hukou) in the rich province.

Despite debate regarding the provincial college allocation mechanism, modifying the allocation to be more equitable or efficient across regions does not produce significant aggregate effects. Welfare effects both overall and within each region are much smaller from these policies than Hukou or Public Spending policies, with the most significant impact on children born to rich parents whose educational outcomes are most likely to change due to policy. Notably, allocating college seats away from the rich province, where parents have resources to smooth over losses in college opportunities through childhood investments, may produce small positive welfare effects for most children born in the region by incentivizing parental investments in children's human capital.

There are additional empirical and theoretical extensions to consider that would improve our work. We could employ longer panel data sets in China to estimate the returns to child

investment. Currently, our human capital accumulation equation relies on the specification and parameter values from Kotera-Seshadri, who estimate the parameters for the United States. While this is a useful approximation, there is reason to believe China may differ in these parameters. If peer effects matter, this could change our results because outcomes would be affected by regional human capital distributions, intensifying "brain drain" effects as living in the rich province would produce positive spillovers. Finally, we acknowledge the possibility of general equilibrium effects in labor markets, public spending, and social insurance that are not in the model. For example, changes in the educational composition of the regional labor force due to changing college quotas or reducing migration restrictions could affect wage premia for college-educated workers. Lifting Hukou restrictions could impose a substantial cost to the government if individuals at the bottom of the income distribution move into the richer cities to benefit from the region's social insurance, or it could cause schooling quality to decrease by diffusing public K - 12 educational resources across too many children if educational tax revenues cannot sustain current levels of per-student spending. It could also be interpreted as unfair to individuals in rich provinces, who would be effectively subsidizing the education of those in poor provinces.

Still, our model produces a number of interesting results for policymakers to consider, and the richness of the model enables us to consider a wide range of other interesting interventions such as affirmative action policies that allocate college seats to individuals from different parts of the income distribution rather than the human capital distribution. One could also consider combining multiple policies discussed in the paper to study the marginal contributions of each policy. For example, combining Merit and Public Spending or Merit and Hukou and observing changes in human capital thresholds relative to Merit alone could help identify to what extent public spending or Hukou restrictions cause a divergence in human capital distributions across provinces. Combining Public Spending and Hukou both equates public schooling quality across provinces and provides free access to the rich province, which could lead to interesting sorting effects in equilibrium. Lastly, there is the consideration of optimal policy; for example, what is the welfare-optimizing allocation of college seats and public education spending across regions? How would results change if weights in a social welfare function depend on parent's income quartile? We leave these extensions to future work.

Appendix

2.A.1 Rich vs. Poor Provinces

Table 2.A.1: Rich vs. Poor Provinces

Poor	Population (10,000s)	GDP per Capita	Public K-12 spending	4-year (%)	Tier 1 (%)
Guizhou	3,469	16,438	3,785	11.4	2.4
Yunnan	4,631	19,204	4,187	13.4	3.3
Gansu	2,564	19,579	4,492	17.4	3.3
Guangxi	4,645	25,233	4,464	11.8	2.6
Anhui	5,968	25,638	4,866	16.4	4.8
Jiangxi	4,488	26,073	4,190	17.2	3.7
Sichuan	8,050	26,120	4,468	11.4	2.5
Henan	9,388	28,687	3,365	15.4	3.3
Hainan	877	28,754	6,963	15.2	4.8
Qinghai	568	29,400	7,365	16.2	5.6
Hunan	6,596	29,822	4,307	17.1	3.3
Xinjiang	2,209	29,927	8,656	13.8	4.6
Shanxi	3,593	31,276	5,363	18.1	3.8
Heilongjiang	3,834	32,817	6,199	25.8	5.8
Ningxia	639	32,875	5,396	20.4	5.6
Shaanxi	3,743	33,432	6,456	22.3	6.7
Hebei	7,241	33,859	4,821	20.8	5.4
Hubei	5,758	34,099	4,138	16.8	3.8
Chongqing	2,919	34,297	5,141	13.6	4.1
Jilin	2,749	38,440	7,314	23.5	5.0
Aggregate	63.9%	28,588	4,693	16.2	3.9
Rich	Population (10,000s)	GDP per Capita	Public K-12 spending	4-year (%)	Tier 1 (%)
Shandong	9,637	47,071	6,162	26.8	8.9
Fujian	3,720	47,205	6,281	21.3	6.4
Guangdong	10,505	50,653	5,031	13.6	2.4
Liaoning	4,383	50,711	7,669	48.9	16.9
Zhejiang	5,463	59,160	8,174	61.0	13.3
Jiangsu	7,899	62,174	8,807	45.0	8.4
Beijing	2,019	80,511	22,039	32.4	16.7
Shanghai	2,347	81,772	19,685	26.4	9.7
Tianjin	1,355	83,449	15,058	42.7	15.3
Aggregate	36.1%	56,319	7,439	24.8	6.7

Income and public spending data from NBS 2011; provincial admissions rates from 2011 administrative data

Table 2.A.1 categorizes provinces in China into "rich" and "poor" based on GDP per capita data from the 2011 Chinese Census, along with their per-student public spending

levels and college allocations. [17] As noted, students applying from rich provinces enjoy higher public spending and better 4-year and elite college prospects than students in poor provinces.

2.A.2 College Seat Allocation

Access to higher education in China is highly regulated by the Ministry of Education. The Ministry organizes a hierarchy of college education ranked by school quality, starting with first-tier colleges at the top, then provincial and local four-year colleges, and finally vocational colleges. A school's ranking determines its level of public funding and subsequently the quality of students who apply to each college. The MoE assigns quotas that govern the number of students enrolling in each tier from each province, thereby controlling each province's access to college education.

Within each province's allocation, college applicants are assigned to schools based on their preferences and their scores on the National College Entrance Examination (NCEE), or *Gaokao*. Applicants choose one of two tracks, social science or natural science. The exam consists of six sections, which depend on the chosen track, worth a total of 750 points. All students must take Chinese (150 points), mathematics (150), and a foreign language (150; most commonly English). Social science students then also take history (100), geography (100), and political science (100), while natural science students take physics (110), chemistry (100), and biology (90). All exams are written and graded within province so that scores are comparable *within-province*, but not *between provinces*.

Students receive scores around two weeks after the exam. The Ministry determines and announces a cutoff total score for each province, college tier, and track based on the score distribution and desired provincial quotas.[18] Students' scores must exceed the cutoff within their track in order to apply for schools within the tier, though this does not guarantee admission. Students then apply to at most 5 colleges (and 3-5 majors within each college) by ranking them on a preference form. If a student is not admitted to any schools on their list, they can indicate on the application that they would be willing to matriculate at other schools with openings.

[17] We define "rich" as greater than 40,000 in GDP per capita, which is a natural cutoff given the large gap between Jilin and Shandong as well as consistent with previous literature classifying rich and poor provinces by Whalley and Zhang (2007).

[18] Cutoff scores are set well below the applicable admissions quotas to ensure that colleges receive enough applications to fill vacant seats.

Once students have supplied their preferences, colleges begin admitting students based on highest NCEE scores during mid-July –August.[19] The top tier admits students first, and the second tier does not begin admitting students until the first tier is complete. Students are admitted to at most one college; they may either accept the offer and matriculate, or decline the offer and not enroll in any college that year.

2.A.3 Steady-State Solution Process

First, to obtain the baseline steady state, we set housing prices P_ℓ and P_h based on empirical housing expenditure shares across provinces and set regional public spending levels $\bar{x}_\ell$ and $\bar{x}_h$ equal to per-student K - 12 spending in the data for each province. Then, we guess equilibrium provincial college human capital thresholds $\bar{H}_{2,\ell}, \bar{H}_{3,\ell}, \bar{H}_{2,h}$, and $\bar{H}_{3,h}$ and poor province location preference mean $\mu_{M,\ell}$. Given these values, we solve for value functions $V_r(H_2, \xi_y, \xi_r, m)$ across all possible combinations of state variables on a grid. We initialize the economy with a large number of individuals and an arbitrary distribution across state variables, then we use the value functions to interpolate each individual's decision rule in each period. Once we obtain the steady state from these value functions, we adjust the guess for the human capital thresholds in each province toward the implied human capital thresholds from the new steady state, and we increase or decrease $\mu_{M,\ell}$ depending on whether province ℓ is too small or too large, respectively, in the attained steady state. Finally, we re-calculate the value functions given these new human capital thresholds and location preference mean and repeat the process. We continue iterating until provincial enrollment and province size match 2011 with negligible tolerance.

To obtain counterfactual steady states, the process is largely the same. However, $\mu_{M,\ell}$ and housing constants A_r are held fixed from baseline, and we guess location size and human capital thresholds in each iteration. We achieve steady state when housing prices, public spending levels, and schooling enrollments (which adjust with location size as described in the paper) are consistent with the location size and human capital threshold guesses.

2.A.4 Solving for the Transition Path

We use the transition path between baseline and counterfactual steady state to compute the effects of counterfactual policy on a fixed initial distribution (namely, the baseline steady state) in Section 5.2 by comparing outcomes in the first period of the counterfactual transition

[19] There are also additional considerations for special talents or for ethnic minorities similar to Affirmative Action in the United States, but admission is solely based on test score for most of the population.

path relative to another generation of the baseline policy. To solve for the transition path, we begin by setting the baseline and counterfactual steady states as the start and endpoints of the path, respectively, and assuming the transition path takes at most T generations to reach the counterfactual steady state. This pins down housing prices and human capital thresholds from period T onward. Then, we guess housing prices and human capital thresholds for each college and province in each period along the transition path. Starting from period $T-1$, we use backward induction to solve for value functions in each preceding period, using guessed prices and thresholds in the current period and value functions in the next period to solve for current-period value functions. Finally, we use these value functions to simulate the transition path and check that equilibrium conditions on housing and college attainment are consistent with the optimal decision rules, i.e. population sizes and college capacities are consistent with the guesses for housing prices and human capital thresholds. If they are not, we adjust the guesses accordingly and repeat the process until equilibrium is achieved.

We show transition paths for average income and location size for each counterfactual in the figures below. For the purposes of this exercise, the policy is enacted at the *beginning* of period 2 so that the starting point in period 1 is the baseline steady state. This implies that location sizes in period 1 and period 2 (which are measured at the beginning of the period) will be identical since individuals affected by the policy do not migrate until the end of the period in which the policy changes. We find that the transitions are relatively smooth and monotone. Location size in the Equity counterfactual appears to fluctuate, but the magnitude of these fluctuations is small, so we attribute them to noise.

From Figures 2.A.1 and 2.A.2, it appears that the economy takes only a few generations to roughly converge from the baseline to the counterfactual steady state in most cases. We choose $n = 9$ as an upper bound, and convergence appears to occur well before then.

2.A.5 Calibration: Non-Targeted Moments

We verify the performance of our model in matching intergenerational quartile transition matrices for the entire population and each province, moments which are not directly targeted by the calibration. Our model suggests slightly higher income persistence for children born to either end of the parental income distribution, particularly at the upper end, but otherwise matches well.

