**Title**
Firm Learning, Unemployment, and Self-Employment in Growth and Development



**Author**
Feng, Ying

**Publication Date**
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

# Firm Learning, Unemployment, and Self-Employment in Growth and Development

A dissertation submitted in partial satisfaction of the

requirements for the degree

Doctor of Philosophy

in

Economics

by

Ying Feng

Committee in charge:

Professor James Rauch, Chair
Professor David Lagakos, Co-Chair
Professor Ruixue Jia
Professor Tommaso Porzio
Professor Krislert Samphantharak

2019

The dissertation of Ying Feng is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California San Diego

2019

DEDICATION

To Zhigang Feng and Yuanfen Jiang.

EPIGRAPH

*Now this is not the end.*

*It is not even the beginning of the end.*

*But it is, perhaps, the end of the beginning.*

—Winston Churchill

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest appreciation to my advisors, Professor James Rauch, and Professor David Lagakos, for their patient guidance, insightful advice, continuous encouragement, and enormous support. I feel truly lucky and honored to have two great advisors, Jim and David. Their enthusiasm for research ignites my passion, and their rigorous attitude towards research shapes my principles.

I would like to thank my committee members, Professor Tommaso Porziom, Professor Ruixue Jia, and Professor Krislert Samphantharak, for their time and valuable suggestions on my research. I also thank all the faculty members at the Department of Economics and the School of Global Policy and Strategy, who have created and maintained the positive, supportive, and interactive professional environment.

My sincere gratitude also goes to my friends and fellow graduate students at UCSD for their continuous help, emotional support, and sincere friendship. I also thank the staff at the Department of Economics for their support and help during my time at UCSD.

Finally, I would like to dedicate this dissertation to my parents for their unconditional love and support throughout my life. I would not be able to finish the Odyssey without the support of my family.

Chapter 2, in part, is currently being prepared for submission for publication of the material. It is coauthored with David Lagakos and James E. Rauch. The dissertation author was a primary investigator and author this material.

Chapter 3 is coauthored with Lindsay Rickey. The dissertation author was the primary investigator and author of the unpublished material.

## VITA

| | |
|---|---|
| 2013 | B.A. in Economics, Wuhan University, China |
| 2013 | B.Sc. in Mathematics, Wuhan University, China |
| 2014 | M.A. in Economics, University of California San Diego |
| 2019 | Ph. D. in Economics, University of California San Diego |

## WORKING PAPERS

Ying Feng, "Firm Life-cycle Learning and Misallocation", UCSD working paper, 2019.

Ying Feng, David Lagakos, and James Rauch, "Unemployment and Development", NBER Working Paper No. 25171.

Ying Feng, and Lindsay Rickey, "Development and Selection into Necessity versus Opportunity Entrepreneurship", UCSD working paper, 2016.

Ying Feng, and James Rauch, "The Impact of Entrepreneurial Risk Aversion on Wages in General Equilibrium", NBER Working Paper No. 20992.

## FIELDS OF STUDY

Macroeconomics, Growth and Development

ABSTRACT OF THE DISSERTATION

**Firm Learning, Unemployment, and Self-Employment in Growth and Development**

by

Ying Feng

Doctor of Philosophy in Economics

University of California San Diego, 2019

Professor James Rauch, Chair

Professor David Lagakos, Co-Chair

Differences in average income levels across countries are vast. This dissertation investigates the interaction between heterogeneous firms or workers' decisions and economic growth.

Chapter 1 of this dissertation studies firm life-cycle learning and misallocation. Misallocation is one of the most prominent theories of Total Factor Productivity in recent years. Specifically, this study focuses on misallocation of resources across producers. Dispersion in marginal revenue products of capital (MRPK) across firms may lower aggregate productivity through misallocation. Using firm-level panel data from China, I document that MRPK dispersion decreases substantially with firm age, particularly before age five. Building on this fact, I provide a new interpretation of MRPK dispersion as firm life-cycle learning. I formalize this idea in a dynamic model, in which firms learn about their fundamental productivity as they age and choose capital inputs in a frictional market based on their priors. Within each cohort of firms, imprecise priors lead firms to differ in their ex-post MRPK even in the absence of firm-level distortions. As firms learn over time and adjust their capital stocks, possibly through exiting the market, dispersion in MRPK decreases. Quantitative analysis of the model shows that omitting firm life-cycle learning leads to sizable overestimation of the aggregate productivity losses from misallocation.

Chapter 2 of this dissertation asks: How does the average unemployment rate change with GDP per capita? This chapter draws on household survey data from countries of all income levels to measure how unemployment varies with income. We document that unemployment is increasing with GDP per capita. Furthermore, we show that this fact is accounted for almost entirely by low-educated workers, whose unemployment rates are strongly increasing in GDP per capita, rather than by high-educated workers, whose unemployment rates are not correlated with income. To interpret these facts, we build a model with workers of heterogeneous ability and two sectors: a traditional sector, in which self-employed workers produce output without reward for ability; and a modern sector, in which firms hire in frictional labor markets, and output increases with ability. Countries differ exogenously in the productivity level of the modern sector. The model predicts that as productivity rises, the traditional sector shrinks, as progressively less-able workers enter the modern sector, leading to a rise in overall unemployment and in the ratio of low-educated to high-educated unemployment rates. A calibrated version of the model accounts for some, but not all, of the cross-country patterns we document.

Chapter 3 of this dissertation proposes a universal division of different types of self-employment. It is well-known that self-employment rate declines with GDP per capita (Gollin, 2008). However, when dividing self-employment into employers and own-account workers (self-employed without employees), this paper documents that the labor share of employers increases with income levels, and the share of own-account workers decreases. Using household surveys from countries of all income levels, we show these facts are robust across main industries and educational categories. We also show nearly universal negative selection on ability into own-account status, and positive selection into employer and wage earning statuses in our data. We develop a simple two-sector model to explain these facts. In general equilibrium, agents with ability below a threshold become own-account workers in the traditional sector, and agents with ability higher than the threshold enter the modern sector, becoming wage workers or employers. Higher aggregate productivity is driven by higher returns to ability in the modern sector due to skill biased technological change, which reduces the threshold ability level. By distinguishing between own-account workers and employers consistently across 56 countries, our database and model help reconcile diverse findings about development and entrepreneurship.

# Chapter 1

# Firm Life-Cycle Learning and Misallocation

I am grateful to my advisors David Lagakos and James Rauch for their guidance and support. For helpful comments, I thank Jim Hamilton, Pete Klenow, Richard Rogerson, Tommaso Porzio, and seminar/conference audiences at Cal State Fullerton, CityUHK, HKU, ITAM, New Structural Economics Conference, NUS, SAIF, Tsinghua, and UCSD. I also benefit from numerous discussions with Emilien Gouin-Bonenfant and Xiao Ma. All potential errors are my own.

## 1.1 Introduction

Differences in average income levels across countries are vast. Development accounting points to differences in total factor productivity (TFP) as an important proximate cause of cross-country income differences ([36]). Yet the determinants of TFP are still not well understood. A prominent theory of TFP emphasized in the recent literature is misallocation. Two influential papers, [73] and [118], interpret dispersion in marginal revenue products of capital (MRPK) across firms as the result of firm-level ‘distortions’ that cause misallocation. They argue that misallocation leads to large TFP losses in developing countries. [13] provide evidence that size-dependent tax, a form of distortions, are more prevalent in low-income countries.

However, the literature is still very much undecided about how to interpret dispersion in MRPK across firms ([119]). A large body of work has provided alternative interpretations. For example, [10] emphasize the role of capital adjustment costs under volatility of productivity, and [45] emphasize the role of uncertainty in contemporaneous productivity. Both channels lead to MRPK dispersion but do not imply misallocation from distortions. [44] further develop a quantitative framework to decompose sources of MRPK dispersion and conclude that, while these channels are present, a large share of dispersion still results from firm-level distortions. An open question in the literature remains: What are the sources of MRPK dispersion?

This paper provides a new source of MRPK dispersion, building on a new pattern I document in the data. Following firm cohorts using firm-level panel data from China for the period 1998 - 2007, I document that MRPK dispersion across firms decreases substantially with firm age, particularly before age 5. The magnitude of this life-cycle decrease is similar to the difference in MRPK dispersion between China and the US as reported in [73]. Furthermore, for young firms, MRPK dispersion decreases at a decreasing rate with firm age.

Yet the challenge is that identifying the age effects separately without any additional assumptions is impossible, because the age, year, and cohort indicators are collinear. In particular, during my sample period, China experienced massive privatization reforms so that revenue share of the state-owned firms in the industrial sector declined by 20 percent ([77]). One can expect large year effects as China underwent such reforms and opened up to international trade. Hence, the decrease in MRPK dispersion over a firm cohort's life cycle could be the result of year effects, rather than age effects. Similarly, one can expect that each successive cohort may be founded with less MRPK dispersion across firms, as they entered the market more for economic reasons rather than political reasons. Thus, including controls for year effects and cohort effects is crucial in any reasonable attempt to identify the age effects on MRPK dispersion.

My preferred identification approach imposes the testable assumption of a linear trend in the age effects at older ages. For example, consider a special case of linear effects as no trend in the age effects on MRPK dispersion after firm age 10. Then year effects can be identified by following the same firm cohorts aging from age 10 because all the changes over time are only due to year effects in the absence of cohort and age effects. I can subsequently identify the age effects and cohort effects after knowing the year effects. Specifically, in the preferred approach, I imposed three plausible trends of age effects at older ages for identification. I also provide two alternative identification approaches in the paper. All three estimation results show negative age effects on dispersion in MRPK across firms. In particular, the estimated profile of the standard deviation of log MRPK within a firm cohort always decreases by more than 0.2 before age five, which accounts for 13% of the initial dispersion at firm entry.

Building on the facts I document, I provide a new interpretation of MRPK dispersion as resulting from firm life-cycle learning. It reflects informational frictions over the firm cohort's life cycle when firms learn about their own fundamental productivity, as in [79]. Within each cohort of firms, differences in the precision of priors lead firms to differ in their ex-post MRPK even in the absence of firm-level distortions. I formalize this idea in a dynamic model in which firms learn over time and choose capital inputs based on their priors in a frictional market with firm-specific distortions. Qualitatively, as priors of the firm cohort improve through learning over time, firms with too much or too little capital stock adjust, and the less productive firms within a cohort exit. Hence, the model predicts that MRPK dispersion within the firm cohort decreases as firms age.

The main quantitative experiment is to compute the model's predictions about MRPK dispersion within a firm cohort as the firms age. To do so, I take the joint distribution of productivity and capital stocks among firm entrants as given in the data. I calibrate the model to match three key moments in the data, namely, the exit rate of firm entrants, the correlation between productivity and capital investment, and the autocorrelation of capital investments. As a result, for the first ten years of the firm cohort's life cycle, the calibrated model accounts for around two thirds of the decrease in MRPK dispersion in the data.

To understand the quantitative role of learning, I decompose changes in MRPK dispersion over the firm cohort's life cycle by sequentially adding mechanisms in the model. If the firms adjust capital stocks without updating their priors and without exiting the market, MRPK dispersion barely decreases with firm age. If firms Bayesian update their priors while adjusting capital stocks, but still do not have the exit option, the dispersion in MRPK decreases around half as much as the benchmark model prediction. Further adding endogenous firm exit under the life-cycle learning accounts for the other half of the benchmark model prediction.

What, then, are the consequences of firm life-cycle learning for aggregate TFP, rather than for a firm cohort? Taking into account the firms' age distribution in the stationary equilibrium, I compare the benchmark model predictions to a hypothetical baseline where young firm cohorts had already completed their learning process as older firms. This comparison suggests that informational frictions from firm life-cycle learning lead to a 10 percent loss in aggregate TFP. I conduct the same analysis in the model after removing firm-level distortions, which suggests that distortions and firm life-cycle learning together result in a 19 percent loss in aggregate TFP. Therefore, omitting the contribution of firm life-cycle learning to MRPK dispersion causes more than half of TFP losses to be incorrectly attributed to distortions. I regard these estimates as lower bounds of TFP losses because the quantitative analysis assumes MRPK dispersion across firms remains constant after age 10.

Before concluding, I present plant-level panel data from Colombia and Chile for an earlier period (around the 1980s). I ask whether MRPK dispersion (measured by standard deviations of log MRPK) decreases with firm-cohort age. I find that, in both countries, MRPK dispersion decreases by around 0.4 through the first five years of the firm cohort's life cycle, which accounts for 29% and 24% of the initial dispersion across age-zero firms in Colombia and Chile, respectively. I conclude that data from other developing countries broadly show decreasing life-cycle MRPK dispersion, similar to the data from China.

**Related Literature.** Most existing work focuses on the aggregate level of MRPK dispersion across firms and does not consider its dynamics over the firm cohort's life cycle. For example, [98] and [90] study financial frictions, [81] combine financial frictions and adjustment costs to investigate MRPK dispersion across plants within the same firms, [72] explore the variation in markups and returns to scale, and [130] consider markup dispersion, adjustment costs, and measurement errors. In addition, all the models above are silent on endogenous firm entry and exit. [140] studies the effects of distortions on firm entry but assumes exogenous exit. [51] emphasizes that, in theory, endogenous exit may offset the effects of distortions on long-run TFP, but does not consider informational frictions. This paper is the first to look at life-cycle MRPK dispersion and the first to interpret MRPK dispersion as resulting from firm learning.[1]

[1] See [119] for an in-depth literature review on the causes and costs of misallocation. Other studies focus

This paper also relates to the literature in macroeconomics that makes cross-country comparisons of average firm sizes over firms' life cycles. [75] show that plants stay much smaller in Mexico and India than in the US over the plants' life cycles. [21] use data from more countries to argue that severe distortions in developing countries discourage investments, leading to smaller average firm sizes. [4] and [40] emphasize the importance of delegation frictions and lack of selection in explaining smaller average plant sizes over the plants' life cycles in developing countries. My results pertain to MRPK dispersion across firms rather than average firm size. The fact that the dispersion decreases with firm age implies considerable improvement in how efficiently resources are allocated across firms over their life cycles.

The idea of firm life-cycle learning is built on the classic model of [79]. By adding capital to his original model, I bring in frictional capital markets, including adjustment costs and fire-sale discounts upon exit. These frictions are important to match the key pattern of life-cycle MRPK dispersion within a firm cohort. In addition, other studies, for example, [11], emphasize that exits of low-productivity firms contribute to aggregate productivity growth. By focusing on MRPK dispersion across firms, this paper can draw further implications of the consequences of firm life-cycle learning and exit for aggregate productivity through reallocating resources across firms.

Finally, this paper adds to the vast literature on the theories of TFP, aiming at advancing our understanding of income differences across countries and across time. For example, [63] consider the macroeconomic implication on reductions in output of size-dependent policies. [33] quantify the role of financial frictions in economic development. [44] emphasize distortions accounts for a larger share of misallocation among Chinese manufacturing firms and adjustment costs are more salient for large US firms. [35] argue that resource misallocation has played a sizable role in slowing down Italian productivity growth. This paper points to the potentially

---

on misallocation over the business cycle: [8] consider the reallocation of products, and [128] emphasizes the rising uncertainty at the start of the Great Recession. [14] consider misallocation in an open economy with trade liberalization.

important role of firm life-cycle informational frictions and learning.

The rest of this paper is structured as follows. Section 1.2 describes the data from China and presents the features of life-cycle MRPK dispersion across firms without any additional assumptions. Section 1.3 reports the estimated profile of MRPK dispersion with firm age while controlling for cohort effects and year effects. Section 1.4 presents the model with learning and its qualitative predictions. Section 1.5 discusses the quantitative analysis. Section 1.6 provides evidence from Colombia and Chile. Section 1.7 concludes.

## 1.2 MRPK Dispersion across Firms over Their Life Cycles

In this section, I describe the data and present cross-sectional evidence on the pattern of MRPK dispersion over the firm cohort's life cycle. When tracking each firm cohort over time, I find that the dispersion in MRPK across firms always decreases with firm age. The younger cohorts tend to have smaller MRPK dispersion than older cohorts, and the aggregate MRPK dispersion in a year also decreases during my sample period (1998 to 2007). I also find that, for firm cohorts before age five, MRPK dispersion decreases at a decreasing rate.

### 1.2.1 Data and Measurement

I use the Annual Industry Surveys for 1998 - 2007 conducted by the National Bureau of Statistics of China. The survey covers all the state-owned firms in the manufacturing sector, as well as non-state-owned manufacturers with sales revenue above 5 million RMB (around 0.7 million USD). I follow the procedure used by [28] to construct the panel data. I start by matching the firms over time by registration ID. When firms changed their registration ID due to restructure or acquisition, I use company name, phone number, and address to identify the same firms. Note that ownership change will not cause false exits in the data, because those firms will still be identified over time through address and name. Throughout the paper, I focus on the firm cohorts founded after 1978, when the "opening-up reform" started. I drop firms founded in a planned economy before the economic reform because they may operate under very different systems. In addition, those firms cannot be observed at ages younger than 20, and are thus less relevant for studying the life cycles of firms. The remaining panel data have an average of around 180,000 firms per year, growing from 106,000 firms in 1998 to over 298,000 firms in 2007. In addition, I use the 4-digit Chinese Industry Code (CIC), birth year, wage, employee benefits, value-added, and capital stock. [2]

Let $i$ denote an individual firm. The firm age $j$ is calculated as the survey year minus the reported birth year. Therefore, the age-one firms are operating for a full year. Let $y_{it}$ denote the revenue output, $k_{it}$ the capital input, and $n_{it}$ the labor input. Then $y_{it}$ is measured as value added, $k_{it}$ is measured as the book value of fixed capital net of depreciation of the year, and $n_{it}$ is measured as the total of wage payments and employee benefits. The employee benefits include unemployment insurance, old care insurance, medical insurance, housing compensation, travailing compensation, and union expenses, but availability of the specific variable varies across years. Hence, I inflate the labor share to match those reported in the annual national accounts as [73] did. This procedure assumes the imputed values of missing benefits are a constant fraction of labor income.To summarize dispersion of the key variables over the firms' life cycles, Appendix Figure 1.14 plots the standard deviation of log value-added ($y_{it}$), log capital input ($k_{it}$), log labor input ($n_{it}$), and log employment by firm age. I find the dispersion of value added and labor input across firms increases with firm age until age 15, while dispersion of capital input and employment across firms increases very marginally with firm age.

[2]The share of firms younger than age 10 is around 72% in every year of my sample.

Let the production function be Cobb-Douglas $y_{it} = e^{z_{it}} k_{it}^{\alpha_1} n_{it}^{\alpha_2}$. I assume decreasing returns to scale, that is, $\alpha_1 + \alpha_2 < 1$. I also allow the capital and labor input share to vary across industries but not over time as in [73]. Following their work, I use the NBER-CES Manufacturing Industry Database to calculate $\alpha_1$ as the average values of capital share at 4-digit SIC level during the period 1987-2011, and then match them to CIC at the 2-digit level. In the empirical analysis, I set $\alpha_1 + \alpha_2$ to be the standard 0.85.[3] By definition, the MRPK of firm $i$ at time $t$ is $\frac{\partial y_{it}}{\partial k_{it}} = \alpha_1 \frac{y_{it}}{k_{it}}$. I measure total factor revenue productivity (TFPR), marginal revenue product of labor (MRPL), and MRPK in log terms throughout the paper:

$$tfpr_{it} = log(y_{it}) - \alpha_1 log(k_{it}) - \alpha_2 log(n_{it}) \tag{1.1}$$

$$mrpk_{it} = log(\alpha_1) + log(y_{it}) - log(k_{it}). \tag{1.2}$$

$$mrpl_{it} = log(\alpha_2) + log(y_{it}) - log(n_{it}). \tag{1.3}$$

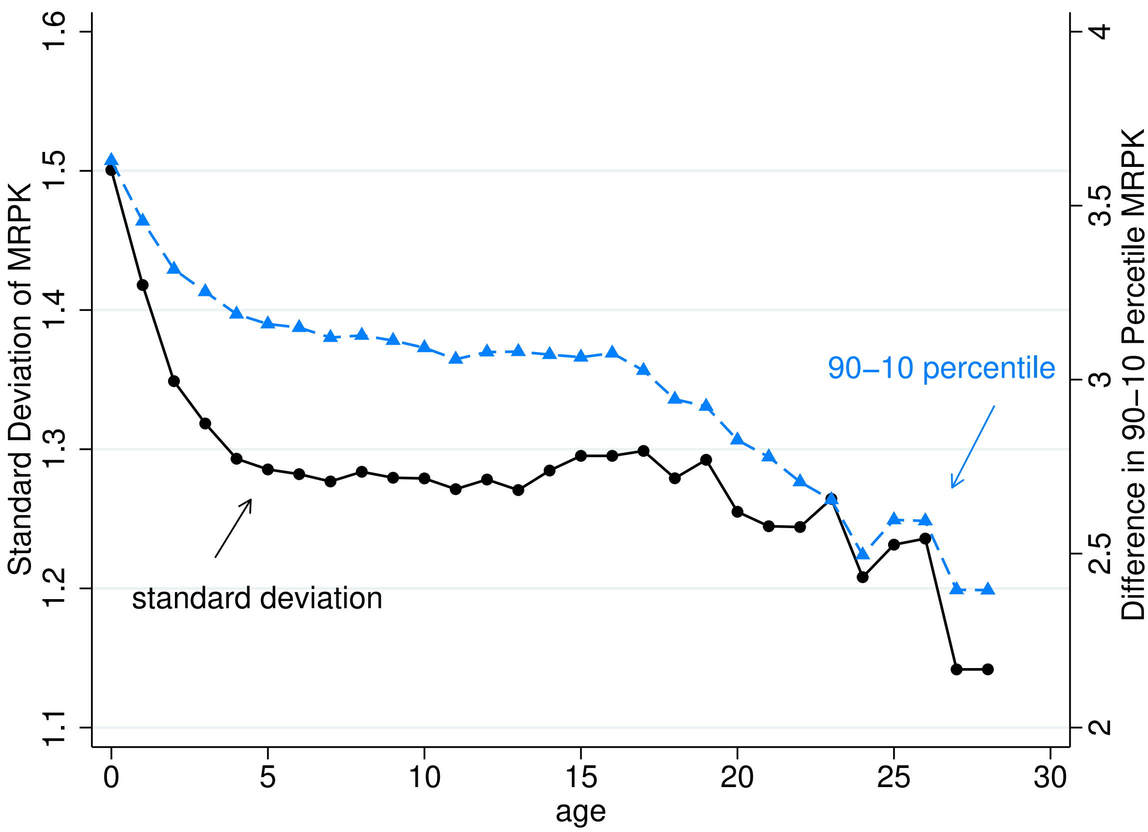
I drop the observations with missing values and trim the 1% tails of measured MRPK or TFPR in each industry-year-age group. The remaining data have an average of around 169,000 firms per year, consisting of more than 480,000 unique firms recorded during the sample period. Around 48% of the unique firms survived for at least four years.

### 1.2.2 Dispersion of Marginal Products by Firm Age

Consider an industry-year-age bin, denoted as $stj$, consisting of firms observed in calendar year $t$ at firm age $j$ in the 4-digit industry $s$. To measure MRPK dispersion within a $stj$ bin, I use the standard deviation of $mrpk_{it}$, denoted as $\sigma_{mrpk,stj}$, and the 90th minus the 10th percentile

[3] I conduct the same analysis assuming constant returns to scale as in [73], letting $\alpha_2 = 1 - \alpha_1$, and get similar results.

$mrpk_{it}$, denoted as $D^{90-10}_{mrpk,stj}$. Note that, by construction, MRPK dispersion across firms within an industry-year-age bin is always measured within the same firm cohort $c$ of firms, which are founded in year $t-j$. When summarizing MRPK dispersion at a give age or in a given year, I will always weight $\sigma_{mrpk,stj}$ by $N_{stj}$, the number of firms in an industry-year-age bin.




Note: This figure plots the weighted average standard deviation of MRPK ($\bar{\sigma}_j$) and the weighted average value of the 90th minus the 10th percentile ($\bar{D}^{90-10}_j$) by firm-cohort age.

**Figure 1.1**: Dispersion of MRPK by Firm Age

Define the weighted average dispersion in MRPK at a given firm-cohort age $j$ in both measures

$$\bar{\sigma}_j \equiv \sum_s \sum_t \sigma_{mrpk,stj} \cdot \omega_{st}$$

$$\bar{D}^{90-10}_j \equiv \sum_s \sum_t D^{90-10}_{mrpk,stj} \cdot \omega_{st},$$

where the weight $\omega_{st} = \frac{N_{stj}}{\sum_s \sum_t N_{stj}}$. Figure 1.1 then plots $\bar{\sigma}_j$ (black solid line with round markers) and $\bar{D}_j^{90-10}$ (blue dashed line with triangle markers) by firm-cohort age. The weighted average standard deviation of MRPK, $\bar{\sigma}_j$, deceases substantially from 1.5 by almost 0.4 points by age 28. Note that this change is huge, and has the same magnitude as the difference between China and the US reported in [73].[4] The plotted average log differences, $\bar{D}_j^{90-10}$, show the average ratio of the 90th to the 10th percentile MRPK decreases from more than 33 ($e^{3.5}$) to only 12 ($e^{2.5}$) as the firm cohort ages from zero to around 25. Similar to the patterns of MRPK standard deviations, the 90-10 percentile difference in MRPK decreases substantially with firm age. I will focus on the standard deviation of MRPK ($\sigma_{mrpk,stj}$) as the dispersion measure for the rest of the paper because it has been used more broadly in the literature. More importantly, I will show later that $\sigma_{mrpk,stj}$ translates directly to TFP losses.[5]

Meanwhile, how does the dispersion of MRPL across firms change as firms age? Appendix Figure 1.15 plots the weighted average standard deviation of MRPL and the weighted average value of the 90th minus the 10th percentile MRPL by firm age. Both measures of dispersion in MRPL decrease very marginally before age 5, and they increase slightly afterward. Therefore, this paper focuses on MRPK dispersion and abstracts from the discussion of MRPL dispersion.

A natural explanation is that, MRPK dispersion decreases with firm age because the less productive firms within a cohort learn about their type and exit gradually over their life cycles. This implies that TFPR dispersion also decreases at a decreasing rate with firm age. As learning and the selection in firm exits becomes less pronounced, the dispersion in both MRPK and TFPR does not decrease further. Denote the weighted average standard deviation of TFPR at age $j$ as $\bar{\sigma}_{tfpr,j}$. Appendix Figure 1.16 shows that $\bar{\sigma}_{tfpr,j}$ decreases at a decreasing rate from around 0.99

[4] Table 2 in their paper reports the difference of 0.14 in the standard deviation of $tfpq$ between China and the US in 2005. It implies a MRPK dispersion difference of 0.4 points in their model under the standard capital share of 0.3.

[5] As an alternative measure of dispersion in MRPK, the average ratio of the 75th to the 25th percentile MRPK decreases monotonically from 6 to 4 as the firm cohort ages from zero to around 25.

at entry to around 0.9 at age five, and fluctuates between 0.85 and 0.95 afterward. Alternatively, if one thinks of the production process as the less productive firms catching up with the most productive firms due to stable innovation investments or spillover effects, one should expect TFPR dispersion to continuously decrease at older ages, which is not observed in the data.[6] Furthermore, Appendix Figure 1.17 plots MRPK dispersion for the balance panel, which consists of firms that I can observe every year during the sample period of 1998-2007. It shows that, when firm exits are shut down, MRPK dispersion decreases with firm age with a smaller magnitude.

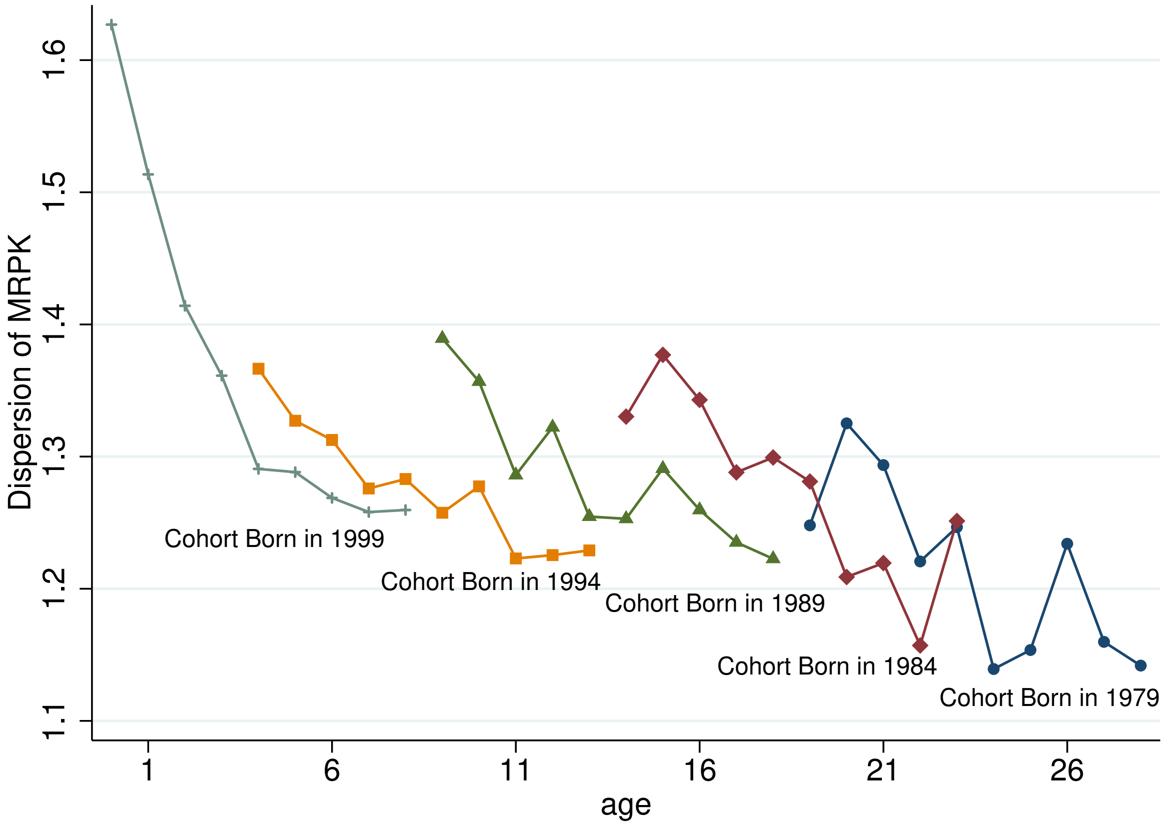
However, the summary statistics by firm-cohort age presented above is the result of a combination of age, cohort, and year effects. In a fast-changing economy like China, one may expect large variations across firm cohorts born in different years. For example, as China moves from an economy of state-owned enterprises to one with mostly private enterprises, each successive cohort of firms may be founded with a smaller dispersion in MRPK. To investigate the life-cycle pattern within each firm cohort $c = t - j$, define the weighted average dispersion across firms at age $j$ as

$$\bar{\sigma}_{jc_{t-j}} \equiv \sum_{s} \sigma_{mrpk,stj} \cdot \omega_{sc},$$

where the weight $\omega_{sc} = \frac{N_{stj}}{\sum_s N_{stj}}$.

Figure 1.2 then plots $\bar{\sigma}_{j,c_{t-j}}$ against firm-cohort age by following each cohort born in 1979, 1984, 1989, 1994, and 1999, respectively. For all the cohorts, we see a general decrease in MRPK dispersion with firm-cohort age. The older firm cohorts tend to have a larger dispersion in MRPK than the younger cohorts at the same ages. In particular, for the firm cohort founded in 1999, MRPK dispersion declines by almost 0.4 from age zero to age eight. Furthermore, Figure 1.3 plots MRPK dispersion for the nine firm cohorts founded between 1998 and 2006, which can be tracked from age 0 at the entry year. Similar to the 1999 firm cohort in Figure 1.2, the

[6]Note that $\bar{\sigma}_{tfpr,j}$ is at a lower scale than $\bar{\sigma}_{mrpk,j}$, due to a large dispersion in $k_{it}$, which is not offset by the empirical correlation between $y_{it}$ and $k_{it}$, thus being reflected in $\bar{\sigma}_{mrpk,j}$.



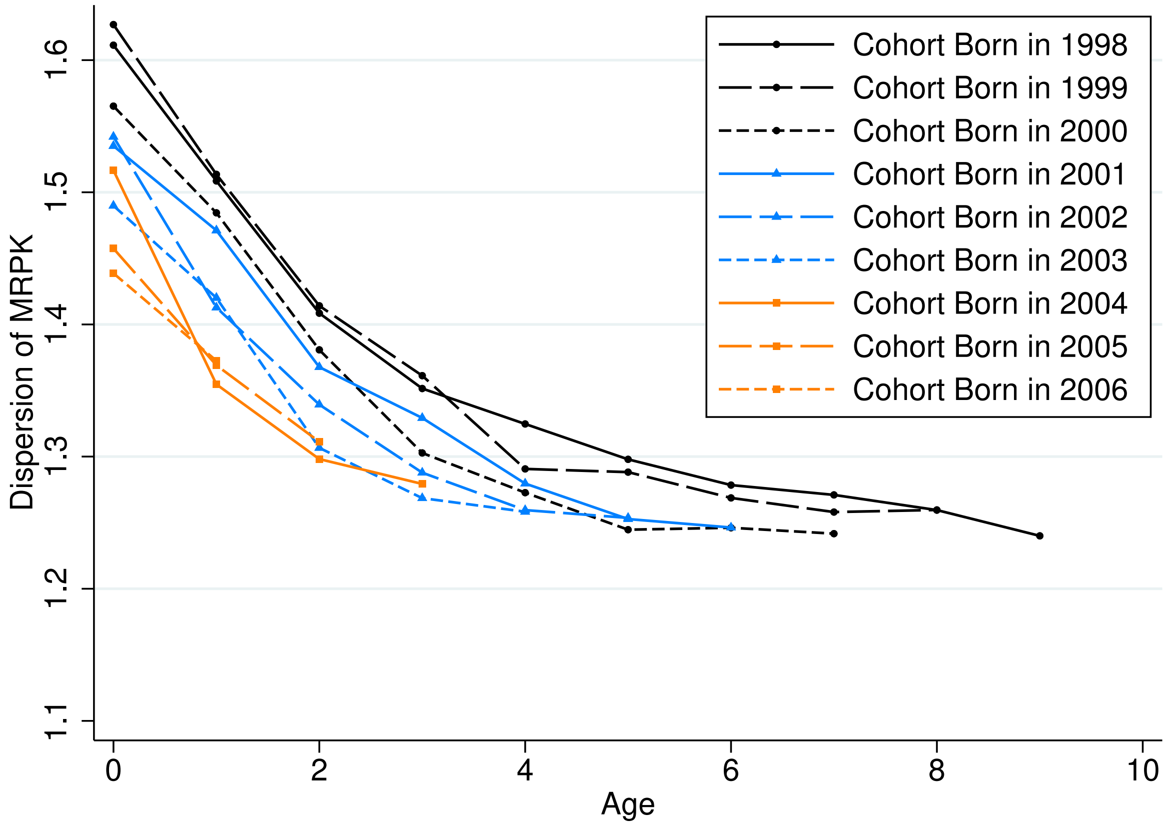

Note: This figure plots the weighted average dispersion, $\bar{\sigma}_{j,c_{t-j}}$, by firm age within each cohort born in year 1979, 1984, 1989, 1994, and 1999, respectively.

**Figure 1.2**: Dispersion of MRPK by Firm Age and Cohort

decline in MRPK dispersion across firms is substantial through the first five years of the firms' life cycles for all nine cohorts. In addition, the older cohorts among the nine again tend to have a larger MRPK dispersion, as we see the black lines are above the blue and the blue are above the orange.[7]

Similarly, it is difficult to imagine the Chinese economy with little year effects as the privatization reforms deepen over time. Define the weighted average dispersion in a given year

[7]Appendix Figure 1.21 plots the exit rates by firm-cohort age of the same firm cohorts in Figure 1.2 after removing zero-sum year effects using the identification approach in Section 1.3.1. Patterns of the exit rates resemble those of the MRPK dispersion.




Note: This figure plots the weighted average dispersion, $\bar{\sigma}_{j,c_{t-j}}$, by firm age within each of the nine cohorts born during 1998 to 2006.

**Figure 1.3**: Dispersion of MRPK by Firm Age, Young Cohorts

$t$ as $\bar{\sigma}_t \equiv \sum_s \sum_j \sigma_{mrpk,stj} \cdot \omega_{sj}$ , where the weight $\omega_{sj} = \frac{N_{stj}}{\sum_s \sum_j N_{stj}}$. Appendix Figure 1.18 plots $\bar{\sigma}_t$ during my sample period. The average aggregate dispersion in MRPK decreases from 1.4 in 1998 to less than 1.3 in 2007.

In summary, by following each firm cohort over time, I find that MRPK dispersion decreases substantially with firm-cohort age. However, the decrease is not necessarily the result of age effects, because it reflects both age effects and year effects. Instead, the substantial decrease could be the result of potentially sizable year effects, because China underwent its reforms and opened up over time. Therefore, in any reasonable attempt to identify age effects, controlling for cohort effects and year effects is crucial.

### 1.2.3 Industry Variations

In different industries, MRPK dispersion decreases at different rates with respect to firm-cohort age. I use this variation to shed light on the mechanisms of decline in MRPK dispersion over the firm cohort's life cycle. Define the weighted average dispersion in MRPK across firms at age $j$ in industry $s$ as $\bar{\sigma}_{sj} \equiv \sum_t \sigma_{mrpk,stj} \cdot \omega_t$, where the weight $\omega_t = \frac{N_{stj}}{\sum_t N_{stj}}$. I investigate how the correlation between $\bar{\sigma}_{sj}$ and firm age $j$ varies across industries, and discuss how it relates to the industry characteristics.

Within each industry, I use $\kappa_s$ in the linear model below to summarize MRPK dispersion over the firm cohort's life cycle:

$$\sigma_{mrpk,stj} = \kappa_{0s} + \kappa_s age_{stj} + \varepsilon_{stj}. \tag{1.4}$$

Estimate $\widehat{\kappa_s}$ describes how $\bar{\sigma}_{sj}$ changes with firm-cohort age under the linear specification. The average value of $\widehat{\kappa_s}$ across industries is -.014, with a standard deviation of 0.026. Therefore, in the majority of industries, MRPK dispersion decreases with firm age, and on average, $\bar{\sigma}_{sj}$ decreases more than 1% per age.

To utilize the standard industry-level characteristics commonly used in the trade literature, I mapped the 4-digit 2003 CIC to the 6-digit US Input-Output classification, and used the industry indexes from [7]. Table 1.1 reports the results of regressing estimated $\widehat{\kappa_s}$ on various industry characteristics with bootstrap standard errors. It shows that when the capital-labor ratio increases by 1 log point, the decrease of $\sigma_{mrpk,stj}$ per age is 0.005 points larger, which is more than one third of the average value 0.014 across all industries. The significant positive coefficient on log

**Table 1.1**: Industry Characteristics and Life-Cycle $\sigma_{mrpk,stj}$

| $\widehat{\kappa_s}$ | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log(Capital Per Worker) | $-.005^{***}$ | $-.005^{*}$ | $-.006^{**}$ | $-.006^{**}$ |
| | (.002) | (.002) | (.003) | (.003) |
| Log(R&D/Sales) | | $-.002^{**}$ | $-.002^{**}$ | $-.002^{*}$ |
| | | (.001) | (.0008) | (.0009) |
| Contractibility | | | -.006 | $-.006^{**}$ |
| | | | (.005) | (.003) |
| Financial Dependence | | | -.004 | .003 |
| | | | (.003) | (.016) |
| Input Substitutability | | | -.00004 | -.00004 |
| | | | (.0002) | (.0003) |
| Log(Capital per worker)*Financial Dependence | | | | -.002 |
| | | | | (.004) |
| Obs. | 423 | 408 | 408 | 408 |
| $R^2$ | .015 | .034 | .044 | .044 |

Note: ***, **, and * indicate statistical significance at the 1-percent, 5-percent, and 10-percent levels, respectively. Independent variable Log(Capital per worker) is the industry average calculated during sample period 1998 - 2007 in China. The source of the industry indexes is [7]: industry-level Log(R&D/Sales) is from Nunn-Trefler (US, 2000-2005); upstream Contractibility from Nunn (2007) based on liberal classification; Financial Dependence is measured as the External Capital Dependence from Rajan-Zingales (1997) calculated using 1980s Compustat data. Input Substitutability is measured as the Import Demand Elasticity (based on SITC33).

capital per worker suggests that industries with higher capital shares are better at decreasing MRPK dispersion over the firm cohort's life cycle. One possible explanation could be that the costs related to capital, such as storage and maintenance costs or adjustment costs, push firms to adjust capital more responsively, reallocating resources to the more productive incumbent firms. Meanwhile, industries that larger innovation expenditure shares are better at decreasing MRPK dispersion over firm-cohort age. Additionally, a slightly significant positive correlation exists between resource reallocation and contractibility: Industries in which it is easier to contract the sales of their capital inputs at less discounted values upon exit also experience a faster decrease in MRPK dispersion with firm-cohort age. I will build these ideas formally in the dynamic firm model in section 1.4 to assess their explanatory power.

## 1.3 Life-Cycle MRPK Dispersion: Controlling for Cohort and Time

The previous section reports the average MRPK dispersion by firm-cohort age simply. Although understanding the data with minimal structure and assumptions is useful, this exercise does not address certain issues. The identification of first-order age effects is a well-known challenge due to the collinearity between age, year, and cohort indicators. In this section, I address the identification issues.

Though Figures 1.2 and 1.3 track the same firm cohorts over time and find a consistent trend of decreasing dispersion with firm-cohort age, they still leave open the possibility that the trend is driven by year effects rather than age effects. For instance, one may expect large negative year effects as China deepens its privatization reforms and shuts down the inefficient state-owned enterprises. The year effects, which are cohort-neutral, could lead to decreasing life-cycle MRPK dispersion within every cohort. In this scenario, year effects lead to a spurious relationship between MRPK dispersion and firm-cohort age for all the cohorts.

The goal of this section is to estimate flexible versions of the MRPK dispersion profile with firm age. The specifications take the following form:

$$\sigma_{mrpk,stj} = \alpha_0 + \sum_{j \in J} \phi_j D_j + \chi_c + \psi_t + \theta_s + \varepsilon_{stj}. \tag{1.5}$$

$D_j$ is a dummy equal to 1 if firms in the industry-year-age bin $stj$ are observed at age $j$. $\psi_t$ captures year fixed effects, $\chi_c$ captures cohort fixed effects, $\theta_s$ captures industry fixed effects, and $\varepsilon_{stj}$ is a mean-zero error term.

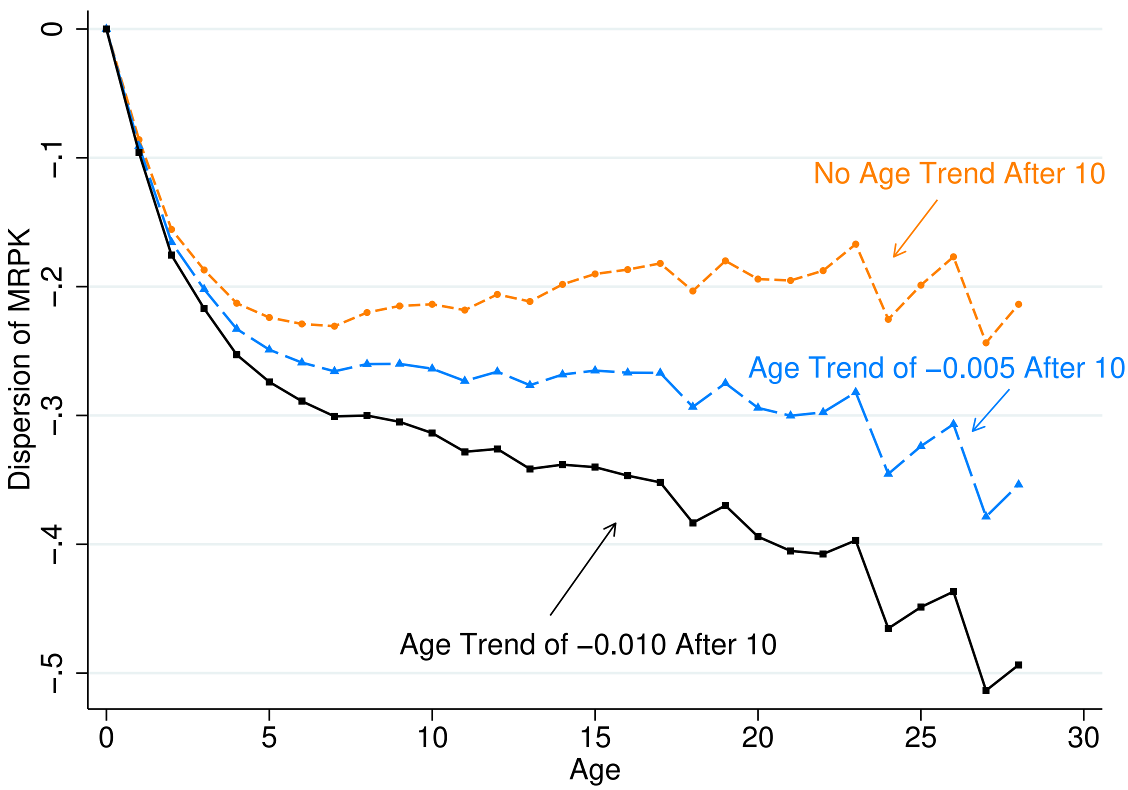
### 1.3.1 Three Approaches to Identifying Age Effects

The main challenge to estimat the age effects on MRPK dispersion is that age indicators are correlated with cohort indicators and year indicators. Therefore, identifying the age effects separately without any additional assumption is impossible. This section uses three approaches to identify age effects while controlling for cohort and year effects.

To resolve the difficulty of collinearity, I follow [46] and imposes one additional linear restriction on the set of cohort and time effects to estimate equation (1.5). Consider the decrease in aggregate MRPK dispersion over time, as plotted in Figure 1.18, which reflects the combined result of cohort-neutral year effects and effects of changes in the composition of firm cohorts in a calendar year. To identify age effects, I need to discipline the relative role of year effects and cohort effects in the decrease in aggregate MRPK dispersion over time.

**Preferred Approach.** My preferred identification approach assumes a linear trend in age effects on MRPK dispersion after age 10. For example, consider the assumption of no trend in the age effects after firm-cohort age 10 as a special case of the linear effects. Then year effects can be identified by following the firm cohorts older than age 10 because all the changes over time are only due to year effects in the absence of cohort and age effects. This assumption is actually also in accord with the empirical findings of [66]: Mature firms in the US have stable dynamics compared to the younger firms.

Furthermore, I can also test this assumption, because the second derivatives of age effects, which inform the curvature of age effects, are always identified as shown in [95]. I find the age effects are convex for the young firms, meaning MRPK dispersion decreases at a decreasing rate with firm cohort age. In addition, I cannot reject the null hypothesis that the age effects on MRPK dispersion are linear after age four. I describe the econometric details in section 1.9. The test results provide econometric foundations for identifying first-order age effects by assuming a linear trend in the age effects at older ages.




Note: This figure plots the estimated profile of MRPK dispersion by firm-cohort age in equation (1.6) using the second identification approach, which assumes (a) no trend in the age effects on MRPK dispersion after age 10 (dashed orange line with circle markers); (b) a small decreasing trend of 0.005 points per age after age 10 (long-dashed blue lines with triangle markers); (c) a moderate decreasing trend of 0.01 points per age after firm age 10 (solid black lines with square markers).

**Figure 1.4**: Estimated MRPK Dispersion by Firm Age, Preferred Approach

Figure 1.4 plots the estimated profile of MRPK dispersion by firm-cohort age using the second identification approach. In particular, I impose three different plausible magnitudes of the linear trend in the age effects on MRPK dispersion after firm-cohort age 10: (a) no trend in the age effects after age 10 (dashed orange line with circle markers); (b) a small decreasing trend of

0.005 points per age after age 10 (long-dashed blue lines with triangle markers); (c) a moderate decreasing trend of 0.010 points per age after age 10 (solid black lines with square markers). The three different trends imposed on age effects after firm-cohort age 10 all yield a substantial and convex decline of dispersion in MRPK for the young firms. In particular, MRPK dispersion decreases 0.22 to 0.25 point before age five.

**Alternative Approach One.** Instead of picking a plausible magnitude of the trend in age effects, the alternative identification approach estimate it by imposing the assumption that two consecutive old cohorts are the same, as in [64]. In the context of my sample between 1998 and 2007 in China, this assumption is based on the background that old firm cohorts founded in the late 1970s are similar because they were founded at the beginning of the economic reform and were adapting gradually, whereas the young firm cohorts could be drastically different because they are founded in different years in the fast-changing economy as China deepened its privatization reforms and largely opened up. The assumption in this approach is also a relaxed constraint of the assumption in [64] (p.248), who assumed all vintages have the same cohort effects to identify the age effects on the prices of used trucks.

This assumption identifies the slope of the linear trend in age effects by observing the old cohorts in the same years. Consider the two consecutive old firm cohorts founded in 1979 and 1980, both observed in the year 1998: one at age 18 and the other at age 19. Because they are observed in the same year, there is no difference in the year effects. The cohorts effects are the same as well under the assumption; hence, the difference in MRPK dispersion is only due to the different age effects at age 18 and age 19. In total, they are observed for ages 19-28 and 18-27, respectively, during 1998-2007. The average difference across years then gives the least-squares estimate of the age effect per year. In addition, by following all firm cohorts over time, this assumption can now help identify the trend in year effects given the age trend.

I implement this approach by estimating equation (1.5) in the framework as described in section 1.11. In practice, I assume every two adjacent cohorts founded during 1979 to 1983 are the same. They are observed from age 15 to 28. For each of the four pairs of adjacent cohorts, I calculate the difference in MRPK dispersion in each year between 1998 and 2007. Then I take the average of the differences of the four pairs as the trend in age effects after age 15, which turns out to be .009. I force the cohort effects of firms born in 1979 - 1983 to be the same.

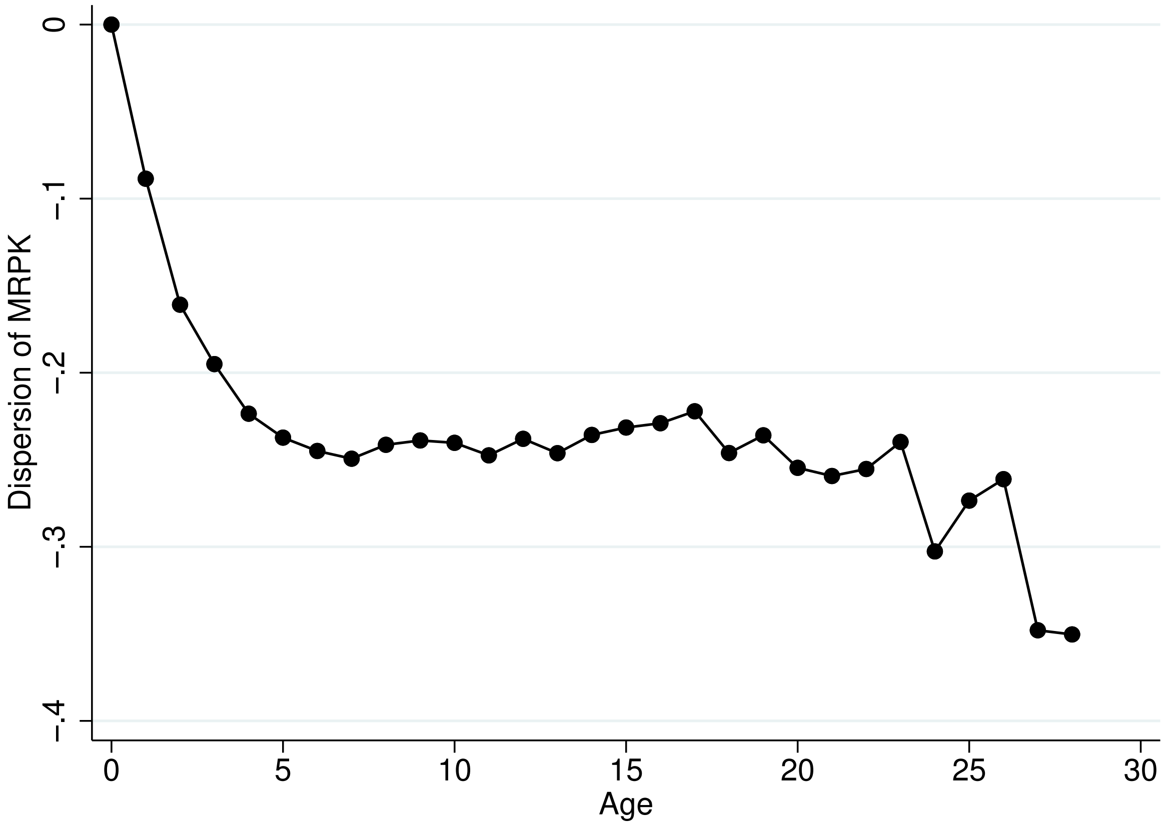
Figure 1.5 plots the estimated profile of MRPK dispersion by firm-cohort age using this approach: MRPK dispersion decreases substantially through the first five years of a firm cohort's life cycle, accumulating to more than -0.24 points. It further decreases after age five, though at a slower rate, accumulating to -0.35 points at age 28 compared to age zero.