Figure 2.A.1: Transition Paths: Location Size

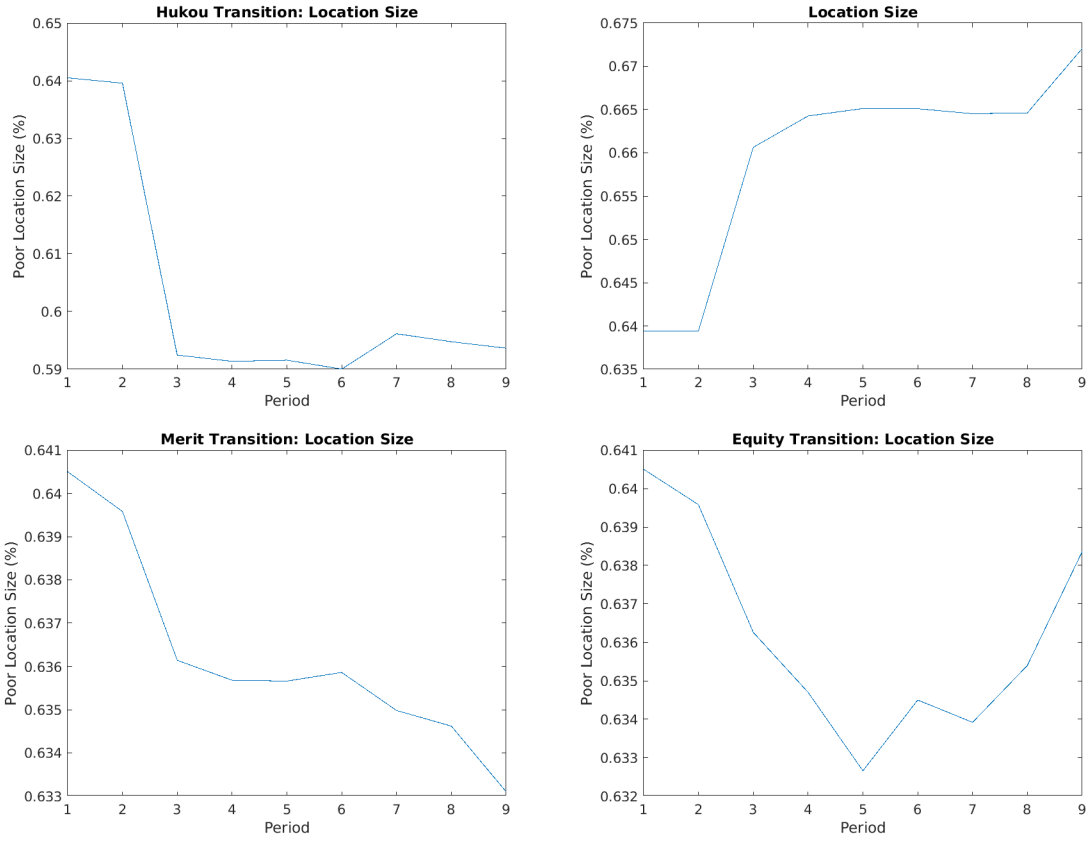


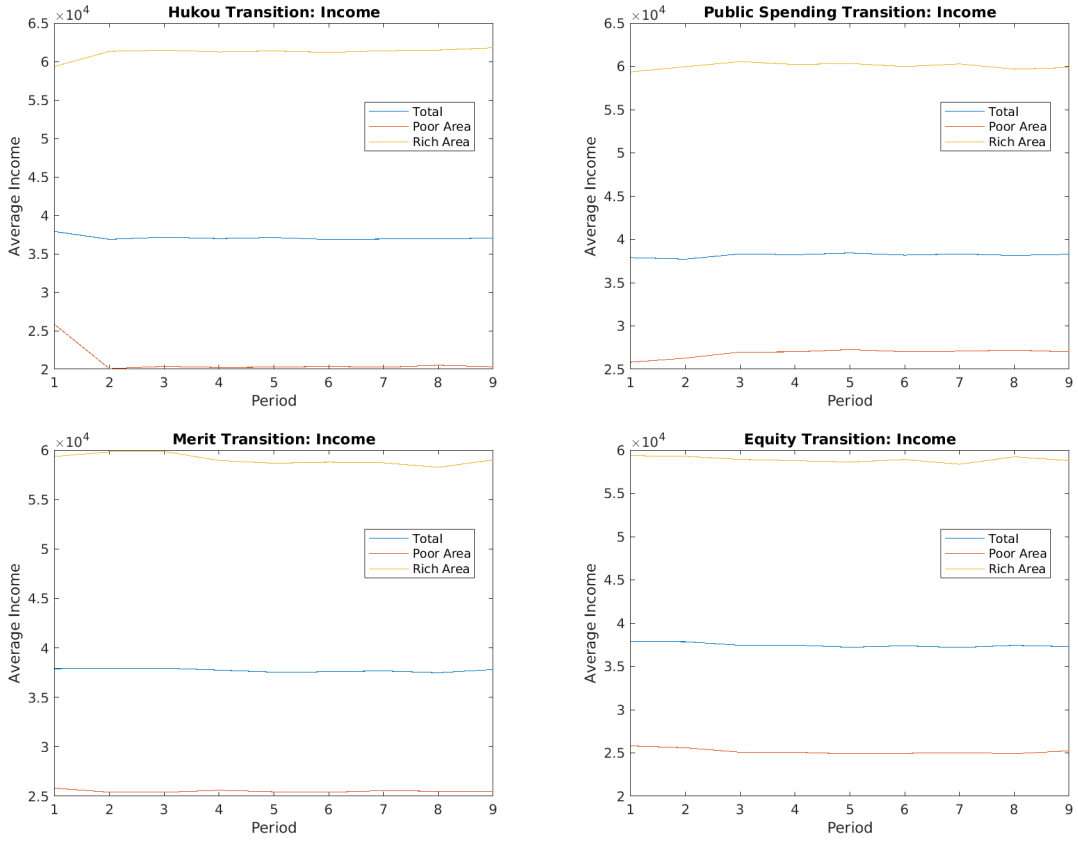
Table 2.A.2: Quartile-Quartile Transition Matrix

		Child Quartile			
	CFPS Data	1	2	3	4
Parent Quartile	1	35.7%	28.5%	19.9%	16.0%
	2	28.4%	27.2%	25.3%	19.1%
	3	19.3%	27.2%	28.7%	24.8%
	4	16.7%	17.6%	25.5%	40.1%
	Model	1	2	3	4
Parent Quartile	1	41.1%	26.0%	19.5%	13.4%
	2	34.8%	26.6%	19.9%	18.7%
	3	30.0%	24.6%	22.1%	23.3%
	4	18.9%	20.2%	20.6%	40.3%

2.A.6 Combining Hukou and Public Spending Policies

The Hukou and Public Spending counterfactuals address two different sources of regional inequality: access to regional amenities through migration and the level of an important

Figure 2.A.2: Transition Paths: Income



regional amenity, respectively. What, then, would happen if both occurred at the same time? Not only are individuals able to sort efficiently

Interactive effects on mobility

Table 2.A.3: Steady State Differences across Policies

	Baseline	Hukou	Public Spending	H + PS
Average Income	37,988	36,966	38,369	37,672
Population size (%), province ℓ	63.9	59.0	67.3	63.1
IGE	0.401	0.384	0.374	0.364
Gini	0.564	0.543	0.563	0.552
Welfare	21.743	21.820	21.779	21.845
Δ Welfare in % Consumption (CE %)	-	8.0	3.7	10.7

Table 2.A.4: Hukou + Public Spending: Effects on Child's Expected Welfare

H + PS	Province ℓ		Province h	
	CE%	% Better Off	CE%	% Better Off
Overall	19.5	95.3	-13.8	12.9
Q1	22.7	98.6	-7.4	28.4
Q2	21.5	97.6	-14.1	14.6
Q3	18.2	95.1	-15.1	15.0
Q4	12.0	85.5	-13.6	7.2

Table 2.A.5: Hukou + Public Spending: Effects on Children's Human Capital, College Attendance %, and Log(Income)

	Province ℓ		Province h	
Human Capital	Baseline	H + PS	Baseline	H + PS
Overall	106.3	111.1	173.0	164.7
Q1	80.9	86.0	73.4	68.7
Q2	95.8	101.0	113.6	105.1
Q3	111.8	116.4	144.8	135.5
Q4	166.7	170.2	250.8	242.5
College Attendance (%)				
Overall	16.4	16.4	24.6	24.8
Q1	7.8	8.2	1.9	1.8
Q2	12.6	12.9	8.5	7.9
Q3	17.5	17.4	15.9	15.8
Q4	38.3	37.0	45.6	46.6
Log(Income)				
Overall	9.72	9.79	10.30	10.21
Q1	9.44	9.52	9.39	9.33
Q2	9.64	9.71	9.88	9.76
Q3	9.80	9.86	10.15	10.04
Q4	10.30	10.35	10.84	10.78

Table 2.A.6: Hukou + Public Spending: Differences Between Winners vs. Losers from Policy

Province ℓ	Higher HC			Higher Edu			Migration			Elite Migration		
	W	L	Diff	W	L	Diff	W	L	Diff	W	L	Diff
Overall	99.8	98.1	1.7	0.4	0.0	0.4	25.1	79.1	-54.0	0.5	55.9	-55.4
Q1	99.7	97.5	2.2	0.5	0.0	0.5	19.6	89.4	-69.8	0.6	17.4	-16.7
Q2	100.0	98.5	1.5	0.5	0.0	0.5	24.1	78.5	-54.4	0.8	40.5	-39.7
Q3	100.0	99.4	0.6	0.2	0.0	0.2	28.0	81.9	-54.0	0.3	58.1	-57.9
Q4	99.6	97.5	2.0	0.0	0.0	0.0	35.8	76.1	-40.3	0.0	66.3	-66.3

Province h	Higher HC			Higher Edu			Migration			Non-Hukou Parent		
	W	L	Diff	W	L	Diff	W	L	Diff	W	L	Diff
Overall	23.7	0.6	23.1	5.0	0.0	5.0	85.6	41.4	44.2	86.0	14.0	72.0
Q1	52.1	8.4	43.7	0.0	0.0	0.0	89.9	58.8	31.1	71.0	29.0	42.0
Q2	22.3	0.3	22.1	0.8	0.0	0.8	90.9	58.4	32.5	93.2	6.8	86.4
Q3	15.0	0.1	14.9	0.7	0.0	0.7	91.1	46.5	44.6	97.0	3.0	94.0
Q4	20.0	0.0	20.0	21.0	0.0	21.0	66.6	25.6	41.0	100.0	0.0	100.0

2.A.7 Scaling College Allocation Counterfactuals

Overall college attendance rates in China have drastically increased in recent years. However, we are unable to obtain admissions data to different levels of college at the province level past 2011. In this section, we adjust for changes in college attendance by scaling the overall baseline college attendance rates in the model to match those in the United States.[20]

Effects are similar to the non-scaled case, with small changes in steady-state welfare and income. One deviation is that the low-income province shrinks in the scaled Equity counterfactual, as significantly more people from the low-income province attend the elite college and can move without facing Hukou costs. The short-run welfare effects are similarly distributed as well, with the largest improvements for the children of high-income parents in the high-income province in the Merit counterfactual.

[20]Since college attendance in China continues to grow currently, we use the United States as a benchmark for the long-run steady state of college attendance.

Table 2.A.7: Steady State Differences across College Allocation Policies (Scaled)

	Baseline	Merit	Equity
Average Income	45,561	46,183	45,260
Population size (%), province ℓ	63.3	63.4	60.7
IGE	0.435	0.459	0.431
Gini	0.560	0.570	0.556
Δ Welfare in % Consumption (CE %)		0	1.3

Table 2.A.8: Merit: Effects on Child's Expected Welfare (Scaled)

Merit	Province ℓ		Province h	
	CE%	% Better Off	CE%	% Better Off
Overall	-14.7	31.4	30.8	94.9
Q1	-14.0	46.1	13.1	85.3
Q2	-14.9	33.4	28.0	90.3
Q3	-15.7	20.3	32.4	95.3
Q4	-15.0	12.5	31.7	99.0

Table 2.A.9: Equity: Effects on Child's Expected Welfare (Scaled)

Equity	Province ℓ		Province h	
	CE%	% Better Off	CE%	% Better Off
Overall	2.1	31.4	–3.0	94.9
Q1	.3	46.1	4.8	85.3
Q2	2.1	33.4	3.1	90.3
Q3	4.0	20.3	-1.7	95.3
Q4	4.3	12.5	-5.9	99.0

Chapter 3

U.S. Employment and Opioids: Is There a Connection?

3.1 Introduction

Many observers have decried the effects of the U.S. opioid epidemic on drug overdoses and mortality. The epidemic is particularly shocking since the majority of users start taking opioids that are prescribed by their physicians, even if they later progress to illicit or illegal opioid use. Case and Deaton (2015) point to the opioid crisis as an important cause of recent increases in mortality among middle-aged non-Hispanic white Americans. And deaths may be viewed as the tip of the iceberg in that for every person who dies, many more are suffering the debilitating effects of addiction. Krueger (2017) documents that in a recent survey of prime-aged white men who were out of the labor force, 50% report chronic pain and daily use of opioid pain medications.

These observations beg the question of whether the opioid crisis is a consequence of unemployment and economic dislocation among less-skilled American workers, or whether the indiscriminate prescription of opioids has promoted economic dislocation by transforming workers with curable and chronic injuries into addicts. A fundamental barrier to answering these questions is that areas such as Appalachia, which historically have low employment-to-population rates, have also been hardest hit by the opioid epidemic. But it is not clear whether this relationship is causal or reflects omitted factors such as local variation in physician prescribing behavior.

This paper uses quarterly county-level data to examine the relationship between per capita opioid prescription rates and employment-to-population ratios. We have data on all

prescriptions of opioids from 2006 to 2014 from IQVIA which can be aggregated to the county-gender-age group-quarter level. These data are linked to data on employment from the Quarterly Workforce Indicators and to information on county population from the U.S. Census.