**Alternative Approach Two.** This approach makes econometric assumptions to split the decreasing trend of dispersion over time between time effects and cohort effects as in [46], and does not make assumptions on the curvature of age effects. This approach also illustrates the econometric difficulty in disentangling the three effects.[8]

In practice, I implement two ways to split the decline in aggregate MRPK dispersion over time: One version attributes all the decline to cohort effects, and the other version attributes all the decline to cohort-neutral year effects. I show in Appendix section 1.10 that the two restrictions provide the lower and upper bounds of age effects if all three effects of age, year, and cohort on MRPK dispersion have non-positive trends. The condition of a non-negative trend in all three effects is a plausible case because the patterns of MRPK dispersion decrease with firm age, calendar year, and the birthyear of a firm cohort, as shown in section 1.2.2.

Specifically, I estimate equation (1.5) under restrictions. The first version attributes all the

[8]This methodology is commonly used in the literature on individuals' life-cycle consumption and income dynamics (e.g., **?**), and was recently used for firms in [102] and [8].




Note: This figure plots the estimated profile of MRPK dispersion by firm-chort age when assuming the two adjacent firm cohorts founded in 1979 - 1983 have the same cohort effects.

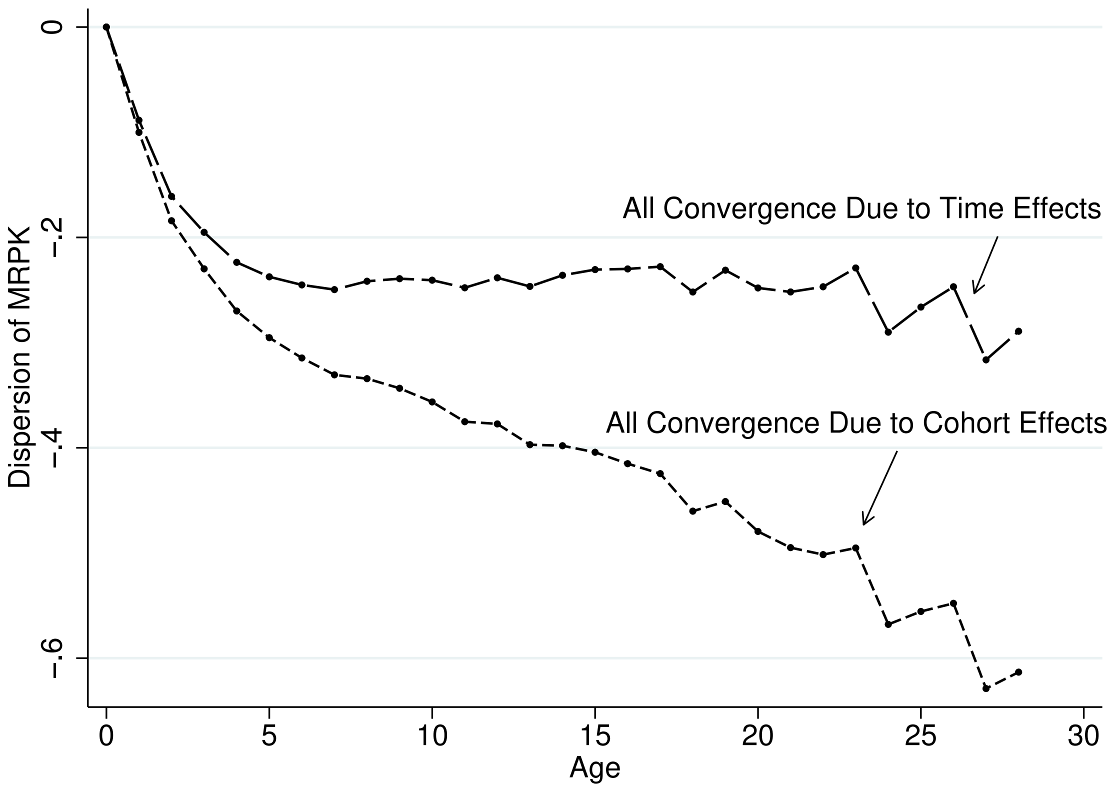
**Figure 1.5**: Estimated MRPK Dispersion by Firm Age, Alternative Approach One

decline over time to cohort effects. It makes the same assumption as in the original analysis in [46] and uses year dummies to capture only cyclical fluctuations. In practice, the first age dummy and the first cohort dummy are omitted as the benchmark reference, and the time dummies are transformed to meet the restriction that the time effects are orthogonal to a time trend. The second version is the opposite extreme case and attributes all decline to time effects. In this version, I assume the cohort effects are orthogonal to a time trend. See Appendix 1.11 for a formal description of this approach and the details of implementing it.

The long-dashed line in Figure 1.6 plots MRPK dispersion with firm-cohort age estimated under the assumption in the first version that all decline is driven by cohort effects. It provides the lower bound of the age effects, if all three effects of age, cohort, and year have non-positive trends, as shown in Appendix section 1.10. In this version, MRPK dispersion decreases substantially with firm age, accumulating to a magnitude of 0.6 points at age 28. The dashed line in Figure 1.6 plots the profile of MRPK dispersion estimated in the second version, where I assume all the decline over time is driven by year effects. In the second version, MRPK dispersion again decreases with firm age; note the decrease is most substantial before age five and flattens out afterward. This version also provides the upper bounds of age effects, as shown in Appendix section 1.10.

In conclusion, the first alternative approach shows MRPK dispersion decreases substantially with firm age, both in the lower and upper bounds. In particular, MRPK dispersion within a firm cohort decreases more than 0.04 points per age on average until age five, though the slope of MRPK dispersion at older ages is sensitive to the restrictions used for identification. When I attribute all the decline in MRPK dispersion over time to year effects, the estimated profile of MRPK dispersion closely resembles the results in the preferred approach.

In summary, although identifying the first-order age effects directly without any additional assumptions is impossible, the three identification approaches in this section establish a substantial decrease in MRPK dispersion with firm-cohort age. The estimation result of the three cases in the preferred approach lies between the upper and lower bounds (estimated in alternative approach two). The result from alternative approach one is also consistent with the upper and lower bounds, with estimates closer to the lower bounds. All three approaches conclude a substantial decrease in MRPK dispersion within the firm cohorts at young ages, accumulating to a magnitude of 0.2 to 0.3 by age five. In addition, the estimated profiles of MRPK dispersion through the first five years of the firm cohort's life cycle are convex, as McKenzie tests predict. The age effects on MRPK dispersion after age five are generally negative across the three approaches, though the magnitude




Note: This figure plots the estimated MRPK dispersion by firm-cohort age using the first approach. The long-dashed line plots MRPK dispersion by firm-cohort age estimated using equation (1.5), under the assumption that all the decline in MRPK dispersion over time is driven by cohort effects. The dashed line plots MRPK dispersion by firm cohort age estimated using equation (1.5), under the assumption that all the decline over time is driven by year effects. See Appendix 1.11 for a detailed description of implementing this methodology.

**Figure 1.6**: Estimated MRPK Dispersion by Firm Age, Alternative Approach One

is sensitive to the specific identification assumptions.

### 1.3.2 Robustness

This section assesses the robustness of the fact I document that MRPK dispersion decreases substantially with firm-cohort age. In particular, I consider other plausible factors that can affect the profile of life-cycle MRPK dispersion within a firm cohort: exit selection, time-series volatility of productivity, firm ownership, financial frictions, firm size, and measurement errors.

### Controlling for the Volatility of Productivity

In this section, I investigate the life-cycle MRPK dispersion after controlling for the volatility of productivity at each age of the firm cohort. [10] argue the dispersion in MRPK can largely be explained by capital adjustment costs in an environment with productivity volatility, where firms choose the capital stocks in the current period, taking into consideration that the volatility of productivity in the future, thus resulting in ex-post static MRPK dispersion in the current period. If volatility in productivity decreases as the firm cohort ages and matures, the older firms will tend to have less dispersion in the ex-post MRPK than the younger firms. In this case, MRPK dispersion may decrease over the firms' life cycles due to decreasing productivity volatility with firm-cohort age. Therefore, not controlling for the volatility of productivity could overstate the magnitude of negative age effects.

I define the time-series productivity volatility to be $\sigma_{\Delta z,stj}$, as in [10], which measures the standard deviation of productivity changes, $(z_{it} - z_{it-1})$, from one period to the next. The index $stj$ indicates the standard deviation is taken across firms within the same industry-year-age bin. Adding $\sigma_{\Delta z,stj}$ as a control variable in equation (1.5), I use the second identification approach to estimate

$$\sigma_{mrpk,stj} = \alpha_0 + \alpha_{vol} \cdot \sigma_{\Delta z,stj} + \sum_{j \in J} \phi_j D_j + \theta_s + \chi_c + \psi_t + \varepsilon_{stj}.^{9} \tag{1.6}$$

Consistent with [10], I find higher productivity volatility is correlated with a higher level of dispersion in marginal capital products. A one unit increase in the volatility of productivity

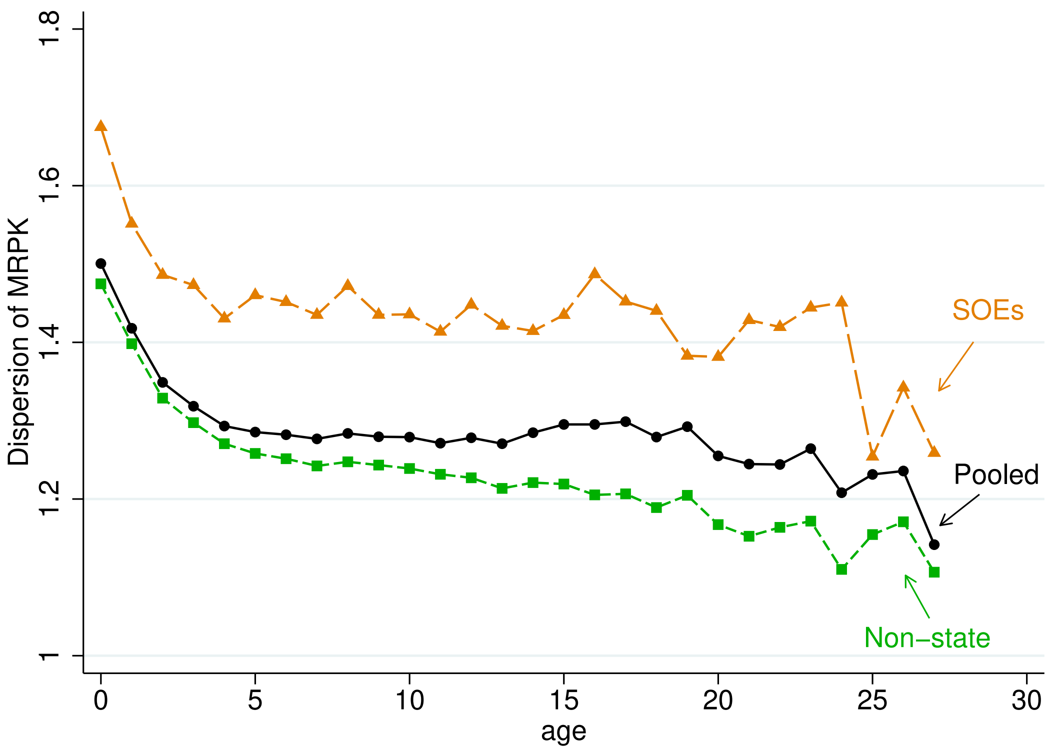
[9] The third approach, which assumes two adjacent old cohorts have the same cohort effects, becomes less straightforward here, because the two cohorts have different volatility of productivity even when observing in the same year.

predicts a 0.32-point increase in the cross-sectional MRPK dispersion, significant at the 1-percent level. The estimated profile of MRPK dispersion, after controlling for volatility of productivity, also decreases with firm age. As before, the decrease is most substantial through the first few years of a firm cohort's life cycle, though at a slightly smaller magnitude, accumulating to -0.18 to -0.22 points by age five compared to entry. I conclude that the decrease in MRPK dispersion with firm-cohort age cannot be explained by declining volatility of productivity as a firm cohort ages. See the estimation results plotted in Appendix Figure 1.22.

**State-owned and Non-state Firms**

The misallocation of capital between the state-owned and the non-state-owned firms is a salient feature in the Chinese economy (see, e.g., [27, 14, 26]). One may worry that the life-cycle production of state-owned enterprises (SOEs) in China responds largely to government policies and do not reflect the market outcomes. The SOEs might drive the pattern of MRPK dispersion over the firm cohort's life cycle. This section reports the life-cycle MRPK dispersion by firm ownership.

I define the firm as a SOE if more than half of its assets is owned by the state, and define the firm as a non-state firm otherwise. Figure 1.7 plots the weighted average MRPK dispersion for the SOEs, the non-state firms, and the pooled aggregate sample. The life-cycle MRPK dispersion of the non-state firms closely resembles that of the pooled sample, which decreases from 1.5 to around 1.2 between entry and age 27. Dispersion in MRPK across SOEs within a cohort also decreases with firm age. In addition, it constantly remains at a higher level than the dispersion among non-state firms, which may reflect larger informational frictions or less learning among SOEs. I conclude that MRPK dispersion robustly decreases with firm-cohort age, both for the SOEs and the non-state firms.




Note: This figure plots the weighted average MRPK dispersion for the SOEs, the non-state firms, and the pooled sample.

**Figure 1.7**: MRPK Dispersion by Firm Ownership

In addition, Table 1.2 reports the difference in the dispersion of MRPK at an older age relative to entry, for the full sample and for only the non-state firms. The t-test results of the equal means show that all the differences are strongly significant. For the full sample, the dispersion in MRPK decreases by 0.4 points until age 27. This decline through entry to age 27 is slightly larger for the non-state firms. Both for the full sample and for the non-state firms, MRPK dispersion decreases substantially before age five. In particular, it drops by almost 0.2 points until age five compared to entry, which accounts for around half of the decrease in MRPK dispersion during firm entry to age 27.

To further identify age effects, I estimate the dispersion in MRPK with firm-cohort age after restricting the sample to only non-state firms, using the second identification approach

**Table 1.2**: Dispersion of MRPK by Firm Age

| $\bar{\sigma}_j$ | Full Sample | Non-state Firms |
|---|---|---|
| Dispersion at Age 1 | -.04*** (.01) | -.03** (.01) |
| Dispersion at Age 5 | -.18*** (.01) | -.19*** (.01) |
| Dispersion at Age 10 | -.23*** (.01) | -.25*** (.01) |
| Dispersion at Age 20 | -.30*** (.02) | -.37*** (.02) |
| Dispersion at Age 27 | -.41*** (.04) | -.44*** (.04) |

Note: This table reports $\bar{\sigma}_j$, the average standard deviation of MRPK at a given age, in the data compared to entry with the estimate of standard error in parentheses. Row 1 uses the full sample, and Row 2 uses only the non-state firms.

and controlling for volatility of productivity. It yields the same coefficient on volatility (0.32), significant at 1%, as the full sample. Dispersion in MRPK across non-state firms decreases 0.20 to 0.25 points by age five compared to age zero, which has a slightly larger magnitude than the estimates using the pooled sample. See the estimation results plotted in Appendix Figure 1.23.

### Financial Frictions

An alternative explanation for the pattern I document is financial frictions, which could generate MRPK dispersion if they were high for some firms and low for others. Then they could gradually go away for various reasons, such as internally generated funds, or learning by banks. If young firms overcome financial constraints over time, they will start with high marginal product, and then decrease it.

However, financial frictions cannot explain why some firms start out with low marginal product and then raise it. In the data, 56% of the firms have higher MRPK than the previous year. In particular, 67% of the survived firm entrants have higher MRPK at age 1, and 59% of the survived age-one firms have higher MRPK at age 2; this percentage fluctuates between 54% to 56% from age 3. Because financial frictions cannot reconcile MRPK dynamics over time of these many firms in the data, I conclude financial constraints are not the driving force of the decline in MRPK dispersion with firm age.

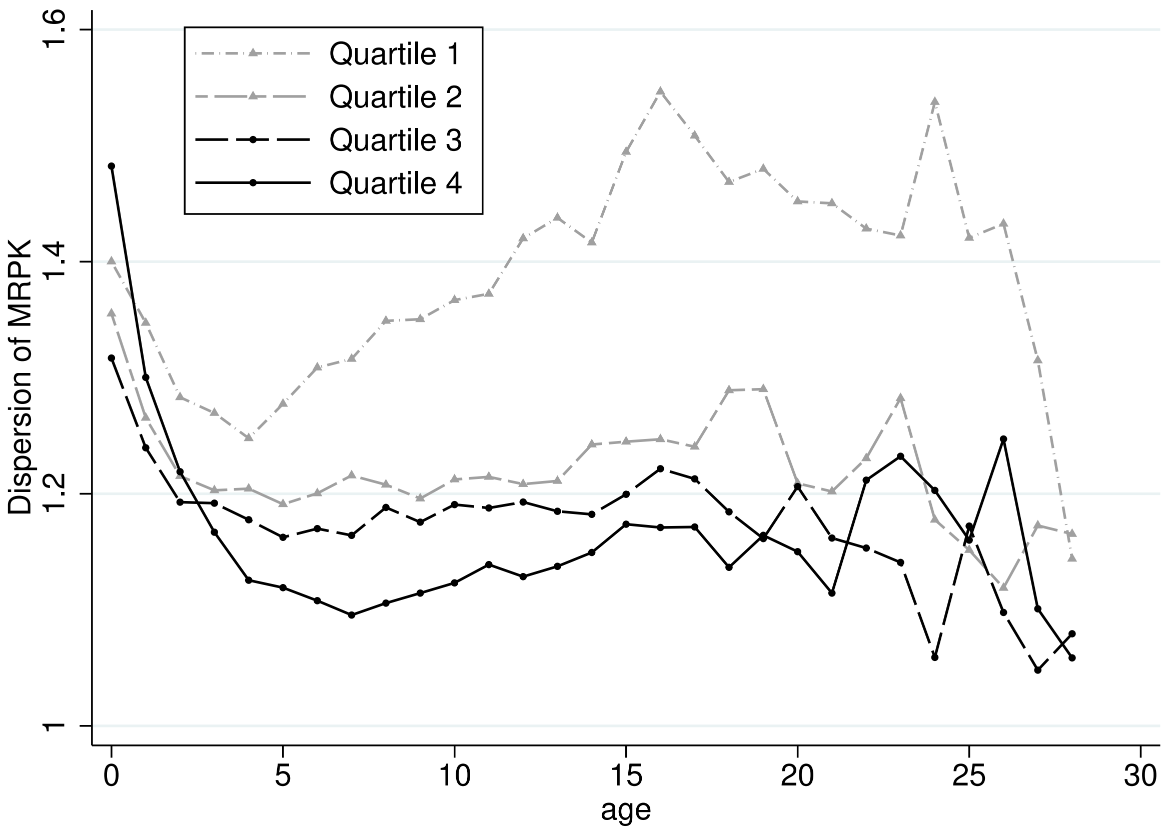
**Firm Size and Measurement Error**

Because average firm size and firm-cohort age are strongly and positively correlated, one may worry about whether the fact that I document is robust across firm groups with different average sizes. This section assesses the robustness of the decreasing MRPK dispersion with firm-cohort age to firm size.

Figure 1.8 plots the weighted average MRPK dispersion by firm size. The bottom-quartile firms have around 45 employees on average, and the top quartile firms on average have more than 166 employees. For firms in quartile two and three, and the top-quartile of firm-size distribution, MRPK dispersion decreases substantially with firm age, particularly for young firms. For firms in the bottom quartile, MRPK dispersion within a firm cohort decreases through the first five years but increases afterward. I conclude the age effects, particularly before age five, are robust to firm size. This is also consistent with the finding in [66] that effects of firm size become insignificant after controlling for firm age.

Measurement error has been an important and challenging concern for the misallocation literature and, more broadly, for measuring capital stocks and revenue outputs using firm-level data.[10] For this paper in particular, one may worry that measurement errors are larger for young

[10] [124] argue that the editing strategies used for U.S. Census of Manufactures may largely decrease the measured




Note: This figure plots the weighted average MRPK dispersion by firm-size quartile.

**Figure 1.8**: MRPK Dispersion by Firm Size

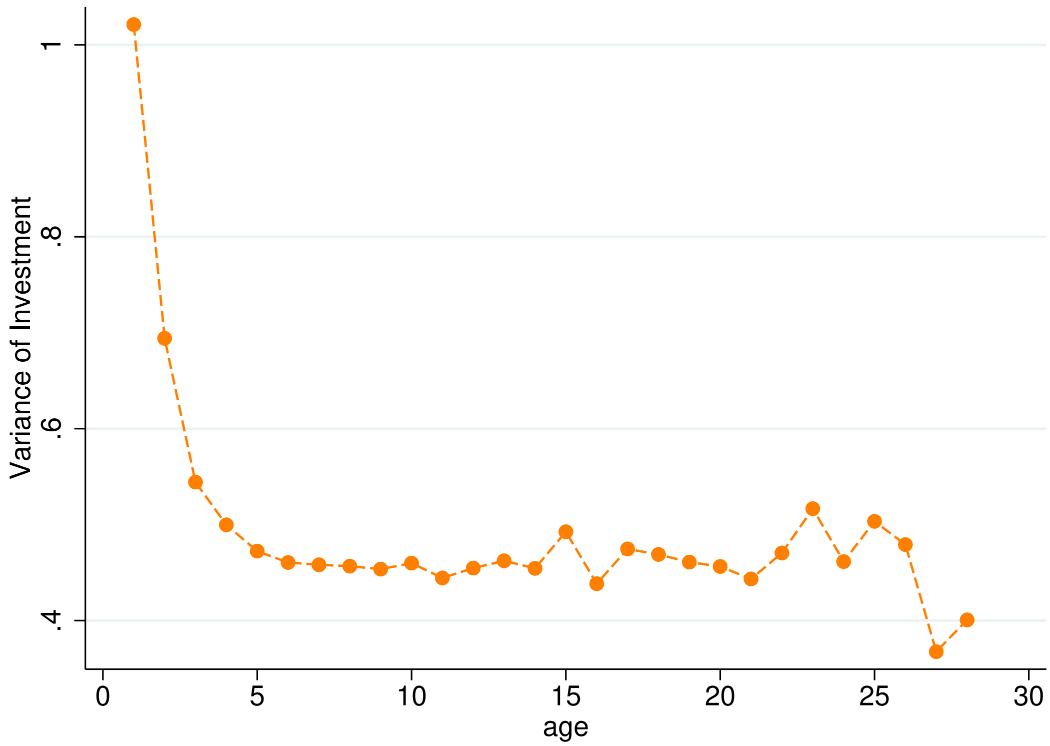
and small firms because they lack the resources and experiences to report measurements precisely. If one thinks of larger firms as those that are competent in corporate finance and accounting, and thus have relatively small measurement errors in reported revenue output and capital stocks, then there is less concern about measurement errors when we look at MRPK dispersion within large firms. The black lines in Figure 1.8 show that MRPK dispersion decreases robustly with firm-cohort age for the third-quartile and top-quartile firms, which have an average of 120 and 520 employees, respectively. That MRPK dispersion across large firms decreases over the firm cohort's life cycle provide indirect evidence of the pattern's robustness to measurement error.

MRPK dispersion in the cleaned dataset, leading to lower MRPK dispersion in the US than in India. However, because I focus on the firm panel-data within one country, differential data editing strategies across countries is much less concerning.

In addition, I follow the approach in [23] to estimate how much additive measurement errors in the revenue output and capital input for firms at each age accounts for observed MRPK dispersion ($\sigma^2_{mrpk}$). This approach essentially involves estimating the following regression: $\Delta log(y_{it}) = \phi_{err} mrpk_{it} + \Psi_{err} \Delta log(k_{it}) - \Psi_{err}(1-\lambda_{err}) mrpk_{it} \cdot \Delta log(k_{it}) + D_{st} + \varepsilon_{it}$, where $\Delta log(y_{it})$ and $\Delta log(k_{it})$ denote changes in log revenue output and capital, and $D_{st}$ denotes the industry-year fixed effects. They show that $(1-\lambda_{err})$ represents the contribution of measurement error to observed variance in MRPK under certain assumptions. The estimates of $1-\lambda_{err}$, using samples restricted to firm cohorts at age one to 28, respectively, have an average value of 0.02, which suggests measurement errors contribute to only 2% of the observed MRPK dispersion on average. Regressing the estimated $1-\widehat{\lambda}_{err}$ on age $j$, one will find a positive and insignificant coefficient of 0.15 with a P-value of 0.19, thus suggesting that the additive measurement error does not contribute differently to MRPK dispersion within firm cohorts at different ages.

### 1.3.3 Empirical Evidence of Firm Life-Cycle Learning

The decline in the variance of firm growth with firm age is evidence of firm life-cycle learning ([50]). Corresponding to the focus of this paper on MRPK, I use capital investment, the difference in the capital between two consecutive years, as the measure of firm growth. Figure 1.9 then plots the weighted average variance of investments across industry-year-age bins by firm age. It shows that investment dispersion decreases substantially with firm age, particularly for young firms. This is consistent with the basic Bayesian learning mechanism: firms enter with imprecise beliefs about their true productivity and they learn over time by observing revenue output realizations. Therefore, young firms face larger uncertainty about their productivity and revise their beliefs and investments relatively more, compared to the older firms, who are better informed with more observations. Furthermore, the pattern of decreasing variance of investment with firm age highly resembles that of MRPK dispersion, which implies the decline in MPRK




Note: This figure plots the weighted average variance of investments by firm age.

**Figure 1.9**: Variance of Investment by Firm Age

dispersion with firm age is likely to be associated with learning. [11]

## 1.4 Model of Firm Life-Cycle Learning

In this section, I develop a general equilibrium model to match the convexly downward sloping profile of MRPK dispersion with firm-cohort age in the data. The model features firm life-cycle learning and endogenous exit as in [79]. By adding capital to his original model, I bring in capital adjustment costs and capital fire-sales upon exiting the market, which are important to match the fact that firms scale down their capital stocks prior to exiting the market. Furthermore,

[11]Using firm-level panel data from Japan, [41] show that older firms make less forecasting errors about their idiosyncratic demand than younger firms, which is more direct evidence of firm learning.

the model with capital input choices in this paper generates losses in aggregate productivity due to informational frictions and capital market frictions over the firms' life cycles.

To conduct quantitative analyses, I build the model with multiple market frictions and firm-level distortions that can contribute to MRPK dispersion. Firms choose inputs facing (i) informational frictions, in the form of imperfect signals about their own fundamental productivity as well as contemporaneous uncertainty due to idiosyncratic shocks in each period, (ii) exit frictions, in the form of discount value from capital fire-sale on exit, (iii) technological frictions, in the form of quadratic capital adjustment costs, and (iv) a generic class of idiosyncratic firm-level distortions as in [73]. The key mechanism is that as firms learn over time, those with too much or too little capital stock adjust and the less productive firms within a cohort exit over time, leading to decreases in MRPK dispersion over a firm cohort's life cycle.

### 1.4.1 Environment and Equilibrium

Consider a discrete-time, infinite-horizon economy, populated by a representative household. The household inelastically supplies a fixed amount of labor $N$ and has a preference over consuming the final good. The household discounts time at rate $\beta$. I deliberately keep the household side of the economy simple because of its limited role in the analysis.

**Distribution of fundamentals.** The distribution of firm fundamental productivity $x_i$ is log-normally distributed, that is, $x_i \sim N(\mu_x, \sigma_x^2)$. In each period, with probability $\lambda \in (0,1)$, a firm $i$ carries over the same fundamental to the next period, and with probability $1-\lambda$, the firm exits exogenously.

**Production.** At the beginning of every period, each firm draws a productivity $z_{it}$, which combines its fundamental and an idiosyncratic transitory shock $e_{it} \sim N(0, \sigma_e^2)$. That is, $z_{it} =$

$x_i + e_{it}$. I assume the standard Cobb-Douglas production function, where output $y_{it}(k_{it}, n_{it}; z_{it}) = e^{z_{it}} k_{it}^{\alpha_1} n_{it}^{\alpha_2}$ with $\alpha_1 + \alpha_2 < 1$. Recall that $k_{it}$ denotes capital input and $n_{it}$ labor input. Note the assumption of decreasing return to scale is equivalent to an alternative environment in which firms produce differentiated products and face downward-sloping demand curves due to decreasing marginal utility of consumption. In that setup, $z_{it}$ can be interpreted as an idiosyncratic demand shifter. In this paper, I will refer to $z_{it}$ as the productivity specific to firm $i$ at time $t$.

**Learning.** Firms learn about their own fundamental productivity by experimenting and observing realized outputs in the previous periods, as in [79] and [82]. The firms' beliefs about their fundamentals are summarized in expected mean $\widehat{x}_{it}$ and expected variance $\widehat{\sigma}_{it}^2$. At the beginning of the firm-entry period, where $t = 0$ and no realizations of productivity $z_{it}$ arrive yet, firms have a common prior belief about their fundamental technology as $N(\widehat{x}_{i0}, \widehat{\sigma}_{i0}^2) \equiv N(\mu_x, \sigma_x^2)$. In every period $t$, they use the noisy signal $z_{it}$ to update and form a posterior belief about their fundamental productivity $x_i$ as $N(\widehat{x}_{i,t+1}, \widehat{\sigma}_{i,t+1}^2)$. Bayesian updating is based on the following equations:

$$\widehat{x}_{i,t+1} = \frac{\widehat{\sigma}_{it}^2 z_{it} + \sigma_e^2 \widehat{x}_{it}}{\widehat{\sigma}_{it}^2 + \sigma_e^2} \tag{1.7}$$

$$\widehat{\sigma}_{i,t+1}^2 = \frac{\widehat{\sigma}_{it}^2 \sigma_e^2}{\widehat{\sigma}_{it}^2 + \sigma_e^2}. \tag{1.8}$$

**Fixed and Input Costs.** Firms pay a fixed operation cost $f_o$ in every period they produce. Labor is hired period by period in a spot market with the competitive wage $w$. With capital depreciation rate $\delta$ and quadratic adjustment costs parameter $\xi$, the total cost of choosing capital stock $k_{i,t+1}$ is given by

$$\Phi(k_{it}, k_{i,t+1}) = k_{i,t+1} - (1-\delta)k_{it} + \frac{\xi}{2}\left(\frac{k_{i,t+1}}{k_{it}} - (1-\delta)\right)^2 k_{it}. \tag{1.9}$$

I also consider other factors that affect capital stock choices in addition to the fundamental productivity or demand. These factors include, for example, government policies, such as size-dependent taxes, or institutional environment, such as legal forms. As in [73] and [44], to capture these factors, I introduce a class of idiosyncratic firm-level "distortions" that appear in the firm's problem as proportional taxes on capital. I leave out the wedges on hiring decisions for simplicity. I allow distortions on capital to covary with contemporaneous productivity, that is, taxes $T_{it}^{k} = e^{z_{it}\tau_k}$, where $\tau_k$ denotes the correlation that determines the extent to which the capital price comoves with the contemporaneous productivity. If $\tau_k$ is positive, distortions discourage the investment of firms with stronger fundamentals while protecting those with weaker fundamentals, which is arguably the empirically relevant case ([76, 21]). The opposite incentive is true if $\tau_k$ is negative.

**Firm's problem.** At the beginning of each period $t$, firms choose whether to exit permanently or continue operating the business, and choose capital stocks $k_{i,t+1}$ if they continue operating. When exiting the market, firms turn to fire sales for their capital stocks and retain discounted values of $\gamma k_{it}$, as in [116]. A firm's state variables, or information set, includes the capital stock $k_{it}$, the observation of a noisy signal in productivity $z_{it}$, and the belief about the their fundamentals, summarized in $\widehat{x}_{it}$ and $\widehat{\sigma}_{it}^2$. Because $\widehat{\sigma}_{it}^2$ has a deterministic path over firm age $j$, I make the $j$ an explicit state variable instead of $\widehat{\sigma}_{it}^2$. Therefore, the value of an incumbent firm at age $j$ is given by $V(k_{it}, z_{it}, \widehat{x}_{it}, j) = \max_{D\in\{0,1\}} \left\{V^E(k_{it}, z_{it}, \widehat{x}_{it}, j), V^C(k_{it}, z_{it}, \widehat{x}_{it}, j)\right\}$, where the dummy variable $D$ denotes the exit choice, $V^C(k_{it}, z_{it}, \widehat{x}_{it}, j)$ denotes the value of continuing operation, and $V^E(k_{it}, z_{it}, \widehat{x}_{it}, j) = \gamma k_{it}$ is the value of exit. Writing the value of continuation in the recursive form yields

$$\begin{aligned} V^C(k_{it}, z_{it}, \widehat{x}_{it}, j) = \max_{k_{i,t+1}, n_{it}} \{ & e^{z_{it}} k_{it}^{\alpha_1} n_{it}^{\alpha_2} - w n_{it} - T_{it}^{k}\Phi(k_{it}, k_{i,t+1}) - f_o + \\ & \beta\big((1-\lambda)\mathbb{E}V(k_{i,t+1}, z_{i,t+1}, \widehat{x}_{i,t+1}, j+1) + \lambda V^E(k_{i,t+1}, z_{i,t+1}, \widehat{x}_{i,t+1}, j+1)\big)\}, \end{aligned}$$

where $\mathbb{E}$ denotes the firm's expectation of the value in $t+1$ conditional on the current information set $\{k_{it}, z_{it}, \widehat{x}_{it}, j\}$. Maximizing over the choice of labor inputs yields $n_{it}(z_{it}, k_{it}) = \left(\alpha_2 \frac{e^{z_{it}} k_{it}^{\alpha_1}}{w}\right)^{\frac{1}{1-\alpha_2}}$. After substituting the optimal choice of labor inputs, the value of continuation becomes

$$V^C(k_{it}, z_{it}, \widehat{x}_{it}, j) = \max_{k_{i,t+1}} \{GAk_{it}^{\alpha} - T_k\Gamma(k_{it}, k_{i,t+1}) - f_o + \tag{1.10}$$

$$\beta\lambda\gamma k_{i,t+1} + \beta(1-\lambda)V(k_{i,t+1}, z_{i,t+1}, \widehat{x}_{i,t+1}, j+1)\},$$

where $G \equiv (1-\alpha_2)\left(\frac{\alpha_2}{w}\right)^{\frac{\alpha_2}{1-\alpha_2}}$, $A = e^{z\frac{1}{1-\alpha_2}}$, and $\alpha \equiv \frac{\alpha_1}{1-\alpha_2}$ is the curvature of revenues net of wages.

**Stationary equilibrium.** We can now define a stationary equilibrium as follows: (i) a wage $w$; (ii) a set of value and policy functions of the firm: $V(k_{it}, z_{it}, \widehat{x}_{it}, j)$, $D(k_{it}, z_{it}, \widehat{x}_{it}, j)$, and $k_{i,t+1}(k_{it}, z_{it}, \widehat{x}_{it}, j)$; and (iii) a joint distribution of $\Omega(k_{it}, z_{it}, \widehat{x}_{it}, j)$ such that (a) taking wages and the law of motion for information set as given, the value and policy functions solve the firm's optimization problem, (b) the labor market clears as labor demand equals labor supply: $\int n_{it}(z_{it}, k_{it}) d\Omega(k_{it}, z_{it}, \widehat{x}_{it}, j) = N$, and (c) the joint distribution is the fixed point through time.

### 1.4.2 Intuitions of the Firm's Problem

Intuitively, without distortions (i.e., $\tau_k$=0), the choice of the next period's capital $k_{i,t+1}$ should be weakly increasing in the three state variables $k_{it}$, $z_{it}$, and $\widehat{x}_{it}$ at any age $j$. However, sufficiently large distortions, which disincentivize investment of more productive firms, may lead to less investment of more productive firms. Although $k_{i,t+1}$ is always weakly increasing in $k_{it}$ given the other state variables, it is not necessarily increasing $\widehat{x}_{it}$ given the other state variables. In section 1.5, I discuss the relevant case of distortions and investment decisions.

Figure 1.10 plots two examples of one firm's state variables over the firm's life cycle, using the same parameter values as in section 1.5. The left panel plots a firm with a low fundamental, in which case the firm chooses to downsize its capital stock in the next period after updating its belief of $x_i$ from zero to negative at age two. When a large negative shock arrives at age three, this firm chooses to exit. The right panel plots a firm with a high fundamental, in which case the firm updates its belief of $x_i$ upward smoothly and accumulates the capital stably through the first 10 years of its life cycle. This figure shows that less productive firms endogenously exit the market over time, whereas more productive firms stay and grow larger. Because initial capital stocks at entry may not match the firms' fundamentals for various reasons, including imprecise priors and large shocks, MRPK dispersion is large within the firm cohort at entry. This dispersion decreases over time as firms learn over time, adjust their capital stocks, and some firms exit the market.

Now I turn to a formal expression of computing the effects of MRPK dispersion on aggregate productivity. As shown in Appendix 1.12, combining the firm's optimal labor choice with the labor market and capital market clearing condition gives the expression of aggregate productivity as

$$z = z^* - \frac{1}{2}\frac{\alpha_1(1-\alpha_2)}{(1-\alpha_1-\alpha_2)}\sigma^2_{mrpk}, \tag{1.11}$$

where $z^*$ is the TFP in the frictionless and undistorted economy without any dispersion in MRPK, and $\sigma_{mrpk}$ is the aggregate standard deviation of MRPK. Taking the partial derivative of equation (1.11) reveals the relationship between MRPK dispersion ($\sigma^2_{mrpk}$) and productivity losses ($z - z^*$):

$$\frac{dz}{d\sigma^2_{mrpk}} = -\frac{1}{2}\frac{\alpha_1(1-\alpha_2)}{(1-\alpha_1-\alpha_2)}.$$

This expression provides a natural way to quantify the effects of changes in $\sigma^2_{mrpk}$ on aggregate productivity. In Section 1.5.2, I use this strategy to decompose the quantitative contribution of each factor to MRPK dispersion and TFP losses.

## 1.5 Quantitative Analysis

Throughout the analysis, I focus on dynamics of MRPK dispersion over the firm cohort's life cycle. Consider an economy with exogenous firm entry: Every period, the firm cohort with a joint distribution over $\{k_{i0}, x_i, e_{i0}\}$ enters. The model can predict the firm cohort's distribution over its life cycle by solving the firms' optimization problems. The joint distribution over $\{\widehat{x}_{it}, z_{it}, k_{it}\}$ is then fixed over time for any given firm-cohort age $j$. Therefore, stationary equilibrium must exist given the distribution of the firm cohort at entry, as long as exogenous exit rate $\lambda$ is positive.

### 1.5.1 Parameterization

I begin by directly assigning parameter values in the production function based on aggregate moments in the Chinese economy. I set the capital share $\alpha_1$ to 0.28, which is the weighted average capital share in the manufacturing sector, and set the labor share to 0.53 as in the Annual National Accounts. These two numbers lead to decreasing returns to scale as $\alpha_1 + \alpha_2 = 0.82$, which is in line with the standard value in the literature. The discount factor is set to 0.97 based on an interest rate of 10-year China government bonds of 3% during my sample period. I set the discount rate of capital fire-sale upon exit to 0.5, as used in [116]. I set the depreciation rate to 0.1, which is close to the median ratio of reported current-year depreciation value to capital stock. I use an exogenous firm exit rate of 0.04, which is close to the average exit rate of old firms in the US. I normalize $\mu_x$, the mean of the firm cohort's fundamentals at entry, to be zero, which is also the common initial belief of expected fundamentals. Regarding the productivity process in the model with time-invariant fundamental $x_i$, the dispersion of fundamentals and transitory shocks, $\sigma_x^2$ and $\sigma_e^2$, are exactly identified by the TFPR variance of the entrants, $Var(z_{it}|j=0)$, and the variance of time-series TFPR changes, $Var(z_{i,t+1}-z_{it})$.

Treating entry as exogenous in the model, I take directly the joint distribution of capital stocks and TFPR of all 60,972 age-zero firms in my sample as the initial distribution of $\{k_{i0}, z_{i0}\}$ among the firm cohort. I back out the fundamentals $x_i = z_{i0} - e_{i0}$, using randomly generated $e_{i0} \sim N(0, \sigma_e^2)$. Now, given the initial joint distribution of $\{k_{i0}, x_{i0}, e_{i0}\}$, a unique stationary equilibrium always exists.

I calibrate the remaining three parameters to jointly match three key moments in the data. The three parameters are the correlated distortion $\tau_k$, the fixed operating cost every period $f_0$, and the parameter in quadratic adjustment cost $\xi$. Let capital investment be $i_{it} = k_{i,t+1} - k_{it}$. The three moments are the exit rate of the firm entrants (11%), the autocorrelation of firm investments (-0.21), and the correlation of investment and productivity (0.17).[12]

Table 1.3 reports each parameter I used in the calibration. In particular, the calibrated value of correlated distortion $\tau_k$ is 0.5. The positive value is consistent with the positive correlation between distortion and fundamental in the literature ([140, 44]). In addition, [21] and [51] show evidence of stronger correlation in poorer countries than in richer countries. Because a large correlation can potentially offset the positive correlation between capital investment and productivity, it is helpful to get a sense of the magnitude of $\tau_k$ in the calibration. Appendix Figure 1.24 plots the policy functions of $k_{i,t+1}$ in the calibrated model, which shows $k_{i,t+1}$ is always increasing in $k_{it}$, as I discussed earlier. In addition, the intuition that firms with lower

[12]Because the China Annual Industry Surveys keep the non-state firms only if their revenues are above 5 million RMB, exit in the survey does not necessarily mean the firm goes out of business. To get a more precise measure of firm exit rates, I searched the operating status of a random sample of firms that exited from the survey during my sample period on the "National Enterprise Credit Information Pulicity System". Among the 528 firms I did find a record, 58% of the firms did shut down and unregistered. Therefore, I calibrate the model to target the adjusted exit rate of 11% rather than the 19% attrition rate for the firm entrants in the survey. If I nonetheless targets an exit rate of 19% in the calibration, the model then over-explains 10% of the data dynamics of MRPK dispersion.

**Table 1.3**: Parameter Values

| Parameter | Value |
|---|---|
| **Panel A: Pre-assigned Parameters** | |
| $\alpha_1$ - Capital share | 0.28 |
| $\alpha_2$ - Labor share | 0.53 |
| $\beta$ - Discount factor | 0.97 |
| $\delta$ - Depreciation rate | 0.1 |
| $\gamma$ - Exit discount in capital | 0.5 |
| $\lambda$ - Exogenous Exit Rate | 0.04 |
| $\mu_x$ - Mean of fundamentals | 0 |
| **Panel B: Exactly Matched Parameters** | |
| $\sigma_x^2$ - Dispersion of fundamentals | 0.70 |
| $\sigma_e^2$ - Dispersion of transitory shocks | 0.33 |
| **Panel C: Calibrated Parameters** | |
| $\tau_k$ - Correlated distortion | 0.50 |
| $f_0$ - Fixed operating costs | 0.41 |
| $\xi$ - Adjustment cost | 7.20 |

capital stock and lower idiosyncratic productivity are more likely to exit carries to the calibrated model with distortions. However, firms with the strongest beliefs of fundamentals and the highest contemporary productivities choose smaller capital stocks in the next period than firms with weaker beliefs and fundamentals due to severe distortions, as shown in the bottom-right panel in Figure 1.24. Therefore, correlated distortions in the model calibration under $\tau_k = 0.5$ are substantial, which strongly disincentivise more productive firms.

Table 1.4 reports the targeted moments in the data and model, which match decently. Although the three parameters are disciplined jointly by three moments, some useful intuitions apply. As in standard firm models, fixed operation costs positively relate to exit rates; and capital adjustment costs are most informative about the autocorrelation of investments. The correlated

**Table 1.4**: Moments Targeted in the Model and Data

| **Moments** | **Target** | **Model** |
|---|---|---|
| $\rho(i,z)$ | 0.17 | 0.17 |
| Exit rate of the entrants | 11.0 | 12.1 |
| $\rho(i,i')$ | -0.21 | -0.18 |

distortions captured in $\tau_k$ are informative about the correlation between capital investments and productivity. This is because a larger $\tau_k$ mitigate the investment responses to the productivity signals for the young firms, but asymptotically $\rho(i,z)$ always goes to 1 for the mature firms, independent of the value of $\tau_k$. Hence, $\tau_k$ is negatively associated $\rho(i,z)$.

### 1.5.2 Quantitative Predictions

I take the initial distribution of the firm cohort at entry as given in the data and report predictions of the calibrated model on the dynamics of MRPK dispersion over the firm cohort's life cycle. This section focuses on the model predictions over the first 10 years of the firm cohort's life cycle, where the data MRPK dispersion decreases robustly with firm-cohort age. In addition, the model MRPK dispersion stabilizes after age 10.

Figure 1.11 plots MRPK dispersion by firm-cohort age in the model and data. As the cohort of firms learn over time and adjust their capital stocks, the model predicts a decrease in MRPK dispersion by 0.15 points until age 10, compared to 0.22 in the data. Hence, the decrease in the model accounts for around two thirds of the magnitude in the data. Accordingly, within the firm cohort, $\sigma^2_{mrpk}$ decreases by 0.43 (that is, $1.50^2 - 1.35^2$) from age zero to age 10, corresponding to 15% TFP gains, based on equation (1.11). The sizable TFP gains over the firms' life cycles suggest considerable improvements in how efficiently resources are allocated across firms within the firm cohort.

**Table 1.5**: Second Derivatives $\widetilde{\phi}_j$ in the Model and Data

| | Age 0 | Age 1 | Average of Age 2 - 9 |
|---|---|---|---|
| Data 95% CI | (0.005, 0.05) | (0.02,0.05) | (-0.01, 0.02) |
| Model | 0.01 | 0.03 | -0.005 |

Furthermore, the model correctly predicts the convex relationship between MRPK dispersion and firm-cohort age, matching the curvature in the data without targeting it directly. Table 1.5 reports second derivatives of the age effects in the model, which are 0.01 at age zero and 0.03 at age one, respectively. These estimates fall right in the confidence interval of second derivatives in the data, as plotted in Figure 1.19. The average second derivative for firms between two and nine years of age is close to zero in the model, consistent with the insignificant values in the data. That MRPK dispersion decreases at a decreasing rate with firm age both in the model and data is consistent with the theory of firm life-cycle learning. For young firms, the number of observations is small, which limits the precision of firm priors. Hence, marginal gains of learning are larger at younger ages, which leads to larger decreases in MRPK dispersion.

To emphasize the selection in exit over the firms' life cycles, Figure 1.12 plots the distribution of firm productivity at age 0, 1 and 5. As in the data, the model predicts that the productivity distribution shifts to the right (i.e., the average productivity increases) as less productive firms exit over time. The growth rate of average productivity from age zero to age one is around 5.7% in the model, which matches the growth rate of 5.4% in the data, without targeting it directly.

In order to understand the quantitative contribution of each mechanism in the decrease in MRPK dispersion, I simulate how MRPK dispersion changes with firm-cohort age by sequentially adding mechanisms in the model. In the basic version, I shut down the exit channel by setting the fixed operation cost and capital fire-sale value to zero, and shut down the learning channel by solving the optimization problem when the firms never updated their beliefs. I find the dispersion barely decreases over age in this scenario. Specifically, $\sigma_{mrpk}$ decreases 0.007 points by age 10 in a model without firm life-cycle learning and an endogenous exit option, which is only 5% of the 0.15-point decrease in the benchmark model. Next, I add Bayesian updating to firms' beliefs about their fundamental productivity but still do not allow endogenous exits. The model then predicts a decrease of 0.08 points in MRPK dispersion by age 10, which accounts for as much as 54% of the decrease in the benchmark model. Further adding endogenous exit brings the model back to the benchmark version and accounts for the remaining half of the life-cycle decrease in MRPK dispersion as plotted in Figure 1.11.[13]

What are then the consequences of life-cycle learning for aggregate TFP rather than for one firm cohort? I begin with a hypothetical baseline in which all firms have completed their life-cycle learning. In particular, I assume MRPK dispersion within each firm cohort remains constant after age 10 in the stationary equilibrium. This assumption is consistent with quantitative predictions in the calibrated model. In effect, I regard the firm cohorts age 10 and older as having learned sufficiently about their fundamental productivities that they cannot reduce their levels of MRPK dispersion by further learning.

Consider the model predictions on two moments: the age distribution of firms, and MRPK dispersion at each age. The aggregate MRPK dispersion is given by the average MRPK dispersion across all firm ages weighted by the number of firms at each age in the equilibrium, that is, 1.46 in the model. Meanwhile, aggregate MRPK dispersion in the hypothetical baseline is calculated by replacing the model MRPK dispersion across firms at ages zero to nine with the dispersion of

[13] If I consider the decomposition of aggregate capital stock within the firm cohort by age, from age zero to age 10, as in [108], the covariance between capital stock and market share (defined by revenue output share) increases by 72%, from 0.11 to 0.19. This increase in covariance with firm age is consistent with the theory of firm life-cycle learning, but unlike the quantitative analysis of my model, it cannot estimate the contribution of life-cycle capital adjustments separately from learning.

age-10 firms, while keeping the age distribution of firms the same as in the benchmark model predictions. Mechanically, the aggregate MRPK dispersion is lower in the hypothetical baseline than in the model, because of the absence of firm life-cycle adjustments. I can use equation (1.11) to compute the implied TFP losses, $\Delta z$, for any given model prediction on $\sigma_{mrpk}$ relative to the hypothetical baseline.

In the first column of Table 1.6, I report the differences in aggregate MRPK dispersion and in log TFP between predictions of the benchmark model and its hypothetical baseline. Aggregate MRPK dispersion in the hypothetical baseline is 0.11 points lower. This difference shows firm life-cycle adjustments accounts for 7% of MRPK dispersion across firms in the economy, which lead to a 10 percent loss in TFP.

**Table 1.6**: Consequences of Firm Life-Cycle Learning in the Model

| | Learning | Distortions + Learning |
|---|---|---|
| $\Delta\sigma_{mrpk}$ | 0.11 | 0.21 |
| $\frac{\Delta\sigma_{mrpk}}{\sigma_{mrpk}}$ | 7% | 14% |
| $\Delta z$ | 0.10 | 0.19 |

To consider the consequences of firm-level distortions for aggregate TFP, I conduct the counterfactual experiment of removing firm-specific distortions by setting $\tau_k$ to 0 in the benchmark model. The standard deviation of MRPK ($\sigma_{mrpk}$) across firms at age 10 becomes 1.25, which is 0.1 points smaller than in the benchmark model. As reported in the second column of Table 1.6, in the corresponding hypothetical baseline, which removes both distortions and firm life-cycle adjustments, the aggregate MRPK dispersion would decrease 0.21 points, from 1.46 to 1.25. Hence, distortions and firm life-cycle learning together account for 14% of MRPK dispersion in the economy, which leads to a 19 percent loss in TFP. Omitting learning over the firm cohort's life cycle will attribute all changes in MRPK dispersion in the hypothetical baseline to distortions, which causes more than half of the TFP losses to be incorrectly attributed to distortions.

I conclude that the model featuring firm life-cycle learning explains around two thirds of the life-cycle MRPK dispersion. Without targeting the curvature of age effects and the productivity growth over the firm cohort's life cycle directly, the model correctly matches these moments in the data. Through the lens of this model, omitting firm life-cycle learning leads to a sizable overestimation of TFP losses from misallocation.

## 1.6 Evidence from Colombia and Chile

In this section, I report patterns of MRPK dispersion over the firms' life cycles using older data from the manufacturing sectors in Colombia and Chile.[14]

The Colombia Industrial Surveys during the period 1977 - 1991 cover around 6,600 plants per year on average. I measure the capital stock ($k_{it}$) as the book value of fixed assets, and measure revenue output ($y_{it}$) as value added constructed by subtracting intermediate inputs from the sum of the value of production, inventory changes, and sales tax ([121]). Again, I use the industry-level capital share from the NBER-CES database and equation (1.3) to calculate MRPK in log terms. To keep sufficient observations to measure dispersion, I calculate the standard deviation of MRPK across plants within the same year-age bins, rather than the same industry-year-age bins.

Figure 1.13 reports the average MRPK dispersion with firm cohorts measured in two ways as firms age from zero to 10 in Colombia. The standard deviation of MRPK decreases from almost 1.4 to 1 by age five, and remains below 1.1 until age 10. The log difference of MRPK between the 90th and the 10th percentile plant decrease from 3.4 to around 2.6 by age five and stays at around 2.8 till age 10. That is, the ratio of the 90th to the 10th percentile MRPK drops one half, from 30 to around 15, during the first five years of the firm cohort's life cycle. In

[14] I thank Mark Roberts for sharing his data with me.

addition, similar to the convex age effects estimated using Chinese data, both measures of MRPK dispersion in Colombia decrease at a decreasing rate before age 10.

The data I have on the manufacturing sector in Chile cover plants with at least 10 employees during the period 1979 - 1986. Though the year of plant entry is not reported in the survey, based on the panel structure of the data, I can identify the year of plant entry $t$ if one plant does not have a record in year $t-1$ but shows up in year $t$. Hence, the oldest plant cohort with a well-defined plant age is established in 1980 and can be observed until age five. The final sample size grows from 226 plants in 1980 to 1,037 plants in 1986.

Using the older and much smaller dataset from Chile, I measure revenue output ($y_{it}$) as value added, and measure capital stocks ($k_{it}$) by summing up the annual investments in buildings, machinery, and vehicles net of depreciation since birth year ([121]). Then I calculate MRPK in log terms using equation (1.3) as before, and I calculate the standard deviation of MRPK across plants within the same year-age bins. I find that, between firm-cohort age zero and five, the average standard deviation of MRPK in Chile decreases from 1.7 to less than 1.2, and the average log difference of MRPK between the 90th and the 10th percentile plant decrease from 4.5 to around 2.5. Note the decrease in MRPK dispersion in the Chilean data is larger than that in China and Colombia during the same age range. Because of the large confidence intervals due to the small number of firms in Chile, I report the t-test results of the differences of average MRPK dispersion between age 0-1 and age 2-5 firms. Table 1.7 shows that both the standard deviation and 90-10 percentile difference of MRPK are significantly larger for young firms in Chile.

I conclude that evidence from Colombia and Chile is in accord with my finding using Chinese data that MRPK dispersion decreases over the firm cohort's life cycle. In Colombia, MRPK dispersion decreases substantially before age five and at a decreasing rate. As in the Chinese data, this pattern is consistent with the theory of firm life-cycle learning, which has larger

Table 1.7: Chile: MRPK Dispersion by Firm Age Group

| | Age 0-1 | Age 2-5 | Difference |
|---|---|---|---|
| Average $\sigma_{mrpk,tj}$ | 1.60 | 1.48 | -0.12*** |
| Average 90-10 | 4.00 | 3.62 | -0.38*** |
| Obs. of firms | 1,935 | 989 | |

impacts at younger ages.

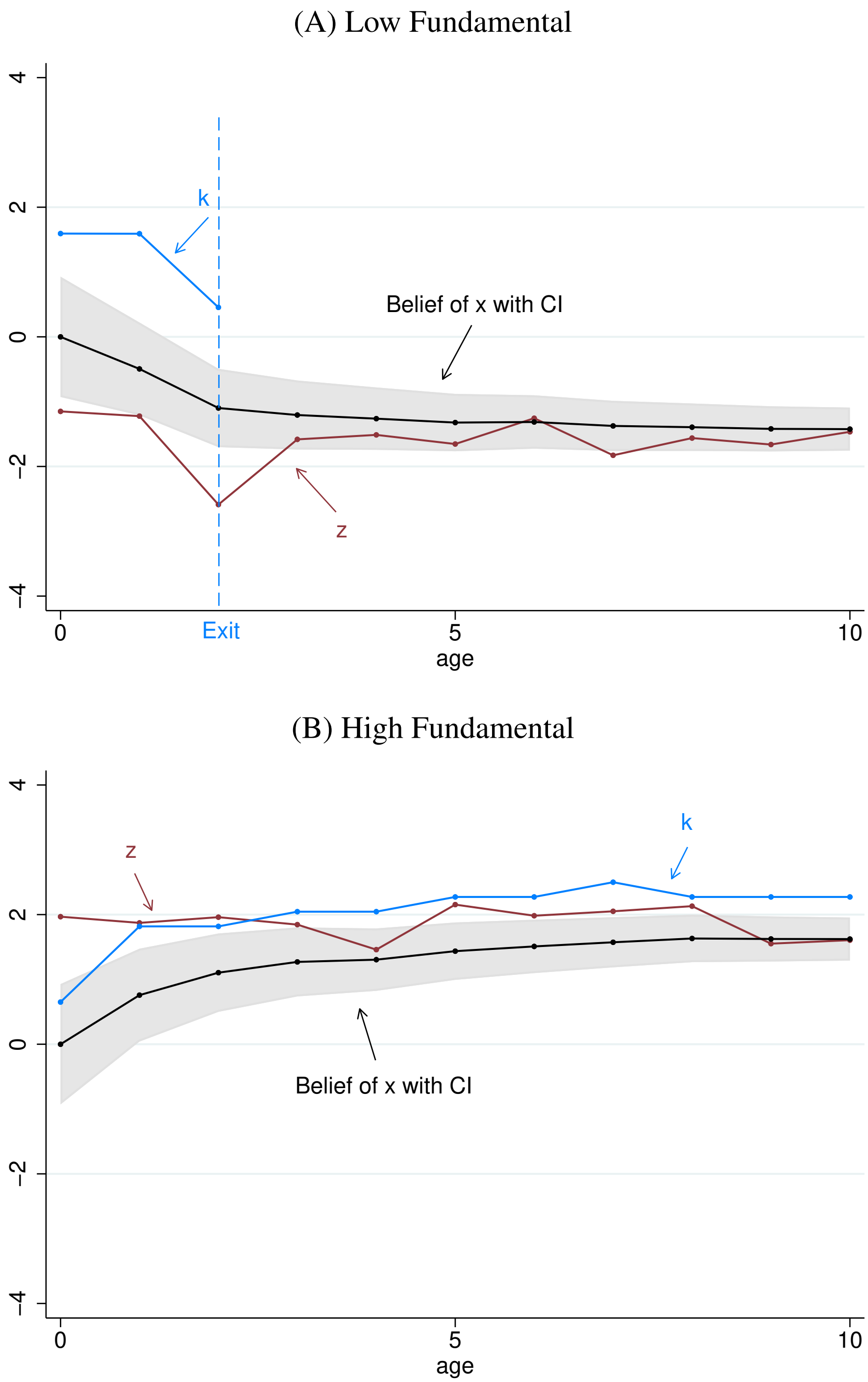
## 1.7 Conclusion

This paper provides a new interpretation of MRPK dispersion as firm life-cycle learning. I draw on the panel firm-level data in China to document substantial decreases in MRPK dispersion with firm-cohort age. In addition, for young firm cohorts, MRPK dispersion decreases substantially and at a decreasing rate. The pattern also holds broadly for data on the manufacturing sectors in Colombia and Chile. Building on the new facts, I develop a dynamic model featuring informational frictions over the firm cohort's life cycle as the firms learn about their own fundamental productivity. The model predicts that as firms learn over time and adjust their capital stocks, possibly through endogenously exiting the market, MRPK dispersion decreases over their life cycles. I highlight the importance of firm life-cycle learning to ex-post aggregate MRPK dispersion. Quantitative analysis suggests that omitting this dimension leads to sizable overestimation of the TFP losses due to misallocation. In addition, TFP losses resulting from firm life-cycle learning to overcome informational frictions is an optimal constrained equilibrium, which may not be fixed by policy interventions.

Though direct measurements of firm- or individual-level information learning is scarce, [133] provide empirical evidence that more productive Japanese firms make more accurate forecasts about the macro economy. Their findings suggest learning may be endogenous: firms can pay costs to learn better information. Although the learning process in this paper is essentially mechanical and homogeneous across firms, I leave the discussion of richer learning models to future research.

This paper shows that data from developing countries generally show decreasing MRPK dispersion over the firm cohort's life cycle. Further exploration of the profiles of life-cycle MRPK dispersion in developed countries would be worthwhile. Comparing economies at different income levels can potentially shed light on the theory of cross-country TFP.

## 1.8 Appendices




Note: This figure plots two examples of a firm's state variables over the firm's life cycle. The maroon line plots realizations of productivity $z_{it}$ over time; the black line plots corresponding beliefs of the fundamental $\widehat{x}_{i,t+1}$ with a 95% confidence interval based on $\widehat{\sigma}^2_{i,t+1}$; and the blue line plots capital stock $k_{it}$.

**Figure 1.10**: Examples of One Firm's Life Cycle in the Model

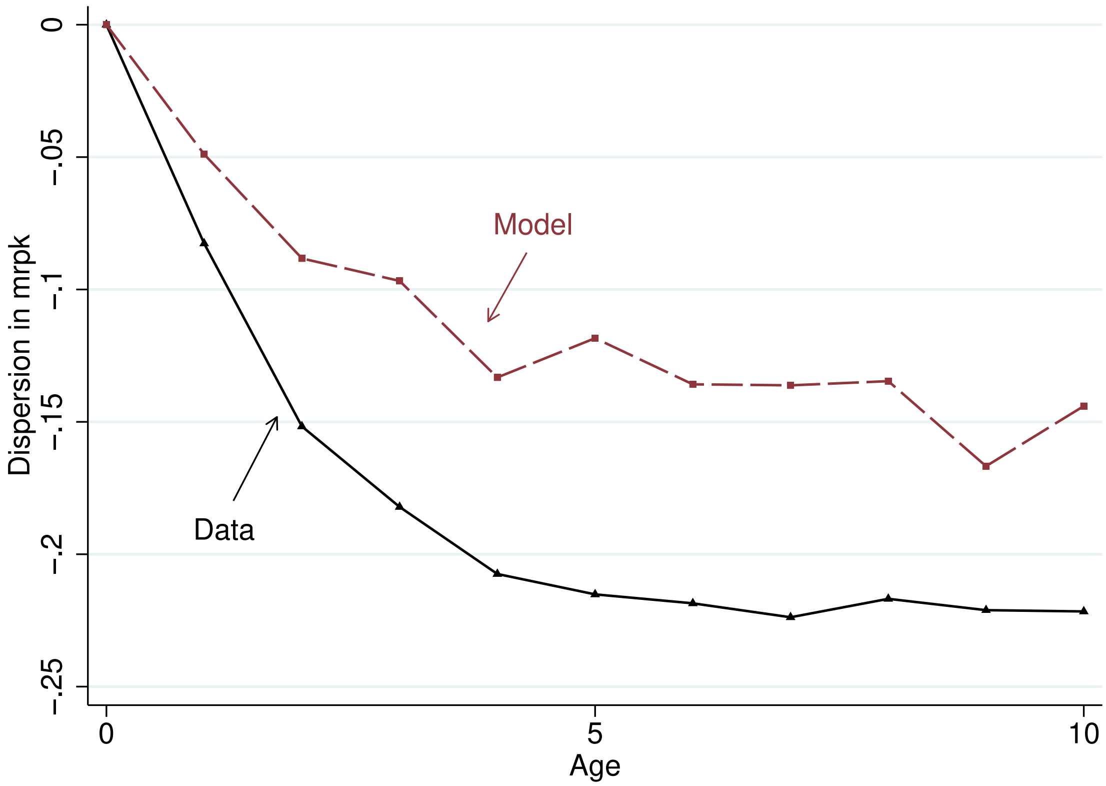
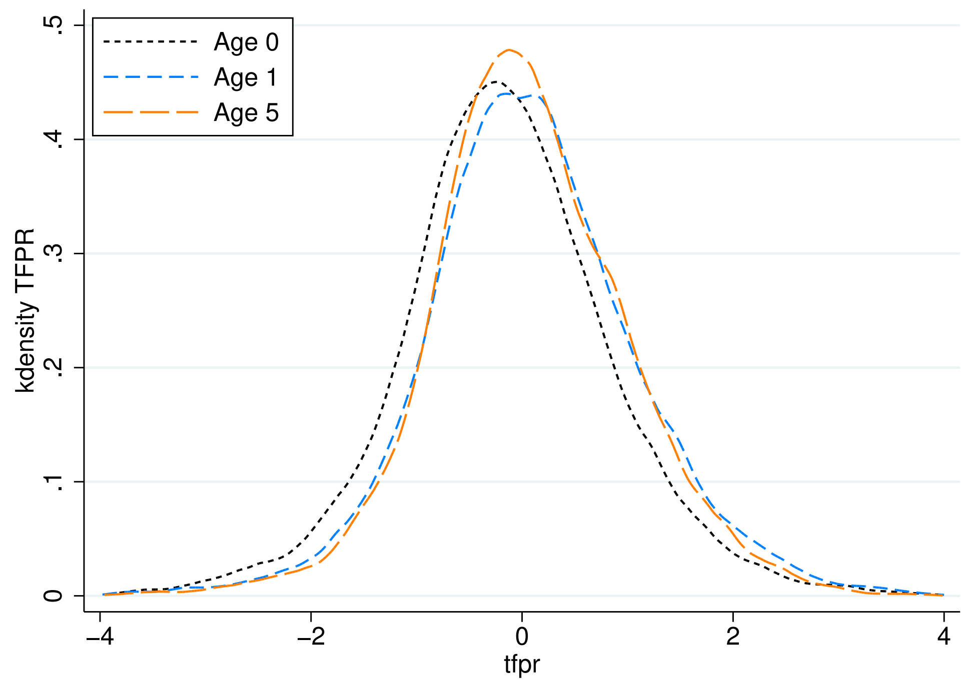



**Figure 1.11**: MRPK Dispersion ($\sigma_{mrpk}$) in the Model and Data

a) Life-Cycle Productivity in the Data




b) Life-Cycle Productivity in the Model

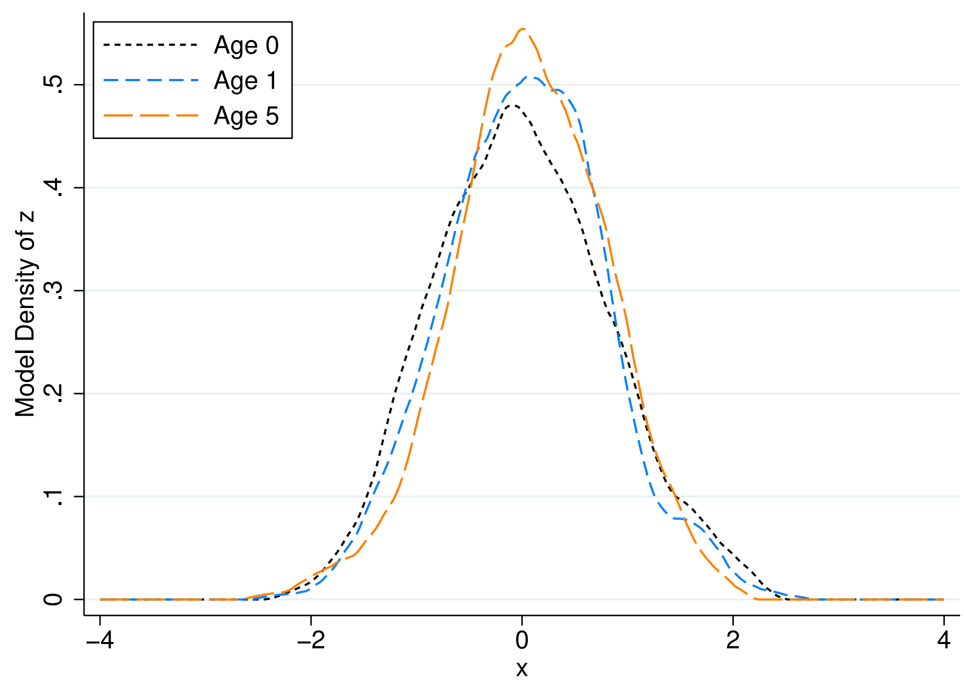



**Figure 1.12**: Distributions of Productivity in Model and Data

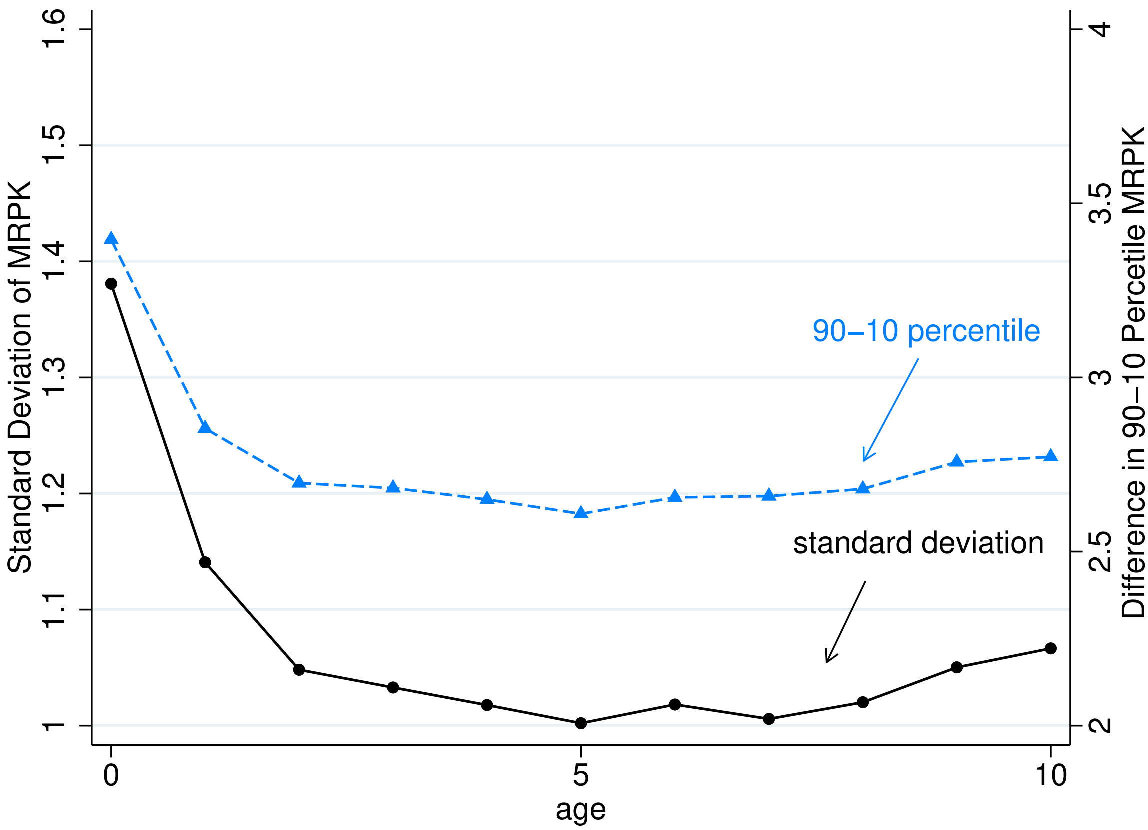
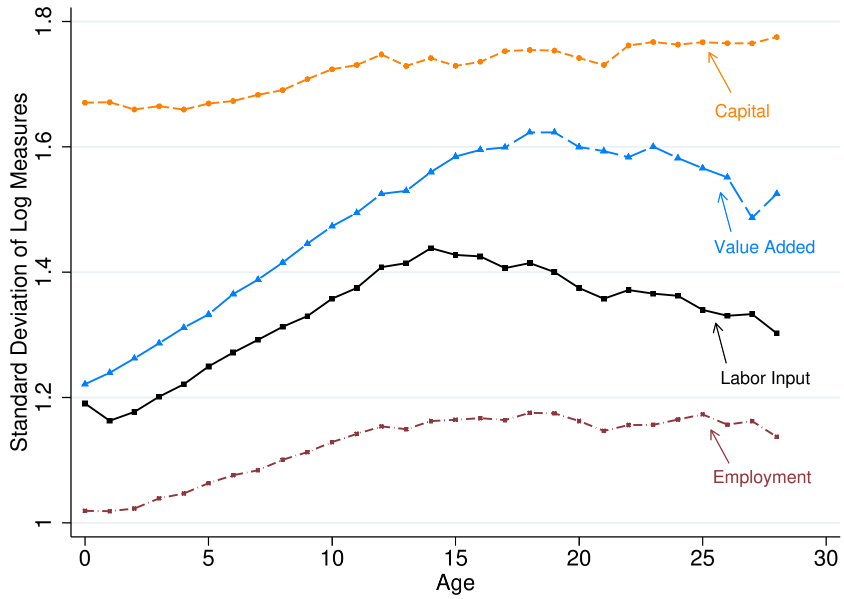
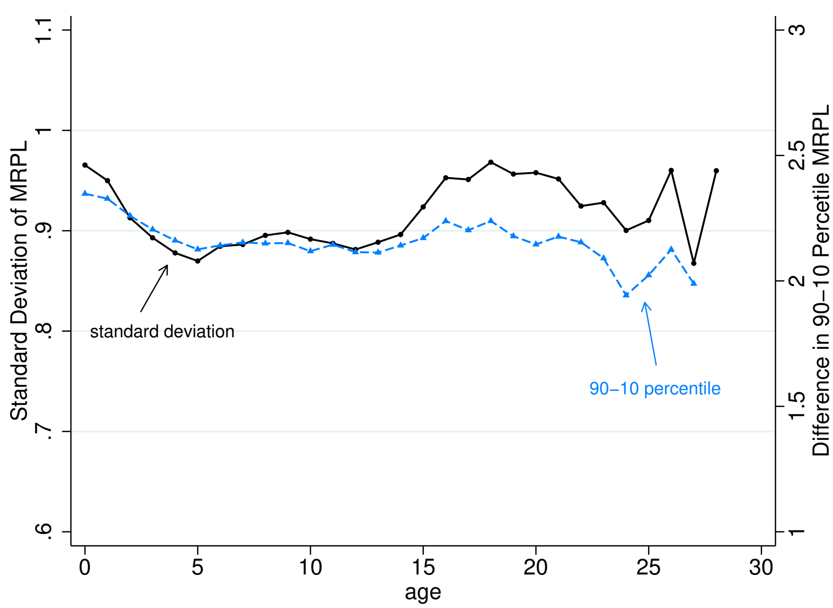



**Figure 1.13**: MRPK Dispersion by Firm Age, Colombia



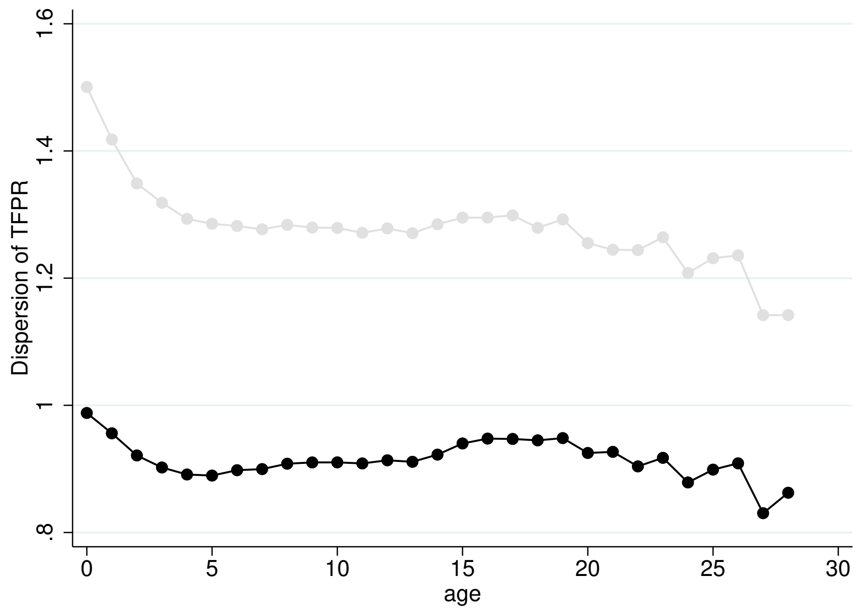

Note: This figure plots the standard deviation of log value-added ($y_{it}$), log capital input ($k_{it}$), log labor input ($n_{it}$), and log employment by firm age.

**Figure 1.14**: Dispersion of Key Variables by Firm Age



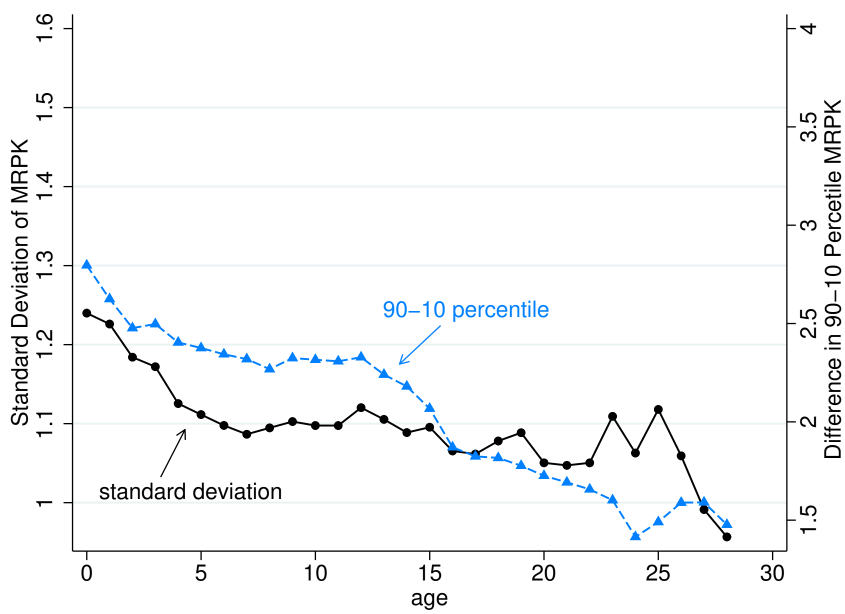

Note: This figure plots the weighted average standard deviation of MRPL and the weighted average value of the 90th minus the 10th percentile MRPL by firm age.

**Figure 1.15**: Dispersion of MRPL by Firm Age



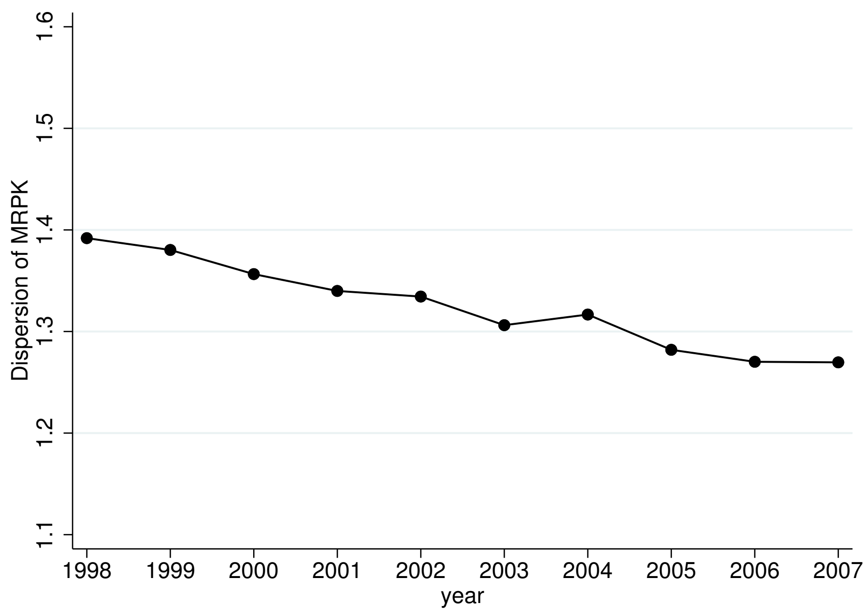

Note: This figure plots the average standard deviation of TFPR ($\bar{\sigma}_{tfpr,j}$) over age, weighted by the number of firms in industry-year-age bins. The gray line replicates the dispersion in MRPK as Figure 1.1 for reference.

**Figure 1.16**: Dispersion of TFPR ($\bar{\sigma}_{tfpr,j}$) by Firm Age




Note: This figure plots the weighted average standard deviation of MRPK ($\bar{\sigma}_j$) and the weighted average value of the 90th minus the 10th percentile ($\bar{D}_j^{90-10}$) by firm-cohort age, for the firms are recorded every year during the sample period 1998 - 2007.

**Figure 1.17**: Dispersion of MRPK by Firm Age, Balanced Panel




Note: This figure plots the weighted average $\sigma_{mrpk,stj}$ during 1998 - 2007.

**Figure 1.18**: Dispersion of MRPK by Year

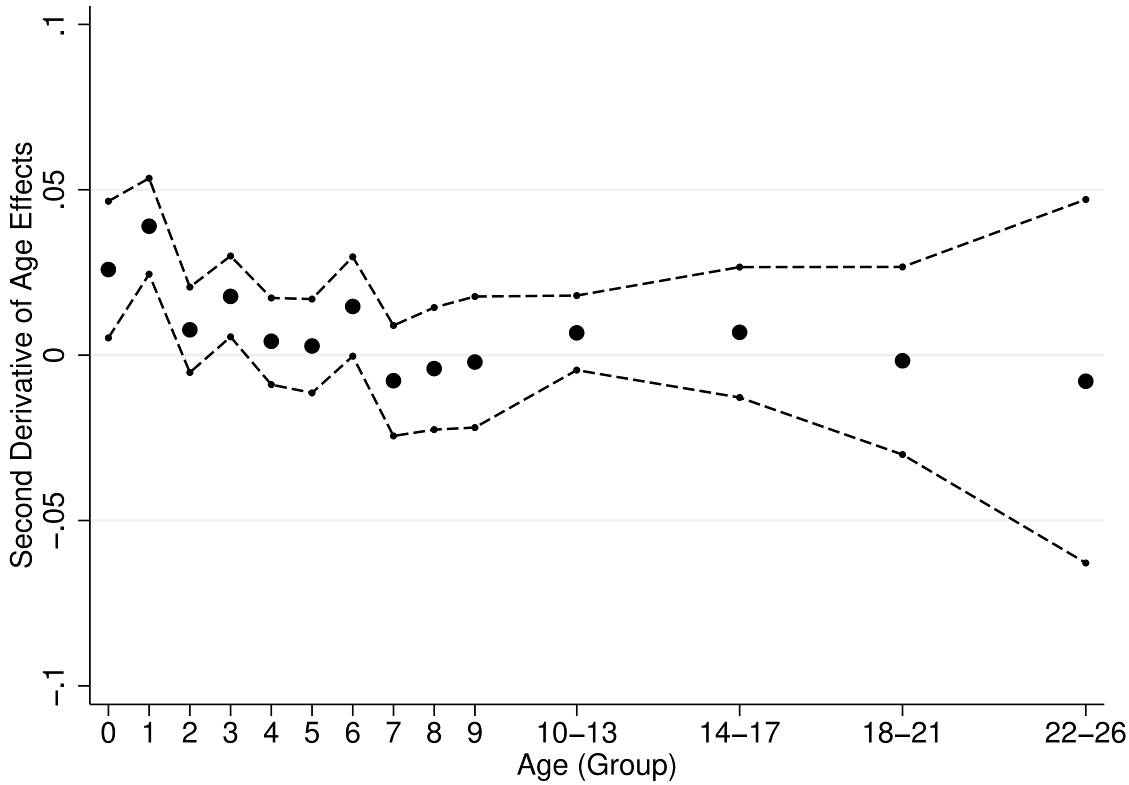
## 1.9 Curvature of the Age Effects

I use the second derivatives of age effects to test the curvature of age effects. Though none of the first-order effects of age, cohort, or time can be identified separately, their second derivatives are always identified ([95]). Recall that the cohort of firms aged $j$ in time period $t$ is denoted as $c_{t-j}$. Consider equation (1.5) for the cohort $c_{t-j}$ observed in year $t$ and $t+1$. To eliminate cohort effects, taking the first difference yields the sum of the first-order age effect between $j$ and $j+1$ and the year effect between $t$ and $t+1$: $\Delta\sigma_{mrpk,stj} \equiv \sigma_{mrpk,s,t+1,j+1} - \sigma_{mrpk,stj} = (\phi_{j+1} - \phi_j) + (\psi_{t+1} - \psi_t) + \Delta_c \varepsilon_{stj}$, where $\Delta_c \varepsilon_{stj} \equiv \varepsilon_{s,t+1,j+1} - \varepsilon_{stj}$. Consider an older firm cohort $c_{t-j-1}$, observed at age $j+1$ and $j+2$ in the same year $t$ and $t+1$. Again, we can identify the sum of the first-order age effect and year effect: $(\phi_{j+2} - \phi_{j+1}) + (\psi_{t+1} - \psi_t)$. Now taking the difference of the two first-order effects gives the second derivative of age effects:

$$\widetilde{\phi}_j \equiv (\phi_{j+2} - \phi_{j+1}) - (\phi_{j+1} - \phi_j).$$

The second derivative of age effects $\widetilde{\phi}_j$ is the difference between two slopes: one slope of MRPK dispersion between age $j+2$ and age $j+1$, and the other slope between age $j+1$ and age $j$. If $\widetilde{\phi}_j = 0$, that is, if the two slopes are the same, the age effects between age $j$ and $j+2$ are linear. If $\widetilde{\phi}_j > 0$, the profile of MRPK dispersion is convex between age $j$ and $j+2$. Therefore, I can estimate $\widetilde{\phi}_j$ to inform the curvature of age effects.

I estimate $\widetilde{\phi}_j$ for each age $j$ between zero and nine, and put the older ages into groups for tighter confidence intervals. Figure 1.19 plots the second derivatives of age effects with 95% confidence intervals. It shows the second derivatives are significantly positive at age zero, one, and three, and become near zero and insignificant after age five. Based on the point estimates, firm age has convex effects on MRPK dispersion through the first five years of the firm cohort's life cycle; that is, MRPK dispersion decreases at a decreasing rate before age five.




Note: This figure plots the estimates of the second derivatives of age effects $\widetilde{\phi}_j$ with 95% confidence intervals. It is estimated for each age between zero and nine, and for each four-age group afterward in order to get tighter confidence intervals.

**Figure 1.19**: Second Derivatives of Age Effects

**Table 1.8**: McKenzie Test of Linear Age Effects

| | | | |
|---|---|---|---|
| $H_0$: Linear Range | Age 0-5 | Age 5-10 | Age 10-28 |
| P-value | 0.00 | 0.30 | 0.15 |
| $H_0$: Linear Range | Age 5-28 | Age 4-28 | Age 3-28 |
| P-value | 0.12 | 0.14 | 0.03 |

Note: This table reports the p-value of the McKenzie test of linear age effects over several age ranges.

Table 1.8 further reports p-values of the Mckenzie tests on linear age effects. It is essentially a formal Wald test for the null hypothesis $H_0$ of $\widetilde{\phi}_j$ being jointly zero for a set of $j$

values. Jointly zero second derivatives imply the corresponding age effects are linear. The tests in the first row strongly reject the hypothesis that the age effects are linear between age zero and age five, but cannot reject they are linear between age 5 to 10 or 10 to 28. The tests in the second row show one cannot reject the null hypothesis of linear age effects between age four or five and age 28. But the McKenzie test rejects the linear age effects with a p-value of 0.03 if one extends the age range to between three and 28. Based on these results, I will assume a linear trend in age effects after age 10 in the second alternative approach to identify the first-order age effects.

## 1.10 Lower and Upper Bounds of Age Effects

In this section, I show the two restrictions that I impose in the second alternative approach provide the upper and lower bounds of age effects if all three effects of age, year, and cohort on MRPK dispersion have non-positive trends.

Consider the case of a linear trend in the three effects of age, year, and cohort: $\phi_j = g_\phi j + u_{\psi,j}$, $\psi_t = g_\psi t + u_{\psi,t}$, and $\chi_c = g_\chi c + u_{\chi,c}$. The condition that all three effects of age, year, and cohort have non-positive trends on MRPK dispersion gives $g_\phi, g_\psi, g_\chi \leq 0$. I show below that (i) $g_\psi = 0$ (attributing the entire decline in MRPK dispersion over time to year effects) yields the upper bounder of $g_\phi$, and (ii) $g_\chi = 0$ (attributing the entire decline in MRPK dispersion over time to cohort effects) yields the lower bounder of $g_\phi$.

Substituting the identity of cohort birth year $c = t - j$ into the observed result, which is the sum of three effects:

$$\phi_j j + \psi_t t + \chi_c c = (g_\phi - g_\chi) j + (g_\psi + g_\chi) t + u,$$

where $u = u_{\psi,j} + u_{\psi,t} + u_{\chi,c}$. Denote $g_{M^*} = g_\phi - g_\chi$ and $g_M = g_\psi + g_\chi$. The unobserved negative trend in age effects $g_\phi$ can be expressed as $g_{M^*} + g_\chi$. Note that $g_M$ is negative by definition; thus, the trend in cohort effects satisfies $g_\chi \in [g_M, 0]$, given the condition of three non-positive trends. Therefore, $g_\phi$ is bounded between $g_{M^*} + g_M$ and $g_{M^*}$.

The first restriction, which attributes the entire deline in MRPK dispersion over time to cohort effects, $g_\chi = 0$ is now equivalent to $g_\chi = g_M$. Hence, it gives the lower bound of the negative $g_\phi$, that is, $g_{M^*} + g_M$. Similarly, the second restriction, $g_\chi = 0$, yields the upper bound of the negative $g_\phi$, that is, $g_{M^*}$. Figure 1.6 shows the first restriction indeed yields a much steeper profile of MRPK dispersion with firm-cohort age.

## 1.11 Details of Alternative Approach Two

Here I explain the details of estimating equation (1.5) under the framework of imposing one additional linear restriction as in [46]. In particular, I describe the two different linear restrictions I impose for results in section 1.3.1 and how to implement them in practice.

To derive the restrictions, consider the weighted average dispersion of marginal products in year t:

$$SD_t = \sum_{c \in C_t} \omega_{stj} \cdot SD_{stj}(mrpk_{it}),$$

where $\omega_{stj}$ is a weight defined as the number of firms in an industry-age-year bin divided by the total number of firms. Let $CIC$ denote the set of all 4-digit industry codes. Substituting in

$SD_{stj}(mrpk_{it})$ from equation (1.5), it is easily shown that the weighted average can be written as

$$\begin{aligned}
SD_t &= \alpha + \psi_t + \bar{X}_t + \bar{\Phi}_t \qquad (1.12)\\
\bar{X}_t &= \sum_{c \in C_t} \frac{\Phi_{ct}}{\bar{\Phi}_t} \chi_c \\
\bar{\Phi}_t &= \sum_{c \in C_t} \Phi_{ct}, \text{and } \Phi_{ct} = \sum_{s \in CIC} \omega_{stj}(\phi_j D_j + \varepsilon_{stj}).
\end{aligned}$$

We see in Figure 1.18 that the weighted average dispersion of marginal products of capital (or $SD_t$) declines from one year to the next. equation (1.12) shows clearly that the decline of dispersion has three sources: the decline due to the time effects $\psi_t$, the decline due to the aggregate cohort effects captured in $\bar{X}_t$, and the decline due to composition of firms at different ages captured in $\bar{\Phi}_t$. The restrictions will be imposed on the term

$$\Omega_t = \alpha + \psi_t + \bar{X}_t. \qquad (1.13)$$

This term $\Omega_t$ captures the year-specific aggregate effects. It changes over time as a result of two effects: (i) cohort-neutral effects captured in $\psi_t$, and (ii) effects due to the changes in the composition of active cohorts operating, captured in $\bar{X}_t$. For example, if younger cohorts are born with a small dispersion of marginal products, the observed aggregate dispersion can decease over time only because young cohorts enter and older cohorts exit the market.

The basic idea of this approach is to decompose the time series of $\Omega_t$ into a trend component and a cyclical component. To identify cohort and year effects in addition to age effects, this approach makes assumptions on the relative role of time and cohort effects in the trend component.

In practice, the first step of implementation is to transform the time dummies as equation (2.94) in [46] such that two restrictions are satisfied: (i) the year dummies add to zero: $\sum_{t=0}^{T} t = 0$, and (ii) the normalization of all year effects adding up to zero: $\frac{1}{T}\sum_{t=0}^{T} \psi_t = 0$. I also want to normalize the cohort effects $\bar{X}_t$ such that $\frac{1}{T}\sum_{t=0}^{T} \bar{X}_t = 0$. I do so by appropriately choosing the constant term $\alpha$ in equation (1.13). Second, the time series of $\psi_t$ and $\bar{X}_t$ can be decomposed into a trend component and a cyclical component:

$$\psi_t = g_\psi t + u_{\psi,t}, \ \ \bar{X}_t = g_\chi t + u_{\chi,t}, \tag{1.14}$$

where $g_\psi = \frac{\sum_{t=0}^{T} \psi_t t}{\sum_{t=0}^{T} t^2}$ and $g_\chi = \frac{\sum_{t=0}^{T} \bar{X}_t t}{\sum_{t=0}^{T} t^2}$. Intuitively, the estimates are simply regressing $\psi_t$ and $\bar{X}_t$ on time, thereby decomposing each time series into a trend component and the cyclical component orthogonal to time. It is the same method as proposed in [67]. Finally, substituting equation (1.14) into equation (1.13) gives

$$\Omega_t = \alpha + g_M t + u_{M,t},$$

where $u_{M,t} = u_{\psi,t} + u_{\chi,t}$ and recall that $g_M = g_\psi + g_\chi$. The restrictions I used in Section 1.3.1 simply make assumptions on how $g_M$ is split between $g_\psi$ and $g_\chi$.

I can also use the McKenzie test, as described in section 1.9, to test the linearity restriction on the series of $\psi_t$. In practice, I first take the difference of MRPK dispersion of the same cohort observed in the two adjacent years to eliminate the cohort effects $\Delta_c \mathrm{SD}_{stj}$. Then I take the second difference for observations of the same age but in two adjacent years: $\Delta_c \Delta_a SD_{stj} = \Delta_c \mathrm{SD}_{stj} - \Delta_c \mathrm{SD}_{st'j}$. Therefore, I can test the hypothesis that the second derivative of time effects, $(\psi_{t+2} - \psi_{t+1}) - (\psi_{t+1} - \psi_t)$, is zero. As a result, I cannot reject the linear hypothesis except for $t$ equal to 2003 and 2004, meaning linear specifications are good enough to estimate the time effects at all other sample years. This McKenzie test result is intuitive by looking at Figure 1.18. We cannot reject the linear hypothesis at year 2003 and 2004, because the MRPK dispersion deviates from the linear fit in 2004, thus decreasing relatively significantly between year 2004 and 2005. Actually, even at this outlier, the deviation from the linear trend is only around 0.01 point in 1.18. This deviation from the linear trend is relatively small compared to the age or time effects I estimated, which have magnitudes around 10 times larger. So I conclude that the linear restriction in the first approach is a reasonable approximation.

Specifically, the two restrictions I use to get the results in Figure 1.6 are the following:

**Restriction 1 (All Decline due to Cohort Effects):**

$$g_\psi = 0, \;\; g_\chi = g_M$$

By the definition of $g_\psi$, this restriction implies $\sum_{t=0}^{T} \psi_t = 0$, meaning that the year effects $g_\psi$ only capture the cyclical variations and are orthogonal to the time trend. This restriction is the same as illustrated by [46, pp. 123 - 127].

**Restriction 2 (All Decline due to Time Effects):**

$$g_\psi = g_M, \;\; g_\chi = 0$$

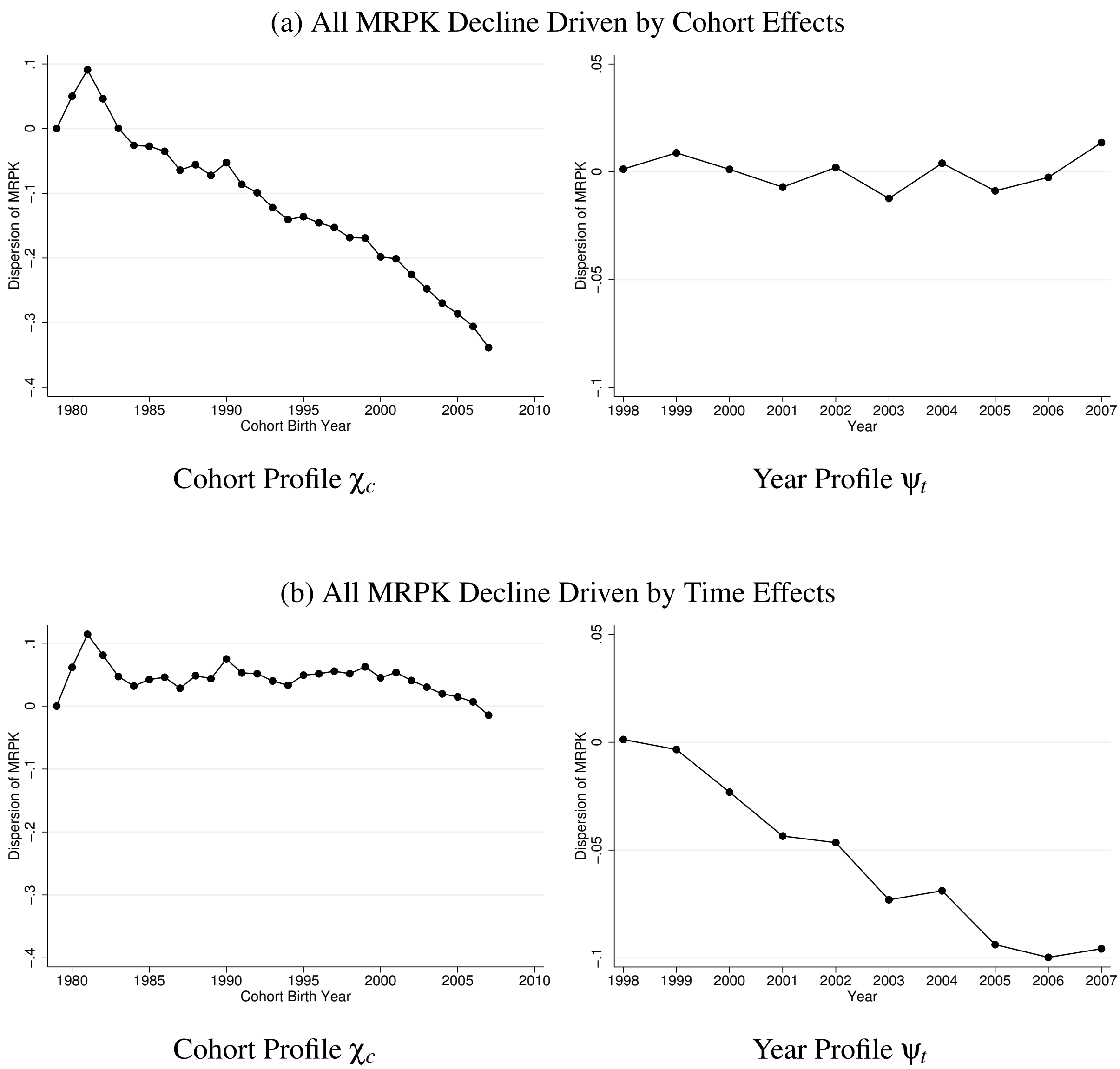
This restriction actually implies the linear restriction $\sum_{t=0}^{T} \bar{X}_t t = 0$, or

$$\sum_{t=0}^{T} \sum_{c \in C_t} \frac{\Phi_{ct}}{\bar{\Phi}_t} \chi_c t = 0.$$

Note the term $\Phi_{ct}$ enters this restriction, which requires estimating equation (1.5). In practice, I use an iterative algorithm to meet this restriction.

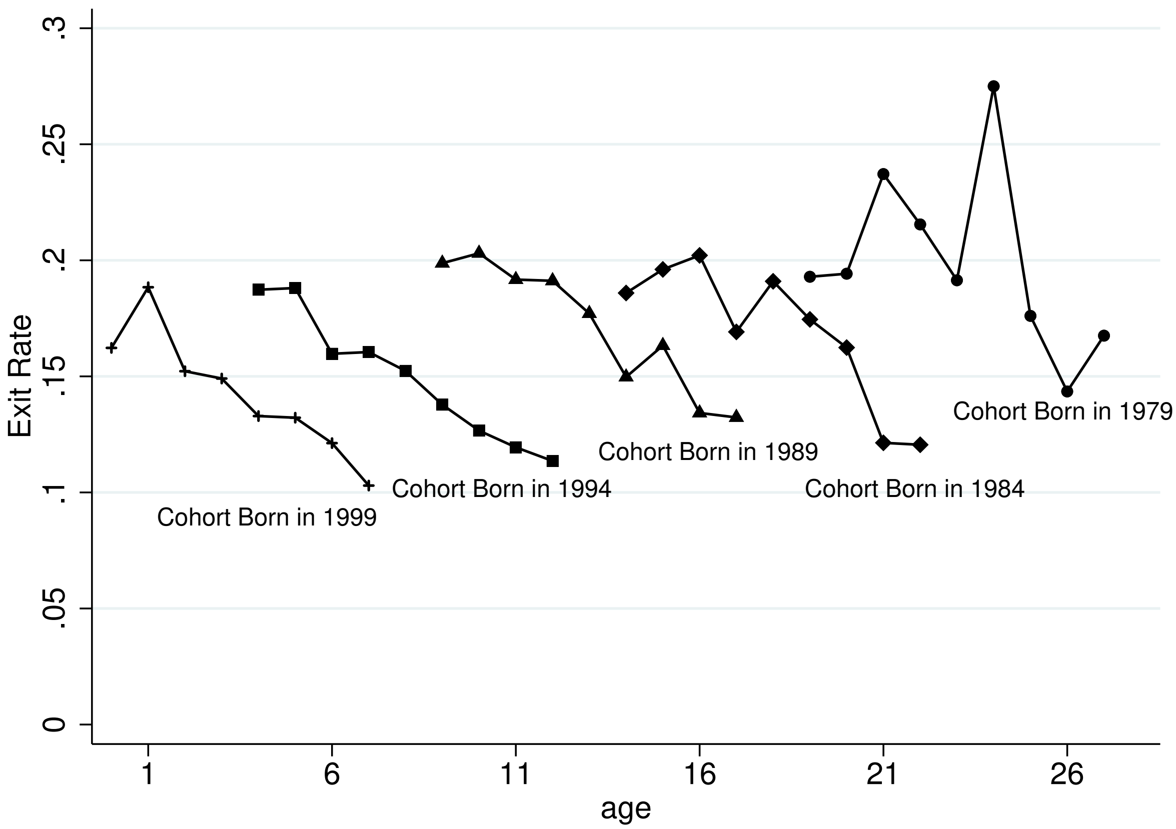
Figure 1.20 plots the estimates of the cohort and time effects under the two restrictions above. The top-panel results impose Restriction 1, so we see a declining trend in the cohort effects, but the time effects are relatively flat. The bottom-panel results impose Restriction 2, so the cohort effects are relatively flat but the time effects have a declining trend. Note the cohort and time variations are large, with the largest magnitudes at -0.3 for cohort effects and -0.1 for the year effects.