We first estimate models with employment-to-population ratios as the dependent variable and lagged opioid prescriptions as the independent variable. In order to address potential confounding, per capita opioid prescriptions for younger ages are instrumented using per capita opioid prescriptions to the elderly. Places with high per capita opioid prescription rates among the elderly also tend to have high prescription rates among people of working age. The identifying assumption in these models is that prescriptions to the elderly should not have a direct effect on employment among younger people. We find that the estimated effect of opioids on employment-to-population ratios is positive but small for women, while there is no relationship for men. Specifically, a 100% increase in opioid prescribing would lead to increases in employment of 3.8% among women in counties with above-median education and 5.2% among women in counties with below-median education.

In order to evaluate the hypothesis that economic dislocation causes dependence on opioids, we also estimate models with per capita opioid prescriptions as the dependent variable and lagged employment-to-population ratios as the independent variable of interest. In these models we use a Bartik-style shift share instrument for employment shocks in which the composition of county employment in a base year (2005) is used to predict the impact of national industry-level employment fluctuations on employment-to-population ratios at the local level (Bartik (1991)). Because the sample period covers the Great Recession and the recovery, there is a lot of local variation in employment to exploit. Arguably, the spike in unemployment, subsequent recovery, and its differential effect on local economies could not easily have been forecast in 2005 given knowledge about county employment composition.

In these specifications, our results are more ambiguous. We find some evidence that higher employment-to-population ratios reduce per capita opioid prescriptions among young workers, though this effect is only statistically significant in the instrumental variable (IV) specifications and only in counties with above-median education. We conclude that the relationship between opioid prescribing and employment is actually quite weak and that factors other than economic dislocation must underlie the sharp increase in deaths due to opioids. The rest of the paper is laid out as follows: We discuss the background literature in Section 1 and the data in Section 2. Section 3 presents an overview of our methods, while results and robustness are discussed in Sections 4 and 5. Section 6 offers a discussion and conclusions.

3.2 Background

Drug overdoses involving opioids rose 200% between 2000 and 2014, becoming an acknowledged public health crisis, and calls for a public policy response (Chen et al. (2014); Dart et al. (2015); Rudd et al. (2016)). Unlike in past drug epidemics, currently many overdoses and deaths involve legal opioids prescribed by a physician. Since 1999, the clinical use of opioids has quadrupled in the U.S., leading to increases in overdoses, emergency room visits, and admissions for drug treatment in addition to the surge in deaths. Although many opioids are prescribed for chronic pain, opioids are not effective for this indication over the long term because patients build up dependence (Frieden and Houry (2016)). Until very recently, prescriptions of opioids continued to rise despite the rising toll (for Behavioral Health Statistics and Quality (2013); Meara et al. (2016)).

In a groundbreaking paper, Case and Deaton (2015) coined the phrase "deaths of despair," arguing that the worsening economic position of less-educated whites in the US had fueled increases in deaths due to suicide, alcohol, and drug addiction. This idea suggests that people are dying in large part because of behaviors that are themselves a response to worsening economic status.

This hypothesis follows naturally from a large body of literature showing that economic dislocation has health consequences for affected individuals. For example, Bergemann et al. (2011), Black et al. (2015), Browning and Heinesen (2012), Eliason and Storrie (2009b,a), and Sullivan and Von Wachter (2009) all find negative effects of individual job displacement on health outcomes. This literature suggests that some of these negative effects are generated by changes in health behaviors, such as increases in smoking (c.f. Black et al. (2015); Falba et al. (2005)).

In contrast to the effects of individual job loss, the broader literature linking general economic conditions to heath and health behaviors suggests that health is countercyclical. Ruhm (2000, 2005) argues that recessions are actually good for people's health because, for example, they have more leisure time for physical activity and other health seeking behaviors. Ásgeirsdóttir et al. (2014) find that the economic crisis of 2008 was associated with many improvements in health behaviors in Iceland. Therefore, it is by no means certain that areas with declining employment will also suffer deteriorations in health behaviors.

Recent studies examining the association between economic conditions and opioid abuse find mixed results. While Hollingsworth et al. (2017), Carpenter et al. (2017), Pierce and Schott (2020), and Charles et al. (2019) conclude that adverse macroeconomic conditions—including rising unemployment, detrimental trade shocks, and declining manufac-

turing employment—have led to increases in opioid use and mortality, Ruhm (2018) finds that variation in economic conditions can explain at most one-ninth of the variation in the growth in overall drug-related mortality rates. In support of Ruhm's findings, Case and Deaton (2017) paint a more nuanced portrait than their earlier paper and highlight that deaths of despair were rising even in the early 2000s—a time of great economic growth and before the Great Recession. Notably, none of these papers other than Pierce and Schott (2020) use methods that address the possibility of reverse causality.

Attributing the rise in opioid abuse to the deterioration of economic conditions is an appealing narrative, but any causal relationship between opioids and employment could just as easily run in the opposite direction. In fact, in its most recent report on employment in the U.S., the Organization of Economic Cooperation and Development states that the opioid epidemic is responsible for recent declines in labor force participation in the U.S. (for Economic Cooperation and Development (2018)). However, the evidence for this proposition is primarily descriptive and cross-sectional.

In descriptive work, Krueger (2017) demonstrates using cross-sectional data that places with the most opioid prescribing in 2015 experienced the largest declines in male labor force participation between 1999 and 2015, though he acknowledges that labor force participation was declining in many of the same places prior to the opioid epidemic. Krueger notes that if the correlation is causal, then increased opioid use could explain up to 20% of the observed decline in labor force participation over the period. Using data for 10 states from 2013 to 2015, Harris et al. (2017) similarly find a negative effect of opioid prescription rates on labor market outcomes, although their sample variation is too limited to allow for county fixed effects.

To our knowledge then, our paper is the first to examine both the hypothesis that lack of employment has a causal effect on opioid use and the hypothesis that opioid use has a causal effect on employment in a similar framework. Our work further builds on previous studies by using IV strategies to address reverse causality, by using a relatively long panel that allows us to control for fixed differences across counties, and by incorporating detailed prescriptions and employment data that allows for county-level analyses within different age and gender groups.

3.3 Data

We purchased prescription data from IQVIA, a public company specializing in pharmaceutical market intelligence. This data set contains the number of prescriptions filled for opioid

analgesics at U.S. retail pharmacies in each year from 2006 to 2014. In addition to the number of prescriptions, the IQVIA data contain information on the patient's age group and gender and the address of the prescriber. These data show a continuous increase in the number of opioid prescriptions from 2006 to 2012 followed by a slight moderation. To calculate opioids per capita, we divide opioid prescriptions within a county from IQVIA by county-level population counts from the 2010 Census.

We construct quarterly, county-level employment data from 2006 to 2014 using information from the Quarterly Workforce Indicators (QWI). The QWI is a unique jobs-level data set that is publicly released by the U.S. Census Bureau. Covering 95% of employment in the US, the data set combines information from a number of sources including administrative employment records, Social Security data, and Federal tax records. Because the QWI is based on administrative data, the county-level employment measures are more accurate than employment counts from surveys such as the Current Population Survey. Moreover, firm- and worker-level data are linked so that the QWI is the only data set from which it is possible to obtain employment numbers by gender, age group, county, and two-digit NAICS industry codes across the US.

The concept of employment in the QWI is not exactly the same as in either worker-based surveys or employer-based counts for a number of reasons. First, since the QWI is a jobs-level data set, it can yield counts that are higher than the number of workers who hold any job due to multiple job holding. Second, a job-worker link is considered to exist at the beginning of the quarter if the job existed in the quarter before and in the quarter following. This way of counting could lead to underestimates of the number of workers with a job when workers switch jobs between quarters. Finally, since people may live and work in different counties, some areas may have employment-to-population ratios greater than one when computed using the QWI. Despite these measurement differences, the QWI accurately captures changes in the number of jobs over the business cycle and across areas (Bureau (2017)).

Two concerns arise when merging the opioid and employment data that are worth noting. First, as shown in Appendix Table 3.A.1, the age groups provided in the prescription and employment data do not align perfectly. When constructing the age groups that we predominately focus on—young adults and prime-aged workers—we therefore use slightly different age ranges in each data set. In particular, the young adult group consists of individuals aged 19-44 in the employment data and 20-39 in the prescription data, while the prime-age group consists of individuals aged 45-64 in the employment data and 40-64 in the prescription data

An additional source of potential discrepancy between the opioid and the employment data is that individuals may not work and seek medical care in the same county. For example, an individual who lives in Princeton, New Jersey and obtains a prescription from a doctor close to home but who commutes to New York City for work would cause an inconsistency in the count of individuals involved when matching employment to opioid prescription data.

Much of the previous work on opioids focuses on individuals with low educational attainment. Although we have access to employment data by education level in the QWI, we do not know the education level of the patient in the prescription database and therefore cannot match prescriptions to employment by education. In order to explore this aspect of the opioid-employment relationship, we rank each county based on the proportion of individuals within the county with highest educational attainment of high school or less as of the 2000 Census. We split the data set in half based on this criterion and estimate identical regression models on each subset of the data.

Finally, because health insurance is closely linked to employment in the US, we also estimate models controlling for the percent of people aged 18 to 64 who are insured in each county. These data come from the U.S. Census Bureau's Small Area Health Insurance Estimates (SAHIE), which is the only source of data for single-year estimates of health insurance coverage status for all counties. The SAHIE estimates are based on data from the American Community Survey, tax returns, administrative data from the Supplemental Nutrition Assistance Program, Medicaid, the Children's Health Insurance Program, and the 2010 Census. Estimates are available for males and females (see https://www.census.gov/data-tools/demo/sahie/sahie.html, accessed October 30, 2017).

3.3.1 Summary Statistics

Table 3.1 provides means of the employment-to-population ratios and the per capita opioid prescription rates. The means are shown by gender for all workers and for the two broad age groups that we consider: young adults and prime-age adults. We also show estimates for counties with above- and below-median education.

The employment-to-population ratios are higher among men and among workers aged 19 to 44. The most striking difference, however, is between counties with higher and lower levels of education. On average, counties with a less-educated population have a male employment-to-population ratio of 0.525 compared to 0.720 in counties with a more-educated population. Turning to the per capita opioid prescription rates, we see a quite different pattern. Per capita opioid prescription rates are highest for women and for older workers. Moreover, the

Table 3.1: Mean Employment to Population and Prescriptions Rates

Mean Employment/Population Ratio				
	Age	All Counties	High Education Counties	Low Education Counties
Males	All	0.624	0.720	0.525
	19-44	0.628	0.722	0.529
	45-64	0.618	0.718	0.521
Females	All	0.610	0.684	0.531
	19-44	0.621	0.697	0.541
	45-64	0.595	0.672	0.521
Mean Prescriptions/Population Ratio				
	Age	All Counties	High Education Counties	Low Education Counties
Males	All	0.180	0.207	0.192
	20-39	0.088	0.090	0.087
	40-64	0.310	0.324	0.298
Females	All	0.243	0.267	0.257
	20-39	0.138	0.138	0.138
	40-64	0.385	0.397	0.375

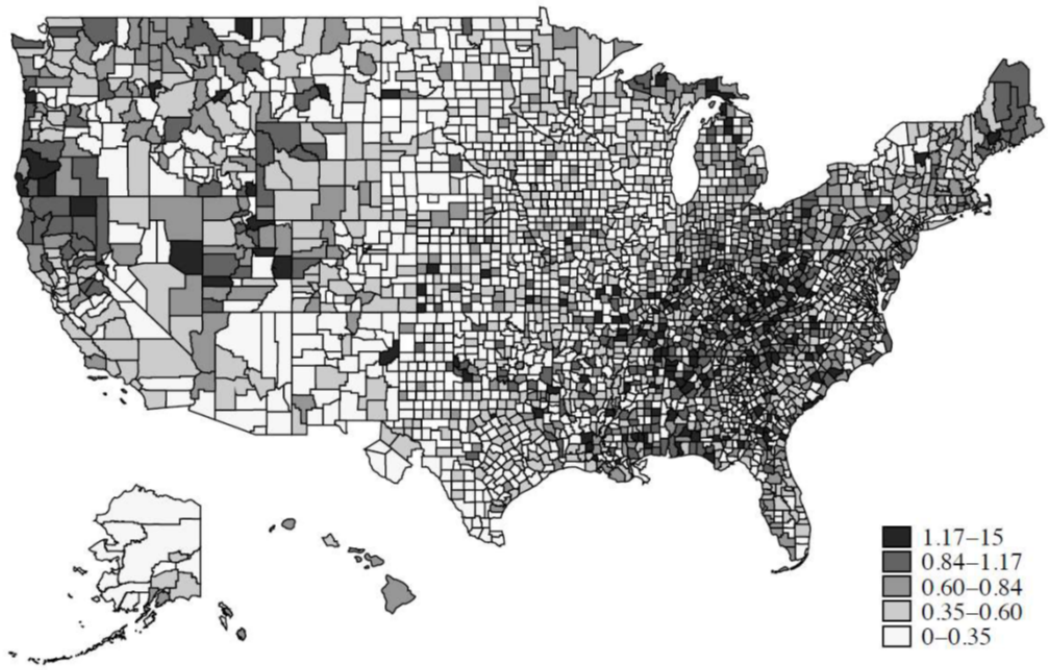
per capita opioid prescription rates are slightly higher in high-education counties, although the differences by education are dwarfed by the differences across gender and age.

Figure 3.1 depicts a heat map of opioid prescriptions per capita. The figure illustrates the large geographic variation in scripts per capita across the country, as well as increases in opioid prescribing over time which one can see by comparing 2006 and 2014. Areas that were harder hit initially tended to be places with higher-than-average unemployment such as Appalachia, Maine, and rust belt states such as Michigan and Northern California. By 2014, prescribing had further intensified in these areas, but much of the rest of the country had begun to catch up, despite the improving economic conditions in most parts of the country.

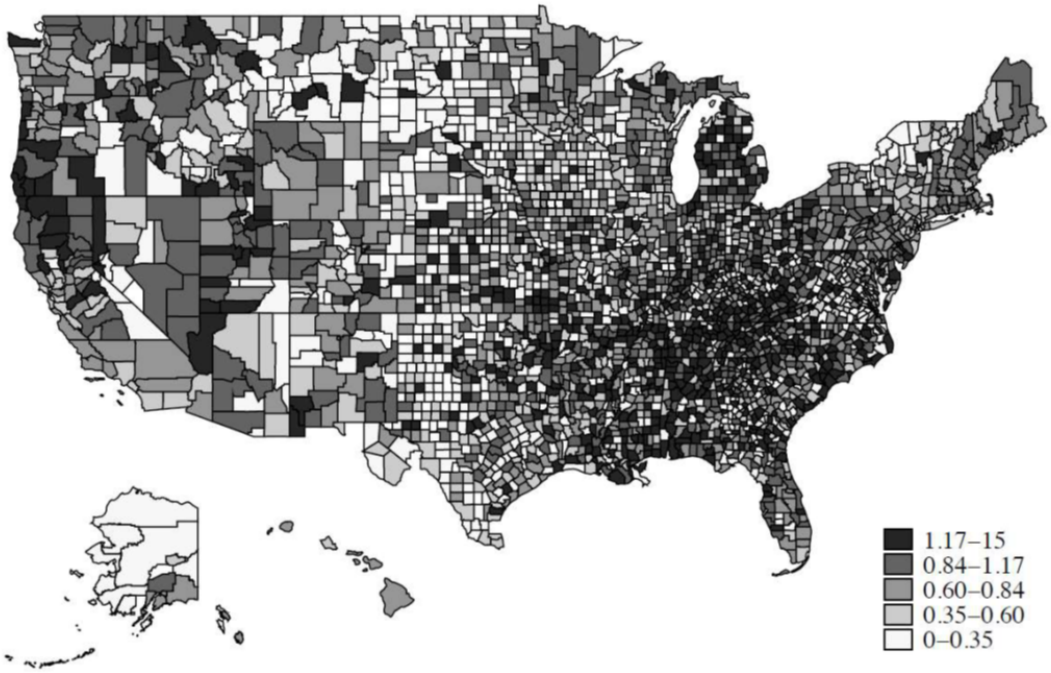
Figure 3.2 shows employment-to-population ratios for workers aged 19 to 64 in the same two years. These figures also show considerable variation across locations, but much less variation over time than the heat maps for opioid scripts shown in Figure 3.1.

Figure 3.2 compares the employment-to-population ratios computed using the QWI to those computed using the Quarterly Census of Employment and Wages (QCEW), which comes only from employer reports for counties with over 100,000 population in 2010. As the figure indicates, the distributions are substantially similar, though the QWI has a somewhat thicker right tail. We also estimate separate regressions using QCEW instead of QWI employment data and obtain similar results, shown in Appendix Table 3.A.3.

Figure 3.1: Opioid Prescriptions per Capita Among Individuals Aged 20-64

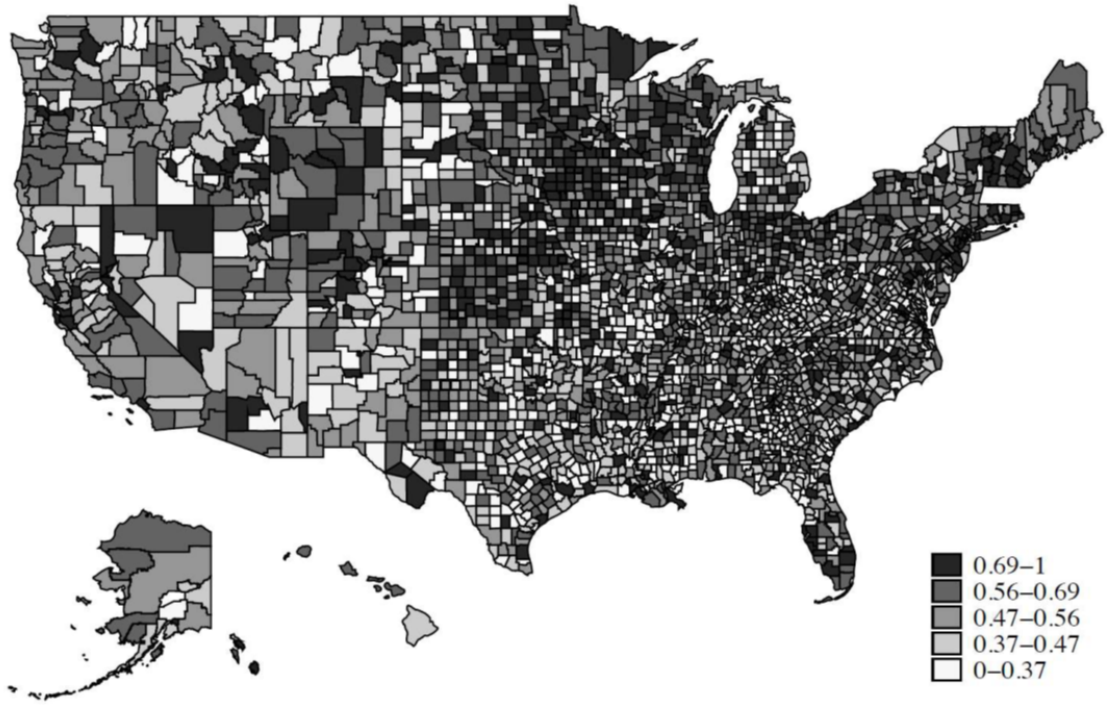


(a) 2006

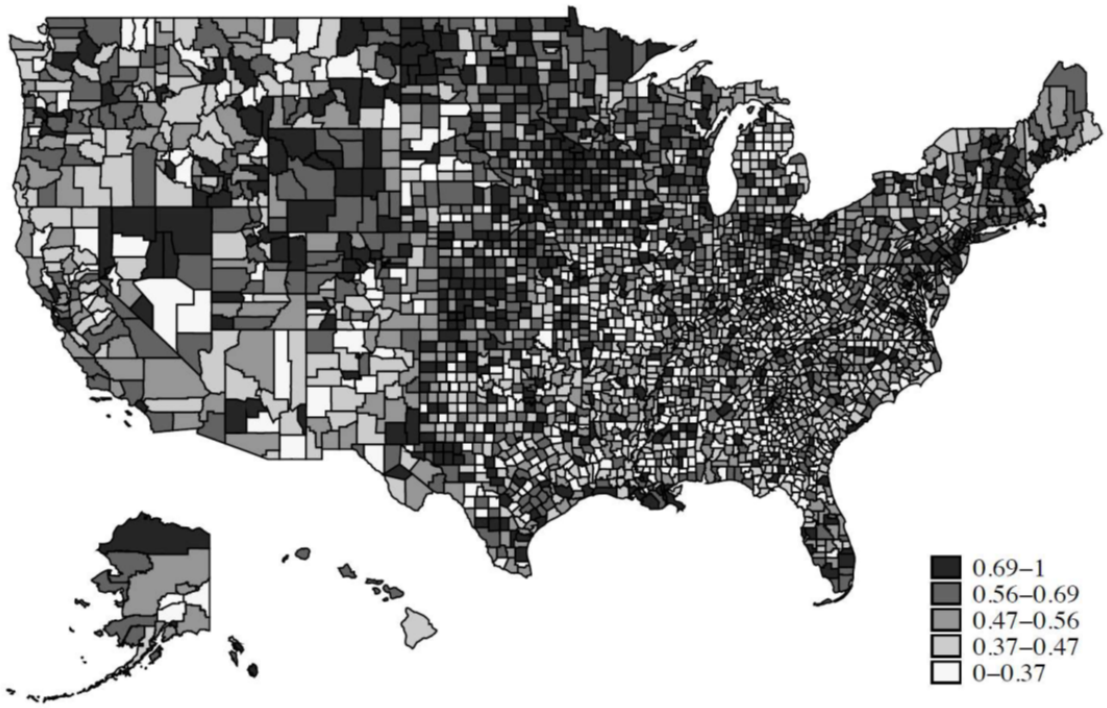


(b) 2014

Figure 3.2: Employment-to-Population Ratios Among Individuals Aged 19-64



(a) 2006



(b) 2014

Figure 3.3: QWI vs. QCEW

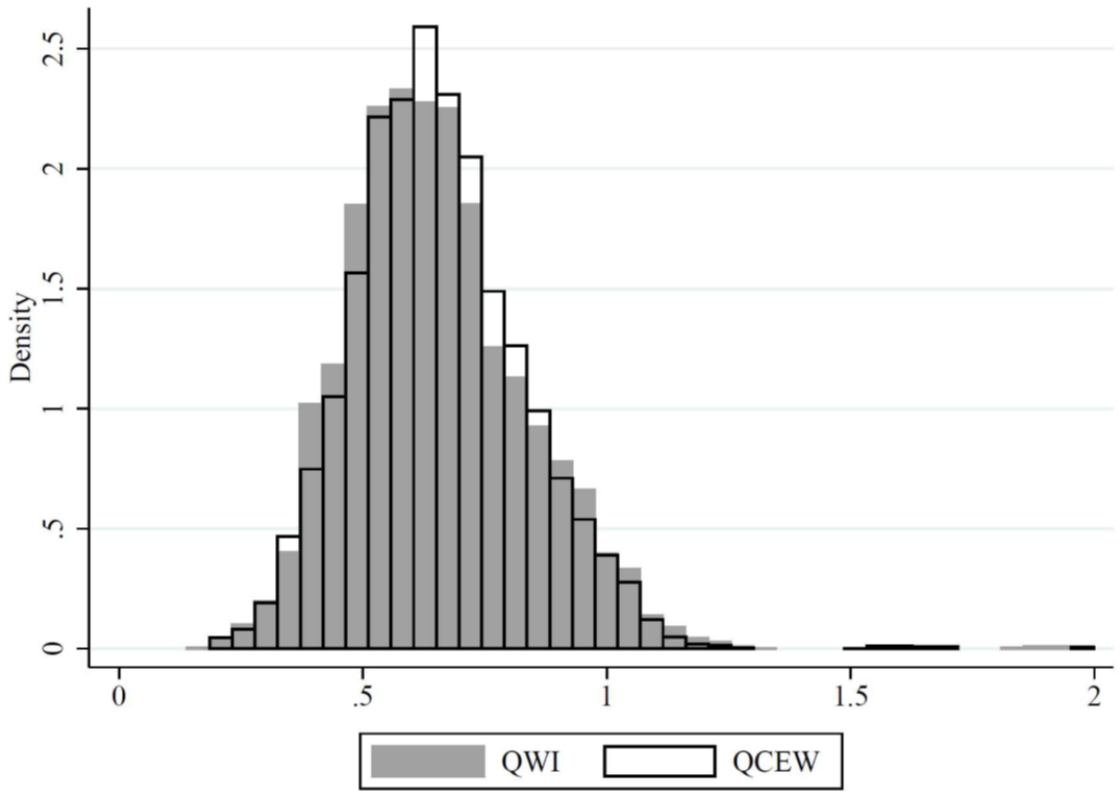
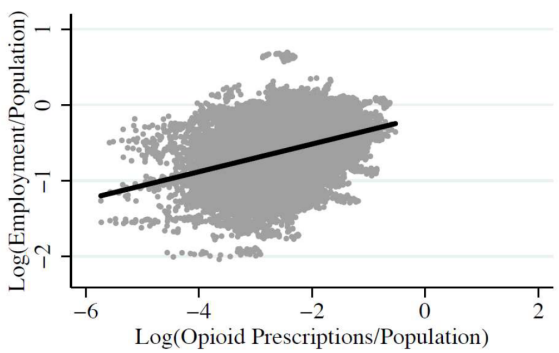


Figure 3.4 shows the contemporaneous relationship between the log of per capita opioid prescriptions and the log of employment-to-population ratios for the four age-sex groups we consider. Perhaps surprisingly, the relationship is positive, as indicated by the fitted regression lines through the scatter plots. Figure 3.4 shows that the data is quite noisy at the county level. In order to minimize the impact of this noise in what follows, we keep all counties with over 100,000 in population in 2010 and then aggregate the other counties in a state into one "rest of state" area.