In addition, Figure 1.21 plots the exit rates by firm-cohort age of the same firm cohorts in Figure 1.2 after removing zero-sum year effects using this methodology.



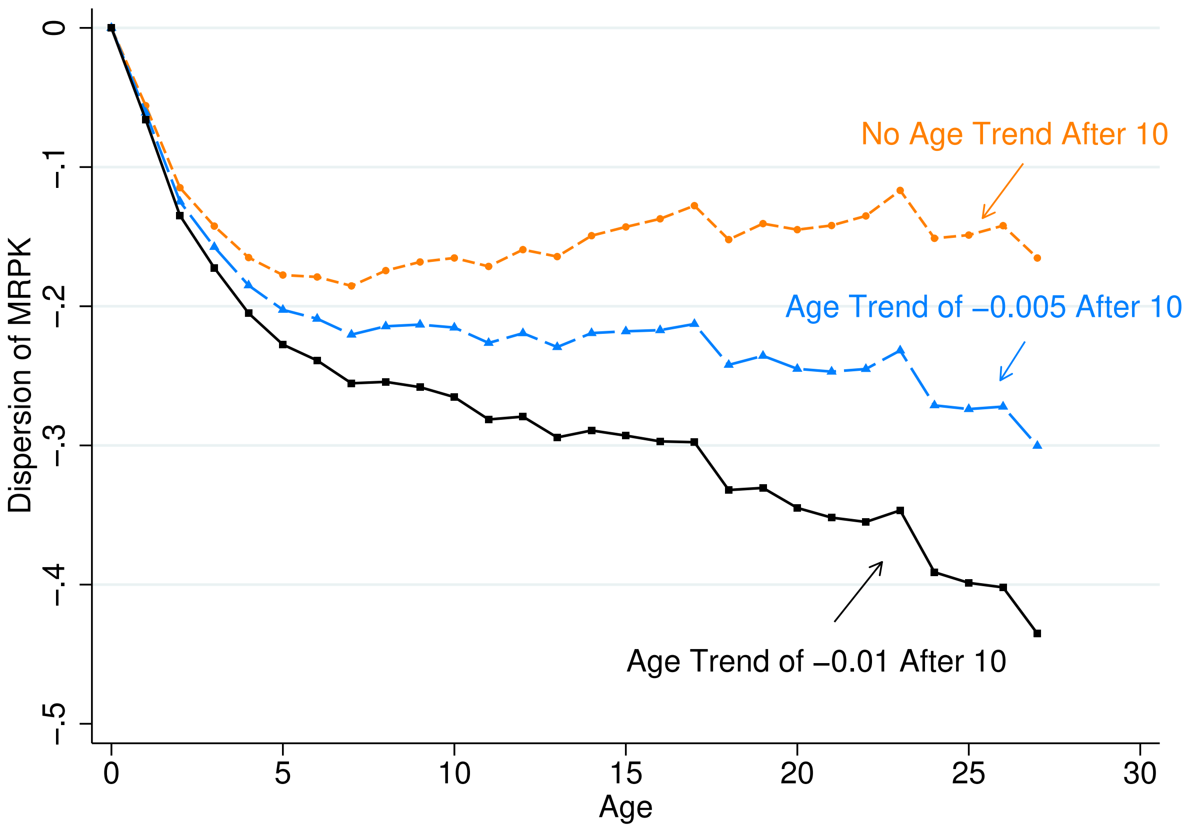

Note: This figure plots the MRPK dispersion by birth year of the firm cohorts and by calendar year estimated using the first alternative approach. The top panel shows the dispersion-cohort and dispersion-year profiles estimated in equation (1.5) using Restriction 1: $g_\psi = 0$. The bottom panel shows the dispersion-cohort and dispersion-year profiles estimated in equation (1.5) using Restriction 2: $g_\chi = 0$.

**Figure 1.20**: MRPK Dispersion by Cohort and Year in Alternative Approach One



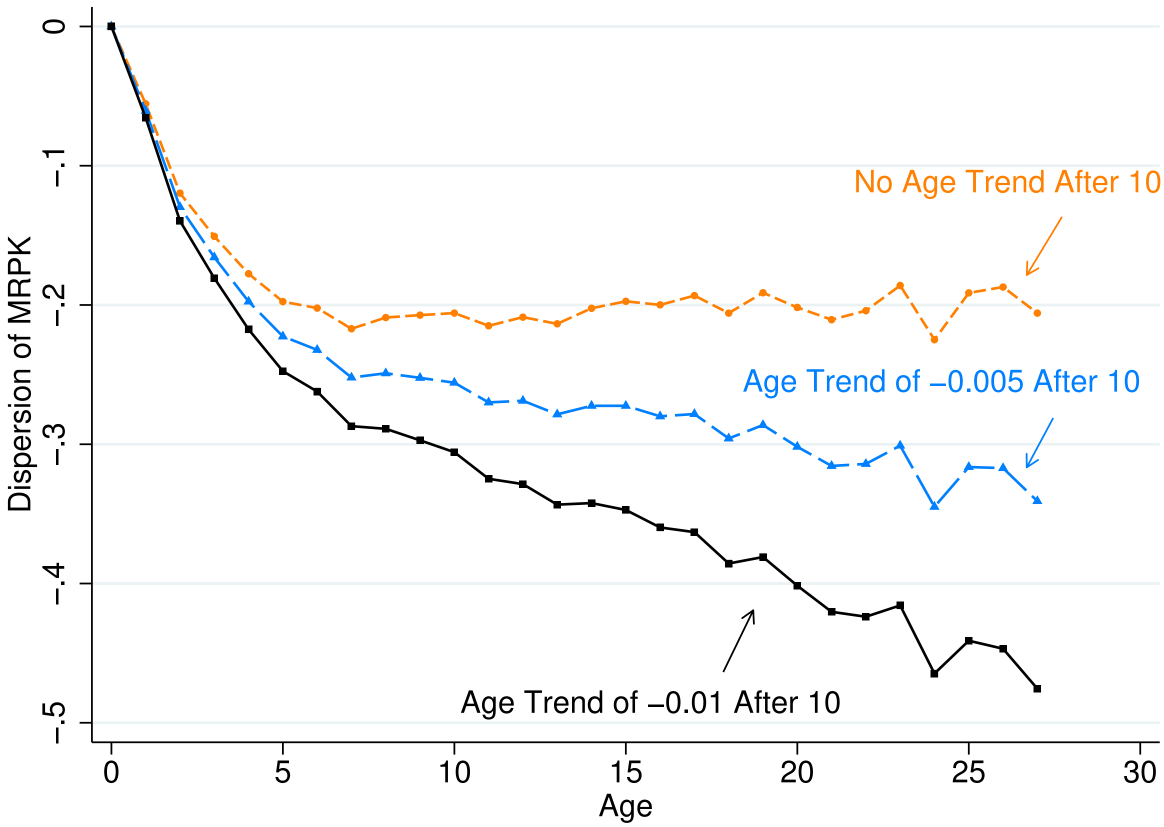

Note: This figure plots the exit rates by cohorts born in 1979, 1984, 1989, 1994, and 1999, respectively, after removing the zero-sum year effects following [46].

**Figure 1.21**: Exit Rates by Cohort




Note: This figure plots the estimated profile of MRPK dispersion by firm-cohort age in equation (1.6) using the second approach, which assumes (a) no trend in the age effects on MRPK dispersion after age 10 (dashed orange line with circle markers), (b) a small decreasing trend of 0.005 point per age after age 10 (long-dashed blue lines with triangle markers), (c) a moderate decreasing trend of 0.01 points per age after age 10 (solid black lines with square markers).

**Figure 1.22**: Dispersion Profiles over Age, Robustness with Volatility of Productivity




Note: This figure plots the estimated profile of MRPK dispersion by firm-cohort age in equation (1.6) using the second approach when restricting the sample to only non-state firms. It plots the estimation results assuming (a) no trend in the age effects on MRPK dispersion after age 10 (dashed orange line with circle markers), (b) a small decreasing trend of 0.005 point after age 10 (long-dashed blue lines with triangle markers), (c) a moderate decreasing trend of 0.005 point after age 10 (solid black lines with square markers).

**Figure 1.23**: Dispersion Profiles over Age, Robustness with Only Non-state Firms

## 1.12 Aggregate Productivity and MRPK Dispersion

Substituting the optimal labor choice of $n_{it}(z_{it},k_{it}) = \left(\alpha_2 \frac{e^{z_{it}} k_{it}^{\alpha_1}}{w}\right)^{\frac{1}{1-\alpha_2}}$ into the production function gives

$$y_{it} = \left(\frac{\alpha_2}{w}\right)^{\frac{\alpha_2}{1-\alpha_2}} e^{z_{it}\frac{1}{1-\alpha_2}} k_{it}^{\frac{\alpha_1}{1-\alpha_2}}. \tag{1.15}$$

Meanwhile, the labor market clearing condition requires that the fixed labor supply equals the aggregate labor demand $N = \int n_{it} di = \left(\frac{\alpha_2}{w}\right)^{\frac{1}{1-\alpha_2}} \int \left(e^{z_{it}} k_{it}^{\alpha_1}\right)^{\frac{1}{1-\alpha_2}} di$, so that $\left(\frac{\alpha_2}{w}\right)^{\frac{1}{1-\alpha_2}} = \frac{N}{\int \left(e^{z_{it}} k_{it}^{\alpha_1}\right)^{\frac{1}{1-\alpha_2}} di}$. Substituting this expression in $y_{it}$ gives

$$y_{it} = \frac{e^{z_{it}\frac{1}{1-\alpha_2}} k_{it}^{\frac{\alpha_1}{1-\alpha_2}} N^{\alpha_2}}{\left(\int \left(e^{z_{it}} k_{it}^{\alpha_1}\right)^{\frac{1}{1-\alpha_2}} di\right)^{\alpha_2}}.$$

Further taking derivative with respect to $k_{it}$ yields $MRPK_{it} = \frac{\alpha_1}{1-\alpha_2} \frac{e^{z_{it}\frac{1}{1-\alpha_2}} k_{it}^{\frac{\alpha_1+\alpha_2-1}{1-\alpha_2}} N^{\alpha_2}}{\left(\int \left(e^{z_{it}} k_{it}^{\alpha_1}\right)^{\frac{1}{1-\alpha_2}} di\right)^{\alpha_2}}$, which can be rearranged to express $k_{it}$ in terms of $MRPK_{it}$:

$$k_{it} = \left(\frac{\frac{\alpha_1}{1-\alpha_2} e^{z_{it}\frac{1}{1-\alpha_2}}}{MRPK_{it}}\right)^{\frac{1-\alpha_2}{1-\alpha_1-\alpha_2}} \cdot \left(\frac{N}{\int \left(e^{z_{it}} k_{it}^{\alpha_1}\right)^{\frac{1}{1-\alpha_2}} di}\right)^{\frac{\alpha_2(1-\alpha_2)}{1-\alpha_1-\alpha_2}}.$$

Meanwhile, capital market clearing condition implies

$$K = \int k_{it} di = \left(\frac{\alpha_1}{1-\alpha_2}\right)^{\frac{1-\alpha_2}{1-\alpha_1-\alpha_2}} \left(\frac{N}{\int \left(e^{z_{it}} k_{it}^{\alpha_1}\right)^{\frac{1}{1-\alpha_2}} di}\right)^{\frac{\alpha_2(1-\alpha_2)}{1-\alpha_1-\alpha_2}} \int \left(\frac{e^{z_{it}\frac{1}{1-\alpha_2}}}{MRPK_{it}}\right)^{\frac{1-\alpha_2}{1-\alpha_1-\alpha_2}} di.$$

Cancelling out the term with $N$ in the last two expressions yields

$$k_{it} = \frac{\left(\frac{e^{z_{it}\frac{1}{1-\alpha_2}}}{MRPK_{it}}\right)^{\frac{1-\alpha_2}{1-\alpha_1-\alpha_2}}}{\int\left(\frac{e^{z_{it}\frac{1}{1-\alpha_2}}}{MRPK_{it}}\right)^{\frac{1-\alpha_2}{1-\alpha_1-\alpha_2}}di}K.$$

Now substituting $k_{it}$ in terms of $K$ into the expression of $y_{it}$ and rearranging gives

$$y_{it} = \frac{\frac{e^{z_{it}\frac{1}{1-\alpha_1-\alpha_2}}MRPK_{it}^{-\frac{\alpha_1}{1-\alpha_1-\alpha_2}}}{\left(\int\left(\frac{e^{z_{it}\frac{1}{1-\alpha_2}}}{MRPK_{it}}\right)^{\frac{1-\alpha_2}{1-\alpha_1-\alpha_2}}di\right)^{\frac{\alpha_1}{1-\alpha_2}}}}{\left(\frac{\int e^{z_{it}\frac{1}{1-\alpha_1-\alpha_2}}MRPK_{it}^{-\frac{\alpha_1}{1-\alpha_1-\alpha_2}}di}{\left(\int\left(\frac{e^{z_{it}\frac{1}{1-\alpha_2}}}{MRPK_{it}}\right)^{\frac{1-\alpha_2}{1-\alpha_1-\alpha_2}}di\right)^{\frac{\alpha_1}{1-\alpha_2}}}\right)^{\alpha_2}}K^{\alpha_1}N^{\alpha_2}.$$

Finally, aggregating the revenue output $y_{it}$ gives $Y = \int y_{it}di = ZK^{\alpha_1}N^{\alpha_2}$, where the aggregate productivity is

$$Z = \left(\frac{\int e^{z_{it}\frac{1}{1-\alpha_1-\alpha_2}}MRPK_{it}^{-\frac{\alpha_1}{1-\alpha_1-\alpha_2}}di}{\left(\int\left(\frac{e^{z_{it}\frac{1}{1-\alpha_2}}}{MRPK_{it}}\right)^{\frac{1-\alpha_2}{1-\alpha_1-\alpha_2}}di\right)^{\frac{\alpha_1}{1-\alpha_2}}}\right)^{1-\alpha_2}.$$

Taking the log of the expression above gives

$$\begin{aligned} z =&(1-\alpha_2)\left[\ln\left(\int e^{z_{it}\frac{1}{1-\alpha_1-\alpha_2}}MRPK_{it}^{-\frac{\alpha_1}{1-\alpha_1-\alpha_2}}di\right)\right.\\ &\left.-\frac{\alpha_1}{1-\alpha_2}\ln\left(\int e^{z_{it}\frac{1}{1-\alpha_1-\alpha_2}}MRPK_{it}^{-\frac{1-\alpha_2}{1-\alpha_1-\alpha_2}}di\right)\right]. \end{aligned}$$

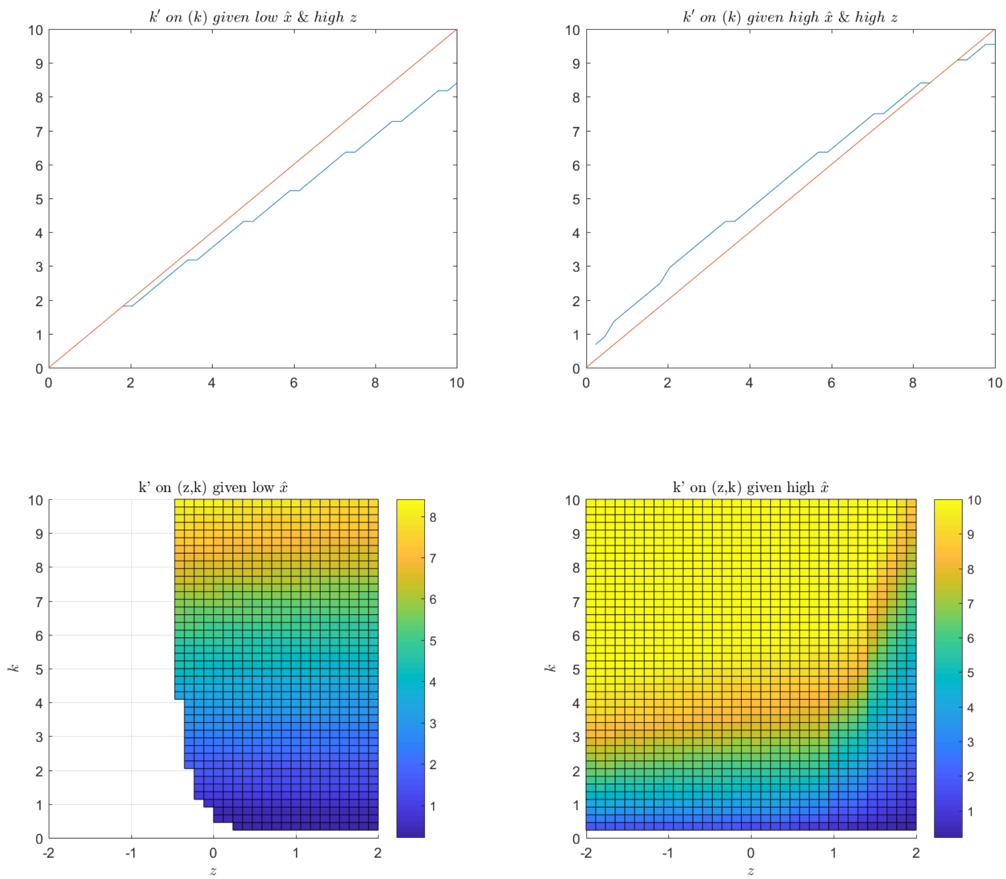
Expanding the two terms in the brackets respectively,

$$\ln\left(\int e^{z_{it}\frac{1}{1-\alpha_1-\alpha_2}} MRPK_{it}^{-\frac{\alpha_1}{1-\alpha_1-\alpha_2}} di\right) = \frac{\bar{z}-\alpha_1\overline{mrpk}}{1-\alpha_1-\alpha_2} + \frac{\sigma_z^2+\alpha_1^2\sigma_{mrpk}^2-2\alpha_1\sigma_{mrpk,z}}{2(1-\alpha_1-\alpha_2)^2},$$

$$\begin{aligned}\ln\left(\int e^{z_{it}\frac{1}{1-\alpha_1-\alpha_2}} MRPK_{it}^{-\frac{1-\alpha_2}{1-\alpha_1-\alpha_2}} di\right) =& \frac{\bar{z}-(1-\alpha_2)\overline{mrpk}}{1-\alpha_1-\alpha_2} \\ &+ \frac{\sigma_z^2+(1-\alpha_2)^2\sigma_{mrpk}^2-2(1-\alpha_2)\sigma_{mrpk,z}}{2(1-\alpha_1-\alpha_2)^2}.\end{aligned}$$

Finally, combining them into the expression of $z$ reveals the relationship between the productivity loss ($z-z^*$) and dispersion in MRPK ($\sigma_{mrpk}^2$):

$$\begin{aligned}z &= (1-\alpha_2)\left[\bar{z}+\frac{\sigma_z^2-\alpha_1\sigma_{mrpk}^2}{2(1-\alpha_1-\alpha_2)}\right] \\ &= z^* - \frac{\alpha_1(1-\alpha_2)}{2(1-\alpha_1-\alpha_2)}\sigma_{mrpk}^2.\end{aligned}$$




Note: This figure plots the policy function of $k_{i,t+1}$ for age-one firms. The top two panels fix the state variable $z_{it}$ at a relatively high level and plot the firm choice of $k_{i,t+1}$ against $k_{it}$ with low belief (left) and high belief (right); the bottom panels plot $k_{i,t+1}$ on the space of $(k_{it}, z_{it})$ with low belief (left) and high belief (right). Blank space in the top- and bottom-left two panels represent missing $k_{i,t+1}$ values when the firm chooses to exit the market.

**Figure 1.24**: Policy Function Given Low and High Beliefs in Model

# Chapter 2

# Unemployment and Development

For helpful comments we thank Gary Fields, Chris Huckfeldt, Ben Moll, Andi Mueller, Tommaso Porzio, Guillaume Rocheteau, Venky Venkateswaran, Mike Waugh, Erin Wolcott, Randy Wright and seminar/conference audiences at Cornell, Harvard/MIT, NYU, Rochester, Midwest Macro (Pittsburgh), SED (Mexico City), the China Conference on Development and Growth (Wuhan), the MacCaLM Workshop (Edinburgh), Trinity College Dublin, UCSD and the West Coast Search Conference (Irvine). All potential errors are our own.

## 2.1 Introduction

No single measure of labor-market performance receives more attention among academics and policy makers than the unemployment rate. It is well known, for example, that average unemployment rates are higher in Western Europe than in the United States and Japan. But there is little systematic evidence about how average unemployment rates vary across the entire world income distribution. Internationally comparable data from the poorest countries of the world are particularly lacking. This lack of data hampers research on the determinants of national average unemployment levels, and on the link between unemployment and development, to name two important topics.

This paper attempts to fill this gap by building a database of national unemployment rates covering countries of all income levels. To do so, we draw on evidence from 199 household surveys from 84 countries spanning 1960 to 2015. The database covers numerous rich countries and around two dozen nations from the bottom quartile of the world income distribution. Since measures of employment and job search vary across surveys, we divide the data into several tiers based on scope for international comparability. We then construct unemployment rates at the aggregate level and for several broad demographic groups, and compare how they vary with average income.

We find, perhaps surprisingly, that unemployment rates are *increasing* in GDP per capita. This finding is present for men and for women, for all broad age groups, within urban and rural areas, and across all comparability tiers of our data. For prime-aged adults, a regression of the country average unemployment rate on log GDP per capita yields a statistically significant positive coefficient of 1.8 percent. Our findings contrast with the (scarce) existing evidence in the literature, and in particular, the work of [36], who finds in an earlier database that unemployment rates do not systematically vary with income per capita.

In addition, we document that unemployment patterns across countries differ markedly by education level. Among high-educated workers (secondary school or more), unemployment rates do not vary systemically with GDP per capita. Among low-educated workers, in contrast, unemployment rates are substantially higher in rich countries. Regressing the country average high-educated unemployment on log GDP per capita yields an insignificant slope coefficient of 0.5 percent, whereas the slope coefficient for the low-educated is a significant 3.2 percent. Our data imply that in rich countries, low-educated workers are more likely than high-educated workers to be unemployed. In poor countries, the opposite is true, and unemployment is concentrated among the high-educated.

To understand these facts, we build a simple two-sector model with frictional labor markets, based on [47] and [103], and heterogeneous workers that sort by ability as in [125]. In the modern sector, labor markets are governed by search frictions, and worker productivity is determined by a worker's ability level. In the traditional sector, workers are self-employed and do not need to search for jobs; however, productivity is independent of ability. Outputs of the modern and traditional sectors are perfect substitutes, and firms operate competitively in the modern sector, with unrestricted entry. Countries differ exogenously in modern-sector productivity, with a single traditional-sector technology available to all countries. This assumption builds on the mounting evidence that cross-country productivity differences are skill-biased, as opposed to skill neutral (see, e.g., [37, 93, 78, 71]).

Our simple model has several main theoretical predictions that are qualitatively consistent with the facts we document. First, as modern-sector productivity increases, the traditional sector shrinks, as progressively less able workers sort into the modern sector. Second, as modern-sector productivity increases, the aggregate unemployment rate increases. This is because as the modern sector expands, a greater fraction of workers now search for jobs in frictional labor markets rather than working in self-employment. Moreover, the job-finding rate falls in equilibrium, since average ability is lower in the modern sector. Third, as productivity increases, unemployment rates rise faster for less able than for more able workers, since a greater share of less able workers are drawn into job search. This third prediction is consistent with the rising ratio of unemployment for low- to high-educated workers with GDP per capita that we document.

To assess the model's quantitative predictions, we extend the simple model in several ways so as to be consistent with salient features of the cross-country data. In particular, we allow modern and traditional sector outputs to be imperfect substitutes, and we allow countries to differ exogenously in both traditional- and modern-sector productivity. We also allow for two education groups, with the distribution of ability for the high-educated group stochastically dominating that of the low-educated. We calibrate the distribution of ability using moments of the U.S. wage distribution, and parameterize other aspects of the model to match key moments of the U.S. labor market—in particular the average unemployment rate and the ratio of the unemployment rate for low- to high-educated workers.

Our main quantitative experiment lowers productivity in the modern and traditional sectors, as well as the fraction of high-educated workers, and then computes how the model's predictions for unemployment – in the aggregate and by education level –vary with GDP per capita. We discipline the cross-country values of modern-sector productivity to match GDP per capita levels across the world income distribution, and we discipline traditional-sector productivity to match the relative prices of traditional goods. We proxy traditional sector employment in the data by the set of workers who are self-employed without paid employees, and who work in low-skilled occupations. Not surprisingly, this share is strongly decreasing in GDP per capita, ranging from around three quarters of the workforce in poor countries to less than three percent in the richest countries.[1]

The calibrated model predicts that unemployment rates are increasing in GDP per capita, as in the data, though the model underpredicts the magnitude of the relationship. Compared to the observed 1.8 percentage-point increase in unemployment for an increase in one log point of GDP per capita, the model predicts an increase of 0.5 percent. For unemployment by education, the model correctly predicts that the ratio of low- to high-educated unemployment is increasing in GDP per capita. Yet it again underpredicts the magnitude of the relationship, with a semi-elasticity of 0.47 in the data compared to 0.25 in the model. We conclude that our mechanism explains 30 percent of the relation between aggregate unemployment and average income, and 53 percent of the relation between the unemployment ratio and average income. Furthermore, the model's predicted share of employment in the traditional sector by GDP per capita corresponds closely with the data. We also show that our results are sensitive to one parameter value in particular: the elasticity of substitution between modern and traditional sector outputs, which governs the strength of our mechanism, the decline of the traditional sector.

As an alternative and complementary theory, we incorporate the less generous unemployment benefits of poor countries relative to richer countries. In the model, lower unemployment benefits in poorer countries discourage search, thus lowering unemployment rates in equilibrium. We find that adding this alternative mechanism increases the explanatory power of our quantitative model from 30 percent to 41 percent of the slope of the aggregate unemployment rate in GDP per capita. On the other hand, it offers little additional explanatory power for the relation between the ratio of low- to high-educated unemployment and income. We conclude that our quantitative model explains a substantial fraction of the cross-country unemployment patterns that

[1]Note that this decrease in the traditional sector after excluding agriculture is of similar magnitude, ranging from around half of the workforce to less than two percent. Thus, the traditional sector is not simply agriculture, but represents the unskilled self-employment that is widespread throughout developing economies (see e.g., ([58, 127, 53]).

we document, but that even including the less generous social security nets of poorer countries, there is a lot left unexplained by the model.

We close the paper by presenting historical data on unemployment from the United States and four other advanced countries for which long time series on unemployment are available: Australia, France, Germany and the United Kingdom. We ask whether unemployment rates are higher now than they were before World War I, which is the earliest period for which unemployment data are available, to our knowledge. We find that for all countries, average unemployment rates are indeed higher now than they were before World War I, and for four of the five countries, the difference is statistically significant. Using the U.S. data, which we have at a more disaggregated level, we ask in addition whether unemployment is particularly higher now for the less-educated. We find that average unemployment has indeed risen faster for the less-educated than for the more-educated, at least since 1940. In 1940, the less-educated were about 1.5 times as likely to be unemployed as the more-educated. Today, the ratio is close to 2.5. We conclude that historical unemployment data are broadly consistent with our cross-country findings, suggesting that unemployment is largely a feature of advanced economies, rather than a by-product of under-development.

**Related Literature.** Most of the literature on average unemployment differences across countries has focused on Europe and the United States (see, e.g., [24, 107, 91]). The few studies that have addressed unemployment across a wider range of income levels have come to contradictory conclusions, most likely due to a lack of comparable cross-country data. [17] compile World Bank unemployment data that suggest a decreasing pattern of unemployment in income per capita, though their data are largely from middle-income and richer countries. Perhaps the most systematic look at aggregate unemployment rates across countries is by [36], who draws on a 1996 World Bank dataset covering 60 countries. These data show no correlation between GDP per capita and average unemployment, though they cover just three countries in the bottom half of the world income distribution. Older studies did not have sufficient data points to draw firm conclusions about cross-country patterns, but tended to find that unemployment rates in developing economies studies were not that different from those of richer economies (see, e.g., [55, 132, 134, 54]). More recently, [112] draws on surveys from 68 countries to study the relationship between self-employment and the ratio of unemployment to wage employment. His explanation emphasizes differences in labor market frictions across countries, whereas our theory emphasizes different forces altogether.

Our paper is closely related to the growing literature on structural change, though our two sectors do not fit neatly into the standard agriculture-manufacturing-services division (used by e.g., [49, 69, 97]).[2] In our modern and traditional sectors, we emphasize skilled wage employment versus unskilled self-employment, both of which can be present within the agriculture, manufacturing, and service sectors. In this way, our sectors are closer to the split between high-educated services and low-educated services taken by [32] and [34], though their models focus on non-homothetic preferences, which play no role in our theory.

By emphasizing the transition from self-employment to wage employment in frictional labor markets, our paper builds on the macroeconomic literature on home production and its role in the development process. This transition to market production with development is a key theme in the model of [106], for example. [60] argue that measured output differences across countries may be overstated due to missing home production in poorer countries. Similarly, [109] show that policies that distort capital accumulation can lead to bigger output losses once a home production sector is introduced into a standard neoclassical growth model, since capital distortions encourage producers to move into self-employment. Empirically, [30] show that the share of household production in total hours decreases with GDP per capita. None of these studies focuses on the

[2] Other multi-sector models in macro split the economy into the consumption vs investment sectors ([120, 74]), goods vs service sectors (e.g., [25]), urban vs rural areas ([38, 141]), or agriculture vs non-agriculture sectors (e.g., [1, 86, 110]). Our modern-traditional division does not correspond cleanly to these splits either.

link between unemployment and development, however.

Finally, our paper builds on the old literature on two-sector models in development, particularly [89] and [68]. However, our model is focused on the determinants of actual measured unemployment (often called "open unemployment"), as opposed to "underemployment" or "disguised unemployment," which corresponds to some extent to our traditional sector. Negative selection into our traditional sector is also quite related to the negative selection into the "informal sector" as characterized by [117], [85, 84] and others. Unlike [68], the urban-rural divide plays no role in our theory; we find similar unemployment patterns in both rural and urban areas and, hence, abstract from them.

## 2.2 Data

This section describes the household survey data that we use to measure unemployment in the aggregate and by demographic group across our set of countries.

### 2.2.1 Data Sources

Our data come from household surveys or censuses that are nationally representative. Many, but not all, are available from the International Integrated Public Use Microdata Surveys (IPUMS) ([101]) or the World Bank's Living Standards Measurement Surveys (LSMS). Tables 2.12, 2.13 and 2.14 in the Appendix list the full set of surveys employed, plus their sources. The key benefit of nationally representative surveys, as opposed to (say) administrative records on unemployment, is that they cover all individuals, including the self-employed. In total, our analysis includes 199 country-year surveys, covering 84 countries, and spanning 1960 to 2015.

Most of our data come from the 1990s and 2000s.[3]

To measure GDP per capita, we divide output-side real GDP at chained PPPs (in 2011 US$) by population, both taken from the Penn World Tables 9.0. Unlike in previous studies, our data have a high representation of the world's poorest countries, with 23 countries from the bottom quartile of the world income distribution, and 27 from the second quartile.

In our main analysis, we restrict attention to prime-aged adults (aged 25-54) of both sexes. We also report our results for males and females separately, for broader age groups, and for urban and rural regions. Throughout, we exclude those with missing values of key variables and those living in group quarters. We use sample weights whenever they are available.

### 2.2.2 Unemployment Definition and Data Tiers

We define an unemployed person as one who (1) is not employed, and (2) has searched recently for a job. We define employment following the U.N. System of National Accounts as "all persons, both employees and self-employed persons, engaged in some productive activity that falls within the production boundary of the SNA" [135]. Thus, we count those working in self-employment as employed. We define the unemployment rate as the ratio of unemployed workers to employed plus unemployed workers.[4]

The key measurement challenge we face is that not all surveys allow us to define unemployment in exactly the same way. To ensure that our cross-country comparisons are as

[3] [48] use surveys from 13 countries to document high-frequency labor market patterns in the urban areas of middle and high income countries. Our paper covers many more low income countries, whereas their study brings in repeated observations from the same individuals.

[4] The BLS *Handbook of Methods* defines an unemployed individual as one who (1) is not employed, (2) has searched recently for a job, and (3) is "available to work" ([136]). However, only 49 of our 199 country-year surveys asked whether the interviewee is "available for work" in some way.

informative as possible, we divide the surveys into tiers, based on their international comparability. Tier 1 has the highest scope for comparability, followed by Tier 2 and then Tier 3. We describe these further below.

In Tier 1 and Tier 2 countries, employment specifically covers all economic activities that produce output counted in the National Income and Product Accounts (NIPA). In other words, employment specifically comprises wage employment, self-employment or work at a family business or farm, whether or not the output is sold or consumed directly.[5] With regard to recent job search, Tier 1 includes surveys in which workers who searched did so either in the last week or the last four weeks. Tier 2 includes surveys in which workers are searching "currently" (without specifying a time frame) or in some time period other than the last week or last four weeks, such as the last two months.

In Tier 3 countries, the employment question has lower scope for comparability. It may, for example, consider those working for their own consumption or those not working for a monetary wage as non-employed. It may include a minimum number of hours worked, or cover only a specific period of time, such as the last seven days. Appendix Table 2.14 lists the way in which each country in Tier 3 has a non-standard employment question. In terms of job search, Tier 3 countries cover any time.

All in all, our dataset consists of 129 Tier 1 surveys, 40 Tier 2 surveys and 30 Tier 3 surveys. In our empirical findings below, we begin with data from all tiers, which maximizes the number of observations available. We then restrict attention to Tier 1 first, followed by Tiers 1 and 2, to explore how our results change when we take into consideration a smaller but more comparable set of countries.

[5]See e.g. [59] for a more detailed treatment of which outputs are covered in NIPA. Not counted is work on home-produced services such as cooking, cleaning or care of one's own children. Home-produced services are not counted in NIPA, and previous studies of time use, such as [3], [115] and [22], treat these categories as "home production" rather than as work.

### 2.2.3 Comparison to ILO and World Bank Data

Two readily downloadable sources of data on unemployment rates at the country level are the “ILO modeled estimates” from the International Labor Office (ILO), and the World Bank’s World Development Indicators (which are in fact derived directly from the ILO). However, many of the ILO’s modeled estimates are, by definition, modeled or imputed rather than computed directly from an underlying survey. Even by the ILO’s own admission, the modeled estimates are fraught with serious non-comparabilities. For example, some estimates cover only main cities or metropolitan areas, while others use non-standard employment definitions that exclude self-employed workers or first-time job seekers.

Acknowledging the lack of international comparability in its full database, the ILO also publishes “ILO-comparable” unemployment rates from 30 countries, which are always based on a household labor force survey ([87]). Unfortunately, the ILO-comparable unemployment rates have very limited coverage of the bottom half of the world income distribution, covering just one such country. Therefore, the ILO-comparable unemployment dataset is ill-suited to answer the question of how average unemployment rates vary between poor and rich countries. In addition, it does not provide disaggregated unemployment rates, such as by education level, which we show are crucial to understanding the aggregate patterns.

If one nonetheless uses these ILO data to estimate how average unemployment rates vary with income per capita, one will find a statistically insignificant or negative relationship. Using the ILO modeled unemployment estimates, a regression of the 2014 unemployment rate on log GDP per capita yields a slope coefficient of 0.02 with a p-value of 0.96. This lack of correlation between unemployment and income is comparable to what [36] found. With the much smaller ILO comparable database, available from 1994 to 2003, the regression coefficient is -3.44 with a p-value of 0.01. Thus, either of the readily available ILO unemployment databases paint a misleading picture of how unemployment rates vary with income level.

## 2.3 Empirical Findings

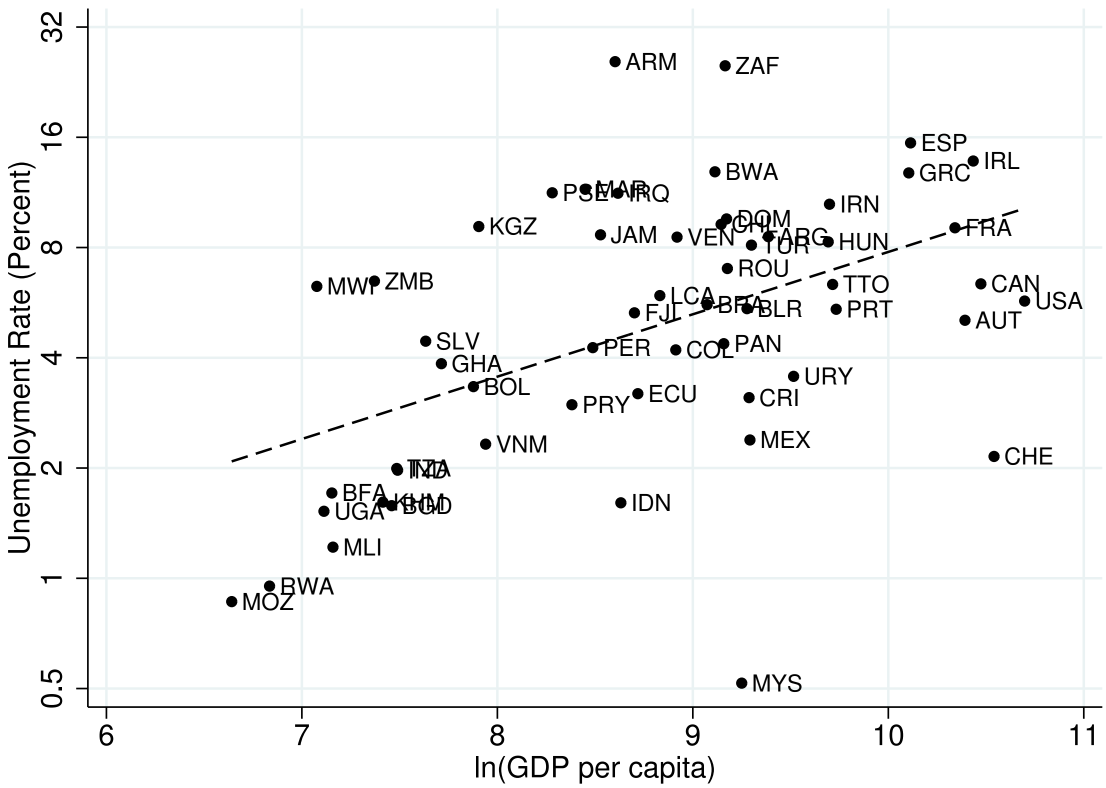
In this section, we report how average unemployment rates vary with GDP per capita. We first compare aggregate unemployment rates, and then look beneath the surface at unemployment by sex, by age group and by rural-urban status.

### 2.3.1 Aggregate Unemployment Rate

Figure 2.1 plots the country average unemployment rate for prime-aged adults (on a log base 2 scale) against log GDP per capita. The figure includes countries from all three tiers with at least two years of data. The dotted black line – the linear regression line – shows a substantial positive slope. The slope coefficient for a regression of the unemployment rate in natural units on log GDP per capita is 1.8 and is statistically significant at the one-percent level. Taking simple averages by country income quartile, the bottom (poorest) quartile has an average unemployment rate of 2.5 percent. By contrast, the top (richest) quartile has an average unemployment rate of 8.7 percent.

Besides the positive slope, Figure 2.1 highlights the large variation in average unemployment rates within each income group. To what extent does this variation simply reflect measurement error? To what extent does the correlation of unemployment rates and GDP per capita survive once we restrict attention to more comparable data?

To help answer these questions, we report the slope coefficient of average unemployment




Note: This figure plots the average unemployment rate for prime-aged adults in each country with at least two observations across all years of data from all tiers.

**Figure 2.1**: Unemployment Rates by GDP per capita

on log GDP per capita using various alternative cuts of the data. The first data column of Table 2.1 reports these slopes. When considering all 199 country-year surveys separately, the slope falls somewhat to 1.1, and is again statistically significant at the one-percent level. When using only Tier 1 surveys, the slope coefficient becomes 1.4, and with Tier 1 and 2 surveys, the slope becomes 1.3. We conclude that the pattern of increasing unemployment is not an artifact of our choice of countries in the main analysis.

### 2.3.2 Unemployment Rate by Education Level

In this subsection, we report our findings by education level, which are helpful in accounting for the aggregate patterns we document above. Later we present results by other demographic groups. We define two education groups, which can be measured consistently across nearly all of our countries. The *low education* group are those that did not finish secondary school. This could mean no school, some or all of primary school completed, some secondary education, or some other specialty education that lasts less than 12 years. The *high education* group are those that completed secondary school or more. This could mean exactly secondary school, some college or university completed, or an advanced degree.

**Table 2.1**: Slope Coefficients of Unemployment Rate on GDP per capita

| | All Workers | N | Low Education | High Education | Ratio |
|---|---|---|---|---|---|
| All surveys | 1.1*** (.3) | 199 | 2.9*** (.4) | -.2 (.3) | .50*** (.03) |
| Country average | 1.8*** (.5) | 55 | 3.2*** (.6) | .5 (.4) | .48*** (.05) |
| Only Tier 1 surveys | 1.4*** (.3) | 127 | 3.2*** (.4) | .4 (.3) | .48*** (.03) |
| Only Tier 1 + 2 surveys | 1.3*** (.3) | 167 | 2.9*** (.4) | -.1 (.3) | .50*** (.03) |

Note: The table reports the slope coefficient from a regression of the prime-age unemployment rate on log GDP per capita and a constant. ***, ** and * indicate statistical significance at the 1-percent, 5-percent and 10-percent levels. The first row includes all surveys in our data. The second row includes one observation per country: the average unemployment rate for those with at least two observations across all years from all tiers. The third row includes only Tier 1 surveys. The fourth row includes only Tier 1 and Tier 2 surveys. Surveys with missing education level data are dropped in the last three columns.

Figure 2.2 plots the unemployment rates for prime-aged adults by education group. As before, we plot the unemployment rates in log base 2 and GDP per capita in natural logs. As one can see, the patterns differ sharply by group. For the low-educated group, unemployment is strongly increasing in GDP per capita. For the high-educated group, unemployment rates are roughly constant across income levels. Again, there is quite a lot of variation in unemployment rates for each income level, though somewhat less than for the aggregate unemployment rates. Taking simple averages by income quartile, for the low-educated workers in the bottom quartile, the average unemployment rate is 2.7 percent. This rises to 8.1 percent in the second quartile, 9.5 in the third and 14.3 in the richest quartile. For the high-educated, the average unemployment rate is not monotonically increasing in income per capita. It rises from 4.9 percent in the bottom quartile to 7.7 in the second, and then falls to 6.2 and 7.3 in the third and fourth quartiles.

The third and fourth data columns of Table 2.1 report the regression coefficients for the low-educated and the high-educated separately. For the low-educated, the coefficient is 2.9 across all surveys, and statistically significant at the one-percent level. When restricted to country averages (i.e., the average across all surveys available for each country), we get a significant slope of 3.2. Across our Tier 1 surveys only, the slope is also 3.2, and when including both Tier 1 and Tier 2 surveys, the slope is 2.9, with statistical significance at the one-percent level in both cases. For the high-educated, in contrast, the slope is statistically insignificantly different from zero in all cases. Across all surveys, the slope coefficient is -0.2 but with a standard error of 0.34. The estimated slopes are noisy and statistically insignificant for country averages, for Tier 1 and for both Tiers 1 and 2, as well.

Figure 2.3 plots the ratio of unemployment for the low-educated to that for the high-educated group. As the figure shows, this ratio is strongly increasing in GDP per capita. It is also less variable across countries within each broad income level than in Figure 2.1, for example. Virtually all of the poorest countries have ratios less than one, meaning that the low-educated workers are *less* likely to be unemployed than the high-educated. All of the richest countries have a ratio above one, meaning that the less-educated are more likely than the high-educated to be unemployed. For the poorest quartile of the world income distribution, the average ratio is 0.52. It rises to 1.1 in the second quartile, 1.5 in the third and 2.1 in the richest quartile. Table 2.1 reports that a regression of this ratio on log GDP per capita yields an estimated slope coefficient always in the ballpark of 0.5 across all surveys, with little variation by data comparability tier.

### 2.3.3 Robustness

In this section, we report how unemployment patterns vary by sex, age, and within rural and urban areas. Table 2.2 presents the slope coefficients from a regression of unemployment rates on log GDP per capita for various disaggregated categories of individuals. We do this separately for the low-education and high-education groups, first over all of our surveys (left panel), and then using only country averages over all available years (right panel).

The first row of Table 2.2 reports the slope for prime-aged males only. Across all surveys and country averages, low-educated prime-aged males have a statistically significant positive slope with GDP per capita, while high-educated ones have an insignificant slope. This pattern is replicated and even stronger in the full sample of households (second row), which includes household members aged 16 to 25, those above age 55, and both sexes. The patterns hold separately for males of all ages (third row) as well, while for females (fourth row), there is even a significant negative trend with GDP per capita among the high-educated. We conclude that our patterns hold for both sexes.

When looking by age group, the low-educated always have a significant and positive relationship with GDP per capita, with the strongest relationship for those aged 16 to 24. The young high-educated have a significant negative slope with GDP per capita, at least across all surveys; the prime-aged have an insignificant negative trend; and the old have a small but

**Table 2.2**: Robustness of Slope Coefficients of Unemployment Rate on log GDP per capita

| | All Surveys | | | All Country Averages | | |
|---|---|---|---|---|---|---|
| | Low Edu. | High Edu. | N | Low Edu. | High Edu. | N |
| Prime males | 2.5*** (.4) | -.3 (.3) | 195 | 2.9*** (.6) | .4 (.3) | 54 |
| Full sample | 3.3*** (.4) | -.5 (.4) | 197 | 3.4*** (.7) | .5 (.6) | 54 |
| Males | 2.9*** (.4) | -.4 (.3) | 197 | 3.1*** (.6) | .4 (.5) | 54 |
| Females | 3.8*** (.4) | -.8* (.5) | 197 | 3.9*** (.8) | .3 (.8) | 54 |
| Age 16-24 | 6.2*** (.7) | -1.2 (.8) | 183 | 6.6*** (1.2) | .5 (1.3) | 52 |
| Age 25-54 | 2.9*** (.4) | -.2 (.3) | 195 | 3.2*** (.6) | .5 (.4) | 54 |
| Age 55+ | 2.0*** (.4) | .5* (.2) | 173 | 2.4*** (.6) | .8* (.4) | 49 |
| Rural | 2.7*** (.6) | -.02 (.7) | 107 | 3.4*** (1.0) | 1.8* (1.0) | 29 |
| Urban | 2.5*** (.9) | -.9 (.6) | 107 | 3.4*** (1.2) | .6 (.8) | 29 |

Note: The table reports the slope coefficients from regressions of the unemployment rate on log GDP per capita and a constant. Observations include aggregate unemployment rates across all Tier 1, 2, and 3 surveys. Country averages are restricted to countries with at least two years' observations. ***, ** and * indicate statistical significance at the 1-percent, 5-percent and 10-percent levels.

significant positive slope. Thus, our patterns are robust across age groups. Finally, we look separately at rural and urban individuals. For both groups, we see the same patterns: strong positive increases in low-educated unemployment with GDP per capita and insignificant slopes for the high-educated. Thus, our findings are present in both rural and urban areas.[6]

[6]One may worry that surveyors in poor countries may systematically avoid times when workers are unlikely to be unemployed, such as harvest times, so as to ensure adequate survey participation. If so, our surveys would overestimate the unemployment rates in the poor countries, thus, underestimating the slope of the relationship

### 2.3.4 Employment, Unemployment, and Not in the Labor Force

Other data sets show that average *employment* rates are lower in rich countries than in poor countries, at least for males (see e.g., [22]). Does this imply that unemployment rates are higher in rich countries? Basic accounting identities show that the answer is no. Those not employed can be either unemployed or not in the labor force. The lower employment rates of rich countries could in principle correspond to lower labor force participation rates, or higher unemployment rates, or both. In practice, we show that the relationship between employment rates, unemployment rates, the percent not in the labor force (NLF), and income per capita varies considerably by gender and education, and cannot be inferred directly from evidence on employment rates alone.

**Table 2.3**: Employment, Unemployment and Not in the Labor Force

| | | Low Education | | | High Education | | |
|---|---|---|---|---|---|---|---|
| | Income Quartile | Bottom | Top | Difference | Bottom | Top | Difference |
| Male | Employed | 87.5 | 72.8 | -14.7*** | 83.1 | 86.3 | 3.2 |
| | Unemployed | 2.3 | 11.2 | 8.9*** | 4.0 | 6.1 | 2.1*** |
| | Not in labor force | 10.2 | 16.0 | 5.8* | 12.9 | 7.6 | -5.3** |
| Female | Employed | 60.4 | 46.0 | -14.4* | 63.1 | 69.7 | 6.6 |
| | Unemployed | 1.9 | 9.1 | 7.2*** | 4.2 | 6.7 | 2.4* |
| | Not in labor force | 37.7 | 44.9 | 7.2 | 32.7 | 23.7 | -9.0* |

Note: This table reports summary statistics of prime age employment, unemployment and percent not in the labor force for the bottom and top quartile countries, by gender and education. The rows present the average of poor countries, the average of the rich countries, and the difference between the poor and rich means, plus the results of a permutation test of the differences in means.

Table 2.3 reports the percent of prime aged adults – by sex and education level – that are employed, unemployed, and not in the labor force, for countries in the bottom and top income quartiles. For low-educated males, employment rates are substantially lower in the richest quartile

between average unemployment and income per capita.

than in the poorest. This reflects a substantially higher percent of low-educated males not in the labor force in the richest quartile, as well as their higher unemployment rates in the richest quartile. A similar pattern also holds for women, though with lower employment levels in both quartiles.

Among high-educated males, employment rates are modestly higher in the richest quartile than in the poorest quartile (though the difference is statistically insignificant). Yet the percent of high-educated males that are unemployed is also modestly higher in the richest quartile. The reason that both are higher in the richest quartile is that, as Table 2.3 shows, the percent not in the labor force is substantially *lower* for high-educated males in the richest quartile. A similar pattern again holds for females, though with larger increases in employment rates and labor force participation rates than for the males. We conclude that there is no simple way one can infer cross-country unemployment patterns by looking solely at data on employment rates, which reflect a margin of labor force participation as well.

## 2.4 A Simple Model of Unemployment and Development

In this section, we build a simple model to qualitatively match the increasing unemployment rate with development, and the patterns of unemployment by education level. Since the main focus of the paper is on unemployment rates, we abstract from the decision of whether to join the labor force. Since our empirical patterns are present for both sexes, all age groups and within both rural and urban areas, we abstract from demographics and regional considerations. In order to match the large decrease in the traditional (low-skilled self-employment) sector that coincides with development, we allow for two sectors in our model.

### 2.4.1 Environment

There is a unit measure of risk-neutral, infinitely-lived workers, each of whom is endowed with efficiency units drawn from a fixed distribution $G(x)$ on $[\underline{x}, \bar{x}]$. We assume that $G(x)$ is differentiable and let $g(x) \equiv G'(x)$ be its probability density function. There is also a continuum of risk-neutral, infinitely-lived firms, each of which can employ one worker. In this simple model, we assume undirected search in the aggregate distribution of ability. Later, in the quantitative version of our model, we relax these assumptions and allow firms to direct their search efforts toward high and low education groups of workers.