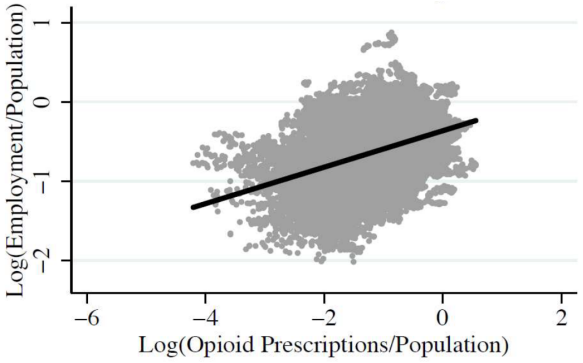
3.4 Methods

We would like to know if more opioid prescribing in a county causes people to lose their jobs, and conversely, whether a lack of employment opportunities causes people to turn to prescription opioids. Posing the questions in this way highlights the potential simultaneity of employment and opioid prescriptions. We deal with this problem in three ways. First, we estimate models in which we regress the dependent variables of interest on lagged values of the independent variables. In so doing, we assume that any effects are not instantaneous and

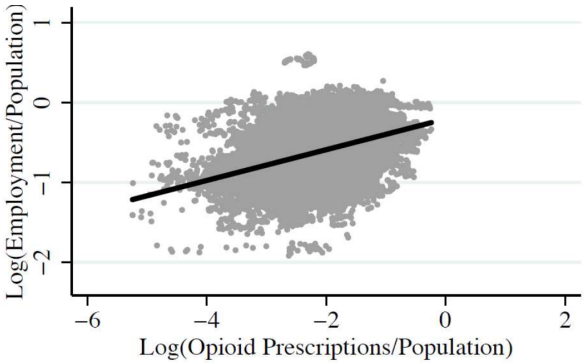
Figure 3.4: Log(Employment to Population) vs. Log(Prescriptions per Capita)



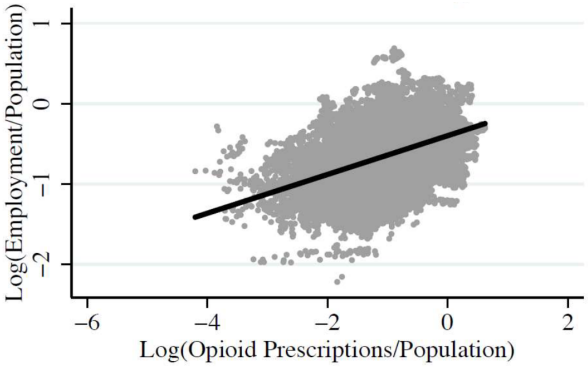
(a) Males, Young Adult



(b) Males, Prime Age



(c) Females, Young Adult



(d) Females, Prime Age

that it is past opioid use that affects employment and vice versa. Second, we estimate models with and without county fixed effects in order to gauge the extent to which any effects that we find are due to constant or long-term characteristics of places rather than short-term fluctuations in either opioid prescribing or employment opportunities. For demographic group j in county i in quarter t, these specifications are given by:

$$\ln\left(\frac{employment_{ijt}}{population_{ij,2010}}\right) = \alpha_0 + \alpha_1 \ln\left[\frac{1}{4}\sum_{q=1}^{4}\left(\frac{prescriptions_{ij,t-q}}{population_{ij,2010}}\right)\right] + \gamma_t + \epsilon_{ijt} \quad (3.1)$$

$$\ln\left(\frac{employment_{ijt}}{population_{ij,2010}}\right) = \alpha_0 + \alpha_1 \ln\left[\frac{1}{4}\sum_{q=1}^{4}\left(\frac{prescriptions_{ij,t-q}}{population_{ij,2010}}\right)\right] + \phi_i + \gamma_t + \epsilon_{ijt} \quad (3.1')$$

$$\ln\left(\frac{prescriptions_{ijt}}{population_{ij,2010}}\right) = \alpha_0 + \alpha_1 ln\left[\frac{1}{4}\sum_{q=1}^{4}\left(\frac{employment_{ij,t-q}}{population_{ij,2010}}\right)\right] + \gamma_t + \epsilon_{ijt} \quad (3.2)$$

$$\ln\left(\frac{prescriptions_{ijt}}{population_{ij,2010}}\right) = \alpha_0 + \alpha_1 ln\left[\frac{1}{4}\sum_{q=1}^{4}\left(\frac{employment_{ij,t-q}}{population_{ij,2010}}\right)\right] + \phi_i + \gamma_t + \epsilon_{ijt} \quad (3.2')$$

where ϕ_i and γ_t denote county-level and time-period fixed effects, respectively, and ϵ_{ijt} is an error term.

Third, we estimate IV models. In models where employment-to-population ratios are the dependent variable, we instrument for per capita opioid prescriptions to working-aged people using per capita opioids prescriptions to people aged 65 and older of the same gender. In the IV version of Equation (3.1'), the first-stage equation is given by:

$$\ln\left(\frac{prescriptions_{ij,t-q}}{population_{ij,2010}}\right) = \beta_0 + \beta_1 ln\left[\frac{1}{4}\sum_{q=1}^{4}\left(\frac{prescriptions_{i65+,t-q}}{population_{ij,2010}}\right)\right] + \phi_i + \gamma_t + \omega_{ijt} \quad (3.3)$$

where $\frac{1}{4}\sum_{q=1}^{4}\left(\frac{prescriptions_{i65+,t-q}}{population_{ij,2010}}\right)$ indicates the average number of per capita prescriptions to those aged 65+ of the same gender as group j in county i in quarters $t-1$ to $t-4$.

One assumption underlying this instrument is that doctors are more likely to prescribe opioids to everyone in some places than in others (Schnell (2017)), so that places where elderly people are more likely to get prescriptions are places where working-age people are also more likely to get them. At the same time, we do not expect prescriptions to the elderly to have a direct effect on the employment of working-aged people.

A second identifying assumption is that there is no third unobserved variable (other than opioid prescriptions for those of working age) that affects both the employment of the working-age population and opioid prescription rates for the elderly. While it is always possible that such a variable exists, it is hard to think of what it might be. Economic conditions should not affect prescription rates among the elderly, since they are largely out of the labor force and have prescription drug benefits under Medicare Part D. High rates of

manual labor and resulting disability might lead to high rates of opioid prescribing among the elderly, but these characteristics of the local labor market would also presumably lead to high opioid prescribing among prime-age workers. To the extent that any such omitted variables are fairly constant over time, they will be captured by county fixed effects in our models.

In regressions where employment-to-population ratios are the independent variables, we instrument for employment using a Bartik-style shift-share instrument. Using the industry composition of the county's employment from the base year of 2005, we calculate the county's predicted employment if the level of employment in each industry had changed in the same ratio as in the rest of the nation in that industry. We sum over industries in each county to calculate the predicted level of employment (the Bartik instrument) for each gender and age group. Thus, the expression for the Bartik instrument in county i for demographic group j in quarter t is given by:

$$Bartik_{ijt} = \sum_{s \in \{sectors\}} \left(employment_{ijs,2005} * \frac{\sum_{k \in \{counties \setminus i\}} employment_{kjst}}{\sum_{k \in \{counties \setminus i\}} employment_{kjs,2005}} \right)$$

In the IV version of Equation (3.2), the first stage is therefore given by:

$$\ln \left[\frac{1}{4} \sum_{q=1}^{4} \left(\frac{employment}{population} \right)_{ij,t-q} \right] = \beta_0 + \beta_1 \ln \left[\frac{1}{4} \sum_{q=1}^{4} \frac{Bartik_{ij,t-q}}{population_{ij,2010}} \right] + \phi_i + \gamma_t + \omega_{ijt} \quad (3.4)$$

Recently Goldsmith-Pinkham et al. (2020) have argued that the Bartik instrument amounts to using interactions between initial local employment shares and national industry employment rates. It is clear that these instruments are much more likely to meet the exogeneity assumption in models that include county fixed effects than in those without, which underscores the importance of examining these relationships using panel data.

3.5 Results

While our primary focus is on models broken down by demographic group using data from the QWI, Appendix Table 3.A.3 shows estimates of Equations (3.1) and (3.2) using all working-age people and comparing estimates obtained using the QWI and the QCEW. We show OLS estimates for all counties and for counties with above- and below-median education.

The estimates are very similar in the two data sources, though somewhat more precisely estimated in the QWI. In what follows, we use the QWI because the QWI allows us to use employment numbers broken down by gender, age group, county, and industry.

Table 3.2 shows estimates for Equation (3.1), in which employment-to-population ratios are a function of per capita opioid prescriptions. We begin with a model without county fixed effects and estimate models separately by gender and age group (young adults and prime-age adults). These estimates should arguably be most similar to past research that also does not include county fixed effects. Because of the log-log formulation, these estimates can be interpreted as elasticities.

The estimates all suggest a positive effect of lagged opioid prescriptions on employment-to-population ratios. The estimates are higher for workers aged 45 to 64 than for younger adults. For older workers, the effects are largest in high-education counties, whereas for younger workers they are largest in low-education counties. IV estimates from the analogous IV specification for Equation (3.1) are shown in columns 4 and 5. The IV estimates are quite similar to the OLS estimates in this formulation. However, the instrument may not be valid without fixed effects as it is possible that prescription rates among the elderly pick up other characteristics of counties.

Table 3.3 shows estimates of Equation (3.1'), which includes county fixed effects in order to control for constant differences across counties. The estimated elasticities are all much smaller than in Table 3.2, but remain positive. They are larger in low-education counties and statistically significant in the OLS models for workers aged 19 to 44 (column 3). The IV estimates show some evidence of significant positive effects in the models for female workers in low-education counties, though the IV estimates for males are small and not statistically significant.

Taken literally, these estimates suggest that opioid prescriptions may help some female workers to stay in the labor force, though the effects are small. In counties that have below the median level of education, a 100% increase in per capita opioid prescriptions is estimated to increase the employment-to-population ratio by 3.3% among women aged 19 to 44 and by 4.3% among women aged 45 to 64. We do not find any statistically significant effect among men.

In Table 3.4 we turn to the opposite question of whether higher employment-to-population ratios discourage reliance on prescription opioids, or Equation (3.2). This table, without county fixed effects, suggests that there are actually large positive effects of higher employment rates on opioid prescribing and that these effects are larger for individuals aged 20-39. For these workers, the estimated effects are once again larger in less-educated than in more-

Table 3.2: Effects of Opioid Prescriptions on Employment: without County Fixed Effects

	(1) All Counties OLS	(2) High Ed. Counties OLS	(3) Low Ed. Counties OLS	(4) High Ed. Counties IV	(5) Low Ed. Counties IV
1. Females, 19-44; Mean = .606					
Log(Prescriptions per Capita)	.188***	.141***	.209***	.205***	.292***
	(.047)	(.039)	(.064)	(.039)	(.050)
R^2	0.113	0.065	0.2244	-	-
First stage F-stat	-	-	-	251.63	310.33
2. Females, 45-64; Mean = .586					
Log(Prescriptions per Capita)	.222***	.253***	.156***	.251***	.205***
	(.038)	(.036)	(.060)	(.037)	(.054)
R^2	0.111	0.149	0.082	-	-
First stage F-stat	-	-	-	633.31	721.88
3. Males, 19-44; Mean = .609					
Log(Prescriptions per Capita)	.193***	.164***	.169***	.223***	.297***
	(.043)	(.033)	(.065)	(.043)	(.049)
R^2	0.105	0.088	0.147	-	-
First stage F-stat	-	-	-	381.49	240.2
4. Males, 45-64; Mean = .599					
Log(Prescriptions per Capita)	.215***	.264***	0.09	.313***	.192***
	(.044)	(.047)	(.061)	(.050)	(.050)
R^2	0.083	0.132	0.032	-	-
First stage F-stat	-	-	-	844.86	486.26
N	20,453	10,142	10,278	10,142	10,278

Notes: First-stage estimates for IV: coefficient on endogenous variable = .944, s.e. = .036, R^2 = .806. The dependent variable in each regression is log(employment/population) for a given demographic group at the county-quarter level. All regressions include year and quarter fixed effects. Standard errors are clustered by county and are shown in parentheses. Counties with fewer than 100,000 people were merged into one aggregate county for each state.

educated counties. The IV point estimates in columns 4 and 5 are slightly larger than the OLS, but not significantly so.

Table 3.5 shows results from Equation (3.2'), the same models as Table 3.4 including county fixed effects. The OLS estimates remain large and positive. The IV estimates are negative for individuals aged 18 to 44 in high-education counties, but large and positive for older women in less-educated counties. Taken at face value, these estimates suggest that younger workers may be less likely to turn to opioids when employment-to-population ratios are higher. This finding might reflect less "despair," but it might also reflect workers' abilities to be more selective about their jobs and avoid jobs that cause them pain or injury. Among

Table 3.3: Effect of Opioid Prescriptions on Employment: with County Fixed Effects

	(1) All Counties	(2) High Ed. Counties OLS	(3) Low Ed. Counties OLS	(4) High Ed. Counties IV	(5) Low Ed. Counties IV
1. Females, 19-44; Mean = .606					
Log(Prescriptions per Capita)	0.024***	0.018*	0.030***	0.017	0.033**
	(.007)	(.010)	(.009)	(.016)	(.016)
R^2	0.986	0.984	0.981	0.984	0.981
First stage F-stat	-	-	-	247.8	116.21
2. Females, 45-64; Mean = .586					
Log(Prescriptions per Capita)	0.014*	0.015	0.019	0.035**	0.043***
	(.008)	(.011)	(.013)	(.017)	(.017)
R^2	0.986	0.985	0.982	0.985	0.981
First stage F-stat	-	-	-	211.53	68.17
3. Males, 19-44; Mean = .609					
Log(Prescriptions per Capita)	0.020**	0.009	0.035***	-0.003	-0.011
	(.008)	(.011)	(.011)	(.017)	(.018)
R^2	0.984	0.982	0.974	0.982	0.973
First stage F-stat	-	-	-	227.74	112.67
4. Males, 45-64; Mean = .599					
Log(Prescriptions per Capita)	0.005	0.0004	0.014	0.012	-0.012
	(.008)	(.011)	(.012)	(.016)	(.016)
R^2	0.989	0.988	0.984	0.988	0.984
First stage F-stat	-	-	-	283	84.82
N	20,453	10,142	10,278	10,142	10,278

Notes: First-stage estimates for IV: coefficient on endogenous variable = .912, s.e. = .033, R^2 = .998. The dependent variable in each regression is log(employment/population for a given demographic group at the county-quarter level. All regressions include year, quarter, and county fixed effects. Standard errors are clustered by county and are shown in parentheses. Counties with fewer than 100,000 people were merged into one aggregate county for each state.

older women, there appears to be a positive relationship between employment-to-population ratios and opioid prescribing, though the lack of significance of the instrument suggests that we should be careful when interpreting this estimate.

3.6 Robustness

There are many potential confounders in any county-level analysis. A potentially important confounder for our study is the prevalence of health insurance in the county, as health insur-

Table 3.4: Effect of Employment on Opioid Prescriptions: without County Fixed Effects

	(1) All Counties OLS	(2) High Ed. Counties OLS	(3) Low Ed. Counties OLS	(4) High Ed. Counties IV	(5) Low Ed. Counties IV
Females, 18-44; Mean = .160					
Log(Employment to Population Ratio)	.575***	.396**	1.034***	.398***	1.137***
	(.161)	(.156)	(.244)	(.149)	(.239)
R^2	0.206	0.191	0.289	-	-
First stage F-stat	-	-	-	98.52	335.49
Females, 45-64; Mean = .433					
Log(Employment to Population Ratio)	.479***	.554***	.497**	.568***	.570***
	(.093)	(.093)	(.196)	(.097)	(.198)
R^2	0.218	0.259	0.187	-	-
First stage F-stat	-	-	-	292.09	246.64
Males, 18-44; Mean = .103					
Log(Employment to Population Ratio)	.479***	.420***	.751***	.416***	.808***
	(.106)	(.112)	(.232)	(.109)	(.252)
R^2	0.195	0.208	0.204	-	-
First stage F-stat	-	-	-	151.89	274.63
Males 45-64, Mean = 0.348					
Log(Employment to Population Ratio)	.355***	.470***	0.242	.512***	.350**
	(.069)	(.071)	(.166)	(.076)	(.175)
R^2	0.196	0.25	0.142	-	-
First stage F-stat	-	-	-	215.76	333.12
N	20,453	10,159	10,278	10,159	10,278

Notes: First-stage estimates for IV: coefficient on endogenous variable = .954, s.e. = .020, R^2 = .933. All regressions include year, and quarter fixed effects. Standard errors clustered by county and shown in parentheses. Counties with fewer than 100,000 people have been merged into one aggregate county for each state.

ance is strongly related to employment in the US. While health insurance varies markedly across counties, it may not be adequately controlled for by including county fixed effects alone since the fraction of Americans with health insurance changed over our sample period. Declining health insurance coverage rates were an important motivation for the Affordable Care Act (ACA). Starting in 2011, the ACA increased coverage for those aged 18 to 26 by allowing them to be covered under their parents' health insurance plans. In 2014, the health insurance exchanges created by the ACA opened, and some states expanded their Medicaid coverage to include childless adults.

Table 3.5: Effect of Employment on Opioid Prescriptions: with County Fixed Effects

	(1) All Counties OLS	(2) High Ed. Counties OLS	(3) Low Ed. Counties OLS	(4) High Ed. Counties IV	(5) Low Ed. Counties IV
1. Females, 18-44; Mean = .160					
Log(Employment to Population Ratio)	0.546***	0.477***	0.641***	-1.143*	-0.093
	(0.088)	(0.130)	(0.121)	(0.620)	(0.448)
R^2	0.93	0.907	0.948	-	-
First stage F-stat	-	-	-	44.7	35.68
2. Females, 45-64; Mean = .433					
Log(Employment to Population Ratio)	.221***	.322***	.263**	15.64	1.553***
	(.071)	(.098)	(.113)	(13.75)	(0.341)
R^2	0.948	0.94	0.958	-	-
First stage F-stat	-	-	-	29.57	29.97
3. Males, 18-44; Mean = .103					
Log(Employment to Population Ratio)	.449***	.337**	.545***	-.798**	-0.881
	(.096)	(.164)	(.100)	(.355)	(.593)
R^2	0.917	0.893	0.937	-	-
First stage F-stat	-	-	-	52.52	10.1
4. Males, 45-64; Mean = 0.348					
Log(Employment to Population Ratio)	.213***	.227**	.307***	0.710	-0.553
	(.065)	(.094)	(.095)	(.630)	(.551)
R^2	0.945	0.936	0.955	-	-
First stage F-stat	-	-	-	48.75	22.2
N	20,453	10,142	10,278	10,142	10,278

Notes: First-stage estimates for IV: coefficient on endogenous variable = .793, s.e. = .196, R2 = .999. All regressions include year and quarter fixed effects. Standard errors clustered by county and shown in parentheses. Counties with fewer than 100,000 people have been merged into one aggregate county for each state.

Tables 3.6 and 3.7 show estimates similar to the models with county fixed effects that were shown in Tables 3.3 and 3.5 but control for the percent insured in each county and gender group. Our intent is not to identify a causal effect of insurance coverage, as to do so would require an additional instrument. Rather, we ask whether the estimates are sensitive to the inclusion of measures of insurance coverage. Because the county-level insurance data are annual, we first estimate models similar to Equations (3.1') and (3.2') but aggregating our data to the annual level. Because estimates of insurance coverage are available only for those aged 18 to 64, we also aggregate the two age categories.

Panels 1 and 3 of Table 3.6 show that our qualitative findings about the effects of opioids on employment-to-population ratios continue to hold in the annual pooled-ages data. Panels 2 and 4 show that when we add the percent insured to the model, the estimated effect of opioid prescriptions on the employment-to-population ratio is unaffected. As before, we find a significant positive relationship between opioids per capita and employment-to-population among women and no relationship among men.

Table 3.7 shows estimates of the effect of employment-to-population ratios on per capita opioid prescriptions using annual pooled-ages data. The estimates are qualitatively similar to those shown in Table 3.5 in terms of the signs. However, none of the OLS estimates are statistically significant for high-education counties and none of the IV estimates are statistically significant. Hence, the estimated effects of employment-to-population ratios on opioid prescribing are more sensitive to specification than the estimated effects of opioids on employment-to-population ratios.

3.7 Discussion and Conclusions

Compared to most previous work, we have the advantage of a long time period that allows us to control for time-invariant differences across counties using county fixed effects. We also have detailed data on opioid prescriptions as well as quarterly employment data that are available at the county level and broken down by age and gender groups. This rich data allows us to explore the extent to which results differ for older and younger workers and for males and females. The estimated effects of prescription opioids on employment-to-population ratios are robust to changes in specification and suggest that there is a positive relationship for women and no significant relationship for men. More specifically, a 100% increase in opioid prescribing would lead to increases in employment of 3.8% among women in counties with above-median education and 5.2% among women in counties with below-median education. Our exploration of the effects of health insurance would appear to rule it

Table 3.6: Effect of Opioid Prescriptions on Employment: controlling for County Fixed Effects and % Insured

	(1) All Counties OLS	(2) High Ed. Counties OLS	(3) Low Ed. Counties OLS	(4) High Ed. Counties IV	(5) Low Ed. Counties IV
1. Females, 18-64, Mean = 0.622					
Log(Prescriptions per Capita)	.025***	0.018	.034***	.038*	.052***
	(.009)	(.013)	(.012)	(.020)	(.017)
R^2	0.988	0.988	0.983	-	-
First stage F-stat	-	-	-	342.27	114.04
2. Females, 18-64, controlling for % insured					
Log(Prescriptions per Capita)	.026***	0.021	.038***	.039**	.058***
	(.009)	(.013)	(.012)	(.020)	(.018)
% insured	.279***	.292***	.334***	.303***	.343***
	(.059)	(.082)	(.086)	(.081)	(.083)
R^2	0.988	0.988	0.984	-	-
First stage F-stat	-	-	-	269.99	103.19
3. Males, 18-64, Mean = 0.632					
Log(Prescriptions per Capita)	.020**	0.009	.038***	0.017	0.007
	(.010)	(.014)	(.012)	(.020)	(.017)
	0.989	0.988	0.983	-	-
	-	-	-	336	139.33
4. Males, 18-64, controlling for %insured					
Log(Prescriptions per Capita)	.022**	0.012	.042***	0.021	0.017
	(.010)	(.014)	(.012)	(.020)	(.016)
% insured	.347***	.391***	.367***	.396***	.358***
	(.061)	(.089)	(.073)	(.086)	(.070)
R^2	0.99	0.988	0.984	-	-
First stage F-stat	-	-	-	300.51	144.81
N	5,585	2,780	2,805	2,780	2,805

Notes: First-stage estimates for IV: coefficient on endogenous variable = .996, s.e. = .031, R^2 = .996. All regressions include year and quarter fixed effects. Standard errors clustered by county and shown in parentheses. Counties with fewer than 100,000 people have been merged into one aggregate county for each state.

out as a major explanation for these findings. An alternative explanation is that although they are addictive and potentially dangerous, opioids nevertheless allow some women to keep working who might otherwise withdraw from the labor force.