Workers can choose to work in one of two sectors: a traditional sector, in which workers are self-employed without returns to ability, and a modern sector, in which firms hire workers subject to matching friction. and production displays constant returns to ability. The technologies in the traditional and modern sectors, respectively, are given by:

$$Y_T = A_T N_T, \quad \text{and} \tag{2.1}$$

$$Y_M = A_M X_M, \tag{2.2}$$

where $Y_T$, $A_T$ and $N_T$ are output, productivity and the number of workers in the traditional sector, and $Y_M$, $A_M$ and $X_M$ are output, productivity and the total number of efficiency units in the modern sector. Countries vary in their level of productivity $A_M$ but not $A_T$, so technological change in our model is skill-biased. Here we assume the outputs of the modern and traditional sectors to be perfect substitutes for simplicity. We relax this assumption and the invariance of $A_T$ in the quantitative model that follows. Our assumption of exogenous modern-sector productivity is abstract, though it may capture more concrete channels that affect firm size and hence the extent of wage employment, such as firm financial frictions (e.g., [62, 15, 33]), or monitoring frictions (e.g., [5, 43]).

We now combine a Diamond-Mortensen-Pissarides model of steady-state unemployment with a Roy model of selection into the modern versus the traditional sector.

**Steady State.** In the steady state, workers will not move between sectors in the absence of shocks. Denote by $x^*$ the efficiency units of the marginal worker who is indifferent between self-employment and entering the modern sector unemployed. We will show below that the value of being unemployed is increasing in $x$; hence, in steady state, workers with $x < x^*$ prefer self-employment in the traditional sector, and workers with $x \geq x^*$ prefer to enter the modern sector as unemployed.

**Modern Sector.** In order to hire a worker, a firm must post a vacancy at flow cost $A_M c$.[7] Let the flow of matches be given by the constant returns to scale function

$$m(u,v) = \eta u^{\alpha} v^{1-\alpha}, \tag{2.3}$$

where $u$ is the endogenous measure of unemployed workers and $v$ is the endogenous measure of vacancies in the economy. Define $\theta \equiv \frac{v}{u}$ as "market tightness." The job-finding rate is then $f(u,v) \equiv \frac{m}{u} = \eta\theta^{1-\alpha}$, and the vacancy hiring rate is $q(u,v) \equiv \frac{m}{v} = \eta\theta^{-\alpha}$.

We assume that workers and firms separate at an exogenous rate $s$. Let $A_M bx$ denote the unemployment flow payoff,[8] where $0 < b < 1$. One rationale for this choice is that unemployment benefits are typically indexed to wages, which we will show scale with $A_M x$ in equilibrium. A second rationale is that job finding rates are approximately constant across skill groups, which is consistent with a model where unemployment benefits scale with the expected wage ([99, 65,

[7] We shall see later that, in equilibrium, wages scale with $A_M$. If the productivity of the vacancy posting process is not affected by $A_M$, the cost of posting a vacancy should also scale with $A_M$.

[8] Our results are qualitatively unchanged if we let $A_M b$ denote the unemployment flow payoff.

104]). Denoting by $\delta$ the rate of time discount for all agents, the values of unemployment and employment for an individual with efficiency units $x$ are given, respectively, by

$$U(x) = A_M bx + \delta\big[fE(x) + (1-f)U(x)\big] \tag{2.4}$$

$$E(x) = w(x) + \delta\big[sU(x) + (1-s)E(x)\big], \tag{2.5}$$

where $w(x)$ is the endogenous flow wage.

Because firms will be matched only with agents in the modern sector, who have efficiency units $x \geq x^*$, we can specify the value of a job to a firm if matched with a worker with efficiency units $x$ and the value of maintaining a vacancy as:

$$J(x) = A_M x - w(x) + \delta\big[sV + (1-s)J(x)\big] \tag{2.6}$$

$$V = -A_M c + \delta\big[q\mathbb{E}\big(J|x > x^*\big) + (1-q)V\big], \tag{2.7}$$

where $\mathbb{E}\big(J|x > x^*\big) = \frac{\int_{x^*}^{\bar{x}} J(x)g(x)dx}{1-G(x^*)}$ is the expected value to the firm of a job match conditional on the workers having entered the modern sector.

Because of the free-entry condition for firms, we have $V = 0$. Let $S(x) \equiv E(x) - U(x) + J(x)$ denote the total surplus of a match, and $\beta \in (0,1)$ be the Nash bargaining power of the worker. The firm then receives $(1-\beta)S(x)$ when a vacancy is filled. Combining this division of the surplus with equations (2.4) to (2.7) allows us to solve for $U(x)$ and $w(x)$, with the former given by:

$$U(x) = \frac{1}{1-\delta}\Big(A_M bx + \delta\eta\theta^{1-\alpha}\frac{\beta}{1-\beta}\frac{A_M x(1-b)(1-\beta)}{\beta\delta\eta\theta^{1-\alpha} + 1 - \delta + \delta s}\Big). \tag{2.8}$$

Equation (2.8) shows that $U(x)$ is increasing, as we asserted previously. We also show in Appendix 2.8.2 that steady state in the modern sector is characterized by the following relationship between $\theta$ and $x^*$:

$$c = \frac{(1-\beta)\delta\eta\theta^{-\alpha}}{\beta\delta\eta\theta^{1-\alpha}+1-\delta+\delta s}(1-b)\mathbb{E}(x|x>x^*). \tag{2.9}$$

Note that market tightness $\theta$ is unaffected by $A_M$ for a given $x^*$. By equation (2.11) below, this implies that unemployment is unaffected by $A_M$ for a given $x^*$. Thus, in the absence of a traditional sector, our model predicts that unemployment remains constant as per capita income increases. If $b$ or $c$ did not scale with $A_M$, $\theta$ would instead decrease with $A_M$ for a given $x^*$, and in the absence of a traditional sector, our model would predict that unemployment decreases as per capita income increases.

**Indifference Condition.** The value of staying in the traditional sector is $\frac{A_T}{1-\delta}$, since any traditional worker produces $A_T$ in every period. The worker with efficiency units $x^*$ is indifferent between staying in the traditional sector and entering the modern sector as unemployed:

$$\frac{A_T}{1-\delta} = U(x^*) = \frac{1}{1-\delta}\left(A_M b x^* + \delta\eta\theta^{1-\alpha}\frac{\beta}{1-\beta}\frac{A_M x^*(1-b)(1-\beta)}{\beta\delta\eta\theta^{1-\alpha}+1-\delta+\delta s}\right). \tag{2.10}$$

**Unemployment Rate.** Letting $u_M$ denote the measure of the modern-sector unemployed and its steady-state value, we can write the change in modern-sector unemployment as $\dot{u}_M = (L_M - u_M)s - u_M f(\theta)$, where $f(\theta) = \eta\theta^{1-\alpha}$ is the steady state job finding rate and $L_M = 1 - G(x^*)$ is the labor that participates in the modern sector. We can then set $\dot{u}_M = 0$ to obtain the measure of steady-state modern sector unemployment, which is the same as the overall unemployment rate, since the overall measure of workers is one and there is no unemployment in the traditional sector:

$$u = \frac{s\big(1-G(x^*)\big)}{s+\eta\theta^{1-\alpha}}. \tag{2.11}$$

Note that the unemployment rate depends on the separation rate, $s$, the (endogenous) market tightness, $\theta$, and the (endogenous) cutoff $x^*$ for working in the modern sector. The fraction

$1-G(x^*)$ represents the measure of workers in the modern sector. The higher is this fraction, all else equal, the higher is the unemployment rate. Similarly, the lower is $\theta$, all else equal, the higher is unemployment.

### 2.4.2 Model Solution and Predictions

We now establish the uniqueness of our model solution, and characterize how the endogenous variables $\theta$ and $x^*$ vary with modern-sector productivity, $A_M$.

**Proposition 1.** *If an interior solution $x^* \in (\underline{x}, \bar{x})$ exists, the model solution $(x^*, \theta)$ is unique, and the cutoff ability $x^*$ decreases as modern-sector productivity $A_M$ increases.*

*Proof.* See Appendix 2.8.2. □

Proposition 1 shows that an increase in modern sector productivity reduces $x^*$, drawing workers out of the traditional sector into the modern sector. This result plays an important role in determining how unemployment rates vary with modern-sector productivity. In particular, we can use it to help establish:

**Proposition 2.** *The aggregate unemployment rate $u$ increases as modern-sector productivity $A_M$ increases.*

*Proof.* See Appendix 2.8.2. □

The intuition for this result is as follows. First, as $A_M$ increases, workers are drawn out of the traditional sector and into search for wage employment in the modern sector, as shown in Proposition 1. Because modern-sector jobs involve regular separations, a larger modern sector means larger steady-state unemployment, all else equal. Second, as $A_M$ increases, market tightness, $\theta$, falls in equilibrium. Because the workers drawn into the modern sector are of lower ability than existing modern-sector workers, the expected value of a match to the firm falls. For the free-entry condition to hold, the job filling rate for a vacancy must rise. This means fewer vacancies per unemployed person, i.e., a smaller $\theta$.

**Proposition 3.** *Let $x^*$ be an interior solution and $x_0 > x^*$ denote a fixed ability level. Then the ratio of the unemployment rate for workers with ability lower than $x_0$ to that of workers with ability higher than $x_0$ increases as modern-sector productivity $A_M$ increases.*

*Proof.* See Appendix 2.8.2. □

In short, this result states that the relative unemployment of less-able to more-able workers increases with development. Intuitively, this occurs because a larger share of less-able workers are drawn into the modern sector as $A_M$ rises. Figure 2.4 illustrates how Proposition 3 works. Denote the "high-ability workers" as those with ability above $x_0$, and those below $x_0$ as the "low-ability workers." The initial cutoff is depicted as $x_1^*$, and hence regions $A$ and $B$ represent the traditional sector, whereas $C$ and $D$ are the modern sector. Once $A_M$ rises, the cutoff falls, by Proposition 1, to a lower cutoff which we denote by $x_2^*$. Region B switches from the traditional to the modern sector. Since these are low-ability workers, and no high-ability workers switch sectors, the ratio of low- to high-ability unemployment increases.

## 2.5 Quantitative Analysis

Though the simple model above is useful for establishing the qualitative properties of our theory, the model is a bit too stylized to use in our quantitative analysis. Thus, in this section we build a richer quantitative version of the model. We then calibrate the model to match features of the U.S. labor market, and compute the model's predictions over the full range of the world income distribution.

### 2.5.1 Quantitative Version of the Model

In our simple model the outputs of the modern and traditional sectors are perfect substitutes, so their relative price cannot change as $A_M$ rises. This is at odds with the well-known tendency for the relative price of non-traded services, in which the traditional sector is intensive, to rise with per capita GDP. With this in mind, we now allow traditional and modern sector outputs to be imperfect substitutes. We specify the following constant-elasticity-of-substitution (CES) aggregate production function:

$$Y = \left(\gamma Y_T^{\sigma} + (1-\gamma) Y_M^{\sigma}\right)^{\frac{1}{\sigma}}, \tag{2.12}$$

where $Y_T$ and $Y_M$ are the aggregate outputs of the traditional and modern sectors, respectively, and the elasticity of substitution between them equals $\frac{1}{1-\sigma}$. Denote the price of traditional-sector output relative to modern-sector output by $P_T$. In a competitive market, the ratio of prices equals the ratio of marginal productivities:

$$P_T = \frac{\partial Y / \partial Y_T}{\partial Y / \partial Y_M} = \frac{\gamma}{1-\gamma} \left(\frac{Y_M}{Y_T}\right)^{1-\sigma}. \tag{2.13}$$

Technological change that is skill-biased across countries is a core assumption of our model. The assumption that technological change in the traditional sector is zero, however, is an oversimplification. In our quantitative exercise we allow for an elasticity of technological change in the traditional sector with respect to technological change in the modern sector that is less than one. Specifically, in our calibration procedure we will assume that $\log(A_T) = \theta_0 + \theta_1 \log(A_M)$, where we expect to find that $\theta_1 < 1$.[9]

Increases in $P_T$ or $A_T$ with $A_M$ cause workers who remain in the traditional sector in rich countries to earn more than traditional sector workers in poor countries. This is more realistic than the prediction of the simple model that earnings of traditional sector workers in rich and poor countries will be the same.

Key predictions of our model concern traditional employment and unemployment by worker ability. Unfortunately, direct measures of ability across many countries are not available. Wage is a linear function of ability in our model, but we cannot observe wages for the self-employed in the traditional sector or the unemployed. Instead, for the purpose of quantifying our predictions regarding traditional self-employment and unemployment by ability, we use education as our proxy for ability. Specifically, we divide the labor force into the two education groups used above, in particular the low education group, which did not finish secondary school, and the high education group, which completed secondary school or more. We incorporate education into our model as a proxy for ability by assuming that the distribution of ability for the high-education group first-order stochastically dominates the distribution of ability for the low-education group: $G_h(x) < G_l(x)$ for all $x \in (\underline{x}, \bar{x})$.[10]

Countries differ exogenously in the fraction $\lambda$ of their workers that are in the low-education group. The remaining $1-\lambda$ are in the high-education group. We assume employers can observe this education credential ex ante and divide the modern sector labor market into two search markets, one for each education level. Finally, we treat the outputs of modern-sector firms that

[9]In our theory, the higher relative output of goods produced by skilled workers that occurs with development results only from increased productivity in the modern sector relative to the traditional sector. In reality, however, development leads to an increase in the relative demand for skill-intensive goods, as richer households demand more skill-intensive products and services ([32, 34]). Our results would still apply, at least qualitatively, if we were to extend our model to include non-homothetic preferences in which higher income causes higher relative demand for skill-intensive goods.

[10]This condition is sufficient, but not necessary, for the results of this subsection. We verified that the distributions calibrated in the next subsection satisfy this condition.

search in the high-education and low-education labor markets as perfect substitutes, and add them to obtain $Y_M$ in equation (2.13).

We also allow for the possibility that the separation rate for high-educated workers is less than for low-educated workers, though this is not necessary to obtain any of our qualitative results: $s_h \leq s_l$. All other parameters are assumed to be the same across the two labor markets.

We can now prove:

**Lemma 1.** *For any interior solution to the model with two labor markets, $x_h^* < x_l^*$.*

*Proof.* See Appendix 2.8.2. □

It follows from Lemma 1 and $G_h(x) < G_l(x)$ that the share of high-educated agents who are self-employed in the traditional sector is lower than the corresponding share of low-educated agents:

**Proposition 4.** *For any interior solution to the model with two labor markets, $G_h(x_h^*) < G_l(x_l^*)$.*

As modern sector productivity $A_M$ increases in our simple model, Proposition 1 states that the share of workers who are self-employed in the traditional sector falls ($x^*$ decreases). Similarly, if increasing $A_M$ dominates increasing traditional sector relative price $P_T$ and traditional sector productivity $A_T$ in our quantitative model, the shares of both high- and low-educated workers who are self-employed in the traditional sector will fall ($x_h^*$ and $x_l^*$ both decrease). The unemployment rates of both high- and low-educated workers must then increase, just as did the aggregate unemployment rate in the simple model (Proposition 2). Here, however, the aggregate unemployment rate does not necessarily increase, despite increases in the unemployment rates for both education groups. The aggregate unemployment rate in the quantitative version of our model is a weighted average of the unemployment rates of high- and low-educated workers, with weights $1-\lambda$ and $\lambda$. In the data, as modern sector productivity and thus GDP per capita increases, the share of low-educated workers $\lambda$ tends to decrease. If the high-educated unemployment rate is smaller than the low-educated unemployment rate, it is possible for the aggregate unemployment rate predicted by the quantitative version of our model to decrease with $A_M$ and GDP per capita.

Whether the ratio of low-educated to high-educated unemployment rates increases with $A_M$ in the quantitative model, which would be the equivalent of Proposition 3 in the simple model, depends on the calibration. However, we can establish a strong presumption that our quantitative model will display this property. The basis for Proposition 3 is that, as $A_M$ increases, participation in the modern sector by workers with low ability increases relative to participation by workers with high ability. We can expect, similarly, that as $A_M$ increases, participation in the modern sector of low-educated workers will increase proportionately faster than participation of high-educated workers. The reason is that low-educated workers' participation in the modern sector must be lower according to Proposition 4, but both participation rates must approach 100 percent as $A_M$ increases. In our quantitative predictions in Subsection 2.5.3 below, participation of low-educated relative to high-educated workers in the modern sector does indeed increase as $A_M$, and thus per capita GDP, increases.

### 2.5.2 Parameterizing the Model

We begin by directly setting some parameter values following the literature. We set the quarterly discount factor to $\delta = 0.99$, consistent with an annual interest rate of around four percent. We set the worker's bargaining weight to $\beta = 0.7$ and the elasticity parameter of the matching function to $\alpha = 0.7$, which are the values used in [56] and are in line with the standard parameter choices used in macro search models. We set the quarterly separation rate for the high-educated workers to $s_h = 0.045$, which is the value estimated in [138]. We use the unemployment benefits replacement rate of 45 percent. This is in line with the 40 percent used by [129] , the 42 percent in [29], and the 50 - 60 percent range in [83]. We also normalize the mean of the ability for low-educated workers to be one.

We set the elasticity of substitution between traditional and modern goods to be 3 in our benchmark calibration, though we explore sensitivity to this parameter, as we describe below. Our elasticity of substitution relates to some extent to the elasticity of substitution between home and market goods that is emphasized by the large literature emphasizing home production in the macroeconomy.[11] Though our model's elasticity is related to these, it is not exactly comparable, and one may imagine that there are greater substitution possibilities between modern and traditional goods than between home and market production, since modern and traditional goods are both purchased in the market. For example, one type of substitution between the modern and traditional sector may be getting older shoes shined and repaired (from a self-employed shoe repairer) rather than purchasing newer shoes (from a modern shoe factory). Another example is buying produce from an informal road-side vendor versus buying produce at a modern supermarket. It is therefore worth looking at alternative evidence on substitution between different categories of purchased goods and services. In a widely cited study, [31] estimate elasticities of substitution across a diverse set of goods varieties, finding a median estimate of around 2.2 to 3.7 across goods categories.[12] Our benchmark value of 3 is right in the middle of their estimates, though since there is not a more precise value suggested by the literature, we explore a lower

[11]See eg. [20, 61, 19, 106, 122, 105, 80]. [9] choose a value of 1.8, and argue that this is close to the midpoint of the range suggested by previous estimates in this literature. For example, [126] use panel data from the PSID with evidence on time spent in home production and market work, and estimate an elasticity of substitution between 1.8 and 2.0. [94] and [39] use U.S. time series data and come up with estimates of 1.5 to 1.8 and 2.3 respectively. [2] draw on detailed household-level data on market goods consumption and time spent on home production, such as cleaning, cooking and home repair. They estimate an elasticity of substitution of 2.1 when considering all home production categories in their data.

[12]We are not aware of any estimates of substitution elasticities between goods with low and high levels of skill inputs. On the production side, the closest estimate would be for the substitution elasticity between high- and low-skilled labor in the aggregate production function; **?**, pg. 11 argues that the "consensus across estimates for the U.S." is that this elasticity is approximately two. [96] estimates an elasticity of substitution of around 6.5 between informal and formal labor, though this is again about production and not final consumption goods.

**Table 2.4**: Calibrated Parameters

| Parameter | Value |
|---|---|
| Panel A: Pre-Assigned Parameters | |
| $\delta$ - Discount factor (quarterly) | 0.99 |
| $\beta$ - Workers' bargaining power | 0.7 |
| $\alpha$ - Matching parameter | 0.7 |
| $s_h$ - Separation rate (quarterly) for high-educated workers | 0.045 |
| $b$ - Unemployed benefits | 0.45 |
| $\frac{1}{1-\sigma}$ - Elasticity of substitution | 3 |
| $A_{T(US)}$ - U.S. traditional-sector productivity | 1 |
| $m_l$ - Mean of ability dfor low-educated workers | 1 |
| Panel B: Calibrated Parameters | |
| $m_h$ - Mean of ability for high-educated workers | 1.66 |
| $v_l$ - Variance of ability for low-educated workers | 0.45 |
| $v_h$ - Variance of ability for high-educated workers | 1.15 |
| $c$ - Vacancy cost | 0.15 |
| $\eta$ - Matching efficiency | 0.85 |
| $\gamma$ - Traditional-sector share in aggregate production function | 0.01 |
| $s_l$ - Separation rate (quarterly) for low-educated workers | 0.112 |
| $max(A_M)$ - Modern-sector productivity for the richest country | 0.04 |

Note: The table reports the values and interpretations of the parameters of the quantitative model under the benchmark calibration.

value of 2.5, closer to the home-production literature, and a higher value of 3.5, close to the upper end of the values estimated by [31].

We calibrate the remaining eight parameters to jointly match eight moments in the data. These parameters are: (i) the mean of the ability distribution for the high-educated workers, $m_h$; (ii) and (iii): the variances of the ability distributions for the low- and high-educated workers, $v_l$ and $v_h$; (iv) the vacancy cost $c$ as a share of the modern-sector productivity for a worker with one unit of ability; (v) the efficiency term, $\eta$, of the matching function; (vi) the traditional-sector share in the aggregate production function, $\gamma$; (vii) the quarterly separation rate for the low-educated workers, $s_l$; and, finally, (viii): the maximum value of $A_M$, which corresponds to the U.S. level.[13]

The eight moments are: (i) the ratio of the average modern-sector wages for the high- over low-educated that we calculated using the 2000 Census 5% sample (1.60); (ii) and (iii) the variances of log wages for the high- and low-educated (0.34 and 0.28), using the same 2000 census; (iv) the vacancy cost of 17 percent of average output in the modern sector as used in [56]; (v) the average U.S. unemployment rate of 5.71 percent in the United States among the 18 samples in our data from 1960 to 2014; (vi) the U.S. expenditure share in the traditional sector, which we conjecture to be smaller than two percent; (vii) the ratio of unemployment for the the low-educated to high-educated (2.31); and (viii) an average employment share of two percent in the traditional sector (as we explain below).

We define the traditional sector as the intersection of own-account (self-employed without employees) workers and occupations with low skill content – in particular, shop and market sales, agricultural and fishery workers, crafts and related trades workers, plant and machine operators and assemblers, and "elementary occupations." Unfortunately, the U.S. data after 1960 distinguish only between incorporated and unincorporated businesses among the self-employed, rather than between own-account workers and employers as in the countries in Figures 2.5 and 2.6 below. Considering that the Canada samples have an average of 2.8 percent prime-aged employment in the traditional sector, which is defined consistently with the other countries, we conjecture that the United States has a smaller share of two percent. As with our benchmark unemployment measures, all traditional sector employment shares reported in this section are calculated for prime-aged workers.

[13]Note that although the absolute value of $A_M$ is smaller than $A_T$, the modern sector is more productive than the traditional sector in value terms. The traditional and modern sectors produce different goods, and the relative price of the traditional good, $P_T$, is around 0.01 in the United States in our calibrated model.

**Table 2.5**: Moments Targeted in the Model vs Data

| Moment | Target | Model |
|---|---|---|
| Ratio of average wage for the high- vs low-educated | 1.60 | 1.61 |
| High-edu log(wage) variance | 0.34 | 0.33 |
| Low-edu log(wage) variance | 0.28 | 0.28 |
| U.S. vacancy cost as % of average output in modern sector | 17 | 16.9 |
| U.S. unemployment rate | 5.71 | 5.69 |
| U.S. % expenditure share of traditional sector | $<2.0$ | 0.67 |
| U.S. ratio of unemployment rates $u_l/u_h$ | 2.31 | 2.32 |
| U.S. traditional sector employment share | 2 | 1.84 |

Note: The table reports the moments targeted in the benchmark calibration of the quantitative model and the model's predictions for each moment.

Table 2.4 reports the value of each parameter used in the calibration. Our calibrated quarterly separation rate for the low-educated is 0.112, similar to the direct estimate of 0.06 - 0.12 during 1980 to 2010 computed by [138] for low-educated workers. Our estimate is also broadly consistent with the separation rate in low-skilled services in the United States. For example, according to the 2017 Job Openings and Labor Turnover Survey, the monthly separation rate in wholesale and retail trade, transportation and utilities is around 3.5 percent. This corresponds to a quarterly separation rate of around 10 percent.

We report each moment and its model counterpart in Table 2.5. Overall, the model matches the desired moments quite well. Although all of the eight parameters reported above jointly discipline all the parameters, it is useful to provide some intuition about which moments are most informative about each parameter. In particular, the mean of the ability distribution for high-educated workers, $m_h$, largely governs the ratio of average wage of the high- to low-educated workers. The variances of the two ability distributions govern the variances of log wages for the low- and high-educated workers. The model vacancy cost and model unemployment benefit are most informative about the relative size of vacancy cost and unemployment benefits to the average output per worker in the modern sector. The matching efficiency parameter $\eta$ mostly informs the average unemployment rate, and the sector share parameter in the aggregate production function mostly informs the expenditure share of traditional-sector output. The quarterly separation rate for low-educated workers is most informative about the unemployment ratio of low- to high-educated workers. Finally, the maximum $A_M$ value governs the traditional sector employment share in the richest country (the United States).

It remains to calibrate the elasticity of traditional sector productivity with respect to modern sector productivity. To do so, we use the fact that greater increases in $A_T$ will result in smaller increases in $P_T$ as GDP per capita increases, all else equal. Specifically, we target the elasticity of the relative price of traditional goods with respect to GDP per capita.

We draw on disaggregated evidence on average national prices for specific products from the 2011 International Comparison Program (ICP). The ICP data are the best available data on the prices of identical (or nearly identical) goods and services around the world, and are available for almost every country in the world. How do we define traditional goods in these data? Consistent with our definition of the traditional sector, we pick goods or services that are have low skill content and are likely to be provided by self-employed workers. We identified eight specific services that plausibly meet these criteria: (i) a shoe repair for women's street shoes; (ii) a shoe repair for men's classic shoes; (iii) a shoeshine; (iv) a 7km taxi ride from the town center; (v) a men's basic haircut; (vi) a ladies haircut with curlers; (vii) a manicure; (viii) a ladies haircut, long hair. Appendix Table 2.15 provides the exact definitions of these eight traditional sector services. Since investment goods largely fit our definition of a modern output, we take the aggregate price level of investment from the Penn World Table as a proxy for our modern sector price. For each traditional-sector service, we then compute the relative price of the service compared to

**Table 2.6**: Slope of Log Relative Prices on log(GDP) in Data

| | | | |
|---|---|---|---|
| Shoe repair - women's street shoes | .39*** (.002) | Men's basic haircut | .61*** (.001) |
| Shoe repair - men's classic shoes | .53*** (.004) | Ladies haircut - curlers | .63*** (.002) |
| Shoeshine | .56*** (.002) | Manicure | .44*** (.003) |
| Local taxi ride | .42*** (.006) | Ladies haircut - long hair | .68*** (.002) |

Note: Data come from the unpublished ICP 2011 disaggregated price data for the Global Core list of goods and services. See Appendix Table 2.15 for the exact definition of each good and service. The table reports the slope coefficient from a regression of the log of the item price relative to the investment goods price on log GDP per capita and a constant. ***, ** and * indicate statistical significance at the 1-percent, 5-percent and 10-percent levels.

investment goods in each country.

Table 2.6 reports the slope coefficient from a regression of the log of the item relative price on log GDP per capita and a constant. As shown in the table, the elasticity of the relative price ranges between 0.39 to 0.68. We target the median of these relative price elasticities, which is around 0.60. Our calibration uses the parameter $\theta_1$, the elasticity of $A_T$ with respect to $A_M$, to target this relative price elasticity. This yields $\theta_1$ = 0.26, with the intercept $\theta_0$ in the equation $\log(A_T) = \theta_0 + \theta_1 \log(A_M)$ determined implicitly by our normalization of $A_T$ to be one in the United States.[14]

[14]Specifically, to match the elasticity of relative price to GDP per capita, we have to solve the full set of countries in the model with potential values of $\theta_1$. In contrast, we only need to solve one country in the model to calibrate the eight U.S. moments.

### 2.5.3 Quantitative Predictions

With the model calibrated to the U.S. data, we then lower $A_M$, $A_T$, and $\lambda$, the fraction of workers that are low-educated. We discipline $\lambda$ directly using data on the fraction of workers with less than high school education across our set of countries (see Appendix Figure 2.9). After solving each economy, we use the equilibrium prices $P_T$ from all economies to compute a single international price, the average of $P_T$ weighted by traditional-sector output in each economy. We use this international price to compute the values of model outputs for all economies, including the U.S., and then scale all output values such that the richest economy matches the U.S. GDP per capita of exp(10.7) or \$44,355.

Figure 2.5 plots the traditional-sector size in the model and data. As GDP per capita decreases from the U.S. level, our model predicts an increase in the traditional-sector size from two percent to almost 60 percent. This is largely in line with our data. Furthermore, our model gets the curvature largely correct – in particular, the convex relationship between traditional-sector share and GDP per capita. This occurs partly because in richer economies almost all high-educated workers in the model are in the modern sector, so when those workers start to switch to the traditional sector, its size increases faster.

To emphasize the mechanisms further, Figure 2.6 plots the traditional-sector shares by education level. As in the data, the model predicts decreasing relationships between the traditional sector shares and per capita GDP for both groups. Crucially, it predicts much higher shares of traditional sector employment for the low-educated in poor countries. As $A_M$ rises, there are more low- than high-educated workers to sort out of the traditional sector, and as a result unemployment rises more for the low-educated (as in the data). This differential rate of exodus from the traditional sector as $A_M$ rises is thus key to our theory, and Figure 2.6 shows that the magnitudes here are largely consistent with the data. Note that the aggregate traditional-sector share in Figure 2.5 is nearly the same as the low-educated traditional sector share in Figure 2.6, because the labor force in poor countries is dominated by low-educated workers.

Figure 2.7 plots the aggregate unemployment level in the model and data. As GDP per capita increases, our model predicts that the unemployment rate will increase from less than 4 percent to the calibrated value of 5.7 percent. This is similar to the magnitudes in the data, though the model somewhat under-predicts the steepness of the relationship. Further, consistent with the data, our model predicts a sharper increase when GDP per capita is lower. This is a result of the faster decrease in the traditional-sector share when GDP per capita is lower.

Figure 2.8 plots the ratio of unemployment for the low-educated to the high-educated in the model and data. The model is calibrated to obtain the correct ratio for the United States. For lower levels of GDP per capita, the model predicts a decline in this ratio, as in the data. Again, the the model underpredicts the steepness of this relationship. The model predicts that this ratio is just above one for the poorest countries, whereas in the data, the ratio is closer to 0.5.

Table 2.8 reports the slope coefficients from regressions of the unemployment rate and other key variables for prime age workers on log GDP per capita and a constant, in our model and in the data. For the aggregate unemployment rate, the model yields a semi-elasticity of 0.5 compared to 1.8 in the data. Thus, the model accounts for around 30 percent (0.5/1.8) of the empirical relationship between unemployment and log GDP per capita. Unemployment rates for the low-educated have a semi-elasticity of 1.7 in the model, compared to 3.2 in the data. The high-educated semi-elasticities are fairly similar, at 0.4 in the model and 0.5 in the data. The ratio of low- to high-educated unemployment rates is 0.5 in the data and 0.3 in the model. Largely consistent with the above discussions, the model yields magnitudes similar to the data but underpredicts the empirical elasticities. Traditional-sector slopes are similar in the model and

**Table 2.7**: Slope Coefficients in Data and Quantitative Model

| | Data | Model |
|---|---|---|
| Aggregate traditional sector share | -15.9 | -13.4 |
| Traditional-sector share for low educated | -16.7 | -12.7 |
| Traditional-sector share for high educated | -4.9 | -5.0 |
| Aggregate unemployment rate | 1.8 | 0.5 |
| Unemployment rate for low-educated | 3.2 | 1.7 |
| Unemployment rate for high-educated | 0.5 | 0.4 |
| Ratio of unemployment rates $u_l/u_h$ | 0.5 | 0.3 |
| Relative price $P_T$ | 0.6 | 0.60 |

Note: The table reports estimated slope coefficients from regressions of the statistics in each row on log GDP per capita. The first data column reports the slopes from our cross-country database, and the second data column reports the slopes from the quantitative model.

data, at -15.9 in the model and -13.4 in the data.

We also calibrated our model using an alternative strategy to discipline the elasticity of technological change in the traditional sector with respect to technological change in the modern sector. We targeted the slope of the aggregate traditional sector share on log GDP per capita instead of the elasticity of the relative price of traditional sector output with respect to GDP per capita. This yields $\theta_1 = 0.19$ when we match the traditional sector share slope of -15.9 precisely. When calibrated this way, the model yields a slightly higher relative price elasticity of 0.67, which is still in the range of the empirical estimates 0.4 to 0.7. Using this strategy, the model accounts for more than 40 percent (0.75/1.8) of the empirical relationship between unemployment and log GDP per capita. It yields the same slope of 0.3 for the ratio of low- to high-educated unemployment as in the benchmark. Appendix Table 2.16 reports all the model slopes when using the alternative calibration strategy.

In our benchmark model, the unemployment benefits replacement rate $b$ is set to 0.45 in all economies. But in reality, the benefits replacement rate is higher in richer countries. To study the quantitative impact of varying $b$ values, we now calibrate the model using increasing $b$ values from 0 in the poorest country to 45 percent in the United States.

Panel A of Table 2.8 reports the slope coefficients from regressions of the unemployment rate and unemployment ratio on log GDP per capita and a constant, in our benchmark model and in the model with varying $b$ values. The model with varying $b$ values predicts an aggregate unemployment rate elasticity of 0.72, compared to 0.52 in the benchmark model. This accounts for 41 percent of the empirical relationship in the data, which is 11 percentage points higher than in our benchmark model. For the unemployment ratio, the model with varying $b$ values has an elasticity of 0.26, very similar to 0.25 in the benchmark model. In addition, Panel B of Table 2.8 reports the difference of average unemployment rates and ratios for the top and bottom income countries, both in the data and in two versions of the model. The top income quartile countries in our sample have an average unemployment rate of 8.0 percent compared to 2.7 percent for the bottom quartile countries. The difference is 5.3 percentage points. The model with varying $b$ values can account for 52 percent of this unemployment rate difference, compared to 42 percent for the benchmark model. For the unemployment ratio difference, the two versions of model have similar explanatory power, 51 percent for the benchmark model and 54 percent for the model with increasing $b$ values.

In summary, an alternative model which includes increasing unemployment benefits with development helps to explain the increase in the unemployment rate with GDP per capita, but not the increase in the unemployment ratio. Thus, although the quantitive model explains a substantial portion of the aggregate unemployment patterns in question, and the higher unemployment benefits in richer countries increase the model's explanatory power, much of the data are left unexplained by the model. Additional forces that help to explain the cross-country relationship between average unemployment and income per capita are a subject for future research.

**Table 2.8**: Benchmark Model and Model with Varying $b$

| Panel A: Slope Coefficients | | | | | |
|---|---|---|---|---|---|
| | Data | Benchmark | Explained | Varying $b$ | Explained |
| Unemployment | 1.76 | 0.52 | 30% | 0.72 | 41% |
| $u_L/u_H$ | 0.47 | 0.25 | 53% | 0.26 | 55 % |

| Panel B: Top Quartile Minus Bottom Quartile | | | | | |
|---|---|---|---|---|---|
| benchmark | Data | Benchmark | Explained | Varying $b$ | Explained |
| Unemployment | 5.3 | 2.2 | 42 % | 2.78 | 52% |
| $u_L/u_H$ | 1.3 | 0.7 | 51% | 0.69 | 54 % |

Note: Panel A reports the slope coefficients from regressions of the unemployment rate and unemployment ratio on log GDP per capita and a constant. Panel B reports the difference between the top and bottom quartiles of the world income distribution. The first data column reports the values from our cross-country database. The second and third data columns report the values from the benchmark model and the percent of the data explained. The fourth and fifth columns report the values from the alternative model, with varying $b$, and the percent explained from that model.

### 2.5.4 Sensitivity Analysis

As noted above, the literature provides us with a range of plausible elasticities of substitution rather than a single firm value. In this section, we explore the sensitivity of our model's predictions to the value for the elasticity of substitution. We compute the model's predictions for elasticities 2.5 and 3.5, in particular, in addition to the benchmark value of 3.

We present the results in Table 2.9. Each row reports the slope coefficient from a regression of the variable on question on log GDP per capita. The first column is the data slope coefficients, the second is that of the benchmark model, and the third and fourth columns are the slope coefficients in the model with the lower and higher values of the substitution elasticities, respectively. For the lower value of 2.5, the model underpredicts the slope of the traditional

**Table 2.9**: Sensitivity Analysis of Model Elasticity of Substitution, $\frac{1}{1-\sigma}$

| | Data | Benchmark | $\frac{1}{1-\sigma}=2.5$ | $\frac{1}{1-\sigma}=3.5$ |
|---|---|---|---|---|
| Aggregate traditional sector share | -15.9 | -13.4 | -9.2 | -17.1 |
| Traditional-sector share for low educated | -16.7 | -12.7 | -8.4 | -16.4 |
| Traditional-sector share for high educated | -4.9 | -5.0 | -2.6 | -7.8 |
| Aggregate unemployment rate | 1.8 | 0.5 | 0.1 | 0.9 |
| Unemployment rate for low-educated | 3.2 | 1.7 | 1.2 | 2.1 |
| Unemployment rate for high-educated | 0.5 | 0.4 | 0.2 | 0.5 |
| Ratio of unemployment rates $u_l/u_h$ | 0.5 | 0.25 | 0.17 | 0.32 |
| Relative price $P_T$ | 0.6 | 0.60 | 0.64 | 0.56 |

Note: This table reports the slope coefficients from regressions of the statistics in each row on log GDP per capita and a constant. The first column is for an elasticity of substitution between modern and traditional output of 2.5, the second column is the benchmark model, and the third column is for an elasticity of substitution of 3.5.

sector shares on log GDP per capita. As a result, the aggregate unemployment rate varies less with GDP per capita (0.1 versus 0.5 in the benchmark model), as do unemployment rates for low-educated workers (1.2 versus 1.7 in the benchmark) and high-educated workers (0.2 versus 0.4 in the benchmark). The ratio of low-to-high unemployment rates also varies less with GDP per capita than in the benchmark (0.17 versus 0.25). The relative price varies more than in the benchmark (0.64 versus 0.60).

For the higher value of 3.5, the model over-predicts the slope of the traditional sector share on log GDP per capita. The unemployment rate varies substantially more with GDP per capita than in the benchmark, both in the aggregate and by education level. The unemployment ratio has a slope of 0.32 compared to 0.25 in the benchmark, and is somewhat closer to the slope of 0.5 in the data.

The intuition for these results is as follows. The change in the level of unemployment is driven by the exodus from the traditional sector, which, in turn, is driven by the increase in the ratio of marginal value products of labor: $\frac{A_M}{P_T A_T}$. The smaller is the elasticity of substitution, the less this ratio changes because the rise in $P_T$ offsets the rise in $A_M$ as we move from the poorest to the richest country. In the benchmark model, the slope of this ratio on log GDP per capita is 0.87, only 0.79 when the elasticity is 2.5, and 0.95 when the elasticity is 3.5. That is why the model predicts so much more change in unemployment when the elasticity is 3.5 than when it is 2.5.

We conclude that the model is sensitive to values of the elasticity of substitution between modern- and traditional-sector output. For our benchmark value of 3 the model explains the traditional-sector employment share across countries quite well, suggesting that this may be a sensible value ex-post. For all three of the values chosen, the model underpredicts the slope of the relationship between unemployment and GDP per capita.

## 2.6 Historical Evidence

In this section, we report historical evidence from countries that have high income per capita today to explore how their average unemployment rates have evolved over the long run with income levels. We first look at aggregate unemployment rates from Australia, France, Germany, the United Kingdom and the United States in the period before World War I compared to the most recent evidence. We then look at more disaggregate evidence from the United States.

**Table 2.10**: Historical Unemployment Rates

| Country | Early Period (source) | Unemployment Early | Unemployment Recent | Difference (p-value) |
|---|---|---|---|---|
| Australia | 1901 - 1913 (Mitchell 1992) | 5.17 | 5.26 | 0.09 (.48) |
| France | 1895 - 1913 (Mitchell 1992) | 7.35 | 8.91 | 1.55*** (.00) |
| Germany | 1887 - 1913 (Mitchell 1992) | 2.37 | 7.55 | 5.18*** (.00) |
| United Kingdom | 1881 - 1913 (UK Central Statistical Office) | 4.71 | 7.29 | 2.57*** (.00) |
| United States | 1869 - 1913 (Vernon 1994, Mitchell 1992) | 5.11 | 6.38 | 1.27*** (.00) |

Note: The table reports the average unemployment rates in the early and recent periods, and the results of a one-sided permutation test of whether the recent period has a larger unemployment rate. The early period is defined as the years before WWI; and the recent period is defined as a corresponding year to 2016 such that we have the same number of years for the two periods in each country; see the text for exact dates.

### 2.6.1 Historical Unemployment Rates

The earliest evidence on unemployment that we can find comes from the late 19th century or early 20th century. For simplicity, we consider two periods: an early period containing all data pre-World War I, and a recent period comprised of the most recently available data covering the same number of years. There are five countries for which we found aggregate unemployment data for at least ten years before WWI started in 1914: Australia, France, Germany, the United Kingdom and the United States. The recent period is then defined as 2004 - 2016 for Australia, 1998 - 2016 for France, 1990 - 2016 for Germany, 1984 - 2016 for the UK, and 1972 - 2016 for the U.S. The recent aggregate unemployment rate data are combined from the World Bank, the U.K. office for National Statistics, and the U.S. BLS.

Table 2.10 reports the average unemployment rates in the early and recent periods for these five countries, the difference between the recent and early periods, and a permutation test of the difference between the recent and early periods. The recent unemployment rate is larger than the early period for all five countries. Among them, Australia's unemployment rate is very similar in the two periods, and the difference is statistically insignificant. For the remaining four, average unemployment is economically and statistically significantly higher in the recent period. France's unemployment is the highest overall in both periods, and rises from 7.4 to 8.9 percent. Germany's unemployment rises from 2.4 to 7.6 percent. The United Kingdom rises from 4.7 to 7.3 percent, and the United States rises from 5.1 to 6.4 percent. All of these countries had very large increases in GDP per capita over this period. We conclude that the historical evidence is consistent with our cross sectional finding that the aggregate unemployment rate increases when GDP per capita increases.

### 2.6.2 Disaggregated U.S. Time Series Evidence

We now turn to evidence from the U.S. time series micro data. These data allow us to go beneath the aggregate unemployment rates and to study what happens to unemployment and traditional sector employment by education group. The data allow us to test our theory's prediction that unemployment rates rose, particularly for the low-educated.[15]

To do so, we draw on the U.S. census every decade from 1910 to 2010 from IPUMS International [101]. To maintain consistency across years, we restrict attention to workers aged 16 and over in all states except Alaska and Hawaii. The first row of Table 2.11 reports the slope coefficients from regressions of the unemployment rates on log GDP per capita and a constant. As the table shows, unemployment rates rose with log GDP per capita on average, particularly for the

[15] Strictly speaking, our theory applies to comparisons across steady states, so the predictions in this section are suggested by our theory rather than directly derived from it.

**Table 2.11**: Slope Coefficients for U.S. Time Series

| | All Workers | Worker Education Group Low | High | Ratio |
|---|---|---|---|---|
| Unemployment rate | 3.3** (1.6) | 10.6*** (2.3) | 3.8** (1.6) | .7** (.3) |
| Traditional sector share | -2.6** (1.0) | -1.6 (1.3) | -.4 (.7) | |

Note: The table reports the slope coefficients from regressions of unemployment rates and the traditional sector share on log GDP per capita and a constant. ***, ** and * indicate statistical significance at the 1-percent, 5-percent and 10-percent levels.

less-educated. The estimated slope of the ratio of low-educated unemployment to high-educated unemployment is 0.7 using these data, compared with 0.5 in the cross-country data. We conclude that disaggregated unemployment rates from historical U.S. data are largely consistent with our theory and our cross-country evidence.

Our theory also predicts that the size of the traditional sector has fallen over time in the United States. To test this prediction, we use the census data from 1960 to 2010 to measure the size of the traditional sector according to our proxy of self-employed workers in low-skilled occupations. The second row of Table 2.11 reports the slope coefficient from a regression of the traditional sector share on log GDP per capita and a constant. As the theory predicts, the traditional-sector share decreases significantly with log GDP per capita, mostly driven by the decrease for the low-education group. We conclude that our theory performs adequately here as well.

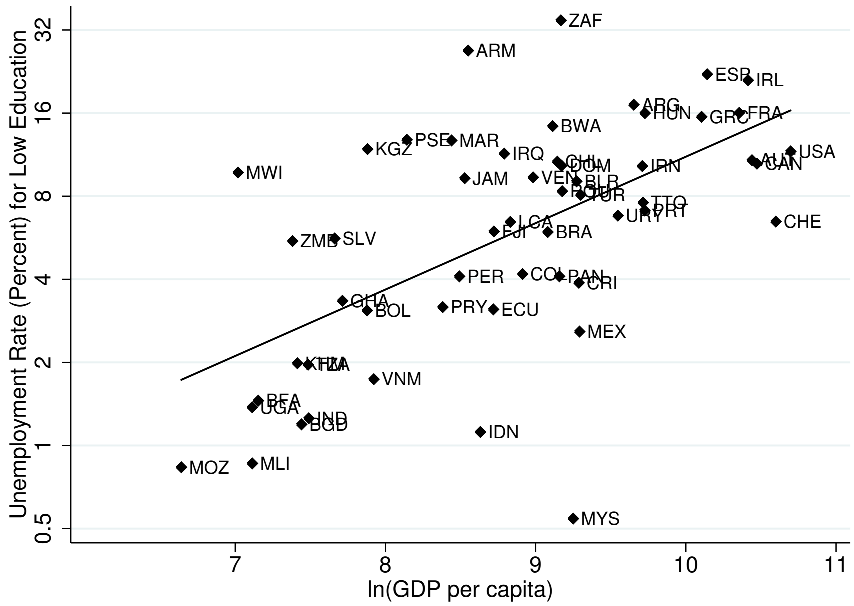
## 2.7 Conclusions

We draw on household survey evidence from around the world to document that unemployment rates are higher, on average, in rich countries than in poor countries. The pattern is particularly pronounced for the less-educated, whose unemployment rates are strongly increasing in GDP per capita, whereas unemployment for the more-educated is roughly constant on average across countries. Our findings imply that the low-educated are more likely to be unemployed than the high-educated in rich countries, whereas the opposite is true in poor countries.

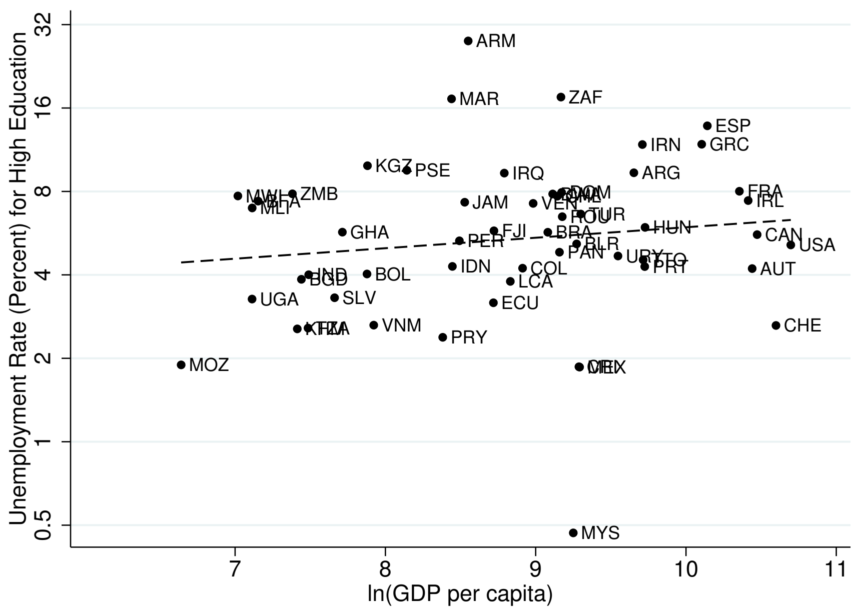
To explain these facts, we build a two-sector model that combines labor search, as in [47] and [103], with a traditional self-employment sector, as in [109]. In our model, countries differ exogenously in the productivity of the modern sector, in which worker productivity depends on ability, and workers offer their services in a labor market with search frictions. All countries have access to an identical traditional sector governed by self-employment and production in which ability plays no role. As such, our model features skill-biased technology differences across countries, as emphasized by, for example, [37]. Workers are heterogeneous and sort as in [125]. As productivity of the modern sector rises, progressively more workers sort into the modern sector. Unemployment levels rise, and particularly so for the less able, as proxied by low education in our empirical findings. A quantitative analysis of the model shows that the model explains a reasonable fraction – on the order of one third – of the cross-country facts that we document.

Our model suggests that at least some rise in unemployment is a natural consequence of the development process, as skilled workers search for jobs, rather than a sign of worsening economic opportunities as countries grow. At the same time, by making unemployment more predictable, we take the first steps toward providing a benchmark against which policy makers can judge the efficiency of their labor markets.

Chapter 2, in part, is currently being prepared for submission for publication of the material. It is coauthored with David Lagakos and James E. Rauch. The dissertation author was a primary investigator and author this material.



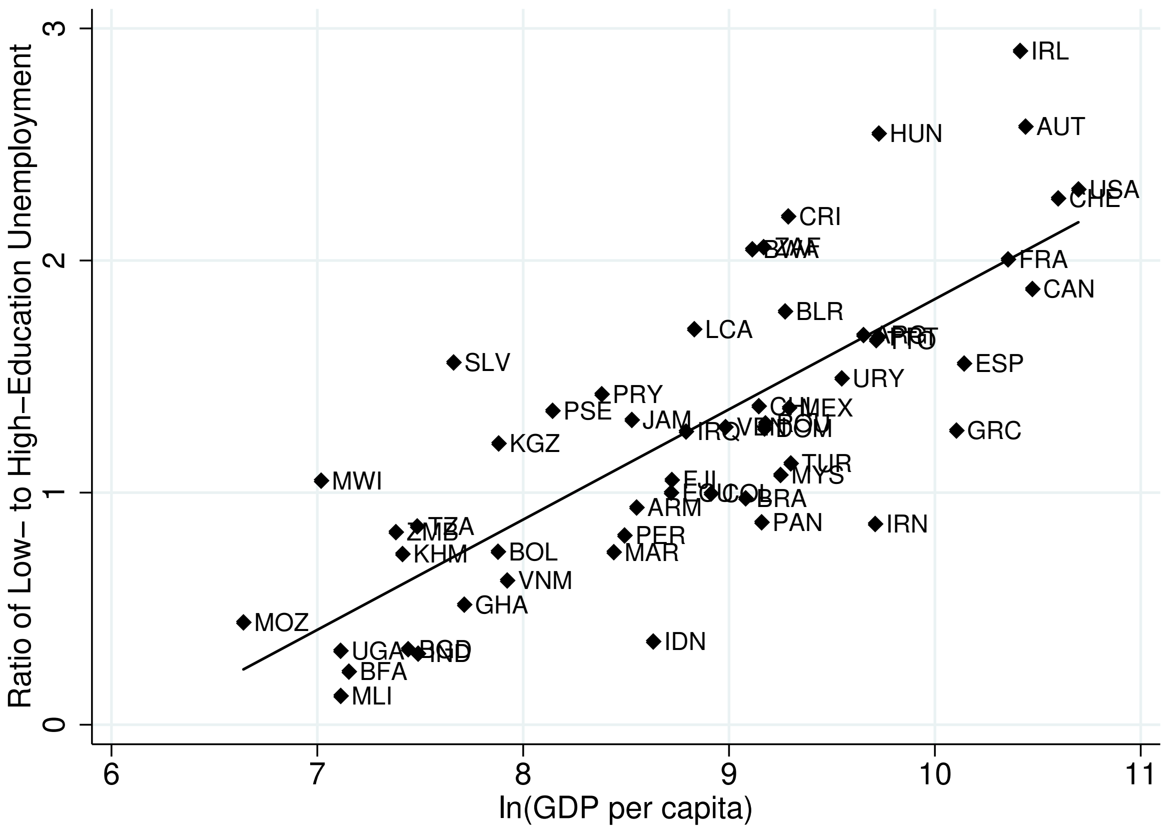

(a) Low-Education Group




(b) High-Education Group

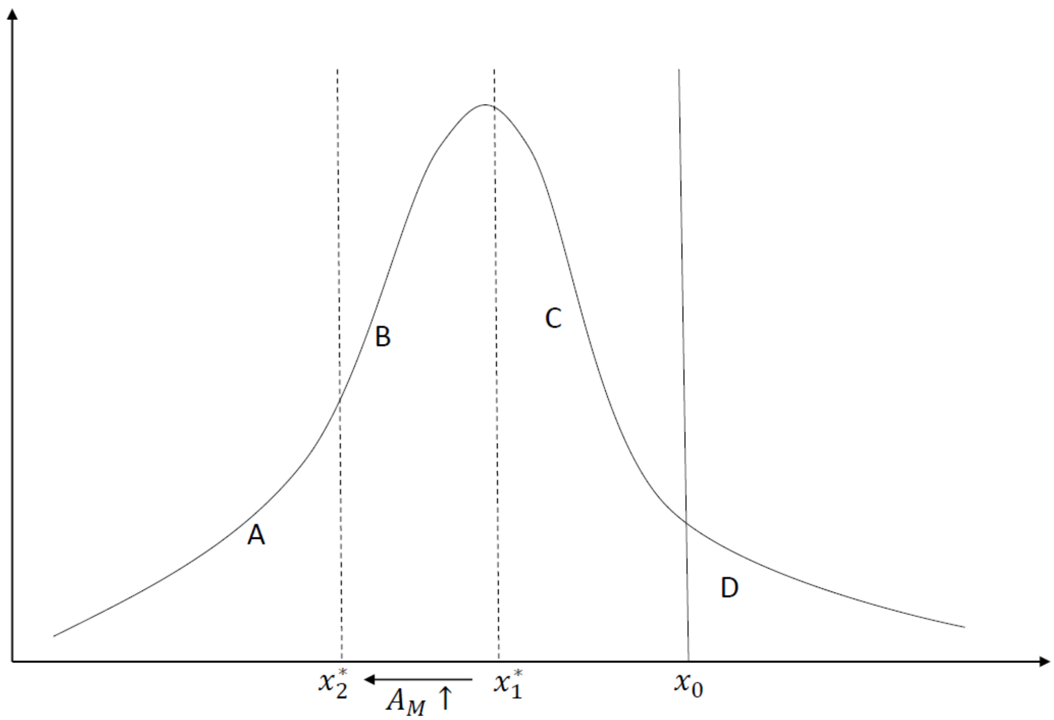
Note: This figure plots the average unemployment rate for prime-aged adults by education level in each country with at least two observations across all years of data from all tiers. Low education means less than secondary school completed; high-education means secondary school completed or more.

**Figure 2.2**: Unemployment Rates by GDP per capita and Education



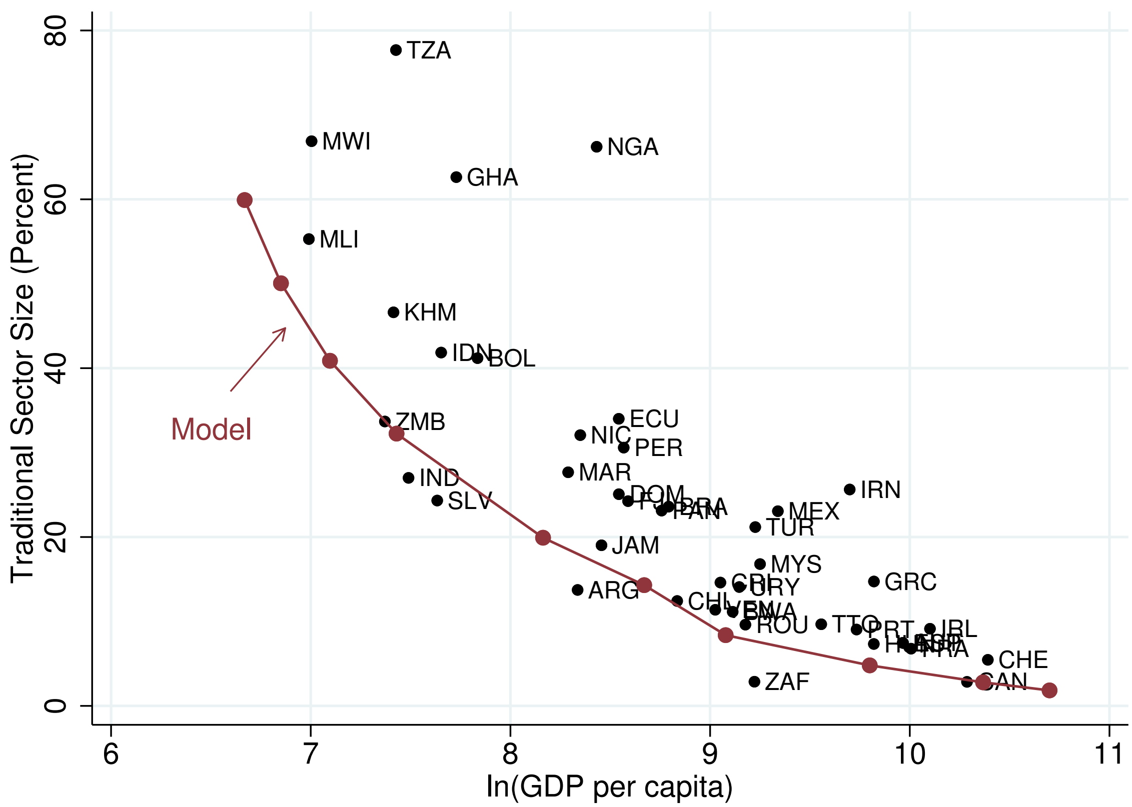

Note: This figure plots the average unemployment ratio of the low-educated workers over the high-educated workers for prime-aged adults across all years of data for each country with at least two years' observations, for Tiers 1, 2 and 3 of surveys. See the Data Appendix for more details.

**Figure 2.3**: Ratio of Unemployment Rates for Low- to High-Educated



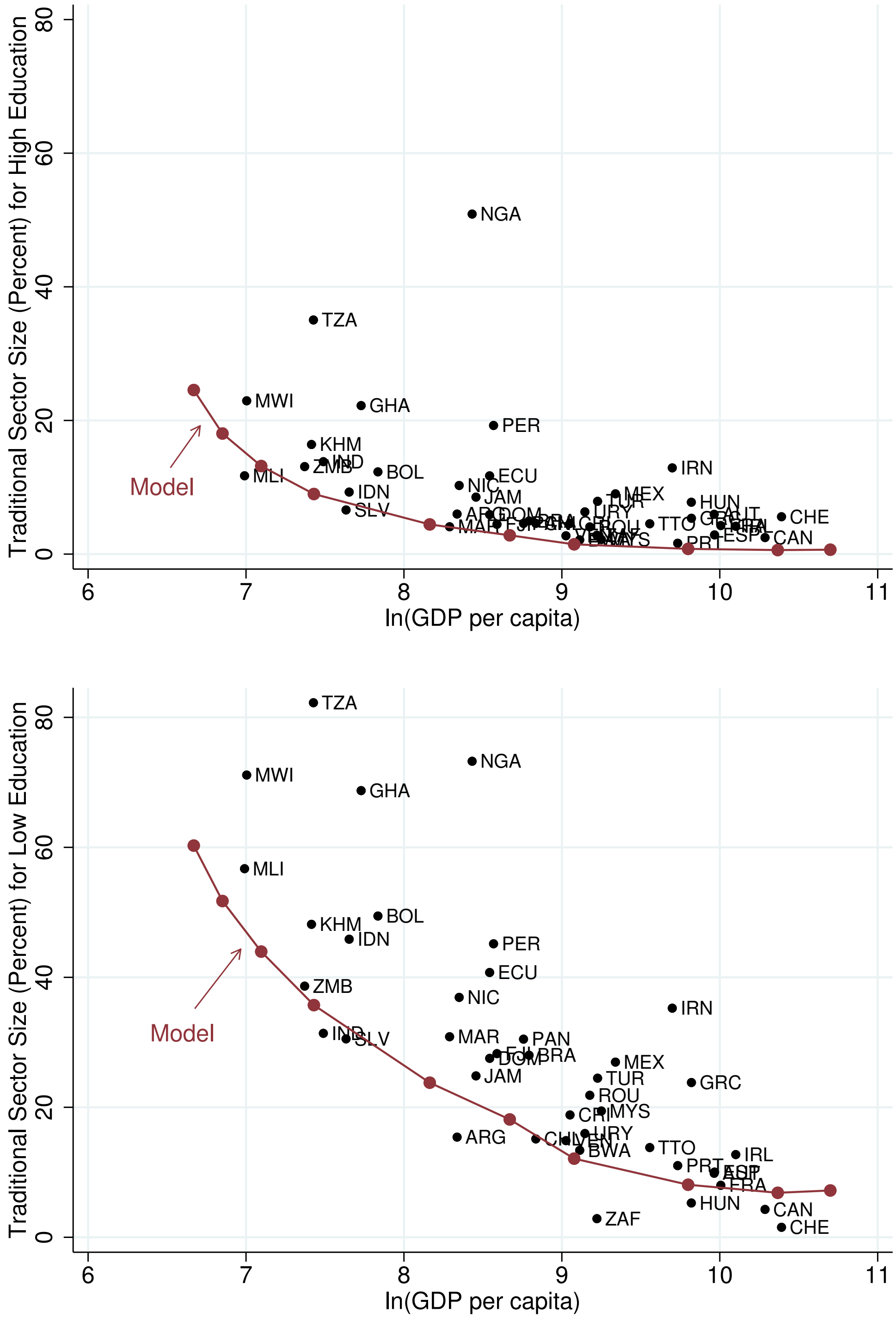

Note: This figure illustrates comparative statics in $A_M$, characterized formally in Propositions 2 and 3.

**Figure 2.4**: Comparative Statics in $A_M$ in Simple Model



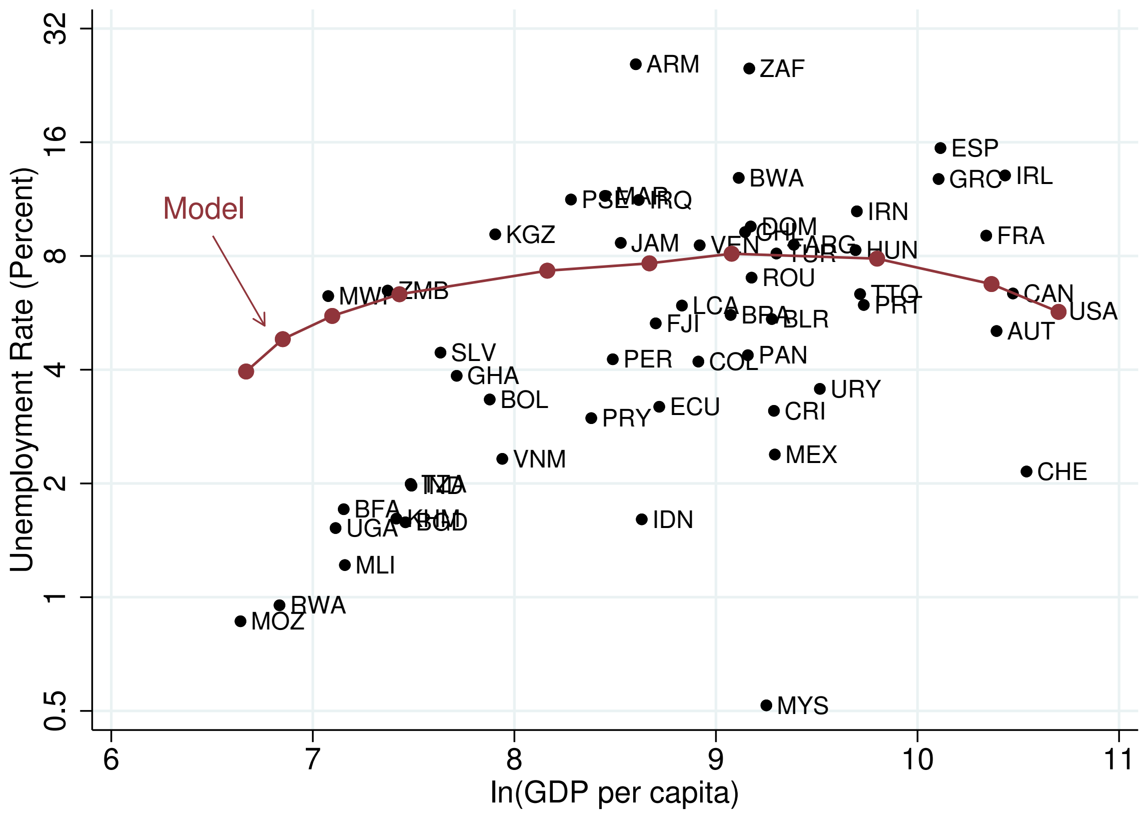

Note: This figure plots the size of the traditional sector against log GDP per capita in the data and model. Each dot represents one country, and the solid line is the prediction of the quantitative model.

**Figure 2.5**: Traditional-Sector Share in Model and Data



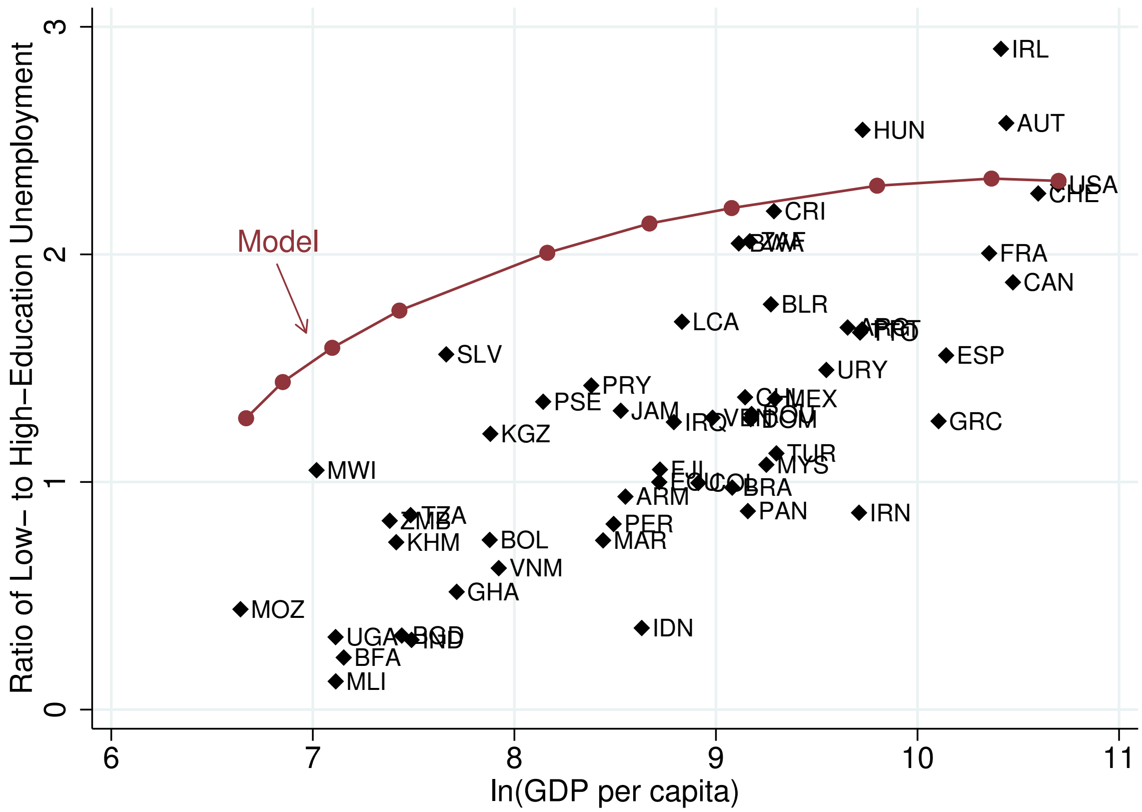

Note: This figure plots the size of the traditional sector against log GDP per capita in the data and model. The top panel is for high-educated workers, and the bottom is for low-educated workers.

**Figure 2.6**: Traditional-Sector Share by Education




Note: This figure plots the aggregate unemployment rate against log GDP per capita. Each dot represents one country in our data, and the solid line is the prediction of the quantitative model.

**Figure 2.7**: Unemployment Rates in the Model and Data




Note: This figure plots the ratio of unemployment for the low-educated to unemployment for the high-educated. Each dot represents one country in our database. The solid line is the prediction of the quantitative model.

**Figure 2.8**: Unemployment Ratio in the Model and Data

## 2.8 Appendices

### 2.8.1 Data Appendix

Among the 199 surveys listed below, there are 11 from earlier than 1990, 59 from the 1990s, 88 from the 2000s, and 41 from 2010 and later. Among the 84 countries, there are 55 for which we have at least two surveys.

**Table 2.12**: Tier 1: Most Comparable Surveys

Tier 1a: Searched for work last week

| Country | Year | Source |
|---|---|---|
| Azerbaijan | 1995 | Survey of Living Conditions |
| Bangladesh | 2000, 2005, 2010 | Household Income-Expenditure Survey (HIES) |
| Bolivia | 1992, 2001 | IPUMS-I |
| Botswana | 2001, 2011 | IPUMS-I |
| Brazil | 2010 | IPUMS-I |
| Burkina Faso | 2014 | LSMS |
| Burkina Faso | 2006 | IPUMS-I |
| Canada | 2011 | IPUMS-I |
| Chile | 1992, 2002 | IPUMS-I |
| Colombia | 1993, 2005 | IPUMS-I |
| Costa Rica | 2000, 2011 | IPUMS-I |
| Cuba | 2002 | IPUMS-I |
| Dominican Republic | 2002 | IPUMS-I |
| Ecuador | 1990, 2001, 2010 | IPUMS-I |
| El Salvador | 1992 | IPUMS-I |
| Fiji | 2007 | IPUMS-I |
| Ghana | 1984, 2000 | IPUMS-I |
| Ghana | 1998 | Living Standards Survey |
| Greece | 1996, 2001, 2011 | IPUMS-I |
| Hungary | 2011 | IPUMS-I |
| India | 1983, 1987, 1993, 1999, 2004 | IPUMS-I |
| Indonesia | 1990, 1995, 2010 | IPUMS-I |
| Indonesia | 2014 | Indonesia Family Life Survey |
| Ireland | 2011 | IPUMS-I |
| Jamaica | 1991, 2001 | IPUMS-I |
| Kenya | 2009 | IPUMS-I |
| Malaysia | 1991, 2000 | IPUMS-I |
| Mexico | 1990, 1995, 2000, 2010, 2015 | IPUMS-I |
| Mongolia | 2000 | IPUMS-I |
| Mozambique | 1997, 2007 | IPUMS-I |
| Nigeria | 2010 | IPUMS-I |
| Pakistan | 1973 | IPUMS-I |
| Panama | 1990, 2000, 2010 | IPUMS-I |
| Paraguay | 1992 | IPUMS-I |
| Peru | 2007 | IPUMS-I |
| Peru | 1994 | Living Standards Survey |
| Philippines | 1990 | IPUMS-I |
| Poland | 2002 | IPUMS-I |
| Portugal | 1991, 2001 | IPUMS-I |
| Romania | 1992, 2002, 2011 | IPUMS-I |
| Rwanda | 2002 | IPUMS-I |
| Saint Lucia | 1980, 1991 | IPUMS-I |
| South Africa | 1993 | Integrated Household Survey |
| South Sudan | 2008 | IPUMS-I |
| Spain | 2011 | IPUMS-I |
| Sudan | 2008 | IPUMS-I |
| Tajikistan | 1999 | LSMS |
| Tanzania | 2002, 2012 | IPUMS-I |
| Trinidad and Tobago | 1970, 1980, 1990, 2000, 2011 | IPUMS-I |
| Uganda | 1991, 2002 | IPUMS-I |
| United States | 1960 | IPUMS-I |
| Venezuela | 2001 | IPUMS-I |
| Zambia | 1990, 2010 | IPUMS-I |

Tier 1b: Searched for work in the last 4 weeks

| Country | Year | Source |
|---|---|---|
| Argentina | 1991 | IPUMS-I |
| Armenia | 2011 | IPUMS-I |
| Belarus | 2009 | IPUMS-I |
| Brazil | 2000 | IPUMS-I |
| Canada | 1991, 2001 | IPUMS-I |
| Dominican Republic | 2010 | IPUMS-I |
| Italy | 2001 | IPUMS- I |
| Jordan | 2004 | IPUMS-I |
| Panama | 2010 | IPUMS-I |
| Paraguay | 2002 | IPUMS-I |
| South Africa | 2007, 2011 | IPUMS-I |
| United States | 1970, 1980, 1990, 2000, 2005 | IPUMS |
| United States | 2001-2014 | American Community Survey (ACS) |
| Bosnia and Herzegovina | 2004 | Living in Bosnia and Herzegovina Survey |
| Brazil | 1997 | Survey of Living Conditions |
| Bulgaria | 2007 | Multi-topic Household Survey |
| Iran | 2011 | IPUMS-I |
| Iraq | 2012 | Household Socio-economic Survey |
| Malawi | 2013 | Integrated Household Panel Survey |
| Serbia | 2007 | LSMS |
| Uganda | 2011 | National Panel Survey |

**Table 2.13**: Tier 2: Comparable Search Questions, Less Comparable Duration Questions

| Country | Year | Source | Seeking window |
|---|---|---|---|
| Armenia | 2001 | IPUMS-I | Current |
| Bangladesh | 1991, 2001 | IPUMS-I | 7 days, main activity |
| Bangladesh | 2011 | IPUMS-I | Current status |
| Brazil | 1980 | IPUMS-I | Current |
| Burkina Faso | 1996 | IPUMS-I | At least 3 days in the last week |
| Cambodia | 1998, 2008 | IPUMS-I | 6 months |
| Egypt | 2006 | IPUMS-I | current |
| El Salvador | 2007 | IPUMS-I | Current/ last week |
| France | 2006, 2011 | IPUMS-I | Current |
| Haiti | 2003 | IPUMS-I | Last month |
| Hungary | 1990 | IPUMS-I | Current |
| Iran | 2006 | IPUMS-I | Past 30 days |
| Iraq | 1997 | IPUMS-I | Current |
| Ireland | 1991, 1996, 2002, 2006 | IPUMS-I | Current |
| Kyrgyz Republic | 1999, 2009 | IPUMS-I | Current |
| Malawi | 2008 | IPUMS-I | Last year |
| Mali | 1998, 2009 | IPUMS-I | 4 weeks |
| Morocco | 1994, 2004 | IPUMS-I | Current |
| Nicaragua | 2005 | IPUMS-I | 2 weeks |
| Portugal | 2011 | IPUMS-I | Current |
| Rwanda | 1991 | IPUMS-I | Most of the week |
| Senegal | 2002 | IPUMS-I | Continuously for at least 3 months |
| Sierra Leone | 2004 | IPUMS-I | 4 weeks |
| South Africa | 1996 | IPUMS-I | Current |
| Switzerland | 2000 | IPUMS-I | Current |
| Turkey | 1990, 2000 | IPUMS-I | Current |
| Uruguay | 2006, 2011 | IPUMS-I | 4 weeks |
| Venezuela | 1990 | IPUMS-I | Current |
| Zambia | 2000 | IPUMS-I | Primary activity, 7 days |

**Table 2.14**: Tier 3: Least Comparable Search or Activity Questions

| Country | Year | Source | Activity | Search |
|---|---|---|---|---|
| Argentina | 2001, 2010 | IPUMS-I | Exclude: for self-consumption | 4 weeks |
| Austria | 1991 | IPUMS-I | A minimum average of 12 hours per week | Current |
| Austria | 2001 | IPUMS-I | 7 days | Only previously employed |
| Austria | 2011 | IPUMS-I | No text | No text |
| Belarus | 1999 | IPUMS-I | Exclude: for self-consumption | Yes |
| Botswana | 2011 | IPUMS-I | 4 Weeks | |
| Cameroon | 2005 | IPUMS-I | 7 Days | Last 7 days for worked before; now for looking for the first job |
| China | 1990 | IPUMS-I | No text | No text |
| Ethiopia | 2007 | IPUMS-I | Standard | No text |
| France | 1990, 1999 | IPUMS-I | Current | Enrollment ANPE |
| Fiji | 1996 | IPUMS-I | Worked for money | Not comparable |
| Ghana | 2010 | IPUMS-I | No text | No text |
| Hungary | 2001 | IPUMS-I | Current | Unemployment benefit |
| India | 2009 | IPUMS-I | Standard | Only 12 months main activity available |
| Liberia | 2008 | IPUMS-I | 12 Months | 12 months |
| Netherlands | 2001 | IPUMS-I | No Text | Not comparable |
| Palestine | 1997, 2007 | IPUMS-I | 7 Days | Included did not seek but want to work |
| Peru | 1993 | IPUMS-I | Not comparable | Not comparable |
| Portugal | 1981 | IPUMS-I | 7 Days | Text not available |
| Slovenia | 2002 | IPUMS-I | Current | Registered as unemployed at the employment service of Slovenia |
| Spain | 1991, 2001 | IPUMS-I | 7 Days | Unemployed, worked previously |
| South Africa | 2001 | IPUMS-I | 4 Weeks | Could not find work |
| Switzerland | 1990 | IPUMS-I | Principal occupation | Current |
| Ukraine | 2001 | IPUMS-I | Status | Unemployment allowances, unemployed |
| Vietnam | 2009, 1991 | IPUMS-I | Earn income | 4 Weeks |

### 2.8.2 Model Derivation and Proofs

#### Model Derivations

In this subsection, we develop the expressions for $U(x)$ and $w(x)$, and show the intermediate steps to develop Equation (2.9). We start by simplifying Equations (2.4) - (2.7) to

$$(1-\delta)U(x) = A_M bx + \delta\eta\theta^{1-\alpha}\big[E(x) - U(x)\big] \tag{2.14}$$

$$(1-\delta)E(x) = w(x) + \delta s\big[U(x) - E(x)\big] \tag{2.15}$$

$$J(x) = \frac{A_M x - w(x)}{1-\delta(1-s)} \tag{2.16}$$

$$\big(1 - G(x^*)\big)A_M c = \delta\eta\theta^{-\alpha}\int_{x^*}^{\bar{x}} J(x)g(x)dx. \tag{2.17}$$

The firm receives $(1-\beta)S(x) = (1-\beta)[E(x) - U(x) + J(x)] = J(x)$ when a vacancy is filled. Combining this division of surplus with equation (2.16) gives

$$E(x) - U(x) = \frac{\beta}{1-\beta}\frac{A_M x - w(x)}{1-\delta(1-s)}. \tag{2.18}$$

Substituting equation (2.18) into equation (2.14) yields

$$U(x) = \frac{1}{1-\delta}\Big(A_M bx + \delta\eta\theta^{1-\alpha}\frac{\beta}{1-\beta}\frac{A_M x - w(x)}{1-\delta(1-s)}\Big). \tag{2.19}$$

We can then solve for $w(x)$ by combining equations (2.19) and (2.18) with equation (2.15):

$$w(x) = \frac{A_M bx}{1+k(\theta)} + \frac{k(\theta)}{1+k(\theta)}A_M x, \text{ with } k(\theta) = \frac{\beta(\delta\eta\theta^{1-\alpha} + 1 - \delta + \delta s)}{(1-\beta)(1-\delta+\delta s)}.$$

Substituting this solution into equations (2.16) and (2.19) gives us, respectively,

$$J(x) = \frac{A_M x(1-b)(1-\beta)}{\beta\delta\eta\theta^{1-\alpha}+1-\delta+\delta s} \tag{2.20}$$

$$U(x) = \frac{1}{1-\delta}\Big(A_M bx + \delta\eta\theta^{1-\alpha}\frac{\beta}{1-\beta}\frac{A_M x(1-b)(1-\beta)}{\beta\delta\eta\theta^{1-\alpha}+1-\delta+\delta s}\Big). \tag{2.21}$$

Equation (2.21) appears as equation (8) in the text. Finally, substituting equation (2.20) into equation (2.17) and dividing both sides by $1-G(x^*)$ yields equation (2.9) that determines $\theta$ for any given level of $x^*$:

$$c = \frac{(1-\beta)\delta\eta\theta^{-\alpha}}{\beta\delta\eta\theta^{1-\alpha}+1-\delta+\delta s}(1-b)\mathbb{E}(x|x>x^*).$$

**Proof of Proposition 1**

Equations (2.9) and (2.10) allow us to solve for unique values of $\theta$ and $x^*$. We first simplify equation (2.10) to

$$\theta^{1-\alpha} = \frac{(A_T - A_M bx^*)(1-\delta+\delta s)}{\beta\delta\eta(A_M x^* - A_T)}. \tag{2.22}$$

Substitute this expression into equation (2.9), yielding a single equation that determines $x^*$:

$$\frac{(A_T - A_M bx^*)^{\frac{\alpha}{1-\alpha}} A_M x^*(1-b)c(1-\delta+\delta s)^{\frac{1}{1-\alpha}}}{(A_M x^* - A_T)^{\frac{1}{1-\alpha}}} = (1-\beta)(\delta\eta)^{\frac{1}{1-\alpha}}\beta^{\frac{\alpha}{1-\alpha}}(1-b)\mathbb{E}(x|x>x^*). \tag{2.23}$$

We assume that a solution $x^* \in (\underline{x}, \bar{x})$ to equation (2.23) exists. Since the existence of this solution implies that $A_M x^* - A_T > 0$, it also implies the existence of a solution $\theta > 0$. Moreover, if the solution $x^*$ is unique, then the solution $\theta$ is also unique.

To demonstrate uniqueness of the solution $x^*$, we first show that the left-hand side of equation (2.23) is decreasing in $x^*$. Inspection of equation (2.23) shows that a sufficient condition is that $A_M x^*/(A_M x^* - A_T)^{\frac{1}{1-\alpha}}$ is decreasing in $x^*$. We have

$$\begin{aligned} sign\left[\frac{d\frac{A_M x^*}{(A_M x^* - A_T)^{\frac{1}{1-\alpha}}}}{dx^*}\right] &= sign\left[(A_M x^* - A_T)^{\frac{1}{1-\alpha}} - \frac{x^*}{1-\alpha}(A_M x^* - A_T)^{\frac{\alpha}{1-\alpha}} A_M\right] \\ &= sign\left[-A_T - \frac{\alpha A_M x^*}{1-\alpha}\right], \end{aligned}$$

which is negative. Since the right-hand side of equation (2.23) is increasing in $x^*$, then the $x^*$ that solves equation (2.23) must be unique.

Having demonstrated that the solution is unique, we turn to comparative statics of an increase in $A_M$. We want to show that the left-hand side of equation (2.23) is decreasing in $A_M$. It is sufficient to show:

$$\begin{aligned} sign\left[\frac{d\frac{A_M x^*}{(A_M x^* - A_T)^{\frac{1}{1-\alpha}}}}{dA_M}\right] &= sign\left[(A_M x^* - A_T)^{\frac{1}{1-\alpha}} - A_M\frac{1}{1-\alpha}(A_M x^* - A_T)^{\frac{\alpha}{1-\alpha}} x^*\right] \\ &= sign\left[(A_M x^* - A_T) - \frac{A_M x^*}{1-\alpha}\right] \\ &= sign\left[-A_T - \alpha\frac{A_M x^*}{1-\alpha}\right] \end{aligned}$$

Thus, we know that the sign of this derivative must be negative. We already know that the left- and right-hand sides of equation (2.23) are decreasing and increasing in $x^*$, respectively, so $dx^*/dA_M < 0$ follows.

**Proof of Proposition 2**

It follows from Proposition 1 that $x^*$ decreases with $A_M$. As $x^*$ decreases, we see from equation (2.9) that $\theta$ decreases. Inspection of equation (2.11) then shows that $u$ must increase.

**Proof of Proposition 3**

The unemployment rate for workers with $x < x_0$ is a weighted average of $\frac{s}{s+\eta\theta^{1-\alpha}}$, for workers with $x^* < x < x_0$, and 0, for workers with $x < x^*$. Therefore

$$\mathbb{E}\big(u|x < x_0\big) = \frac{\frac{s}{s+\eta\theta^{1-\alpha}}\big(G(x_0) - G(x^*)\big) + 0 \cdot G(x^*)}{G(x_0)} = \frac{\frac{s}{s+\eta\theta^{1-\alpha}}\big(G(x_0) - G(x^*)\big)}{G(x_0)}. \tag{2.24}$$

The ratio of this unemployment rate to the unemployment rate for workers with ability higher than $x_0$ is

$$\frac{\mathbb{E}\big(u|x < x_0\big)}{\mathbb{E}\big(u|x > x_0\big)} = \frac{\frac{s}{s+\eta\theta^{1-\alpha}}\big(G(x_0) - G(x^*)\big)}{G(x_0)} \Big/ \frac{s}{s+\eta\theta^{1-\alpha}} = 1 - \frac{G(x^*)}{G(x_0)}. \tag{2.25}$$

This ratio increases with $A_M$ since $x^*$ decreases with $A_M$, as proved in Proposition 1.

**Proof of Lemma 1**

We can solve for market tightness $\theta_h$ and $\theta_l$ and cutoff ability levels $x_h^*$ and $x_l^*$ using the equivalents of equations (2.9) and (2.22) for the high- and low-educated labor markets in the quantitative model:

$$c = \frac{(1-\beta)\delta\eta\theta_h^{-\alpha}}{\beta\delta\eta\theta_h^{1-\alpha} + 1 - \delta + \delta s_h}(1-b)\mathbb{E}_h(x|x > x_h^*) \tag{2.9$h$}$$

$$\theta_h^{1-\alpha} = \frac{(A_T - A_M b x_h^*)(1-\delta+\delta s_h)}{\beta\delta\eta(A_M x_h^* - A_T)} \tag{2.22h}$$

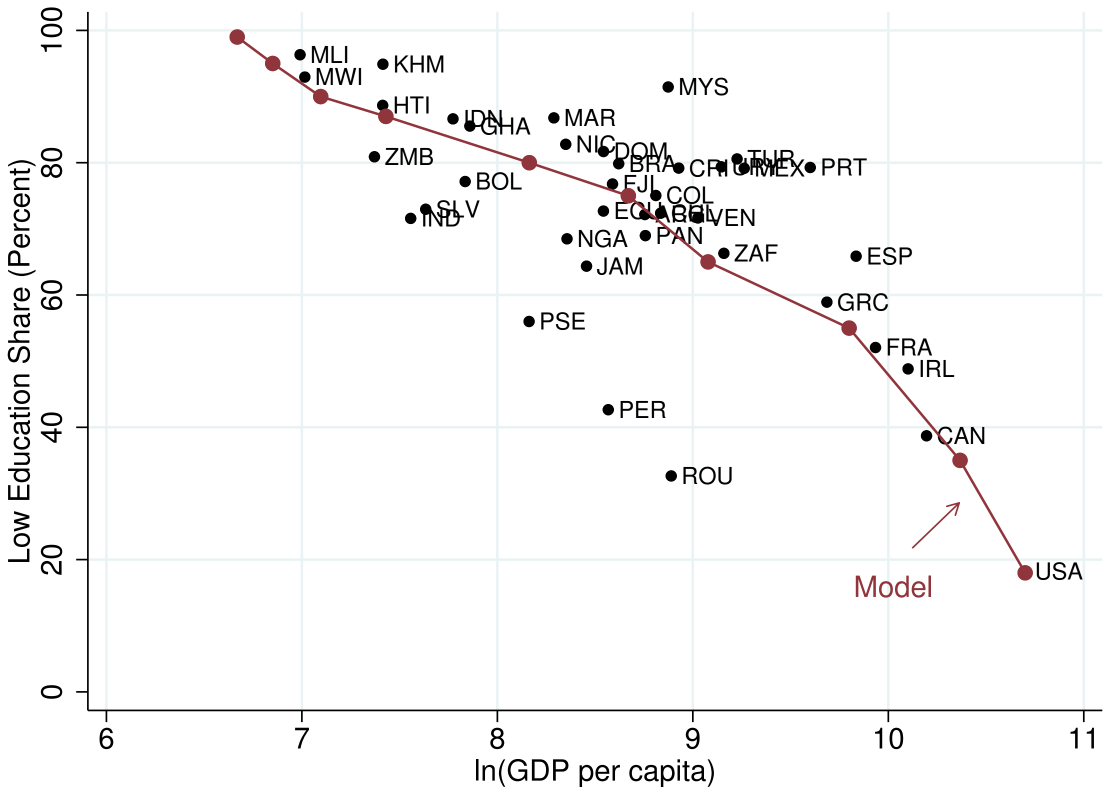
$$c = \frac{(1-\beta)\delta\eta\theta_l^{-\alpha}}{\beta\delta\eta\theta_l^{1-\alpha} + 1 - \delta + \delta s_l}(1-b)\mathbb{E}_l(x|x > x_l^*) \tag{2.9l}$$

$$\theta_l^{1-\alpha} = \frac{(A_T - A_M b x_l^*)(1-\delta+\delta s_l)}{\beta\delta\eta(A_M x_l^* - A_T)}, \tag{2.22l}$$

where $\mathbb{E}_h$ and $\mathbb{E}_l$ are computed using $g_h(x)$ and $g_l(x)$, respectively.

It follows that equation (2.23) that determines $x^*$ can, with appropriate subscripting, determine $x_h^*$ or $x_l^*$. We showed in the proof of Proposition 1 that the left- (right-) hand side of equation (2.23) is decreasing (increasing) in $x^*$. Inspection of the left-hand side of equation (2.23) shows that it is increasing in $s$, hence, any increase in $s$ from $s_h$ to $s_l$ must increase $x_l^*$ relative to $x_h^*$. Inspection of the right-hand side of equation (2.23) shows that it is increasing in $\mathbb{E}(x|x > x^*)$; thus, computing the expectation using $g_h(x)$ relative to $g_l(x)$ must decrease $x_h^*$ relative to $x_l^*$, because $G_h(x)$ first-order stochastically dominates $G_l(x)$.

### 2.8.3 Appendix Figures and Tables




Note: This figure plots the values of λ used in the quantitative experiments of Section 3.5 (solid line), and the percent of the labor force that is low-educated in each of our countries (dots with identifiers). The data come from IPUMS. Low-educated individuals are defined to be those with less than a secondary school education.

**Figure 2.9**: Low-Education Share, λ, in Model and Data

**Table 2.15**: Definition of Traditional Sector Goods

| Item | Details |
|---|---|
| Shoe Repair - Women Street Shoes | Replacement of 2 heels (glued and nailed);<br>While-you-wait in shop service;<br>Heel: Synthetic polyurethane, small heel. |
| Shoe Repair - Men Classic Shoes | Re-soling rubber soles (glued & nailed or stitched);<br>Not "urgent" in shop service. |
| Shoeshine | Cleaning leather shoes with a brush and polishing;<br>Manual work while keeping the shoes on;<br>Exclude service in a shop. |
| Taxi | 7 km in the town center on working days at 3 p.m.;<br>Includes: Possible fixed starting fee + price per km;<br>Excludes: Taxi called by telephone. |
| Men basic haircut | Scissor cut of short hair for male adults;<br>Type of establishment: Common men's barber shop;<br>No shampoo/washing nor styling/fixing products;<br>Full price including tips if any. |
| Ladies haircut - curlers | Hair with curlers cut to medium (basic) for female adult;<br>Shampoo/washing, blow drying, and styling/fixing products;<br>Establishment: Common hairdresser (exclude hair stylist). |
| Manicure | Standard manicure on natural nails by nail technician;<br>Establishment: Professional beautician;<br>Full price including tips if any;<br>Bath, filing, cuticles treatment, one-color varnishing. |
| Ladies haircut - long hair | Long hair cut to short for female adult;<br>Shampoo/washing, blow drying, styling/fixing products;<br>Establishment: Common hairdresser (exclude hair stylist). |

Note: The table reports the definitions of each ICP traditional service used in Table 2.6, and described in Section 2.5.2. The services come from the unpublished ICP 2011 Global Core list of goods and services.

**Table 2.16**: Slope Coefficients in the Alternative Calibration

| | Data | Model | Alternative Cali. |
|---|---|---|---|
| Aggregate traditional sector share | -15.9 | -13.4 | -15.9 |
| Traditional-sector share for low educated | -16.7 | -12.7 | -15.2 |
| Traditional-sector share for high educated | -4.9 | -5.0 | -6.7 |
| Aggregate unemployment rate | 1.8 | 0.5 | 0.7 |
| Unemployment rate for low-educated | 3.2 | 1.7 | 1.9 |
| Unemployment rate for high-educated | 0.5 | 0.4 | 0.4 |
| Ratio of unemployment rates $u_l/u_h$ | 0.5 | 0.3 | 0.3 |
| Relative price $P_T$ | 0.6 | 0.60 | 0.67 |

Note: The table reports slope coefficients from regressions of the statistics in each row on log GDP per capita.

# Chapter 3

# Development and Selection into Necessity versus Opportunity Entrepreneurship

We are grateful to James Rauch for extensive suggestions and support. We also thank David Lagakos, Prashant Bharadwaj, Douglas Gollin, Ruixue Jia, and Markus Poschke for helpful comments. Any potential errors are our own.

## 3.1 Introduction

Entrepreneurship is often seen as the engine of economic growth. At the same time, it is widely recognized that only a minority of entrepreneurs fuel that engine. Efforts to distinguish that minority from the rest have led to labels for entrepreneurs such as "opportunity" versus "necessity" (the Global Entrepreneurship Monitor), "productive" versus "unproductive" ([18]), and "transformative" versus "subsistence" ([127]). Policies that target entrepreneurs indiscriminately risk wasting most of their impact ([16]).[1]

In this paper, we propose a simple division of entrepreneurs into employers (self-employed with paid employees) and own account workers (self-employed without employees). This division has the advantage of being consistently defined across censuses for 56 countries from 162 country-year surveys. To link our work to the earlier literature, we will use the terms "employers" and "opportunity entrepreneurs" interchangeably and the terms "own account" and "necessity entrepreneurs" interchangeably.

To fix ideas, we develop a simple two-sector general equilibrium model of labor force allocation between opportunity entrepreneurs, necessity entrepreneurs, and wage workers. In the traditional sector, necessity entrepreneurs work on their own accounts without rewards to ability; in the modern sector, employers and wage workers produce with rewards to their abilities.[2] In equilibrium, agents with abilities below a threshold become own-account self-employed workers and agents with abilities higher than the threshold enter the modern sector, becoming wage workers or employers. Higher aggregate productivity is driven by higher returns to ability in

[1] [16] find that microfinance increases significantly the profits of the top tercile of businesses that started before the intervention, but its benefits to the rest of the self-employed (the majority of entrepreneurs) are generally indistinguishable from zero.

[2] The assumption of differential returns to ability (as proxied by years of schooling) in the two sectors is consistent with [123], who argues that schooling has little influence on productivity if the tasks are simple, whereas there are higher returns to schooling if the tasks are substantially complex.

the modern sector due to technological progress, which reduces the threshold ability level. In other words, development draws the more able agents from the traditional sector into the modern sector. Our model thus predicts that, at the aggregate level and across education groups (our proxy for fixed ability levels), the shares of employers and wage workers rise with GDP per capita (hereafter GDPPC) at the expense of own-account workers.

Bringing our predictions to the data, we begin with a multinomial probit model of choice between own-account self-employment, wage employment, and employer status. We find that 91 out of 98 country-year observations have strong negative selection on ability (as proxied by years of schooling) into own-account self-employment, and 81 out of 98 have positive selection into employer status. We also find that own-account self-employment decreases from 83% to 6% as GDPPC increases from I$442 (in 2005 International dollars) to I$41,000 across 162 country-year observations, whereas the employers' rate rises strongly from 0.1% to 14.0% over the same range of GDPPC. Moreover, we show that our parameterized model captures the aggregate quantitative patterns of labor force allocation over a cross-section of countries.

Since farming entrepreneurs account for a considerable portion of the self-employed, especially in developing countries, our findings are in line with the literature that shows sorting by unobserved ability/skill between agriculture and non-agriculture ([86]). However, in this paper we focus the data analysis on industries excluding agriculture, fishing and forestry, and we model occupational choices without market frictions. Our empirical results of negative selection into own-account workers, positive selection into employers, and the impact of development on labor force allocation are obtained when restricting the samples to non-agricultural sectors. Thus our work is essentially independent of the literature on sorting into agriculture.

Our findings that the impacts of selection and economic growth work oppositely on necessity and opportunity entrepreneurs confirm the importance of distinguishing between these two types of entrepreneurship rather than grouping them together as self-employed. Otherwise, the numerical dominance of necessity entrepreneurs can lead researchers to extend conclusions for the self-employed to entrepreneurship in general. For example, Van der Sluis, Van Praag, and Vijverberg (2005, p. 248) use entrepreneurship choice and nonfarm self-employment interchangeably and conclude from their meta-analytical review of empirical studies in developing countries that education lowers the likelihood of nonfarm self-employment. Woodruff (2007, p. 55) interprets the model of [92] to imply that entrepreneurship decreases as an economy's income level rises, because increasing income is associated with a higher wage rate that induces the marginal employer to leave self-employment for a wage job.

**Related Literature.** The papers closest to our work distinguishes between two types of entrepreneurs ([111], [113], [88], [131]). However, they do not investigate the pattern of entrepreneurship as a function of variations in income levels. [88] distinguish between "unincorporated" and "incorporated" entrepreneurs in the US. They find that incorporated entrepreneurs are better educated and have better performance than unincorporated business owners in the United States. However, this classification of incorporated and unincorporated entrepreneurship cannot be used to conduct cross-country analysis. First, data are not widely available, especially for poor countries. Second, the costs and benefits of incorporation differ widely across countries depending on legal systems, tax policies, and levels of corruption. In contrast, we use the same classification as that used by [131] for Germany. This division of entrepreneurs into own-account self-employed and employers is consistently defined across countries by whether paid employees are hired. Even though more stringent regulations may make it harder to hire employees in developed countries, this only biases our estimated positive effect of economic growth on the employers' rate downward.

This paper also relates to the macro-development literature. In particular, [58] shows that self-employment declines over development due to productivity differences. However, we differ from his work in two aspects. First, we use more recent micro household surveys that allows analysis by educational groups and industries, whereas [58] used limited national level reports from International Labor Organization in the 1990s. Second, we distinguishes employers from the necessity entrepreneurs. The fact that the share of employers rises with development emphasizes the importance of distinguishes the two types of self-employment. Compared to other papers ([114], [6], [57]) proposes theories of financial access or tax evasion as determinants of informal sector size, we show that the productivity differences can account for most of the differential labor market division across countries.

The rest of this paper is structured as follows. Section 3.2 develops a two-sector general equilibrium model that explains the impact of development on labor force division. In Section 3.3 and Section 3.4, we present empirical findings and robustness checks for by using household surveys of 56 countries from all income levels. Section 3.5 parameterizes the model to evaluate its quantitative predictions. Conclusions are in Section 3.6. The proofs of all propositions and lemmas are in Appendices.

## 3.2 Model

We start by positing a perfectly competitive general equilibrium model with two sectors: a traditional sector where agents are own-account self-employed without returns to ability, and a modern sector where employers hire wage workers for production with constant returns to ability. The production functions per self-employed worker are

$$y_T = A_T \tag{3.1}$$

$$y_M = A_M G(h, L), \tag{3.2}$$

where $y_T(y_M), A_T(A_M)$ are output and productivity per self-employed worker in the traditional (modern) sector respectively, $h$ is the employer's ability measured in efficiency units and $L$ is the labor input measured in efficiency units.

Let the aggregate production function for the economy take the constant elasticity of substitution (CES) form[3]:

$$y = (\gamma y_T^{\sigma} + (1-\gamma) y_M^{\sigma})^{\frac{1}{\sigma}}. \tag{3.3}$$

Because, in equilibrium, our division of sectors into traditional and modern will match the division of less and more skill-intensive sectors in the literature, we assume here the output of the two sectors are imperfect substitutes. Consistent with empirical consensus (e.g. **?** and **?**), we let $0 < \sigma < 1$: although $y_T$ and $y_M$ are imperfect substitutes, the elasticity of substitution between them is high.[4] We normalize the output price in the modern sector to be 1, and let the output price in the traditional sector be $P_T$. In a competitive market, the relative price of the traditional product equals the ratio of marginal productivities:

$$P_T = \frac{\partial y / \partial y_T}{\partial y / \partial y_M} = \frac{\gamma}{1-\gamma} (\frac{y_M}{y_T})^{1-\sigma}. \tag{3.4}$$

Hence, for an own-account self-employed worker in the traditional sector, the payoff is $A_T P_T$. Regarding production in the modern sector, we assume $G(h,L)$ is homogeneous of degree 1, concave in $L$, and $G(0,L) = G(h,0) = 0$. Rewrite the production function as $y_M = A_M h G(1, \frac{L}{h}) = A_M h g(l)$, where $l \equiv \frac{L}{h}$ and $g(x) \equiv G(1,x)$. It follows that $g(\cdot)$ is concave and $g(0) = 0$.

We assume there is a continuum of risk neutral agents with measure 1 in the economy

[3]None of our analysis would change if this were a utility function, but calibration would become more difficult without qualitatively changing our results.

[4]The elasticity between $y_T$ and $y_M$ equals $\frac{1}{1-\sigma} > 1$ under this assumption.

and each individual is endowed with efficiency units $h$, which is distributed according to $F(h)$ on $[\underline{h},\bar{h}]$. We also assume $F(h)$ is differentiable and let $f(h)=F'(h)$ be its probability density function. Let a wage worker with efficiency units $h$ be paid $wh$, where $w$ is the equilibrium wage per efficiency unit. Employers solve the profit maximization problem to obtain

$$\Pi(w,h)=\max_L A_M G(h,L)-wL. \tag{3.5}$$

The first order condition with respect to $L$ gives

$$A_M G_L(h,L)-w=A_M g'(l)-w=0. \tag{3.6}$$

Equation (3.6) determines $L=hl(w)$ as a function of $w$. Substituting it into equation (3.5) gives $\Pi(w,h)=h\pi(w)$, where $\pi(w)=A_M(g(l(w))-l(w)g'(l(w))$. Because $w$ is taken as given by the agents, both wage workers and employers see a linear return to abilities. Therefore, the equilibrium condition requires all agents in the modern sector to be indifferent between a wage job and being an employer, otherwise there will either be no employers or no wage workers.[5] Mathematically, we have

$$\Pi(w,h)=h\pi(w)=wh, \text{ or } \pi(w)=w. \tag{3.7}$$

Since $\pi(w)$ is decreasing in $w$, equation (3.7) uniquely determines $w$. Note that since $\Pi(w,h)$ is linear homogeneous in $A_M$ and $w$, so is $\pi(w)$. Then by equation (3.7) $w$ must increase proportionately to $A_M$. Therefore, equation (3.6) yields an invariant l.[6]

The unique $l$ characterizes the partition of talent allocation between employers and wage

[5] There is an equilibrium ratio of efficiency units between wage workers and employers in the modern sector, but whether an individual becomes a wage worker or employer is indeterminate.