When we examine the effects of employment-to-population ratios on opioid prescriptions, our results are more ambiguous. We find some evidence that higher employment-to-

Table 3.7: Effect of Employment on Opioid Prescriptions: controlling for County Fixed Effects and % Insured

	(1) All Counties OLS	(2) High Ed. Counties OLS	(3) Low Ed. Counties OLS	(4) High Ed. Counties IV	(5) Low Ed. Counties IV
1. Females, 18-64; Mean = 1.085					
Log(Employment to Population Ratio)	.200**	0.171	.341***	-11.59	0.211
	(.082)	(.115)	(.100)	(32.71)	(.433)
R^2	0.986	0.985	0.987	-	-
First stage F-stat	-	-	-	49.68	26.07
2. Females, 18-64; Controlling for %insured					
Log(Employment to Population Ratio)	.212**	0.186	.373***	-15.9	0.251
	(.083)	(.114)	(.101)	(64.62)	(.439)
%insured	-.229	-.411	-.395	2.287	-.359
	(.157)	(.231)	(.234)	(10.83)	(.273)
R^2	0.986	0.985	0.987	-	-
First stage F-stat	-	-	-	44.71	22.31
3. Males, 18-64; Mean = 0.802					
Log(Employment to Population Ratio)	.223**	0.141	.349***	-2.179	-2.786
	(.088)	(.141)	(.088)	(1.767)	(2.296)
R^2	0.984	0.983	0.986	-	-
First stage F-stat	-	-	-	63.76	19.67
4. Males, 18-64; Controlling for %insured					
Log(Employment to Population Ratio)	.239***	0.166	.379***	-2.276	-3.151
	(.090)	(.144)	(.091)	(2.005)	(2.892)
%insured	-.230	-.437**	-.372	.258	.733
	(.153)	(.218)	(.242)	(.619)	(.877)
R^2	0.984	0.983	0.986	-	-
First stage F-stat	-	-	-	66.04	16.23
N	5,585	2,784	2,801	2,784	2,801

Notes: First-stage estimates for IV: coefficient on endogenous variable = .616, standard error = .226, R^2 = .999. All regressions include year, and quarter fixed effects. Standard errors clustered by county and shown in parentheses. Counties with fewer than 100,000 people have been merged into one aggregate county for each state.

population ratios reduce per capita opioid prescriptions among young adults, though this effect is only statistically significant in the IV specification and only in counties with education above the median.

The limited effect of employment-to-population ratios on per capita opioid prescriptions that we document should perhaps not be a surprise. As Case and Deaton (2017) clarify, the

type of despair they describe has more to do with a longer-term unraveling of the social fabric than with short-term variations in employment prospects. Moreover, there is increasing evidence that opioid prescribing patterns depend more on idiosyncratic factors such as the characteristics and preferences of local doctors than on local economies (Schnell and Currie (2018); Schnell (2017)).

Overall, our analysis suggests that the relationship between opioid prescribing and employment is considerably weaker and murkier than popular narratives suggest. The fact that many opioid users are still in the labor market —and are likely having their prescriptions paid for by employer-sponsored health insurance—is one reason the opioids are having such a large impact on American employers. This observation suggests that policy responses should be designed to take into account the fact that many opioid users work. Treatment options that help people retain their connection to the labor market are likely to be necessary to effectively combat the epidemic.

Appendix

Table 3.A.1: Age Groupings in Prescription and Employment Data

Age group classification	Opioid	Employment
1: Adolescents	0-2, 3-9, 10-19	14-18
2: Young adults	20-39	19-21, 22-24, 25-34, 35-44
3: Prime-age workers	40-59, 60-64	45-54, 55-64
4: Retirees	65-74, 75-84, 85+	65-99

Table 3.A.2: Mapping from NAICS Codes to SIC Codes

NAICS industry	SIC industry
11 (Agriculture, Forestry, Fishing, and Hunting)	01-09 (Agriculture, Forestry, and Fishing)
21 (Mining, Quarrying, and Oil and Gas Extraction)	10-14 (Mining)
23 (Construction)	15-17 (Construction)
31-33 (Manufacturing) 51 (Information)	20-39 (Manufacturing)
22 (Utilities) 48-49 (Transportation and Warehousing)	40-49 (Transportation and Public Utilities: Electric, Gas, Communications, Sanitary)
42 (Wholesale Trade)	50-51 (Wholesale Trade)
44-45 (Retail Trade) 72 (Accommodation and Food Services)	52-59 (Retail Trade)
52 (Finance and Insurance) 53 (Real Estate and Rental and Leasing) 55 (Management of Companies and Enterprises)	60-67 (Finance, Insurance, and Real Estate)
54 (Professional, Scientific, and Technical Services) 56 (Administrative and Support and Waste Management and Remediation Services) 61 (Educational Services) 62 (Health Care and Social Assistance) 71 (Arts, Entertainment, and Recreation) 81 (Other Services (except Public Administration))	70-89 (Services)
92 (Public Administration)	91-99 (Public Administration)

Table 3.A.3: Comparison of QWI and QCEW OLS Estimates for Working Age Individuals

	(1) All Counties QWI	(2) All Counties QCEW	(3) High Ed. Counties QWI	(4) High Ed. Counties QCEW	(5) Low Ed. Counties QWI	(6) Low Ed. Counties QCEW
1. Dependent Variable = ln(Employment/Population); Mean = .664 (QWI), .666 (QCEW)						
Log(Prescriptions per Capita)	0.011*	0.011*	0.015*	0.014	0.010*	0.009*
	(0.005)	(0.005)	(0.007)	(0.007)	(0.004)	(0.004)
R^2	0.985	0.986	0.985	0.987	0.972	0.976
N	100,744	100,738	50,156	50,154	50,588	50,584
2. Dependent Variable = ln(Prescriptions/Population); Mean = .286						
Log(Emp. per Capita)	0.158**	0.166**	0.203**	0.220*	0.137*	0.141
	(0.050)	(0.065)	(0.068)	(0.0910	(0.067)	(0.077)
R^2	0.947	0.947	0.957	0.957	0.925	0.925
N	100,517	100,517	50,048	50,48	50,469	50,469

Standard errors in parentheses. These are population weighted regressions including all available county-quarter observations. All regressions include county fixed effects.

Bibliography

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L. Violante, "Education Policy and Intergenerational Transfers in Equilibrium," *Journal of Political Economy*, 2019, *127* (6), 2569–2624.

Abowd, John M and Martha H Stinson, "Estimating measurement error in annual job earnings: A comparison of survey and administrative data," *Review of Economics and Statistics*, 2013, *95* (5), 1451–1467.

Alesina, Alberto F, Marlon Seror, David Y Yang, Yang You, and Weihong Zeng, "Persistence Despite Revolutions," Working Paper 27053, National Bureau of Economic Research April 2020.

Alvarez, Jorge A, "The agricultural wage gap: Evidence from Brazilian micro-data," *American Economic Journal: Macroeconomics*, 2020, *12* (1), 153–173.

Ásgeirsdóttir, Tinna Laufey, Hope Corman, Kelly Noonan, þórhildur Ólafsdóttir, and Nancy E Reichman, "Was the economic crisis of 2008 good for Icelanders? Impact on health behaviors," *Economics & Human Biology*, 2014, *13*, 1–19.

Asher, Sam, Paul Novosad, and Charlie Rafkin, "Intergenerational Mobility in India: Estimates from New Methods and Administrative Data," 2018.

Azevedo, Viniane MR and Cesar P Bouillon, "Intergenerational social mobility in Latin America: a review of existing evidence," *Revista de Análisis Económico–Economic Analysis Review*, 2010, *25* (2), 7–42.

Baker, Michael and Gary Solon, "Earnings dynamics and inequality among Canadian men, 1976–1992: Evidence from longitudinal income tax records," *Journal of Labor Economics*, 2003, *21* (2), 289–321.

Bartik, Timothy J, "Who benefits from state and local economic development policies?," 1991.

Becker, Gary S. and Nigel Tomes, "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 1979, *87* (6), 1153–1189.

Becker, Gary S and Nigel Tomes, "Human capital and the rise and fall of families," *Journal of labor economics*, 1986, *4* (3, Part 2), S1–S39.

Behrman, Jere R, Alejandro Gaviria, Miguel Székely, Nancy Birdsall, and Sebastián Galiani, "Intergenerational mobility in Latin America [with comments]," *Economia*, 2001, *2* (1), 1–44.

Benabou, Roland, "Workings of a City: Location, Education, and Production," *The Quarterly Journal of Economics*, 1993, *108* (3), 619–652.

Bergemann, Annette, Erik Gronqvist, and Soffia Gudbjörnsdottir, "The effects of job displacement on the onset and progression of diabetes," 2011.

Bjorklund, Anders and Markus Jantti, "Intergenerational mobility of socio-economic status in comparative perspective," *Nordic Journal of Political Economy*, 2000, *26*, 3–32.

Black, Sandra and Paul Devereux, "Recent Developments in Intergenerational Mobility," 2011, *4B*, 1487–1541.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "Losing Heart? The Effect of Job Displacement on Health," *ILR Review*, 2015, *68* (4), 833–861.

Boffy-Ramirez, Ernest and Soojae Moon, "The role of China's household registration system in the urban-rural income differential," *China Economic Journal*, 2018, *11* (2), 108–125.

Bosker, Maarten, Steven Brakman, Harry Garretsen, and Marc Schramm, "Relaxing Hukou: Increased labor mobility and China's economic geography," *Journal of Urban Economics*, 2012, *72* (2), 252–266.

Bowles, Samuel, "Schooling and inequality from generation to generation," *Journal of Political Economy*, 1972, *80* (3, Part 2), S219–S251.

Bratsberg, Bernt, Knut Røed, Oddbjørn Raaum, Robin Naylor, Markus Ja¨ ntti, Tor Eriksson, and Eva O¨ sterbacka, "Nonlinearities in intergenerational earnings mobility: consequences for cross-country comparisons," *The Economic Journal*, 2007, *117* (519), C72–C92.

Browning, Martin and Eskil Heinesen, "Effect of job loss due to plant closure on mortality and hospitalization," *Journal of health economics*, 2012, *31* (4), 599–616.

Bryan, Gharad and Melanie Morten, "The aggregate productivity effects of internal migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

_ , Shyamal Chowdhury, and Ahmed Mushfiq Mobarak, "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh," *Econometrica*, 2014, *82* (5), 1671–1748.

Bureau, US Census, "Quarterly Workforce Indicators 101, Employment Dynamics," 2017. https://lehd.ces.census.gov/doc/QWI_101.pdf.

Bénabou, Roland, "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 2002, *70* (2), 481–517.

Capelle, Damien, "Essays on the Roles of Financial and Higher Education Institutions in Macroeconomics." PhD dissertation, Princeton University 2020.

Carpenter, Christopher S, Chandler B McClellan, and Daniel I Rees, "Economic conditions, illicit drug use, and substance use disorders in the United States," *Journal of health economics*, 2017, *52*, 63–73.

Case, Anne and Angus Deaton, "Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century," *Proceedings of the National Academy of Sciences*, 2015, *112* (49), 15078–15083.

_ and _ , "Mortality and morbidity in the 21st century," *Brookings Papers on Economic Activity*, 2017, *2017*, 397.

Charles, Kerwin Kofi, Erik Hurst, and Mariel Schwartz, "The transformation of manufacturing and the decline in US employment," *NBER Macroeconomics Annual*, 2019, *33* (1), 307–372.

Chen, Cao, "Shanghai relaxes hukou rules to attract more talent," 6 2022.

Chen, Li Hui, Holly Hedegaard, and Margaret author Warner, *Drug-poisoning deaths involving opioid analgesics: United States, 1999-2011* number 2014, US Department of Health and Human Services, Centers for Disease Control and . . . , 2014.

Chen, Wen-Hao, Yuri Ostrovsky, and Patrizio Piraino, *Intergenerational Income Mobility: New Evidence from Canada*, Statistics Canada Ottawa, 2016.

Chetty, Raj and Nathaniel Hendren, "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects," *The Quarterly Journal of Economics*, 2018, *133* (3), 1107–1162.

_ , _ , Patrick Kline, and Emmanuel Saez, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

_ , _ , _ , _ , and Nicholas Turner, "Is the United States still a land of opportunity? Recent trends in intergenerational mobility," *American Economic Review*, 2014, *104* (5), 141–147.

China Labour Bulletin, "Migrant workers and their children," 5 2022.

Chow, Gregory C. and Linlin Niu, "Housing Prices in Urban China as Determined by Demand and Supply," *Pacific Economic Review*, 2015, *20* (1), 1–16.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux, "Sorting and local wage and skill distributions in France," *Regional Science and Urban Economics*, 2012, *42* (6), 913–930.

Corak, Miles, "Do poor children become poor adults? Lessons from a cross-country comparison of generational earnings mobility," in "Dynamics of inequality and poverty," Vol. 13, Emerald Group Publishing Limited, 2006, pp. 143–188.

_ , "Income inequality, equality of opportunity, and intergenerational mobility," *Journal of Economic Perspectives*, 2013, *27* (3), 79–102.

_ **and Andrew Heisz**, "The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data," *Journal of Human Resources*, 1999, pp. 504–533.

Couch, Kenneth A and Dean R Lillard, "Nonlinear patterns of intergenerational mobility in Germany and the United States," *Generational income mobility in North America and Europe*, 2004, pp. 190–206.