[6] The existence of this unique $l$ is proved in Lemma 4 in the Appendices.

workers. Because $l$ is invariant across firms and countries, it implies that $L$, the firm size in efficiency units, is growing with $h$, efficiency units of the employer. Note that all agents in the modern sector are indifferent between being a wage worker and an employer, so firms that hire more efficiency units will on average have more workers as well. This is consistent with [92] in the sense that employers with greater talent yield larger firm sizes measured by the number of workers hired. In addition, since all agents in the modern sector are equally likely to be employers, the probability density function of firm sizes will be equivalent to a truncated distribution of $h$ for those workers in the modern sector.

Now denote by $h^*$ the efficiency units of the marginal agent who is indifferent between own-account self-employment and receiving the equilibrium wage. We have

$$P_T A_T = wh^*. \tag{3.8}$$

The necessity entrepreneurs' (own-account) rate is then

$$S_n = F(h^*). \tag{3.9}$$

Since the decomposition of the modern labor force into wage workers and employers is the same as the division in efficiency units, the opportunity entrepreneurs' (employers') rate is $S_o = \frac{1-S_n}{1+l}$ and the wage workers' rate $S_w = \frac{(1-S_n)l}{1+l}$. Hence, the aggregate traditional and modern outputs are:

$$y_T = A_T S_n \tag{3.10}$$

$$y_M = A_M E(h|h > h^*)g(l)\frac{1-S_n}{1+l}. \tag{3.11}$$

**Proposition 5.** *There exists a unique interior solution $h^*$ in equilibrium such that agents with*

*$h \in [\underline{h}, h^*]$ are own-account self-employed workers, and agents with $h \in (h^*, \bar{h}]$ enter the modern sector.*

Preliminary to our main comparative static result, we show

**Lemma 2.** *When $A_M$ increases, $h^*$ falls.*

If we think of the traditional sector as intensive in non-traded services such as haircutting, tailoring or street vending, association of higher $P_T$ with higher $A_M$ is consistent with the well-known tendency for the relative price of such services to rise with GDPPC (given Lemma 3). This also implies that the incomes of own-account workers rise with GDPPC conditional on their efficiency units.

We now show

**Lemma 3.** *The aggregate output value $GDP = y$ is increasing in $A_M$.*[7]

Our main results now follow by combining Lemma 2 and Lemma 3.

**Proposition 6.** *When GDPPC increases due to improvements in $A_M$, the share of own-account workers $S_n$ decreases, whereas both the share of employers $S_o$ and the share of wage workers $S_w$ rise.*

It is worth pointing out that even though an increase in either $A_T$ or $A_M$ will increase aggregate productivity, they have very different implications for the labor market, as shown in the following Proposition.

**Proposition 7.** *The incomes of agents in the traditional sector rise when $A_T$ increases, but the traditional sector expands at the expense of the modern sector.*

Now there are two alternative models: the increase in GDPPC could be driven by a relatively larger increase in $A_M$, or by a larger increase in $A_T$. A dominating increase in $A_M$ will

[7] Since there is a unit measure of agents in the economy, GDP and GDPPC are equivalent in our model.

result in a decrease in the labor force share of own-account self-employed workers, whereas a dominating increase in $A_T$ will lead to an increase in the own-account self-employment share. Empirically, because the own-account self-employment share drops strongly as GDPPC increases, the former model is supported. That is, we infer that increases in GDPPC across countries are primarily driven by increases in $A_M$.

## 3.3 Empirical Findings

In this section, we document that the labor share of employers actually increases with income levels, whereas the share of own-account workers (self-employed without employees) decreases. We show this pattern is robust for employment rates separated by main industries or by education categories. We also find nearly universal negative selection on ability into own-account status, and positive selection into employer and wage earning statuses in our data.

### 3.3.1 Measurement of ability/skill

Unfortunately, it is impossible to find direct measures of (or good instruments for) ability for a wide range of developing and developed countries. Consistent with the macro-development literature, we use schooling as our proxy for ability. [141] and [70] both argue that the sorting of more (less) educated workers into urban areas/non-agriculture (rural areas/agriculture) reflects sorting on underlying ability/skill. Like these papers, we keep in the background the dynamic process by which individuals with different abilities acquire different levels of schooling and concentrate on the static allocation of the labor force. We acknowledge that individuals from richer families or with better educated parents may have more schooling than others despite similar abilities, but also note that ability is intergenerationally correlated.

Specifically, we assume log(Ability) = Schooling + $u$, where the error term $u$ is normally distributed according to $\mathcal{N}(0, \sigma_u)$. Therefore, ability of each agent is drawn from a distribution centered at his education level. As shown in Figure 3.1, our model predicts that agents with ability $h < h^*$ ($h > h*$) have probability one (zero) to work on own account. Considering agents A, B with $schooling_A < schooling_B$ as depicted, the probability that agent A's efficiency units are smaller than the cutoff $h^*$ is larger than that for agent B, because it requires a larger positive draw of $u$ for A to exceed the cutoff. Similarly, an agent with $schooling_C$ has a smaller probability of working on own-account than agent B. Mathematically, individual $i$'s probability of working on own-account is

$$\begin{aligned} Pr(v_i = \text{own-account}) = Pr(h_i < h^*) &= Pr(\text{schooling}_i + u_i < h^*) \\ &= Pr(u_i < h^* - \text{schooling}_i) \\ &= \Phi(\frac{h^* - \text{schooling}_i}{\sigma_u}), \end{aligned} \tag{3.12}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. Therefore, our theory predicts that the probability of working on own-account decreases with years of schooling. By the same logic, the model predicts that the probabilities of being a wage worker and of being an employer increase with years of schooling.

### 3.3.2 Data and Summary statistics

The data we use are the Integrated Public Use Microdata Series, International (IPUMS-I) . Our analysis covers 56 countries from 1960 to 2010, integrating 162 publicly available population censuses from IPUMS-I. This paper focuses on occupational choice of labor market participants, so we restrict the sample to prime age (25-55) male workers.

In Table 3.1, we classify the labor force into necessity entrepreneurs, opportunity entrepreneurs, and salaried/wage workers, which are consistently measured across countries. Opportunity entrepreneurship is defined as employers, with permanent employees. Necessity entrepreneurs are defined as individuals who report their employment status as "working on own-account". Table 3.1 provides detailed sub-categories of the three types of labor force from IPUMS-I. These subcategories differ across countries. For example, only 31 country-year observations distinguish between wage workers for a private employer and for the government.[8] The undefined labor force, such as unpaid workers and trainees, only accounts for a negligible fraction of the labor force. The division between employers, wage workers, and own-account workers provides a universally consistent measure of employment status across countries and time.

Table 3.2 presents summary statistics for our 162 observations, with each observation representing one country in one year. Within one country-year sample, the statistics are constructed for the prime age (25-55) male workers. Our data cover a wide range of GDPPC, from I$442 (Liberia in 2008) to more than I$41,000 (Ireland in 2006). The self-employment rate also differs greatly, ranging from 10.6% to 83.1%. Despite the large variations above, own-account workers always account for more than half of the self-employed labor force. On average, only 4.2% of the labor force are employers. The average primary school completion is 60% and the university completion 6.8%. The number of observations drops from 162 to 156 when we need industry information to omit agriculture, fishing, and forestry. The labor force excluding these three industries has slightly better average education, slightly smaller self-employed and own-account average participation rates, and 4.5% employers on average.

There are 98 country-year observations that have data on years of schooling. These samples cover a much smaller range of GDPPC, from I$442 to I$12,000. Whether or not

[8] We will use these observations to conduct a robustness check of the selection between wage workers and employers in Section 3.3.3.

agriculture, fishing and forestry are included, the self-employed have a lower average schooling (5.6 and 7.0 years, respectively) compared to the salaried workers (7.6 and 8.3 years, respectively), but this is driven by the own-account workers in the self-employed. The employers have an average of 7.9 years schooling, and 8.9 years if we omit agriculture, fishing and forestry. These basic facts foreshadow our subsequent findings of negative (positive) selection into own-account (employers') work.

### 3.3.3 Results

#### Selection in the labor force

In this section, we focus on the labor force excluding agriculture, fishing and forestry. Since this paper models occupational choices without market frictions, it applies less well to the agricultural industry where agents are born on the farm and grow up working as farmers. However, including these agents makes our results on selection stronger.

We use the multinomial probit model to estimate the three unordered labor choice responses. The unobservable utilities of individual $i$ from choosing alternative $j \in \{n, w, o\}$ are given by

$$v_{in}^* = \alpha_n + \beta_n \text{schooling}_i + \eta_n X_i + \varepsilon_{in} \tag{3.13}$$

$$v_{iw}^* = \alpha_w + \beta_w \text{schooling}_i + \eta_w X_i + \varepsilon_{iw} \tag{3.14}$$

$$v_{io}^* = \alpha_o + \beta_o \text{schooling}_i + \eta_o X_i + \varepsilon_{io}, \tag{3.15}$$

where controls in $X$ are age, age squared, and a dummy for native-born; $\varepsilon_{ij}$ is a normally distributed error term; and $n, w, o$ denote necessity (own-account) entrepreneur, wage worker, and opportunity entrepreneur (employer), respectively. Note that the $\varepsilon_{ij}$ are not independently distributed because the unobservable error term $u_i$ is absorbed by $\varepsilon_{ij}$.[9] We hence use the multinomial probit estimation model, which allows for a full correlation structure of $\varepsilon_{ij}$, unlike the multinomial logit model. Letting $v_i$ be the labor choice of individual i, then

$$v_i = j \text{ if } v_{ij}^* = \max\{v_{in}^*, v_{iw}^*, v_{io}^*\}. \tag{3.16}$$

As discussed earlier, taking account of the random error term that connects schooling and efficiency units, our model predicts that $\frac{dPr(v_i=n)}{d \text{ schooling}} < 0$, $\frac{dPr(v_i=w)}{d \text{ schooling}} > 0$, and $\frac{dPr(v_i=o)}{d \text{ schooling}} > 0$, i.e., the marginal effect of schooling on becoming own-account self-employed (wage workers, employers) is negative (positive).

Taking Thailand in 2000 as an example, Table 3.3 reports the average marginal effects (AME) on employment at specific schooling years from the above multinomial probit model. We see strong negative selection into own-account self-employment and positive selection into employers. At the mean schooling of 8.46 years, if a worker has one more year of schooling, the average probability of working on own-account will decrease by 1.7% and the average probability of being an employer will increase by 0.3%. At the 5th percentile (4 years) and the 95th percentile (16 years), if a worker has one more year of schooling, the probability of working on own-account decreases by 2.0% and 1.2%, respectively. This suggests that the impact of schooling on selection into necessity entrepreneurship is greater when the education level is lower.

Table 3.4 reports summary statistics for the AME of schooling on the probability of working on own-account at mean years of schooling within a country-year observation. Among the 98 multinomial probit regressions, there are 91 estimations that have significantly negative selection into own-account workers. On average, an extra year of schooling decreases an

[9] We interpret $\beta_j$ as the marginal returns to schooling rather than ability, so the interaction of coefficient $\beta_j$ and $u$ is not in the residual, which preserves the consistency of our estimator.

individual's probability of being a necessity entrepreneur by 2%. The largest negative AME is a 5.7% decline in the probability to work on own-account if an individual's schooling goes up by 1 year at the mean schooling in Guinea in 1983. The three observations with positive AME of schooling on being own-account self-employed tend to be richer economies, and the mean of these three significant positive AME is only 0.5%. In sum, Table 3.4 provides overwhelming evidence for strongly negative selection into own-account self-employed workers.

Table 3.5 reports summary statistics for the AME of schooling on the probability of being employers at mean years of schooling within a country-year observation. Among the 98 regressions, which includes a wide range of GDPPC country-year samples, 81 have significantly positive selection into employers. On average, an extra year of schooling increases an individual's probability of being an opportunity entrepreneur by 0.3%. In Argentina in 1991, the AME of one more year of schooling at the mean on the probability of being employers is 1%, which is the largest in our sample. There are only 6 out of 98 country-year observations that have significant negative selection into employers, and the absolute value of AME is much smaller than for the positive selections. Table 3.5 shows that employers are positively selected on schooling in the overwhelming majority of country-year observations.

In the multinomial probit model, the average marginal effects of schooling on the three outcomes of being own-account, employers and wage workers add up to 0. We thus conclude there is positive selection into wage workers, because the absolute value of the negative AME of schooling on being own-account (0.02) is larger than the positive AME of schooling on being employers (0.003). Now we examine whether there is selection between employers and salaried workers on ability proxied by education.

Table 3.6 reports a summary of whether the AME of schooling on the probabilities of becoming wage workers and employers are statistically different within each estimation. In around 70% of the 98 multinomial probit estimations, the impact of schooling on selection into salaried workers is statistically greater than the impact on being employers. But in another 22 country-year samples, the impact of schooling on selection into employers is statistically greater. In these 22 samples, at mean schooling, if a worker has one more year's schooling, the probability of being wage workers decreases by 0.2%. The mean level comparisons in Table 3.6 indicate that there might be weak selection in favor of wage workers relative to employers, but this may be driven by government employees subjected to more stringent education requirements than private employees.

To further examine this selection issue, we restrict the samples to private sector workers and re-estimate the multinomial probit model (equations (3.13), (3.14), (3.15)). There are 31 country-year samples where we are able to distinguish between private sector and government workers. Table 3.7 reports a summary of whether, in the private sector, the AME of schooling on the probabilities of becoming wage workers and employers are statistically different within each of the 31 estimations. We find that schooling has a statistically greater impact on being wage workers than employers in half of these samples, but either a statistically equal or smaller impact in the other half of the samples. The results suggest that there is no uniform selection into employers against private sector wage workers.

**The labor force division and development**

Another main prediction of our model is that productivity improvement attracts agents from the traditional sector into the modern sector, so the own-account workers' rate falls and employers' rate rises with GDPPC. In addition, as shown in Proposition 2, the threshold ability level decreases with $A_M$, therefore with GDPPC. In other words, development draws the more able agents from the traditional sector into the modern sector. Using education as our proxy for ability, our model thus also predicts that the own-account self-employment (employers', wage workers') rate for any fixed educational attainment decreases (increases) with GDPPC. We test these predictions in this section.

Figure 3.2 shows that, excluding agriculture, fishing and forestry, participation in own-account self-employed decreases sharply from around 60% to 10% as GDPPC increases while participation in employers increases from nearly 0 to 10% in all of our country-year observations. Figure 3.3 and Figure 3.4 show the country average patterns of own-account rate and employers rate, respectively. Table 3.8 reports the results of the following estimations:[10]

$$\text{Self-employment Rate}_c = \alpha_s + \beta_s \ln \text{GDPPC}_c + \gamma_s X_c + \varepsilon_{s,c} \tag{3.17}$$

$$\text{Own-account Rate}_c = \alpha_n + \beta_n \ln \text{GDPPC}_c + \gamma_n X_c + \varepsilon_{n,c} \tag{3.18}$$

$$\text{Employers' Rate}_c = \alpha_o + \beta_o \ln \text{GDPPC}_c + \gamma_o X_c + \varepsilon_{o,c}, \tag{3.19}$$

where in regressions (7) to (9), controls in $X_c$ are average years of schooling, average age, and average native-born rate. Taking all available census samples from IPUMS-I, the dependent variables self-employment rate, own-account self-employment rate and employers' rate are weighted by "person weight" after restricting the samples to prime age (25-55) males. In regressions (1) to (3) the available datasets cover 56 countries across different years, summing to 161 country-year observations. Dropping samples without industry information, in regressions (4) to (6), there are 55 countries (Bangladesh drops out) left across different years. These impacts of economic development on labor force allocation are robust, because the magnitudes of coefficients stay the same as before when we add controls in regressions (7) to (9).

According to Table 3.8, the strong decline in self-employment against GDPPC is dominated by the decrease in the number of own-account self-employed workers. Including agriculture,

[10] See Table 3.16 for the all country-year observations' regression results rather than the country average results.

fishing and forestry only makes this decline stronger. The employers' rate increases robustly and significantly when the economy's income level rises. Since, in one economy, the participation in own-account, wage earning, and employers' status add up to 1, then we know that the wage workers' rate increases unambiguously and strongly with GDPPC. This is because the drop in the share of own-account workers is larger, in absolute value, than the increase in the share of employers. Consistent with our model, these results show that higher GDPPC pulls agents into the modern sector, such that both the employers' rate and wage workers' rate increase with GDPPC.

Now we test the prediction of decreasing own-account share and increasing employers' share against GDPPC across fixed education groups. We divide the agents into five constant educational attainment groups: no primary school completion (less than 5/6 years of education), primary school completion but not lower secondary school (5/6 to 9 years of education), lower secondary school completion but not higher secondary school (9 to 12 years of education), secondary school completion but not university (roughly 12 to 15 years of education), and university completion. Figure 3.5 presents the plot of own-account rates against GDPPC by educational attainment, excluding agriculture, fishing and forestry. The figures are consistent with our predictions.

Table 3.9 reports the estimation of equation (3.18) when restricting the sample to 5 fixed educational attainment groups with controls for average age, age squared, and native-born rate. The own-account self-employment rate is significantly decreasing in GDPPC for the three lower levels of educational attainment. However, there are no significant effects for the higher educational groups. One interpretation is that because, in the majority of countries, $h^*$ is not high enough to correspond to secondary school completed, setting the fixed education at high schooling levels has little test power. Another interpretation is that the own-account self-employed actually contains two types of agents: the necessity own-account and the "distinguished own-account" such as consultants or authors. In the overall population, the quantity of the "distinguished own-account" is so small that they are negligible in the data. But when we restrict the samples to highly educated individuals, the share of the "distinguished own-account" rises and the share of necessity own-account drops, thus making the negative effect of GDPPC on the labor share of own-account self-employment insignificant. The second interpretation is consistent with [131], who find that the own-account entrepreneurs are, on average, the poorest labor force but have a much larger variation in incomes than the employers.

Table 3.10 reports the parallel estimation of equation (3.19). The employers' rate is always increasing in GDPPC for all of the educational attainment groups. Since the changes in labor force share of own-account workers, wage workers and employers add up to zero, we know the share of wage workers increases with GDPPC by comparing the magnitudes of changes in own-account workers and employers' share. These results confirm our model prediction that higher GDPPC pulls the more able own-account self-employed workers into the modern sector, thus resulting in higher shares of employers and wage workers at constant educational attainment groups.

## 3.4 Robustness checks

### 3.4.1 Does employers' rate rise with GDPPC in different industries?

It may be that the pattern of increasing participation in employers with aggregate income level is a result of industrial transformation rather than general productivity improvement. We have shown in Section 3.3.3 that this prediction holds when including and excluding agriculture, fishing and forestry. Now we examine whether this is true in specific industries.

Table 3.11 and Table 3.12 report the estimations of equations (3.18) and (3.19), respectively, when restricting the sample to the 4 largest industries in the majority of countries: manufacturing, sales, service and construction. The own-account self-employment rate is significantly decreasing in GDPPC in the three out of the four industries but not in the service industry. Consistent with discussions in Section 3.3.3, this could be because most "distinguished own-account self-employed" individuals such as consultants and authors are in the service industry. The employers' rate is always increasing in GDPPC in all of the four industries. This evidence again confirms the mechanism in our model that higher productivity draws agents out of the traditional sector into the modern sector, thus resulting in a higher employers' rate.

### 3.4.2 What happens to one country's labor market as GDPPC increases over time?

Previous empirical sections have used country average level statistics to focus on long-run equilibrium results. Now we examine to what extent the predictions hold within one country's time-varying data as GDPPC changes.

Table 3.13 and Table 3.14 report the results of the following estimations with country fixed effects,

$$\text{Own-account Rate}_{ct} = \alpha_c + \beta_n \ln \text{GDP}_{ct} + \gamma_n X_{ct} + \varepsilon_{n,ct} \tag{3.20}$$

$$\text{Employers' Rate}_{ct} = \alpha_c + \beta_o \ln \text{GDP}_{ct} + \gamma_o X_{ct} + \varepsilon_{o,ct}, \tag{3.21}$$

where controls $X$ are group level average age, age squared, and native-born rate. In the fixed effect specifications, the negative effect of GDPPC on own-account self-employment rate is still significant at the aggregate level but not across education groups; and we do not find significant impact of GDPPC on the employers' rate either at the aggregate level or across education groups.

However, these results are not surprising. Our model predictions work for a substantial development of the economy, which usually takes one country decades or more to achieve. In addition, changes in occupational choice may require a generation or more, even if there is substantial productivity improvement. Therefore, the fixed effect regressions capture more temporal noise and frictions than long-run economic growth's impact on the labor market.

## 3.5 Calibration

We calibrate the model in this section to assess its quantitative performance. Our strategy is to parameterize the model to match the moments of Canada, the benchmark country where we have wage data, and then lower $A_M$ to compute the model's predictions for other countries.

### 3.5.1 Quantitative Version of Model

Our benchmark model delivers two key predictions: selection on occupational choice and sorting according to technological progress. The key mechanism is critically based on the heterogeneous ability of agents in the labor force. We do not have a direct measure of ability in the data, but labor income is a linear function of ability in the modern sector. Therefore, for the purpose of quantifying the model predictions, we use education as our proxy for education. In particular, we divide the labor force into two education groups: workers who did not finish high school and workers who have at least a high school diploma. Thus the ability distribution of high education group $G_h(x)$ and low education group $G_l(x)$ are disciplined by their corresponding wage distributions respectively. Finally, the aggregate ability distribution is a weighted sum of draws from $G_h(x)$ and $G_l(x)$ based on the country specific share of low education workers in the labor force.

### 3.5.2 Parameterizing the Model

In order to quantify the model's qualitative predictions on the division of the labor force in a cross section of countries, we calibrate the two parameters $\gamma$, $\sigma$, and the modern sector production function, and the ability distribution. Our strategy is to parameterize the model to match the moments of a rich country, and then lower $A_M$ to compute the model's predictions for other countries. Using cross-country differences in 1988, [37] estimate the productivity of skilled workers to be strongly increasing with GDP per worker relative to the productivity of unskilled workers. Hence, we set $A_T$ to be fixed and normalized to 1 for all countries and allow $A_M$ to vary across countries.

We choose Canada as the benchmark rich country for target moments, because it is the only country among the IPUMS-I samples for which we can distinguish between own-account workers and employers and which has the earned income and hours worked data needed to compute wages. In particular, we pick the 2001 census of Canada because it is the only available year for which the sample weights in the Individuals File are calculated by Statistics Canada adjusting for sex, age groups, and geographic areas.

The parameter $\gamma$ is related to the share of traditional sector output in aggregate production, so we will calibrate it to match the share of own-account workers in Canada in 2001. We choose the modern sector production functional form to be $G(h,L) = h^{\alpha}L^{1-\alpha}$ for simplicity. The parameter $\alpha$ measures employer's ability share of modern production, so we will use it to match the share of employers in Canada in 2001. The parameter $\sigma$ is related to the elasticity of substitution between traditional and modern output, $\frac{1}{1-\sigma}$. After reviewing the evidence, [12] concluded that the elasticity of substitution between unskilled and skilled labor is very unlikely to fall outside 1 and 2.[11] Since unskilled labor is correlated with traditional output, whereas

[11] See also [37].

skilled labor is correlated with modern output, the two elasticities are also connected. Moreover, for example, vegetables from traditional farmers are more substitutable for modern agricultural products than the unskilled farmers for the skilled farmers who operate agricultural machines. Therefore, we set $\frac{1}{1-\sigma}$ to be 2, the upper bound of the recognized range for the elasticity of substitution between unskilled and skilled labor, for our benchmark calibration.

Finally, we calibrate the parameters of ability distribution to match the wage distribution of Canada in 2001. In the calibration exercise, we divide the labor force into two groups: individuals with only primary school completion (i.e., high school dropouts) as a proxy for the low-skill workers, and individuals with at least secondary school completion as a proxy for the high-skill workers. Then the wage distribution is characterized by the ratio of average wages for these two groups and the variance of aggregate $log(wage)$. To match these two moments, we let the abilities be drawn from two log-normal distributions with the same variance and different means. In particular, the mean of the $log(ability)$ distribution for agents without secondary school completion is normalized to be one, and the mean of those with secondary school completion and the variance of aggregate $log(wage)$ are calibrated to match the wage distribution of Canada in 2001.

In Figure 3.6, the first graph presents the probability density function of the low- and high-mean log-normal ability distributions for all countries after fitting the wage distribution of Canada in 2001. To generate the distribution of ability for the full population in each country, we make the share of ability draws from the mean one log-normal distribution a linear function of $A_M$ to fit the low-education labor force share in data. The second graph of Figure 3.6 presents the fitted share of ability draws from the low-mean distribution in the model and the labor force share of agents who do not complete secondary school in the data. For example, the third graph plots the mapping of raw $A_M$ value to ln(GDPPC), which is imputed from the numeric model total output to match the scale in data; the fourth graph then presents the aggregate ability distribution of the richest country in our model with an almost 11 ln(GDPPC), thus a fitted 54% ability values drawn from the mean one distribution and the 45% from the higher mean distribution in the first graph.

Table 3.15 reports the parameter values that are used to match the data of Canada in 2001. Our model matches the wage distribution and own-account workers' rate accurately, but over-predicts the employers' rate in Canada in 2001. This is partly because Canada has the second lowest employers' rate among all countries in our sample. Figure 3.7 shows that our model can accurately predict the labor force share of own-account entrepreneurs across countries. Figure 3.8 shows that our model also slightly over-predicts the employers' rate across countries. Overall, our parsimonious model makes good quantitative predictions of the labor force division over a cross-section of countries with a wide range of development levels.

## 3.6 Conclusion

Our model and supporting evidence show that entrepreneurs without (with) employees are negatively (positively) selected on ability, and entrepreneurs without employees (hence most entrepreneurs) are negatively selected relative to wage workers. Moreover, economic development increases (decreases) the labor force share of entrepreneurs with (without) employees, at the aggregate level and across constant educational attainment groups (our proxy for fixed ability levels). Improving technology pulls the more able agents from the traditional sector into the modern sector and results in a higher employers' rate despite increasing wages. Predictions regarding the impact of development on labor force division made by a calibrated version of our simple general equilibrium model fit the cross-country data quantitatively well.

This overwhelming evidence suggests that there are two distinct types of entrepreneurship.

Necessity entrepreneurs lack the ability to build promising businesses and become successful employers, whereas opportunity entrepreneurs combine workers with modern technology. Given that own-account workers are mostly operating in the informal sector, our results are consistent with the literature (e.g., **?**, **?**, 2014) that views business owners in the informal sector as "reluctant entrepreneurs" with low productivity that cannot survive economic growth.

It is very unlikely that substantial employers can be fostered by encouraging necessity micro-businesses operated by negatively selected agents. Thus it is not surprising that studies of micro-credit programs find a pattern of modestly positive, but neither transformative nor persistent effects of expanded access to micro-credit on the profits of small businesses. Policy makers wanting to maximize impact of programs designed to help businesses limited by market frictions such as finance constraints should consider focusing on employers.

Chapter 3 is coauthored with Lindsay Rickey. The dissertation author was the primary investigator and author of the unpublished material.

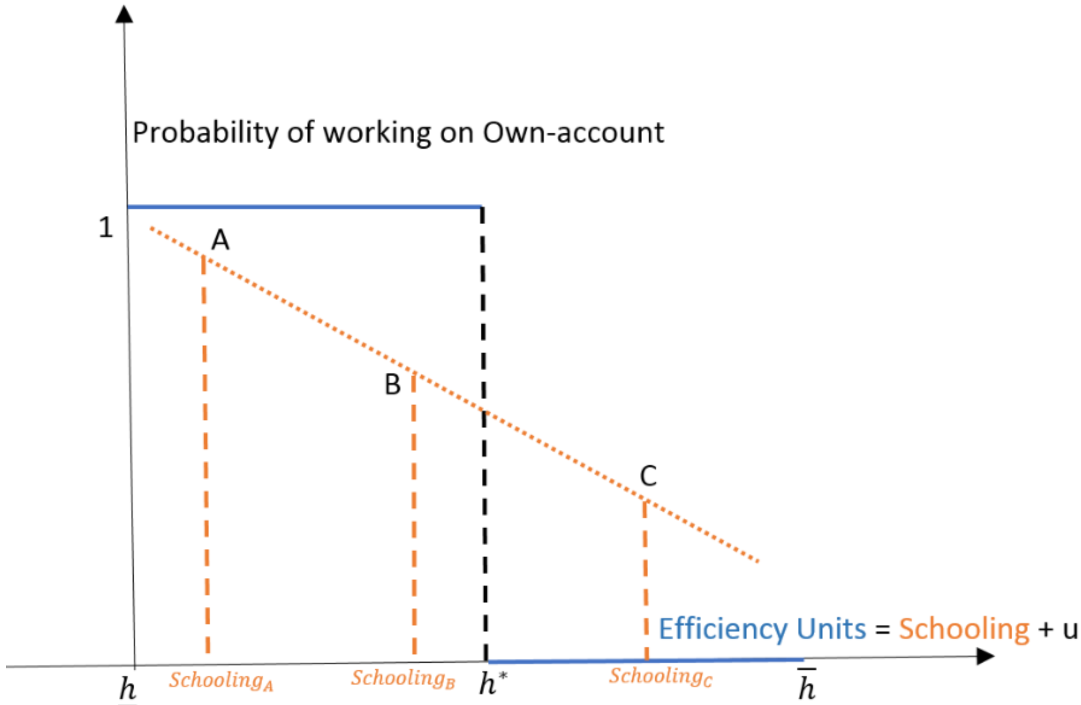



**Figure 3.1**: Probability of being own-account worker and Schooling

**Table 3.1**: Employment Categorization

| Necessity Entrepreneu | Wage/Salary workers | Opportunity Entrepreneurship | Undefined Labor Force |
|---|---|---|---|
| Own account, agriculture<br>Domestic worker, self-employed<br>Subsistence worker, own consumption<br>Own account, without temporary/unpaid help<br>Own account, with temporary/unpaid help<br>Member of cooperative<br>Sharecropper | Management, Non-management<br>White collar (non-manual)<br>Blue collar (manual)<br>Employee, with a permanent job<br>Employee, occasional/temporary/contract<br>Employee without legal contract<br>Wage/salary worker, private employer<br>Wage/salary worker, government<br>work for private household<br>Seasonal migrant<br>Other wage/salary workers | Employer | Unpaid family worker<br>Apprentice or trainee<br>Works for others without wage<br>Other undefined labor force |

## 3.7 Appendices

### 3.7.1 Proofs

*Proof of Proposition 5*

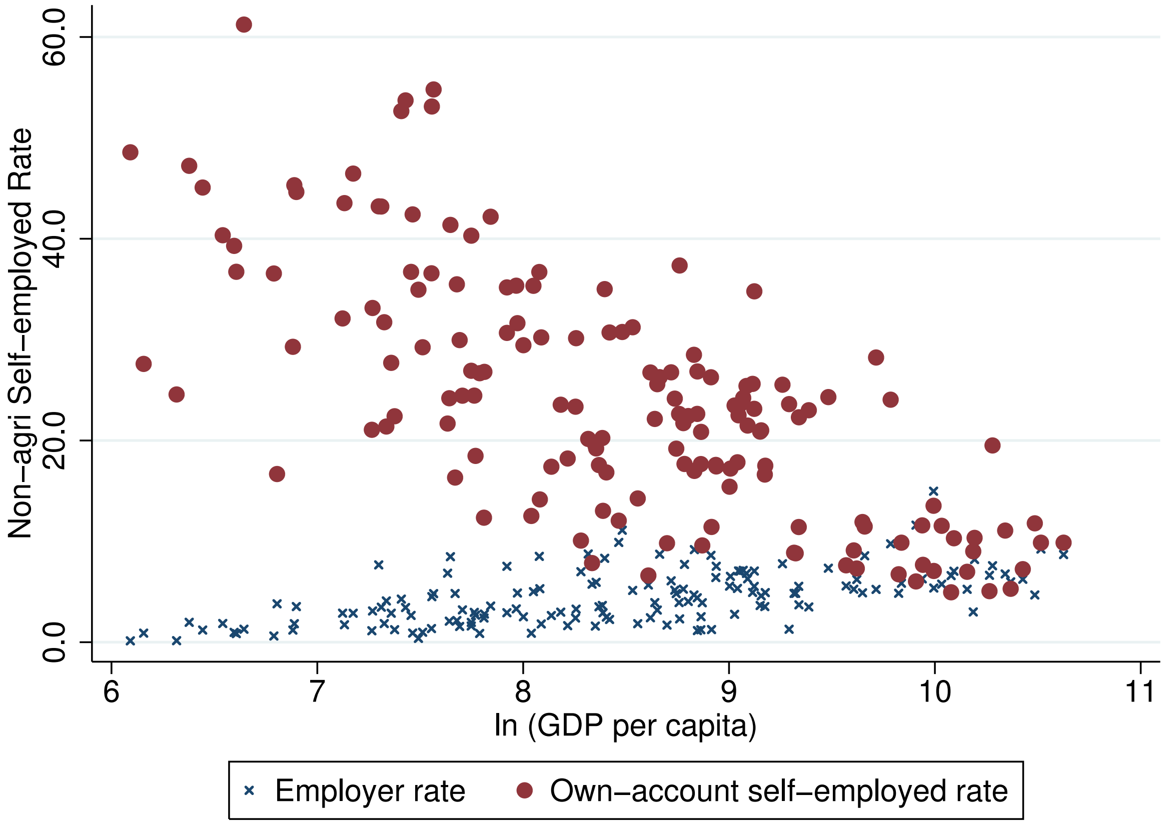



**Figure 3.2**: Self-employment Rate by type


**Figure 3.4**: Employers' Rate

*Proof.* Substituting equation (3.6) into equation (3.8) yields

$$h^* = \frac{P_T A_T}{A_M g'(l)}. \tag{3.22}$$

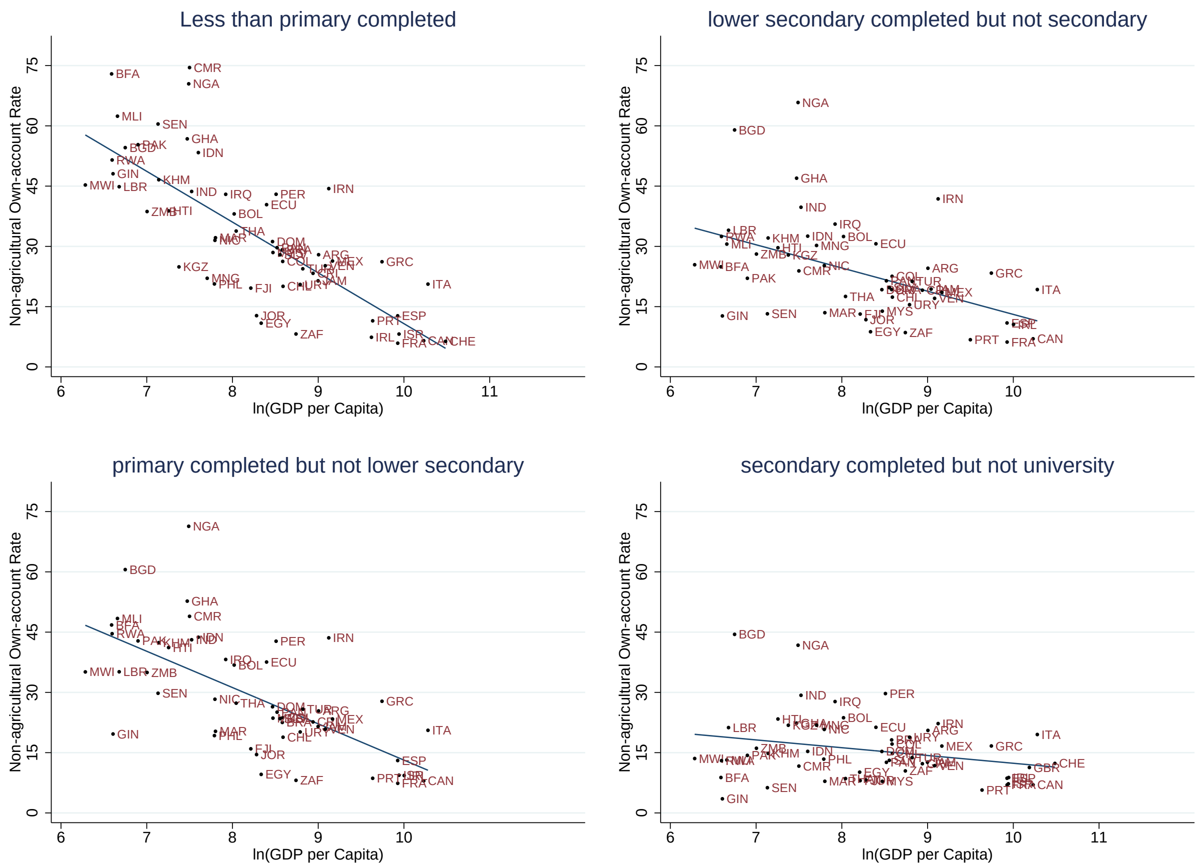



**Figure 3.5**: Own-account self-employment rate by educational attainment

Substituting equations (3.22) into equation (3.9), (3.10) and (3.11), and combining them with equation (3.4) yields

$$P_T = \frac{\gamma}{1-\gamma}\left[\frac{A_M g(l)\int_{\frac{P_T A_T}{A_M g'(l)}}^{\bar{h}} hf(h)dh}{(1+l)A_T F(\frac{P_T A_T}{A_M g'(l)})}\right]^{1-\sigma}. \tag{3.23}$$

Note the left and right hand sides of equation (3.23) are monotonically increasing and decreasing in $P_T$, respectively. Let $P_T \to \frac{\underline{h} A_M g'(l)}{A_T}$ s.t. $\frac{P_T A_T}{A_M g'(l)} \to \underline{h}$, then the right hand side of equation (3.23) goes to infinity, which is larger than the left hand side $\frac{\underline{h} A_M g'(l)}{A_T}$. Let $P_T \to \frac{\bar{h} A_M g'(l)}{A_T}$ s.t. $\frac{P_T A_T}{A_M g'(l)} \to \bar{h}$, then right hand side of equation (3.23) goes to 0 while the left hand side equals $P_T > 0$. Therefore, there exists a unique endogenous $P_T \in (\frac{\underline{h} A_M g'(l)}{A_T}, \frac{\bar{h} A_M g'(l)}{A_T})$ such that equation (3.23) holds. Recall $l$ is also unique, so equation (3.22) defines a unique interior solution $h^* \in (\underline{h}, \bar{h})$ that holds in general equilibrium. $\square$

*Proof of Lemma 2*

*Proof.* We will show that the price of unskilled products increases less than proportionally with $A_M$, i.e., $0 < \frac{dP_T/P_T}{dA_M/A_M} < 1$. The result then follows from inspection of equation (3.22).

Consider equation (3.23). Suppose $A_M$ increases but $P_T$ remains unchanged or decreases, we will have $\int_{\frac{P_T A_T}{A_M g'(l)}}^{\bar{h}} hf(h)dh$ increases and $F(\frac{P_T A_T}{A_M g'(l)})$ decreases, thus the left hand side of equation (3.23) smaller than the right hand side, which is a contradiction. Hence, we have $\frac{dP_T}{dA_M} > 0$. Now rewrite equation (3.23) as

$$(P_T)^{\frac{\sigma}{1-\sigma}} \frac{P_T}{A_M} = \frac{\gamma}{1-\gamma} \frac{g(l) \int_{\frac{P_T A_T}{A_M g'(l)}}^{\bar{h}} hf(h)dh}{(1+l)A_T F(\frac{P_T A_T}{A_M g'(l)})}. \tag{3.24}$$

Recall $\frac{dP_T}{dA_M} > 0$ as shown in Lemma 3. Now suppose $P_T$ increases such that $\frac{dP_T/P_T}{dA_M/A_M} \geq 1$, we will have $\frac{P_T}{A_M}$ increases or remains unchanged. Then the left hand side of equation (3.24) goes up since $0 < \sigma < 1$, and the right hand side decreases or stays unchanged. This is again a contradiction. Therefore, we conclude $0 < \frac{dP_T/P_T}{dA_M/A_M} < 1$. $\square$

*Proof of Lemma 3*

*Proof.* Because there is no market failure, our perfectly competitive model solves the social planner's problem:

$$\max_{h^*} y = (\gamma y_T^{\sigma} + (1-\gamma) y_M^{\sigma})^{\frac{1}{\sigma}}, \tag{3.25}$$

where $y_T = A_T F(h^*)$ and $y_M = A_M \frac{g(l)}{1+l} \int_{h^*}^{\bar{h}} hf(h)dh$. The first order condition has $\frac{dy}{dh^*} = 0$. So by the envelope theorem,

$$\begin{aligned}\frac{dy}{dA_M} &= \frac{\partial y}{\partial A_M} \\ &= (1-\gamma)y_M^{\sigma-1}(\gamma y_T^{\sigma}+(1-\gamma)y_M^{\sigma})^{\frac{1}{\sigma}-1}\frac{g(l)}{1+l}\int_{h^*}^{\bar{h}} hf(h)dh \\ &> 0\end{aligned} \tag{3.26}$$

$\square$

*Proof of Proposition 6*

*Proof.* $S_n = F(h^*)$ decreases with $A_M$ according to Lemma 2. Thus, $\frac{dS_o}{dA_M} = -\frac{1}{1+l}\frac{dS_n}{dA_M} > 0$ and $\frac{dS_w}{dA_M} = -\frac{l}{1+l}\frac{dS_n}{dA_M} > 0$. The result then follows by Lemma 3. $\square$

*Proof of Proposition 7*

*Proof.* Consider equation (3.23). Suppose $A_T$ increases but $P_T$ remains unchanged or increases, then $\int_{\frac{P_T A_T}{A_M g'(l)}}^{\bar{h}} hf(h)dh$ decreases and $F(\frac{P_T A_T}{A_M g'(l)})$ increases, thus the left hand side of equation (3.23) becomes larger than the right hand side, which is a contradiction. Hence, we have $\frac{dP_T}{dA_T} < 0$. Now suppose $P_T$ decreases such that $A_T P_T$ gets smaller or remained unchanged. Rewrite equation (3.23) as

$$(P_T)^{\frac{\sigma}{1-\sigma}} P_T A_T = \frac{\gamma}{1-\gamma} \frac{A_M g(l) \int_{\frac{P_T A_T}{A_M g'(l)}}^{\bar{h}} hf(h)dh}{(1+l) A_T F(\frac{P_T A_T}{A_M g'(l)})}. \tag{3.27}$$

Then left hand side of equation (3.27) will be smaller than the right hand side. This is again a contradiction. Therefore, we conclude that $A_T P_T$ increases with $A_T$, so does $h^*$ and $S_n = F(h^*)$. $\square$

**Lemma 4.** *There exists a unique partition of efficiency units l such that the equilibrium condition holds and that employers solve their profit maximization problems.*

*Proof.* Substituting equation (3.7) into equation (3.6) obtains $H \equiv g(l) - g'(l)l - g'(l) = 0$. Since $\frac{dH}{dl} = -g"(l)(l+1) > 0$ by the concavity assumption, $H$ is increasing monotonically in $l$. As $l \to 0$, $H \to \lim_{l \to 0} g'(l)l - g'(l) < 0$; as $l \to \infty$, $H \to \lim_{l \to \infty} l[\frac{g(l)}{l} - g'(l)] - g'(l) \to \infty$ because $\frac{g(l)}{l}$ - the average productivity of $l$ exceeds $g'(l)$ - the marginal productivity when l approaches infinity by concavity. Therefore there exists a unique $l$ solves equation (3.5) and (3.7). $\square$

### 3.7.2 Tables

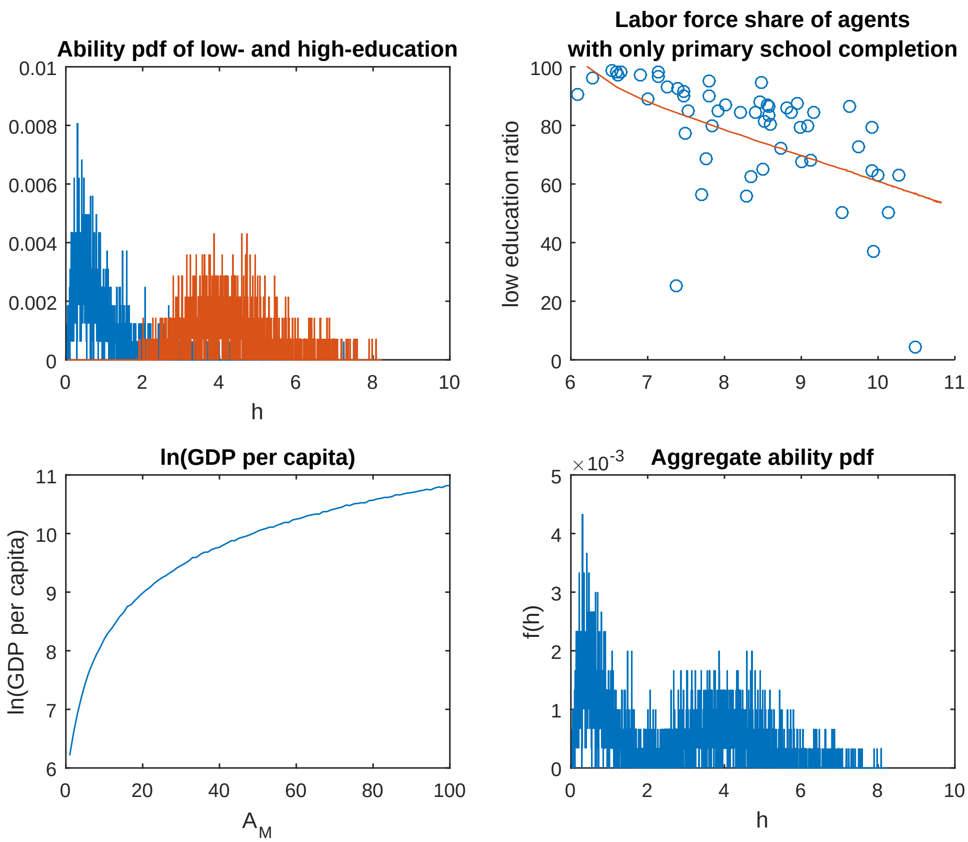



**Figure 3.6**: The ability distribution of calibration input

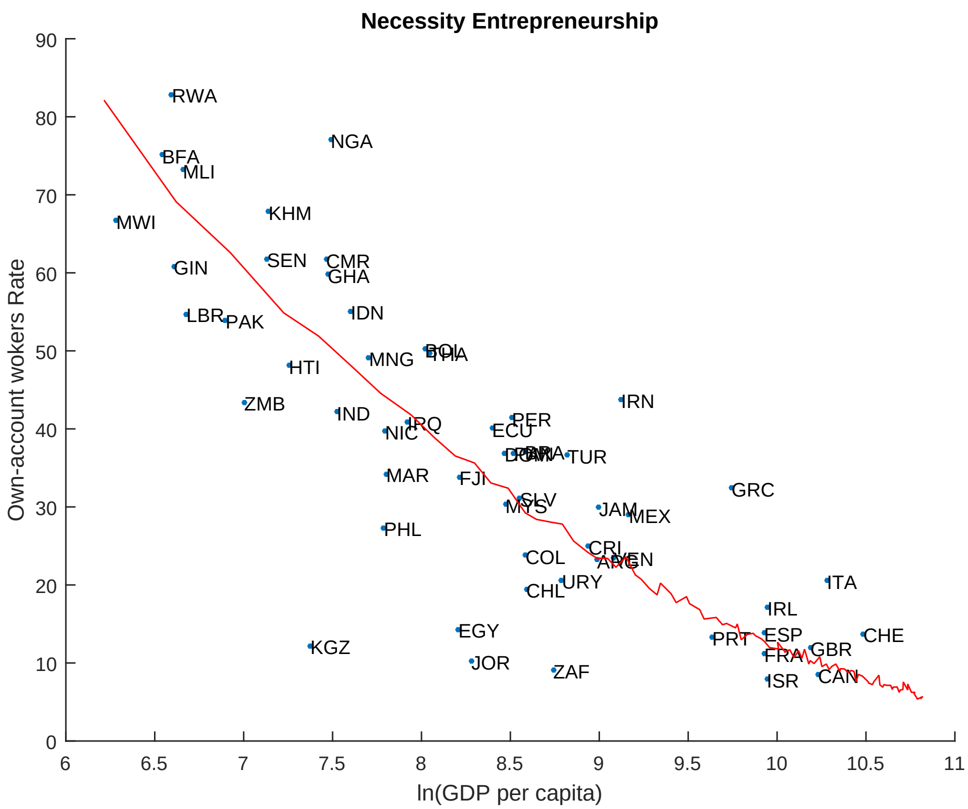



**Figure 3.7**: Data versus model predictions on share of the own-account

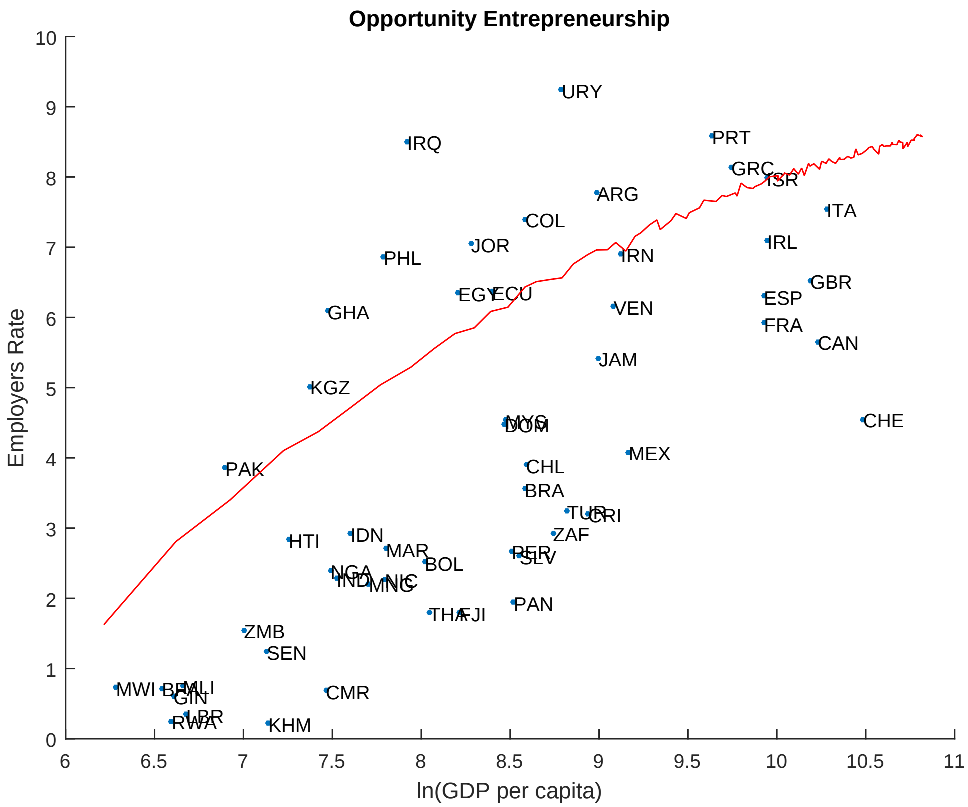



**Figure 3.8**: Data versus model predictions on share of employers

**Table 3.2**: Summary Statistics

| Variable | Mean | Std. Dev. | Min. | Max. |
|---|---|---|---|---|
| **Panel A: All industries** | | | | |
| Self-employed Rate | 40.3 | 18.1 | 10.6 | 83.1 |
| Own-account Rate | 36.2 | 19.7 | 6.2 | 82.9 |
| Employers' Rate | 4.2 | 2.8 | 0.1 | 14.0 |
| Primary school completion | 60.0 | 25.3 | 7.1 | 99.2 |
| Secondary school completion | 25.6 | 20.0 | 1.3 | 92.5 |
| University completion | 6.8 | 5.9 | 0.0 | 31.1 |
| GDPPC | 7822.1 | 8495.2 | 442.2 | 41158.9 |
| year | 1991.3 | 13.2 | 1960 | 2010 |
| Number of individual observations | 287565.6 | 456685.4 | 4946 | 3507673 |
| **Country-year observations** | | 162 | | |
| | | | | |
| **Average years of schooling** | | | | |
| All | 6.4 | 2.3 | 1.2 | 11.5 |
| Salaried | 7.6 | 2.0 | 3.1 | 12.1 |
| Self-employed | 5.6 | 2.5 | 0.2 | 10.9 |
| Own-account | 5.4 | 2.4 | 0.2 | 10.4 |
| Employers | 7.9 | 2.4 | 0.7 | 12.5 |
| GDP per capita | 4678.7 | 3006.1 | 442.2 | 11939.8 |
| year | 1990.7 | 13.8 | 1960 | 2010 |
| Number of individual observations | 220627.4 | 332141.7 | 6807 | 1629695 |
| **Country-year observations** | | 98 | | |
| **Panel B: Omit agriculture, fishing, and forestry** | | | | |
| Self-employed Rate | 28.4 | 11.5 | 11.3 | 62.5 |
| Own-account Rate | 23.9 | 12.1 | 5.0 | 61.2 |
| Employers' Rate | 4.5 | 2.7 | 0.1 | 15.0 |
| Primary school completion | 71.8 | 20.0 | 15.0 | 99.3 |
| Secondary school completion | 33.6 | 19.7 | 2.9 | 92.4 |
| University completion | 9.2 | 6.5 | 0.1 | 31.8 |
| GDP per capita | 8078.6 | 8553.6 | 442.2 | 41158.9 |
| Number of individual observations | 201093.7 | 357384.0 | 2528 | 3373662 |
| **Country-year observations** | | 156 | | |
| | | | | |
| **Average years of schooling** | | | | |
| All | 7.8 | 1.9 | 3.1 | 11.9 |
| Salaried | 8.3 | 1.8 | 4.3 | 12.2 |
| Self-employed | 7.0 | 2.2 | 0.7 | 11.4 |
| Own-account | 6.7 | 2.2 | 0.7 | 11.1 |
| Employers | 8.9 | 2.2 | 1.4 | 12.7 |
| Number of individual observations | 149861.0 | 237657.1 | 3206 | 1174286 |
| **Country-year observations** | | 98 | | |

*Notes*: Table 3.2 reports summary statistics within a country-year observation from IPUMS-I. Samples are restricted to prime age (25-55) male workers, excluding people living within group quarters. All mean values are weighted by personal weight in the census survey. GDP per capita used is from Penn World Table 7.1, the PPP Converted GDP Per Capita (Laspeyres), derived from growth rates of c, g, i, at 2005 constant prices.