Dart, Richard C, Hilary L Surratt, Theodore J Cicero, Mark W Parrino, S Geoff Severtson, Becki Bucher-Bartelson, and Jody L Green, "Trends in opioid analgesic abuse and mortality in the United States," *New England Journal of Medicine*, 2015, *372* (3), 241–248.

Dreger, Christian and Yanqun Zhang, "The Hukou Impact on the Chinese Wage Structure," Technical Report, DIW Berlin, German Institute for Economic Research 2017.

Dulleck, Uwe, Jonas Fooken, and Yumei He, "Hukou Status and Individual-Level Labor Market Discrimination: An Experiment in China," *ILR Review*, May 2020, *73* (3), 628–649.

Durlauf, Steven, "Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality," 1996.

Eckert, Fabian and Tatjana Kleineberg, "Saving the American Dream? Education Policies in Spatial General Equilibrium," 03 2021.

Eliason, Marcus and Donald Storrie, "Does job loss shorten life?," *Journal of Human Resources*, 2009, *44* (2), 277–302.

_ **and** _ , "Job loss is bad for your health–Swedish evidence on cause-specific hospitalization following involuntary job loss," *Social science & medicine*, 2009, *68* (8), 1396–1406.

Falba, Tracy, Hsun-Mei Teng, Jody L Sindelar, and William T Gallo, "The effect of involuntary job loss on smoking intensity and relapse," *Addiction*, 2005, *100* (9), 1330–1339.

Fan, Yi, Junjian Yi, and Junsen Zhang, "Rising Intergenerational Income Persistence in China," *American Economic Journal: Economic Policy*, February 2021, *13* (1), 202–30.

Fernandez, Raquel and Richard Rogerson, "Income Distribution, Communities, and the Quality of Public Education," *The Quarterly Journal of Economics*, 02 1996, *111* (1), 135–164.

_ **and** _ , "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform," *American Economic Review*, 1998, *88* (4), 813–33.

Ferreira, Sergio Guimaraes and Fernando A Veloso, "Intergenerational mobility of wages in Brazil," *Brazilian Review of Econometrics*, 2006, *26* (2), 181–211.

for Behavioral Health Statistics, Center and Quality, "National survey on drug use and health," 2013.

for Economic Cooperation, Organization and Development, *OECD Economy Surveys: United States 2018*, 2018.

Frieden, Thomas R and Debra Houry, "Reducing the risks of relief—the CDC opioid-prescribing guideline," *New England Journal of Medicine*, 2016, *374* (16), 1501–1504.

Friedman, Milton, "The permanent income hypothesis," in "A theory of the consumption function," Princeton University Press, 1957, pp. 20–37.

Galton, Francis, "Typical laws of heredity," 1877.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik instruments: What, when, why, and how," *American Economic Review*, 2020, *110* (8), 2586–2624.

Gollin, Douglas, David Lagakos, and Michael E Waugh, "Agricultural productivity differences across countries," *American Economic Review*, 2014, *104* (5), 165–170.

Grawe, Nathan D, "Reconsidering the use of nonlinearities in intergenerational earnings mobility as a test for credit constraints," *Journal of Human Resources*, 2004, *39* (3), 813–827.

_ , "Lifecycle bias in estimates of intergenerational earnings persistence," *Labour economics*, 2006, *13* (5), 551–570.

_ **and Casey B Mulligan**, "Economic interpretations of intergenerational correlations," *Journal of Economic Perspectives*, 2002, *16* (3), 45–58.

Guo, L, Prashant Loyalka, and Xiaoyang Ye, "Are Regional Quotas Fair? Simulating Merit-Based College Admissions Using Unique Student-Level Data from China," 2018.

Haider, Steven and Gary Solon, "Life-cycle variation in the association between current and lifetime earnings," *American economic review*, 2006, *96* (4), 1308–1320.

Hamory, Joan, Marieke Kleemans, Nicholas Y Li, and Edward Miguel, "Reevaluating agricultural productivity gaps with longitudinal microdata," *Journal of the European Economic Association*, 2021, *19* (3), 1522–1555.

Han, Song and Casey B Mulligan, "Human capital, heterogeneity and estimated degrees of intergenerational mobility," *The Economic Journal*, 2001, *111* (470), 207–243.

Harris, Matthew C, Lawrence M Kessler, Matthew N Murray, and M Elizabeth Glenn, "Prescription opioids and labor market pains," *Knoxville, TN. http://cber. haslam. utk. edu/staff/harris/Opioids_HKMG. pdf*, 2017.

Herskovic, Bernard and Joao Ramos, "Should College Admissions Favor Low-Income Applicants?," 2019.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina, "The inheritance of educational inequality: International comparisons and fifty-year trends," *The BE Journal of Economic Analysis & Policy*, 2008, *7* (2).

Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh B. Paul, "Breaking the Caste Barrier: Intergenerational Mobility in India," *The Journal of Human Resources*, 2013, *48* (2), 435–473.

Hokayem, Charles, Christopher Bollinger, and James P Ziliak, "The role of CPS nonresponse in the measurement of poverty," *Journal of the American Statistical Association*, 2015, *110* (511), 935–945.

Hollingsworth, Alex, Christopher J Ruhm, and Kosali Simon, "Macroeconomic conditions and opioid abuse," *Journal of health economics*, 2017, *56*, 222–233.

Jantti, Markus, Bernt Bratsberg, Knut Roed, Oddbjorn Raaum, Robin Naylor, Eva Osterbacka, Anders Bjorklund, and Tor Eriksson, "American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States," 2006.

Jia, Ruixue and Hongbin Li, "Just above the exam cutoff score: Elite college admission and wages in China," *Journal of Public Economics*, 2021.

Kanbur, Ravi and Xiaobo Zhang, "Fifty Years of Regional Inequality in China: a Journey Through Central Planning, Reform, and Openness," *Review of Development Economics*, 2005, *9* (1), 87–106.

Kotera, Tomoaki and Ananth Seshadri, "Educational policy and intergenerational mobility," *Review of Economic Dynamics*, 2017, *25*, 187–207. Special Issue on Human Capital and Inequality.

Krueger, Alan B, "Where have all the workers gone? An inquiry into the decline of the US labor force participation rate," *Brookings Papers on Economic Activity*, 2017, *2017* (2), 1.

Lee, Sang Yoon (Tim) and Ananth Seshadri, "On the Intergenerational Transmission of Economic Status," *Journal of Political Economy*, 2019, *127* (2), 855–921.

Loury, Glenn C., "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 1981, *49* (4), 843–867.

Loyalka, Prashant, Yue Qu, Jianguo Wei, Binzhen Wu, and Xiaoyang Ye, "Out of the gauntlet and into the fire: How college choices in centralized admissions systems disadvantage disadvantaged students," 2017.

Mari, Tanaka, Narita Yusuke, and Moriguchi Chiaki, "Meritocracy and Its Discontent: Long-run Effects of Repeated School Admission Reforms," Discussion papers 20002, Research Institute of Economy, Trade and Industry (RIETI) January 2020.

Mazumder, Bhashkar, "Earnings mobility in the US: A new look at intergenerational inequality," 2001.

_ , "Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data," *Review of Economics and Statistics*, 2005, *87* (2), 235–255.

Meara, Ellen, Jill R Horwitz, Wilson Powell, Lynn McClelland, Weiping Zhou, A James O'malley, and Nancy E Morden, "State legal restrictions and prescription-opioid use among disabled adults," *New England Journal of Medicine*, 2016, *375* (1), 44–53.

Miguel, Edward and Joan Hamory, "Individual ability and selection into migration in Kenya," 2009.

Mion, Giordano and Paolo Naticchioni, "The spatial sorting and matching of skills and firms," *Canadian Journal of Economics/Revue canadienne d'économique*, 2009, *42* (1), 28–55.

Mulligan, Casey B, "Galton versus the human capital approach to inheritance," *Journal of political Economy*, 1999, *107* (S6), S184–S224.

Ng, Irene, "Intergenerational income mobility in Singapore," *The BE Journal of Economic Analysis & Policy*, 2007, *7* (2).

Nunez, Javier I and Leslie Miranda, "Intergenerational income mobility in a less-developed, high-inequality context: The case of Chile," *The BE Journal of Economic Analysis & Policy*, 2010, *10* (1).

Otero, Sebastián, Nano Barahona, and Cauê Dobbin, "Affirmative Action in Centralized College Admission Systems: Evidence from Brazil," 11 2021.

Patrizio, Piraino, "Comparable Estimates of Intergenerational Income Mobility in Italy," *The B.E. Journal of Economic Analysis & Policy*, October 2007, *7* (2), 1–27.

Pierce, Justin R and Peter K Schott, "Trade liberalization and mortality: evidence from US counties," *American Economic Review: Insights*, 2020, *2* (1), 47–64.

Pulido, José, Tomasz Swiecki et al., "Barriers to mobility or sorting? sources and aggregate implications of income gaps across sectors and locations in indonesia," *Unpublished Working Paper, Vancouver School of Economics*, 2018.

Ravallion, Martin, Dominique Van de Walle, and Banco Mundial, *Cost-of-Living Differences between Urban and Rural Areas in Indonesia*, Vol. 341, Agriculture and Rural Development Department, World Bank, 1989.

Restuccia, Diego and Richard Rogerson, "The causes and costs of misallocation," *Journal of Economic Perspectives*, 2017, *31* (3), 151–174.

Rudd, Rose A, Noah Aleshire, Jon E Zibbell, and R Matthew Gladden, "Increases in drug and opioid overdose deaths—United States, 2000–2014," 2016.

Ruhm, Christopher J, "Are recessions good for your health?," *The Quarterly journal of economics*, 2000, *115* (2), 617–650.

_ , "Healthy living in hard times," *Journal of health economics*, 2005, *24* (2), 341–363.

_ , "Deaths of despair or drug problems?," Technical Report, National Bureau of Economic Research 2018.

Schnell, Molly, "Physician behavior in the presence of a secondary market: The case of prescription opioids," *Princeton University Department of Economics Working Paper*, 2017, *5*.

_ **and Janet Currie**, "Addressing the opioid epidemic: is there a role for physician education?," *American journal of health economics*, 2018, *4* (3), 383–410.

Sieg, Holger, Chamna Yoon, and Jipeng Zhang, "The Impact of Migration Controls on Urban Fiscal Policies and the Intergenerational Transmission of Human Capital in China," Technical Report, National Bureau of Economic Research 2020.

Solon, Gary, "Biases in the estimation of intergenerational earnings correlations," *The Review of Economics and Statistics*, 1989, pp. 172–174.

_ , "Intergenerational income mobility in the United States," *The American Economic Review*, 1992, pp. 393–408.

_ , "Chapter 29 - Intergenerational Mobility in the Labor Market," in Orley C. Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3, Elsevier, 1999, pp. 1761–1800.

Sullivan, Daniel and Till Von Wachter, "Job displacement and mortality: An analysis using administrative data," *The Quarterly Journal of Economics*, 2009, *124* (3), 1265–1306.

Tsang, Mun C., "Financial reform of basic education in China," *Economics of Education Review*, 1996, *15* (4), 423–444. The economics of education in developing countries - The Manchester symposium.

Wang, Chuhong, Mehtap Akgüç, Xingfei Liu, and Massimiliano Tani, "Expropriation with hukou change and labour market outcomes in China," *China Economic Review*, 2020, *60*, 101391.

Whalley, John and Shunming Zhang, "A numerical simulation analysis of (Hukou) labour mobility restrictions in China," *Journal of Development Economics*, 2007, *83* (2), 392–410.

Xu, Chenggang, "The Fundamental Institutions of China's Reforms and Development," *Journal of Economic Literature*, 2011, *49* (4), 1076–1151.

Yan, Weibo and Xiaolan Deng, "Intergenerational income mobility and transmission channels in a transition economy: Evidence from China," *Economics of Transition and Institutional Change*, 2022, *30* (1), 183–207.

Yang, "Place-Based College Admission, Migration and the Spatial Distribution of Human Capital: Evidence from China," 2020.

Zhang, Hongliang, Jere R Behrman, C Simon Fan, Xiangdong Wei, and Junsen Zhang, "Does parental absence reduce cognitive achievements? Evidence from rural China," *Journal of Development Economics*, 2014, *111*, 181–195.

Zhang, Qinghua and Heng-Fu Zou, "Regional Inequality in Contemporary China," *Annals of Economics and Finance*, 2012, *13* (1), 113–137.

Zheng, Angela and James Graham, "Public Education Inequality and Intergenerational Mobility," *American Economic Journal: Macroeconomics*, July 2022, *14* (3), 250–82.

Zhong, Hai, "Returns to higher education in China: What is the role of college quality?," *China Economic Review*, 2011, *22* (2), 260–275.