**Table 3.3**: Average Marginal Effects, Thailand 2000

| Schooling at | (1)<br>d(y=Own-account)/dx | (2)<br>d(y=Employer)/dx |
|---|---|---|
| 5th percentile | -0.0199*** | 0.0024*** |
| | (0.0006) | (0.0001) |
| mean | -0.0175*** | 0.0027*** |
| | (0.0005) | (0.0002) |
| 95th percentile | -0.0123*** | 0.0029*** |
| | (0.0002) | (0.0003) |
| Observations | 50,146 | 50,146 |

*** represents statistical significance at 1%; Standard errors are in parentheses.

*Notes*: Table 3.3 reports the average marginal effects calculated from the multinomial probit model (equation (3.13), (3.14), (3.15) and (3.16) ) through the delta method. The dependent variable is one of the labor choices, own-account self-employed workers, wage workers, or employers; and controls are age, age squared, and a dummy for native-born. Samples are from IPUMS-I, restricted to prime age (25-55) male workers not in agriculture, fishing or forestry.

**Table 3.4**: AME at Mean Schooling on being Own-account

| **Variable** | **Mean** | **Std. Dev.** | **Min.** | **Max.** | **N(/98)** |
|---|---|---|---|---|---|
| **Negative Average Marginal Effects (t $<$ -1.96)** | | | | | |
| AME | -0.02 | 0.013 | -0.057 | -0.001 | 91 |
| GDPPC | 4548.044 | 3037.813 | 442.201 | 11939.771 | 91 |
| **Insignificant Negative Average Marginal Effects** | | | | | |
| AME | 0 | 0 | 0 | 0 | 2 |
| GDPPC | 4221.563 | 737.374 | 3700.161 | 4742.965 | 2 |
| **Insignificant Positive Average Marginal Effects** | | | | | |
| AME | 0 | 0 | 0 | 0 | 2 |
| GDPPC | 7570.448 | 2961.004 | 5476.702 | 9664.194 | 2 |
| **Positive Average Marginal Effects (t $>$ 1.96)** | | | | | |
| AME | 0.005 | 0.001 | 0.004 | 0.005 | 3 |
| GDPPC | 7017.893 | 889.430 | 5995.739 | 7615.497 | 3 |

*Notes*: This table summarizes the AME of schooling on the probability of being own-account self-employed workers at mean years of schooling within each country-year observation from 98 multinomial probit regressions. The dependent variable is one of the labor choices, own-account self-employed workers, wage workers, or employers; and controls are age, age squared, and a dummy for native-born. Samples are from IPUMS-I, restricted to prime age (25-55) male workers not in agriculture, fishing or forestry.

**Table 3.5**: AME at Mean Schooling on being Employers

| Variable | Mean | Std. Dev. | Min. | Max. | N(/98) |
|---|---|---|---|---|---|
| **Negative Average Marginal Effects (t $<$ -1.96)** | | | | | |
| AME | -0.001 | 0.001 | -0.003 | -0.001 | 6 |
| GDPPC | 3556.493 | 2599.423 | 739.945 | 7074.860 | 6 |
| | | | | | |
| **Insignificant Negative Average Marginal Effects** | | | | | |
| AME | 0 | 0 | 0 | 0 | 5 |
| GDPPC | 2507.923 | 2169.525 | 442.201 | 5995.739 | 5 |
| | | | | | |
| **Insignificant Positive Average Marginal Effects** | | | | | |
| AME | 0 | 0 | 0 | 0.001 | 6 |
| GDPPC | 1433.758 | 1139.323 | 553.554 | 3577.978 | 6 |
| | | | | | |
| **Positive Average Marginal Effects (t $>$ 1.96)** | | | | | |
| AME | 0.003 | 0.002 | 0 | 0.01 | 81 |
| GDPPC | 5136.155 | 2971.12 | 471.689 | 11939.771 | 81 |

*Notes*: This table summarizes the AME of schooling on the probability of being employers at mean years of schooling within each country-year observation from 98 multinomial probit regressions. The dependent variable is one of the labor choices, own-account self-employed workers, wage workers, or employers; and controls are age, age squared, and a dummy for native-born. Samples are from IPUMS-I, restricted to prime age (25-55) male workers not in agriculture, fishing or forestry.

**Table 3.6**: Average Marginal Effects on being wage workers versus employers

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| **Statistically AME$_{opp}$ =AME$_{wage}$** | | | | | |
| AME$_{opp}$ | 0.005 | 0.001 | 0.003 | 0.006 | 6 |
| AME$_{wage}$ | 0.005 | 0.002 | 0.003 | 0.008 | 6 |
| GDPPC | 5381.098 | 3337.924 | 2141.621 | 9087.893 | 6 |
| **Statistically AME$_{opp}$ >AME$_{wage}$** | | | | | |
| AME$_{opp}$ | 0.005 | 0.002 | 0 | 0.01 | 22 |
| AME$_{wage}$ | -0.002 | 0.004 | -0.013 | 0.004 | 22 |
| GDPPC | 6944.335 | 2040.807 | 3700.161 | 11379.896 | 22 |
| **Statistically AME$_{opp}$ <AME$_{wage}$** | | | | | |
| AME$_{opp}$ | 0.002 | 0.002 | -0.003 | 0.007 | 70 |
| AME$_{wage}$ | 0.022 | 0.013 | 0.006 | 0.058 | 70 |
| GDPPC | 3906.396 | 2888.611 | 442.201 | 11939.771 | 70 |

*Notes*: Statistically significance means that the hypothesis is not rejected at 5% significance level. The dependent variable is one of the labor choices, own-account self-employed workers, wage workers, or employers; and controls are age, age squared, and a dummy for native-born. Samples are from IPUMS-I, restricted to prime age (25-55) male workers not in agriculture, fishing or forestry.

**Table 3.7**: Average Marginal Effects in the Private Sector

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| **Statistically AME$_{opp}$ =AME$_{wage}$** | | | | | |
| AME$_{opp}$ | 0.004 | 0.003 | 0 | 0.006 | 4 |
| AME$_{wage}$ | 0.005 | 0.001 | 0.003 | 0.006 | 4 |
| GDPPC | 4438.786 | 961.654 | 3101.991 | 5200.413 | 4 |
| **Statistically AME$_{opp}$ >AME$_{wage}$** | | | | | |
| AME$_{opp}$ | 0.01 | 0.003 | 0.004 | 0.013 | 11 |
| AME$_{wage}$ | -0.004 | 0.007 | -0.016 | 0.006 | 11 |
| GDPPC | 7075.147 | 1752.148 | 3844.639 | 9087.893 | 11 |
| **Statistically AME$_{opp}$ <AME$_{wage}$** | | | | | |
| AME$_{opp}$ | 0.003 | 0.002 | 0 | 0.007 | 16 |
| AME$_{wage}$ | 0.013 | 0.006 | 0.007 | 0.031 | 16 |
| GDPPC | 5340.447 | 2982.444 | 588.604 | 10849.332 | 16 |

*Notes*: Statistically significance means that the hypothesis is not rejected at 5% significance level. The dependent variable is one of the labor choices, own-account self-employed workers, wage workers, or employers; and controls are age, age squared, and a dummy for native-born. Samples are from IPUMS-I, restricted to prime age (25-55) male workers not in agriculture, fishing or forestry.

**Table 3.8**: Prime Male Necessity and Opportunity Self-Employment Rates Across Countries

| | All Industries | | | Excluding Agriculture, fishing, and forestry | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Rates by Employment: | Self-Employed | Own-account | Employer | Self-Employed | Own-account | Employer | Self-Employed | Own-account | Employer |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| ln (GDP per capita) | -12.85*** | -14.47*** | 1.619*** | -6.596*** | -8.119*** | 1.524*** | -6.585*** | -7.890*** | 1.305*** |
| | (1.398) | (1.460) | (0.226) | (1.033) | (0.963) | (0.212) | (2.277) | (2.049) | (0.462) |
| schooling | | | | | | | -1.088 | -1.204 | 0.116 |
| | | | | | | | (0.965) | (0.868) | (0.196) |
| age | | | | | | | 1.367 | 0.771 | 0.597 |
| | | | | | | | (2.647) | (2.382) | (0.538) |
| nativity | | | | | | | 22.18 | 19.56 | 2.618 |
| | | | | | | | (38.45) | (34.59) | (7.808) |
| Constant | 148.0*** | 157.2*** | -9.291*** | 84.44*** | 92.85*** | -8.412*** | 20.02 | 52.41 | -32.39 |
| | (11.69) | (12.20) | (1.893) | (8.662) | (8.076) | (1.778) | (95.57) | (86.00) | (19.41) |
| Observations | 56 | 56 | 56 | 55 | 55 | 55 | 33 | 33 | 33 |
| R-squared | 0.610 | 0.645 | 0.486 | 0.435 | 0.573 | 0.493 | 0.390 | 0.538 | 0.457 |

***, * represents statistical significance at 1% and 10% respectively; Standard errors are in parentheses.
*Notes*: Taking all available census samples from IPUMS-I, the dependent variables self-employment rate, own-account self-employment rate, and employers rate are weighted by "person weight" after restricting the samples to prime age (25-55) males. Regressions (7) to (9) include observations with controls for average years of schooling, average age, and average native-born rate.

**Table 3.9**: Own-account self-employment rate by educational attainment

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| own-account rate for | <Primary | Primary | Lower Secondary | Secondary | University |
| ln (GDP per capita) | -12.06*** | -6.861*** | -5.370*** | -2.147 | 1.721 |
| | (2.430) | (2.350) | (1.664) | (1.429) | (1.154) |
| age | 30.72 | 163.4*** | -53.47 | 73.15** | -8.263 |
| | (63.98) | (45.42) | (40.54) | (27.53) | (10.57) |
| $age^2$ | -0.395 | -2.224*** | 0.799 | -0.992** | 0.146 |
| | (0.798) | (0.613) | (0.582) | (0.385) | (0.147) |
| nativity | 9.402 | 25.32 | 37.58 | 29.06 | 13.86 |
| | (34.73) | (56.38) | (39.89) | (19.59) | (8.713) |
| Constant | -475.0 | -2,939*** | 919.4 | -1,339** | 90.92 |
| | (1,274) | (840.2) | (712.6) | (488.7) | (186.2) |
| Observations | 33 | 33 | 31 | 33 | 33 |
| R-squared | 0.576 | 0.449 | 0.383 | 0.348 | 0.455 |

***, ** represents statistical significance at 1% and 5% respectively; Standard errors are in parentheses.
*Notes*: Table 3.9 reports the estimation of equation (3.18) when restricting the sample to 5 fixed educational attainment groups: less than primary school completed, primary school completed but not lower secondary school, lower secondary completed but not upper secondary school, secondary school completed but not university and university completed. The dependent variable is the weighted own-account self-employment rate after taking the average of the multiple years' observations from one country. Samples are from IPUMS-I, restricted to prime age (25-55) male workers not in agriculture, fishing or forestry.

**Table 3.10**: Employers' rate by educational attainment

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| own-account rate for | <Primary | Primary | Lower Secondary | Secondary | University |
| ln (GDP per capita) | 1.217*** | 1.247*** | 1.690*** | 2.533*** | 3.461*** |
| | (0.346) | (0.288) | (0.356) | (0.442) | (0.580) |
| age | 1.480 | 1.236 | -27.46** | -7.426 | -7.623 |
| | (9.325) | (11.54) | (11.65) | (8.775) | (4.981) |
| age$^2$ | -0.0207 | -0.0139 | 0.400** | 0.110 | 0.112 |
| | (0.117) | (0.159) | (0.167) | (0.126) | (0.0696) |
| nativity | -1.010 | -2.813 | 1.066 | -1.288 | 4.913 |
| | (5.676) | (10.23) | (7.433) | (5.607) | (4.897) |
| Constant | -32.14 | -30.63 | 459.3** | 110.5 | 102.3 |
| | (183.0) | (213.5) | (204.2) | (153.3) | (88.19) |
| Observations | 33 | 33 | 31 | 33 | 33 |
| R-squared | 0.302 | 0.337 | 0.591 | 0.504 | 0.627 |

***, ** represents statistical significance at 1% and 5% respectively; Standard errors are in parentheses.

*Notes*: Table 3.10 reports the estimation of equation (3.19) when restricting the sample to 5 fixed educational attainment groups: less than primary school completed, primary school completed but not lower secondary school, lower secondary completed but not upper secondary school, secondary school completed but not university and university completed. The dependent variable is the weighted employers' rate after taking the average of the multiple years' observations from one country. Samples are from IPUMS-I, restricted to prime age (25-55) male workers not in agriculture, fishing or forestry.

**Table 3.11**: Own-account self-employment rate by industry

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Own-account Rate in | Manufacture | Sales | Service | Construction | Manufacture | Sales | Service | Construction |
| ln (GDP per capita) | -18.70*** | -17.40*** | -1.096 | -7.381*** | -18.92*** | -20.14*** | -0.584 | -7.874** |
| | (3.042) | (2.189) | (2.729) | (2.387) | (2.860) | (3.303) | (2.102) | (3.221) |
| age | | | | | 186.1 | 133.0 | -304.7 | -1.465 |
| | | | | | (149.4) | (83.23) | (184.5) | (137.5) |
| age$^2$ | | | | | -2.509 | -1.781 | 4.084 | 0.0257 |
| | | | | | (2.035) | (1.131) | (2.446) | (1.821) |
| nativity | | | | | -42.79 | 19.39 | 87.20 | -63.05 |
| | | | | | (43.00) | (32.34) | (53.84) | (52.56) |
| Constant | 178.3*** | 192.1*** | 27.62 | 88.21*** | -3,228 | -2,287 | 5,620 | 172.0 |
| | (25.72) | (17.88) | (22.87) | (19.89) | (2,732) | (1,515) | (3,437) | (2,554) |
| Observations | 33 | 33 | 33 | 33 | 33 | 33 | 33 | 33 |
| R-squared | 0.630 | 0.612 | 0.009 | 0.230 | 0.658 | 0.648 | 0.161 | 0.263 |

***, ** represents statistical significance at 1% and 5% respectively; Standard errors are in parentheses.

*Notes*: Table 3.11 reports the estimation of equation (3.18) when restricting the sample to 4 specific industries. The dependent variable is the own-account self-employment rate weighted by "person weight" after taking the average of the multiple years' observations from one country. Samples are from IPUMS-I, restricted to prime age (25-55) male workers.

**Table 3.12**: Employers' rate by industry

| Employers' Rate in | (1) Manufacture | (2) Sales | (3) Service | (4) Construction | (5) Manufacture | (6) Sales | (7) Service | (8) Construction |
|---|---|---|---|---|---|---|---|---|
| ln (GDP per capita) | 1.578*** | 3.982*** | 1.705*** | 0.933*** | 1.582*** | 3.864*** | 1.783*** | 1.488*** |
| | (0.342) | (0.610) | (0.257) | (0.300) | (0.344) | (0.517) | (0.329) | (0.318) |
| age | | | | | -20.88 | -92.22*** | -62.53*** | -33.78 |
| | | | | | (26.76) | (18.21) | (12.35) | (29.54) |
| age$^2$ | | | | | 0.286 | 1.273*** | 0.841*** | 0.442 |
| | | | | | (0.365) | (0.252) | (0.164) | (0.389) |
| nativity | | | | | 8.561 | -3.626 | 10.71* | 5.603 |
| | | | | | (6.083) | (5.778) | (5.614) | (8.959) |
| Constant | -8.643*** | -25.84*** | -10.33*** | -4.216 | 364.5 | 1,647*** | 1,141*** | 630.2 |
| | (2.767) | (4.545) | (1.950) | (2.599) | (491.5) | (330.2) | (228.3) | (553.4) |
| Observations | 33 | 33 | 33 | 33 | 33 | 33 | 33 | 33 |
| R-squared | 0.289 | 0.529 | 0.487 | 0.124 | 0.314 | 0.701 | 0.601 | 0.255 |

***, ** represents statistical significance at 1% and 5% respectively; Standard errors are in parentheses.
*Notes*: Table 3.12 reports the estimation of equation (3.19) when restricting the sample to 4 specific industries. The dependent variable is the employers' rate weighted by "person weight" after taking the average of the multiple years' observations from one country. Samples are from IPUMS-I, restricted to prime age (25-55) male workers.

**Table 3.13**: Own-account self-employment rate by educational attainment with Fixed Effects

| own-account rate for | (1) All | (2) <Primary | (3) Primary | (4) Lower Secondary | (5) Secondary | (6) University |
|---|---|---|---|---|---|---|
| ln (GDP per capita) | -7.194*** | 1.819 | -2.581 | 1.981 | 1.981 | -2.286 |
| | (2.174) | (3.039) | (2.760) | (2.559) | (1.596) | (1.413) |
| age | -124.8 | -74.96*** | -8.272 | -6.950 | -11.20 | -20.74 |
| | (77.87) | (27.08) | (21.43) | (30.63) | (15.42) | (14.44) |
| age$^2$ | 1.682 | 0.941*** | 0.138 | 0.116 | 0.191 | 0.297 |
| | (1.049) | (0.339) | (0.292) | (0.430) | (0.216) | (0.193) |
| schooling | 1.606*** | 0.100 | 1.062 | 0.0415 | -0.0368 | 0.237 |
| | (0.575) | (0.226) | (1.963) | (0.270) | (0.114) | (0.178) |
| nativity | 54.76 | 41.84 | 125.2** | 135.3*** | 57.85*** | -5.339 |
| | (34.53) | (37.33) | (47.90) | (39.24) | (15.97) | (8.915) |
| Constant | 2,332 | 1,466*** | 35.90 | -24.77 | 101.0 | 393.0 |
| | (1,452) | (540.9) | (411.8) | (551.5) | (276.9) | (270.3) |
| Observations | 91 | 91 | 91 | 86 | 91 | 91 |
| R-squared | 0.318 | 0.145 | 0.210 | 0.272 | 0.527 | 0.253 |
| Number of cntry | 28 | 28 | 28 | 26 | 28 | 28 |

***, ** represents statistical significance at 1% and 5% respectively; Standard errors are in parentheses.
*Notes*: Table 3.13 reports the estimation of equation (3.20) with and without restricting the sample to 5 fixed educational attainment groups: less than primary school completed, primary school completed but not lower secondary school, lower secondary completed but not upper secondary school, secondary school completed but not university, and university completed.

**Table 3.14**: Employers' rate by educational attainment with Fixed Effects

| Employers rate for | (1)<br>All | (2)<br><Primary | (3)<br>Primary | (4)<br>Lower Secondary | (5)<br>Secondary | (6)<br>University |
|---|---|---|---|---|---|---|
| ln (GDP per capita) | 0.849 | 0.560 | 0.729 | -0.158 | -0.357 | -0.610 |
| | (0.822) | (0.730) | (0.755) | (0.801) | (0.756) | (1.085) |
| age | 92.84*** | -1.401 | 3.360 | 18.61* | 16.34** | 25.96** |
| | (29.44) | (6.505) | (5.862) | (9.586) | (7.303) | (11.09) |
| $age^2$ | -1.247*** | 0.0203 | -0.0470 | -0.259* | -0.229** | -0.343** |
| | (0.396) | (0.0814) | (0.0799) | (0.135) | (0.102) | (0.149) |
| schooling | 0.243 | 0.0539 | 0.766 | 0.117 | 0.0286 | 0.143 |
| | (0.217) | (0.0543) | (0.537) | (0.0844) | (0.0540) | (0.136) |
| nativity | -7.880 | 3.308 | -31.53** | -20.12 | 7.282 | 10.24 |
| | (13.06) | (8.967) | (13.10) | (12.28) | (7.563) | (6.850) |
| Constant | -1,725*** | 18.61 | -37.10 | -309.9* | -290.9** | -489.4** |
| | (549.0) | (129.9) | (112.6) | (172.6) | (131.2) | (207.7) |
| Observations | 91 | 91 | 91 | 86 | 91 | 91 |
| R-squared | 0.195 | 0.081 | 0.122 | 0.160 | 0.088 | 0.132 |
| Number of cntry | 28 | 28 | 28 | 26 | 28 | 28 |

***, ** represents statistical significance at 1% and 5% respectively; Standard errors are in parentheses.
*Notes*: Table 3.14 reports the estimation of equation (3.21) with and without restricting the sample to 5 fixed educational attainment groups: less than primary school completed, primary school completed but not lower secondary school, lower secondary completed but not upper secondary school, secondary school completed but not university, and university completed.

**Table 3.15**: Calibration

| **Target Moments (CAN in 2001)** | **Data** | **Model** | **Parameter** |
|---|---|---|---|
| Ratio of the average wage for high- to low-education | 1.3 | 1.3 | Ratio of the log(ability) distribution mean for high- to low-education = 4.2 |
| Variance of log(wage) | 0.8 | 0.8 | Variance of log(ability) distribution= 1.1 |
| Own-account Rate | 8 | 8 | Sector share $\gamma = 0.52$ |
| Employers' Rate | 5 | 8 | Employers' share in formal production $\alpha = 0.91$ |

**Table 3.16**: Prime Male Necessity and Opportunity Self-Employment Rates Across Country-years

| | All Industries | | | Excluding Agriculture, fishing, and forestry | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Rates by Employment: | Self-Employed | Own-account | Employer | Self-Employed | Own-account | Employer | Self-Employed | Own-account | Employer |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| ln (GDP per capita) | -13.09*** | -14.75*** | 1.656*** | -6.782*** | -8.271*** | 1.489*** | -6.653*** | -7.922*** | 1.269*** |
| | (0.823) | (0.852) | (0.159) | (0.683) | (0.641) | (0.164) | (1.257) | (1.164) | (0.332) |
| schooling | | | | | | | 0.0769 | 0.0575 | 0.0195 |
| | | | | | | | (0.512) | (0.474) | (0.135) |
| age | | | | | | | 1.991 | 1.458 | 0.533 |
| | | | | | | | (1.255) | (1.162) | (0.331) |
| nativity | | | | | | | 50.16** | 47.01** | 3.147 |
| | | | | | | | (23.57) | (21.82) | (6.214) |
| Constant | 150.6*** | 160.4*** | -9.793*** | 85.96*** | 94.11*** | -8.156*** | -38.39 | -8.947 | -29.44** |
| | (6.996) | (7.239) | (1.347) | (5.836) | (5.483) | (1.403) | (49.36) | (45.70) | (13.01) |
| Observations | 162 | 162 | 162 | 156 | 156 | 156 | 98 | 98 | 98 |
| R-squared | 0.612 | 0.652 | 0.405 | 0.391 | 0.519 | 0.348 | 0.284 | 0.399 | 0.264 |

***, * represents statistical significance at 1% and 10% respectively; Standard errors are in parentheses.

*Notes*: Taking all available census samples from IPUMS-I, the dependent variables self-employment rate, own-account self-employment rate, and employers rate are weighted by "person weight" after restricting the samples to prime age (25-55) males. Regressions (7) to (9) include observations with controls for average years of schooling, average age, and average native-born rate.

# Bibliography


[1] Tasso Adamopoulos and Diego Restuccia. "The Size Distribution of Farms and International Productivity Differences". In: *American Economic Review* 104.6 (2014), pp. 1667–97.

[2] Mark Aguiar and Erik Hurst. "Lifecycle Prices and Production". In: *American Economic Review* 97.5 (Dec. 2007), pp. 1533–59.

[3] Mark Aguiar and Erik Hurst. "Measuring Trends in Leisure: The Allocation of Time Over Five Decades". In: *Quarterly Journal of Economics* 122.3 (2007), pp. 969–1006.

[4] Ufuk Akcigit, Harun Alp, and Michael Peters. "Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries". Unpublished Working Paper, University of Chicago. 2016.

[5] Ufuk Akcigit, Harun Alp, and Michael Peters. "Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries". Unpublished Working Paper, Yale University. 2018.

[6] Pedro S Amaral and Erwan Quintin. "A competitive model of the informal sector". In: *Journal of monetary Economics* 53.7 (2006), pp. 1541–1553.

[7] Pol Antras. *Global Production: Firms, Contracts, and Trade Structure*. Princeton University Press, 2015.

[8] David Argente, Munseob Lee, and Sara Moreira. "Innovation and product reallocation in the great recession". In: *Journal of Monetary Economics* 93 (2018). Carnegie-Rochester-NYU Conference on Public Policy held at the Stern School of Business at New York University, pp. 1–20. ISSN: 0304-3932. DOI: https://doi.org/10.1016/j.jmoneco.2017.11.003. URL: http://www.sciencedirect.com/science/article/pii/S0304393217301319.

[9] S. Boragan Aruoba, Morris A. Davis, and Randall Wright. "Homework in Monetary Economics: Inflation, Home Production, and the Production of Homes". In: *Review of Economic Dynamics* 21 (2016), pp. 105–124.

[10] John Asker, Allan Collard-Wexler, and Jan De Loecker. “Dynamic Inputs and Resource (Mis)Allocation”. In: *Journal of Political Economy* 105.1 (2015), pp. 131–171.

[11] Jose Asturias, Sewon Hur, Timothy J. Kehoe, and Kim J. Ruhl. *Firm Entry and Exit and Aggregate Growth*. Working Paper 23202. National Bureau of Economic Research, Feb. 2017. DOI: 10.3386/w23202. URL: http://www.nber.org/papers/w23202.

[12] David Autor, Lawrence F. Katz, and Alan B. Krueger. “Computing inequality: have computers changed the labor market?” In: *The Quarterly Journal of Economics* 113.4 (1998), pp. 1169–1213.

[13] Pierre Bachas, Roberto N Fattal-Jaef, and Anders Jensen. “Size-dependent tax enforcement and compliance: global evidence and aggregate implications”. In: (2018). Unpublished Working Paper, The World Bank.

[14] Yan Bai, Keyu Jin, and Dan Lu. “Misallocation Under Trade Liberalization”. Unpublished Working Paper, University of Rochester. 2018.

[15] Abhijit V. Banerjee and Andrew F. Newman. “Occupational Choice and the Process of Development”. In: *Journal of Political Economy* 101.2 (1993), pp. 274–98.

[16] Abhijit Banerjee, Emily Breza, Esther Duflo, and Cynthia Kinnan. “Do credit constraints limit entrepreneurship? Heterogeneity in the returns to microfinance”. Working Paper, Northwestern. 2015.

[17] Anurag Banerjee, Parantap Basu, and Elisa Keller. “Cross-Country Disparities in Skill Premium and Skill Acquisition”. Unpublished Working Paper, Durham University. 2016.

[18] William J Baumol. “Entrepreneurship: Productive, unproductive, and destructive”. In: *Journal of Business Venturing* 11.1 (1996), pp. 3–22.

[19] Marianne Baxter and Urban J. Jermann. “Household Production and the Excess Sensitivity of Consumption to Current Income”. In: *American Economic Review* 84.4 (1999), pp. 902–920.

[20] Jess Benhabib, Richard Rogerson, and Randall Wright. “Homework in Macroeconomics: Household Production and Aggregate Fluctuations”. In: *Journal of Political Economy* 99.6 (Dec. 1991), pp. 1166–1187.

[21] Pedro Bento and Diego Restuccia. “Misallocation, Establishment Size, and Productivity”. In: *American Economic Journal: Macroeconomics* 9.3 (2017), pp. 267–303.

[22] Alexander Bick, Nicola Fuchs-Schuendeln, and David Lagakos. “How Do Hours Worked Vary with Income? Cross-Country Evidence and Implications”. In: *American Economic Review* 108.8 (2018), pp. 170–99.

[23] Mark Bils, Peter J Klenow, and Cian Ruane. "Misallocation or Mismeasurement?" Unpublished Working Paper, Stanford University. 2017.

[24] Olivier J. Blanchard and Lawrence H. Summers. "Hysteresis and the European Unemployment Problem". In: *NBER Macroeconomics Annual* 1 (1986).

[25] Timo Boppart. "Structural Change and the Kaldor Facts in a Growth Model With Relative Price Effects and Non?Gorman Preferences". In: *Econometrica* 82.6 (2014), pp. 2167–2196.

[26] Loren Brandt, Gueorgui Kambourov, and Kjetil Storesletten. "Firm Entry and Regional Growth Disparities: the Effect of SOEs in China". In: *University of Toronto mimeo* (2016).

[27] Loren Brandt, Trevor Tombe, and Xiaodong Zhu. "Factor market distortions across time, space and sectors in China". In: *Review of Economic Dynamics* 16.1 (2013), pp. 39–58.

[28] Loren Brandt, Johannes Van Biesebroeck, and Yifan Zhang. "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing". In: *Journal of Development Economics* 97.2 (2012), pp. 339–351.

[29] J. Carter Braxton, Kyle Herkenhoff, and Gordon Phillips. "Can the Unemployed Borrow? Implications for Public Insurance". Unpublished Manuscript, University of Minnesota. 2018.

[30] Benjamin Bridgman, Georg Duernecker, and Berthold Herrendorf. "Structural Transformation, Marketization, and Household Production around the World". In: *Journal of Development Economics* 133 (2018), pp. 102–126.

[31] Christian Broda and David E. Weinstein. "Globalization and the Gains from Variety". In: *Quarterly Journal of Economics* 121.2 (2006), pp. 541–585.

[32] Francisco J. Buera and Joseph P. Kaboski. "The Rise of the Service Economy". In: *American Economic Review* 102 (2012), pp. 2450–69.

[33] Francisco J. Buera, Joseph P. Kaboski, and Yongseok Shin. "Finance and Development: A Tale of Two Sectors". In: *American Economic Review* 101.8 (2011), pp. 1964–2002.

[34] Francisco J. Buera, Josephi P. Kaboski, and Richard Rogerson. "Skill-Biased Structural Change". Unpublished Working Paper, University of Notre Dame. 2015.

[35] Sara Calligaris, Massimo Del Gatto, Fadi Hassan, Gianmarco IP Ottaviano, and Fabiano Schivardi. "The productivity puzzle and misallocation: an Italian perspective". In: (2018). Forthcoming, Journal of Population Economics.

[36] Francesco Caselli. “Accounting for Cross-Country Income Differences”. In: *Handbook of Economic Growth*. Ed. by Philippe Aghion and Steven N. Durlauf. 2005.

[37] Francesco Caselli and Wilbur John Coleman. “The World Technology Frontier”. In: *American Economic Review* 96.3 (June 2006), pp. 499–522.

[38] Pedro Ferreira Cavalcanti, Alexander Monge-Naranjo, and Luciene Torres de Mello. “Of Cities and Slums”. Federal Reserve Bank of St. Louis Working Paper 2016-022A. 2016.

[39] Yongsung Chang and Frank Schorfheide. “Labor-Supply Shifts and Economic Fluctuations”. In: *Journal of Monetary Economics* 50 (2003), pp. 1751–1768.

[40] Chaoran Chen, Ashique Habib, and Xiaodong Zhu. “Contracting Frictions with Managers, Financial Frictions, and Misallocation”. Unpublished Working Paper. 2018.

[41] Cheng Chen, Tatsuro Senga, Chang Sun, and Hongyong Zhang. *Uncertainty, Imperfect Information, and Learning in the International Market*. Working Paper. University of Hong Kong, 2018.

[42] Gabriel Chodorow-Reich and Loukas Karabarbounis. “The Cyclicality of the Opportunity Cost of Employment”. In: *Journal of Political Economy* 124.6 (2016), pp. 1563–1618.

[43] Harold L. Cole, Jeremy Greenwood, and Juan Sanchez. “Why Doesn’t Technology Flow from Rich to Poor Countries?” In: *Journal of Political Economy* 84.4 (2016), pp. 1477–1521.

[44] Joel M. David and Venky Venkateswaran. *The Sources of Capital Misallocation*. Working Paper 23129. National Bureau of Economic Research, Feb. 2017. DOI: 10.3386/w23129. URL: http://www.nber.org/papers/w23129.

[45] Joel David, Hugo Hopenhayn, and Venky Venkateswaran. “Information, Misallocation and Aggregate Productivity”. In: *Quarterly Journal of Economics* 131 (2016), pp. 943–1005.

[46] Angus Deaton. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. The World Bank, 1997.

[47] Peter A. Diamond. “Aggregate Demand Management in Search Equilibrium”. In: *Journal of Political Economy* 90.8 (1982), pp. 881–894.

[48] Kevin Donovan, Jianyu Lu, and Todd Schoellman. “Labor Market Flows and Development”. Unpublished Manuscript, University of Notre Dame. 2018.

[49] Margarida Duarte and Diego Restuccia. “The Role of the Structural Transformation in Aggregate Productivity”. In: *Quarterly Journal of Economics* 125.1 (2010), pp. 129–173.

[50] David S Evans. "The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries". In: *Journal of Industrial Economics* (1987), pp. 567–581.

[51] Roberto N. Fattal-Jaef. "Entry and Exit, Multiproduct Firms, and Allocative Distortions". In: *American Economic Journal: Macroeconomics* 10.2 (Apr. 2018), pp. 86–112. DOI: 10.1257/mac.20140075. URL: http://www.aeaweb.org/articles?id=10.1257/mac.20140075.

[52] Robert C. Feenstra, Robert Inklaar, and Marcel P. Timmer. "The Next Generation of the Penn World Table". In: *American Economic Review* 105.10 (2015), pp. 3150–3182.

[53] Ying Feng and Lindsay Rickey. "Development and Selection into Necessity versus Opportunity Entrepreneurship". Unpublished Working Paper, UCSD. 2016.

[54] Gary S Fields. "A Guide to Multi-sector Labor Market Models". Unpublished Working Paper, Cornell University. 2004.

[55] Gary S Fields. "Education and Income Distribution in Developing Countries: A Review of the Literature". World Bank Staff Working Paper No. 402 (pp. 231 - 315). 1980.

[56] Shigeru Fujita and Gary Ramey. "Exogenous Versus Endogenous Separation". In: *American Economic Journal: Macroeconomics* 4.4 (2012), pp. 68–93.

[57] Xavier Giné and Robert M Townsend. "Evaluation of financial liberalization: a general equilibrium model with constrained occupation choice". In: *Journal of development Economics* 74.2 (2004), pp. 269–307.

[58] Douglas Gollin. "Nobody's Business But My Own: Self-Employment and Small Enterprise in Economic Development". In: *Journal of Monetary Economics* 55.2 (2008), pp. 219–233.

[59] Douglas Gollin, David Lagakos, and Michael E. Waugh. "The Agricultural Productivity Gap". In: *Quarterly Journal of Economics* 129.2 (2014), pp. 939–993.

[60] Douglas Gollin, Stephen L. Parente, and Richard Rogerson. "Farm Work, Home Work and International Productivity Differences". In: *Review of Economic Dynamics* 7 (2004), pp. 827–850.

[61] Jeremy Greenwood and Zvi Hercowitz. "The Allocation of Capital and Time Over the Business Cycle". In: *Journal of Political Economy* 99.6 (1991), pp. 1188–1214.

[62] Jeremy Greenwood, Juan M. Sanchez, and Cheng Wang. "Financing Development: The Role of Information Costs". In: *American Economic Review* 100.4 (2010), 1875?1891.

[63] Gustavo Ventura Guner Nezih and Daniel Xu. "Macroeconomic Implications of Size-Dependent Policies". In: *Review of Economic Dynamics* 11.4 (Oct. 2008), pp. 721–744.

[64] Robert E Hall. “The measurement of quality change from vintage price data”. In: *Price indexes and quality change*. Ed. by Z. Griliches. Cambridge: Harvard University Press, 1971. Chap. 8, pp. 240–271.

[65] Robert E. Hall and Andreas I. Mueller. “Wage Dispersion and Search Behavior: The Importance of Non-Wage Job Values”. In: *Journal of Political Economy* (Forthcoming).

[66] John Haltiwanger, Ron S Jarmin, and Javier Miranda. “Who creates jobs? Small versus large versus young”. In: *Review of Economics and Statistics* 95.2 (2013), pp. 347–361.

[67] James D Hamilton. “Why you should never use the Hodrick-Prescott filter”. In: *Review of Economics and Statistics* 0 (2017).

[68] John R. Harris and Michael P. Todaro. “Migration, Unemployment and Development: A Two-Sector Analysis”. In: *American Economic Review* 60.1 (1970), pp. 126–142.

[69] Berthold Herrendorf, Richard Rogerson, and Akos Valentinyi. “Growth and Structural Transformation”. In: *Handbook of Economic Growth*. Ed. by Philippe Aghion and Steven N. Durlauf. Vol. 2. Elsevier, 2014. Chap. 6, pp. 855–941.

[70] Berthold Herrendorf and Todd Schoellman. “Wages, Human Capital, and Structural Transformation”. Unpublished Manuscript, Arizona State University. 2016.

[71] Jonas Hjort and Jonas Poulsen. “The Arrival of Fast Internet and Employment in Africa”. In: *American Economic Review* (forthcoming).

[72] Sui-Jade Ho and Dimitrije Ruzic. “Returns to Scale, Productivity Measurement, and Trends in US Manufacturing Misallocation”. In: *2018 Meeting Papers*. 119. Society for Economic Dynamics. 2018.

[73] Chang-Tai Hsieh and Peter J. Klenow. “Misallocation and Manufacturing TFP in China and India”. In: *Quarterly Journal of Economics* 124 (2009), pp. 1403–1448.

[74] Chang-Tai Hsieh and Peter J. Klenow. “Relative Prices and Relative Prosperity”. In: *American Economic Review* 97.3 (June 2007), pp. 562–585.

[75] Chang-Tai Hsieh and Peter J. Klenow. “The Life Cycle of Plants in India and Mexico *”. In: *The Quarterly Journal of Economics* 129.3 (2014), pp. 1035–1084. DOI: 10.1093/qje/qju014. eprint: /oup/backfile/content_public/journal/qje/129/3/10.1093_qje_qju014/4/qju014.pdf. URL: http://dx.doi.org/10.1093/qje/qju014.

[76] Chang-Tai Hsieh and Benjamin A Olken. *The Missing Missing Middle*. Working Paper 19966. National Bureau of Economic Research, 2014. DOI: 10.3386/w19966.

[77] Chang-Tai Hsieh and Zheng Michael Song. *Grasp the large, let go of the small: the transformation of the state sector in China*. Tech. rep. NBER Working Paper No. 21006, 2015.

[78] Michael Jerzmanowski and Robert Tamura. "Directed Technological Change and Cross Country Income Differences: A Quantitative Analysis". Unpublished Working Paper, Clemson University. 2017.

[79] Boyan Jovanovic. "Selection and the Evolution of Industry". In: *Econometrica* 50.3 (1982), pp. 649–670. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1912606.

[80] Loukas Karabarbounis. "Production, Labor Wedges, and International Business Cycles". In: *Journal of Monetary Economics* 64 (2014), pp. 68–84.

[81] Matthias Kehrig and Nicolas Vincent. "Do firms mitigate or magnify capital misallocation? Evidence from plant-level data". In: (2017). US Census Bureau Center for Economic Studies Paper No. CES-WP-17-14.

[82] William R. Kerr, Ramana Nanda, and Matthew Rhodes-Kropf. "Entrepreneurship as Experimentation". In: *Journal of Economic Perspectives* 28.3 (Sept. 2014), pp. 25–48.

[83] Alan B. Krueger and Andreas Mueller. "Job Search and Unemployment Insurance: New Evidence from Time Use Data". In: *Journal of Public Economics* 94.3 (2010), pp. 298–307. ISSN: 0047-2727. DOI: https://doi.org/10.1016/j.jpubeco.2009.12.001. URL: http://www.sciencedirect.com/science/article/pii/S0047272709001625.

[84] Rafael La Porta and Andrei Shleifer. "Informality and Development". In: *Journal of Economic Perspectives* 28.3 (2014), pp. 109–26.

[85] Rafael La Porta and Andrei Shleifer. "The Unofficial Economy and Economic Development". In: *Brookings Papers on Economic Activity* 2 (2008), pp. 275–363.

[86] David Lagakos and Michael E Waugh. "Selection, Agriculture, and Cross-Country Productivity Differences". In: *The American Economic Review* 103.2 (2013), pp. 948–980.

[87] Ferdinand Lepper. *Comparable Annual Employment and Unemployment Estimates*. Tech. rep. Geneva: Department of Statistics, International Labour Office, 2004.

[88] Ross Levine and Yona Rubinstein. *Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?* Working Paper 19276. National Bureau of Economic Research, Aug. 2013. DOI: 10.3386/w19276. URL: http://www.nber.org/papers/w19276.

[89] W. Arthur Lewis. "Economic Development with Unlimited Supplies of Labor". In: *The Manchester School* 22.2 (1954), pp. 139–91.

[90] Huiyu Li. "Leverage and productivity". In: *Unpublished manuscript: Stanford University* (2015).

[91] Lars Ljungqvist and Thomas J. Sargent. "Two Questions about European Unemployment". In: *Econometrica* 76.1 (2008), pp. 1–29.

[92] Robert E. Lucas, Jr. "On the Size Distribution of Business Firms". In: *Bell Journal* 9 (1978), pp. 508–523.

[93] Hannes Malmberg. "Human Capital and Development Accounting Revisted". Unpublished Working Paper, IIES Stockholm. 2016.

[94] Ellen R. McGrattan, Richard Rogerson, and Randall Wright. "An Equilibrium Model of the Business Cycle with Household Production and Fiscal Policy". In: *International Economic Review* 38 (1997), pp. 267–290.

[95] David J McKenzie. "Disentangling age, cohort and time effects in the additive model". In: *Oxford bulletin of economics and statistics* 68.4 (2006), pp. 473–495.

[96] Kathleen McKiernan. "Welfare Impacts of Social Security Reform: The Case of Chile in 1981". Unpublished Working Paper, Vanderbilt University. 2018.

[97] Martí Mestieri, Diego Comin, and Danial Lashkari. "Structural Change with Long-run Income and Price Effects". Unpublished Working Paper, Northwestern University. 2018.

[98] Virgiliu Midrigan and Daniel Xu. "Finance and Misallocation: Evidence from Plant-Level Data". In: *American Economic Review* 104.2 (2014), pp. 422–458.

[99] Jacob Mincer. "Education and Unemployment". NBER Working Paper No. 3838. 1991.

[100] Minnesota Population Center. *Integrated Public Use Microdata Series, International: Version 6.4*. Minneapolis: University of Minnesota. 2015.

[101] Minnesota Population Center. *Integrated Public Use Microdata Series, International: Version 6.5 [dataset]*. Minneapolis: University of Minnesota. http://doi.org/10.18128/D020.V6.5. 2017.

[102] Sara Moreira. "Firm Dynamics, Persistent Effects of Entry Conditions, and Business Cycles." In: (2017). U.S. Census Bureau Center for Economic Studies Working Paper No. CES-WP-17-29.

[103] Dale Mortensen and Christopher Pissarides. "Job Creation and Job Destruction in the Theory of Unemployment". In: *Review of Economic Studies* 61 (1994), pp. 397–415.

[104] Andreas I. Mueller. "Separations, Sorting, and Cyclical Unemployment". In: *American Economic Review* 107.7 (2017), pp. 2081–2107.

[105] Rachel L. Ngai and Christopher A. Pissarides. “Taxes, Social Subsidies and the Allocation of Work Time”. In: *American Economic Journal: Macroeconomics* 3.4 (2011), pp. 1–26.

[106] Rachel L. Ngai and Christopher A. Pissarides. “Trends in Hours and Economic Growth”. In: *Review of Economic Dynamics* 11.2 (2008), pp. 239–56.

[107] Stephen Nickell, Luca Nunziata, and Wolfgang Ochel. “Unemployment in the OECD Since the 1960s. What Do We Know?” In: *Economic Journal* 115.500 (2004), pp. 1–27.

[108] G. Steven Olley and Ariel Pakes. “The Dynamics of Productivity in the Telecommunications Equipment Industry”. In: *Econometrica* 64.6 (1996), pp. 1263–1297. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/2171831.

[109] Stephen L. Parente, Richard Rogerson, and Randall Wright. “Homework in Development Economics: Household Production and the Wealth of Nations”. In: *Journal of Political Economy* 108.4 (Aug. 2000), pp. 680–687.

[110] Tommaso Porzio and Gabriella Santangelo. “Structural Change and the Supply of Agricultural Workers”. Unpublished Working Paper, University of California San Diego. 2017.

[111] Markus Poschke. *The decision to become an entrepreneur and the firm size distribution: a unifying framework for policy analysis*. Tech. rep. IZA Discussion Paper, 2013.

[112] Markus Poschke. “Wage Employment, Unemployment and Self-Employment Across Countries”. Unpublished Manuscript, McGill University. 2018.

[113] Markus Poschke. “Who becomes an entrepreneur? Labor market prospects and occupational choice”. In: *Journal of Economic Dynamics and Control* 37.3 (2013), pp. 693–710.

[114] Sangeeta Pratap and Erwan Quintin. *The informal sector in developing countries: Output, assets and employment*. 2006/130. Research Paper, UNU-WIDER, United Nations University (UNU), 2006.

[115] Valerie A. Ramey and Neville Francis. “A Century of Work and Leisure”. In: *American Economic Journal: Macroeconomics* 1.2 (2009), pp. 189–224.

[116] Valerie A Ramey and Matthew D Shapiro. “Displaced capital: A study of aerospace plant closings”. In: *Journal of political Economy* 109.5 (2001), pp. 958–992.

[117] James E. Rauch. “Modelling the Informal Sector Informally”. In: *Journal of Development Economics* 35.1 (1991), pp. 33–47.

[118] Diego Restuccia and Richard Rogerson. “Policy Distortions and Aggregate Productivity with Heterogeneous Establishments”. In: *Review of Economic Dynamics* 11.4 (Oct. 2008), pp. 707–720.

[119] Diego Restuccia and Richard Rogerson. “The causes and costs of misallocation”. In: *Journal of Economic Perspectives* 31.3 (2017), pp. 151–74.

[120] Diego Restuccia and Carlos Urrutia. “Relative Prices and Investment Rates”. In: *Journal of Monetary Economics* 47.1 (2001), pp. 93–121.

[121] Mark J Roberts and James R Tybout. *Industrial evolution in developing countries: Micro patterns of turnover, productivity, and market structure*. Oxford University Press, 1996.

[122] Richard Rogerson. “Structural Transformation and the Deterioration of European Labor Market Outcomes”. In: *Journal of Political Economy* 116.2 (2008), pp. 235–259.

[123] Mark R Rosenzweig. “Why are there returns to schooling?” In: *The American Economic Review* 85.2 (1995), pp. 153–158.

[124] Martin Rotemberg and T Kirk White. *Measuring Cross-Country Differences in Misallocation*. Tech. rep. Working Paper, 2017.

[125] A. Roy. “Some Thoughts on the Distribution of Earnings”. In: *Oxford Economic Papers* 3 (1951), pp. 135–46.

[126] Peter Rupert, Richard Rogerson, and Randall Wright. “Estimating Substitution Elasticities in Household Production Models”. In: *Economic Theory* 6 (1995), pp. 179–93.

[127] Antoinette Schoar. “The Divide between Subsistence and Transformational Entrepreneurship”. In: *Innovation Policy and the Economy*. Ed. by Josh Lerner and Scott Stern. University of Chicago Press, 2010. Chap. 3, pp. 57–81.

[128] Tatsuro Senga. *A new look at uncertainty shocks: Imperfect information and misallocation*. Tech. rep. Working Paper, School of Economics and Finance, Queen Mary University of London, 2015.

[129] Robert Shimer. “The Cyclical Behavior of Equilibrium Unemployment and Vacancies”. In: *American Economic Review* 95.1 (Mar. 2005).

[130] Zheng Song and Guiying Laura Wu. “Identifying capital misallocation”. In: (2015).

[131] Alina Sorgner, Michael Fritsch, and Alexander Kritikos. “Searching for Fortune fffdfffdfffd When Does Entrepreneurship Pay?” Economics of Entrepreneurship Workshop, Washington, DC. 2015.

[132] Lyn Squire. *Employment Policy in Developing Countries: A Survey of Issues and Evidence*. New York: Published for the World Bank by Oxford University Press., 1981.

[133] Mari Tanaka, Nicholas Bloom, Joel M David, and Maiko Koga. *Firm Performance and Macro Forecast Accuracy*. Working Paper 24776. National Bureau of Economic Research, June 2018. DOI: 10.3386/w24776. URL: http://www.nber.org/papers/w24776.

[134] David Turnham. *Employment and Development: A New Review of Evidence*. Paris, OECD, 1993.

[135] United Nations. "System of National Accounts, 2008". https://unstats.un.org/unsd/sna1993/WC-SNAvolume2.pdf. 2008.

[136] U.S. Bureau of Labor Statistics. *Handbook of Methods*. Washington, D.C.: U.S. Government Printing Office, 2016.

[137] Justin Van der Sluis, Mirjam Van Praag, and Wim Vijverberg. "Entrepreneurship selection and performance: A meta-analysis of the impact of education in developing economies". In: *The World Bank Economic Review* 19.2 (2005), pp. 225–261.

[138] Erin Wolcott. "Employment Inequality: Why Do the Low-Skilled Work Less Now?" Unpublished Working Paper, Middlebury College. 2018.

[139] Christopher Woodruff. "Self-employment: Engine of growth or self-help safety net?" In: *Employment and Shared Growth* 53 (2007).

[140] Mu-Jeung Yang. "Micro-level Misallocation and Entry Selection". Unpublished Working Paper, University of Washington. 2016.

[141] Alwyn Young. "Inequality, the Urban-Rural Gap and Migration". In: *The Quarterly Journal of Economics* 129.2 (2013), pp. 939–993.