Title
Essays on Workers' Human Capital Accumulation and Wage Experience Profiles

Author
Nakab, Alejandro

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays on Worker's Human Capital Accumulation and Wage Experience Profiles

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Alejandro Nakab

Committee in charge:

Professor David Lagakos, Co-Chair
Professor Valerie Ramey, Co-Chair
Professor Titan Alon
Professor Ruixue Jia
Professor Munseob Lee
Professor Marc-Andreas Muendler
Professor Natalia Ramondo

2021

The dissertation of Alejandro Nakab is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my main advisors, David Lagakos, Valerie Ramey and Natalia Ramondo, for their support. Their guidance has been crucial for my personal development as an Economist.

I would also like to thank Titan Alon, Ruixue Jia, Munseob Lee, Marc Muendler, and Tommaso Porzio for their continuous guidance. I am also grateful to the entire faculty of the Economics department for insightful comments during seminars, classes and meetings.

Chapter 1 and Chapter 2 are parts of the paper "Human Capital Accumulation and Development: The Role of On-the-Job Training" coauthored with Xiao Ma and Daniela Vidart. As of May 2021, this paper is being prepared for submission for publication. The dissertation author was a primary investigator of this material. I would like to thank Xiao and Daniela for being great coauthors.

Chapter 3 is currently being prepared for submission for publication and is coauthored with Xiao Ma. The dissertation author was a primary investigator of this material. I would like to thank Xiao in particular for all the invaluable discussions we had over these years. I learned a lot from our interactions.

All the results and conclusions are ours and not those of Eurostat, the European Commission, the World Bank or any of the national statistical authorities whose data have been used.

VITA

2012	B.A. in Economics, Universidad Torcuato Di Tella
2013	M.A. in Economics, Universidad de San Andres
2015	M.A. in International Economics and Politics, Universidad de San Andres
2021	Ph. D. in Economics, University of California San Diego

ABSTRACT OF THE DISSERTATION

Essays on Worker's Human Capital Accumulation and Wage Experience Profiles

by

Alejandro Nakab

Doctor of Philosophy in Economics

University of California San Diego, 2021

Professor David Lagakos, Co-Chair
Professor Valerie Ramey, Co-Chair

This dissertation consists of three chapters. Chapter 1 and Chapter 2 jointly offer an explanation for why workers in richer countries have faster rates of wage growth over their lifetimes than workers in poorer countries. We propose that workers in richer economies receive more firm-provided training. In Chapter 1, we document two main facts: the share of workers who receive firm-provided training increases with development, and that this is a key determinant of worker human capital investments. In Chapter 2 we build a general equilibrium search model with firm-training investments and frictional labor markets. Our model suggests firm-training accounts for a large share of the cross-country wage growth differences. We find that self-employment is a

key factor explaining the lack of training in the poorest economies, whereas labor market frictions are key to explaining training differences within firms as countries develop. Finally, our model predicts considerable inefficiencies in human capital investments and sizeable aggregate gains from training subsidies to firms, which may be particularly desirable in poor countries where economic environments disincentivize training.

Chapter 3 studies how exporting shapes experience-wage profiles. Using detailed Brazilian employer-employee and customs data, we document that workers' experience-wage profiles are steeper in exporters than in non-exporters. Aside from self-selection of firms with higher returns to experience into exporting, we show that workers' experience-wage profiles are steeper when firms export to industrialized destinations. We propose that this result is likely driven by faster human capital accumulation of workers in firms that export to advanced economies. To support our preferred hypothesis, we use the World Bank Enterprise Surveys and document that exporters are more likely to train workers than non-exporters, especially when they adopt foreign technology.

Part I

Human Capital Investment and Development: The Role of On-the-job Training

Introduction.

Recent papers have shown that workers in richer countries have faster rates of wage growth over their lifetimes than workers in poorer countries (Lagakos et al., 2018*b*; Islam et al., 2019). Different theories of life-cycle wage growth are consistent with this pattern, including differences in human capital accumulation, labor market frictions or long-term contracts. Since these possible drivers have massively different implications for policy and for explaining cross-country income differences, understanding the reasons behind this pattern is a first order question. In this paper, we offer an explanation for this new stylized fact by focusing on one key source of worker human capital accumulation: firm-provided training. To that end, we carefully measure workers' post-schooling human capital accumulation investments, and explore how they differ across countries. Our results explain why post-schooling human capital accumulation is greater for workers in more developed economies.

We present both empirical and quantitative evidence on the link between firm-provided training and the level of development. In the empirical portion of the paper we rely on enterprise surveys covering more than 400,000 firms over 102 countries and worker surveys covering more 600,000 people over 26 countries with detailed information on workers' training investments. These surveys allow us to construct harmonized cross-country representative measures of on-the-job training provisions spanning a broad range of development with PPP-adjusted GDP per capita ranging from \$1,000 to \$60,000. With different measures of on-the-job training, we document two novel facts.

The first fact we document is that the share of workers who receive firm-provided training rises strongly with the level of country per-capita GDP. We show that a key margin mediating this positive correlation in poor economies is the large share of self-employed workers who do not receive employer-provided training. Moreover, focusing only on workers employed by firms, we still find that the share of workers who are offered training rises with country-level GDP per capita. Richer countries exhibit a larger share of firms offering training, along with a larger share of trainees within the firms offering training and a greater share of hours in training relative to total hours worked. In addition, firms in richer countries spend more on training per participant, which potentially reflects training quality.

For the second fact, we provide a detailed description of adults' human capital investments and document that job-related firm-provided training is a key determinant of on-the-job human capital accumulation. This evidence suggests that firms play a substantial role in adult human capital investments, and thus, canonical models *à la* Ben-Porath, which do not include firm-level decisions, provide an incomplete picture of the on-the-job skill acquisition process.

To shed light on the mechanisms giving rise to the positive correlation between training and development, we build a general equilibrium model that explicitly accounts for firm-worker decision-making regarding on-the-job training. The model features two sectors: a self-employment sector and a wage-sector. The self-employment sector has no learning opportunities and no frictions. The wage sector, on the other hand, is characterized by labor market frictions and firm heterogeneity *à la* Burdett and Mortensen (1998). Firms post vacancies and wages and meet workers by random search following Mortensen and Pissarides (1994) and Pissarides (2000). We incorporate training investments based on the theoretical literature on general training investments developed by Acemoglu (1997), Acemoglu and Pischke (1998), and Moen and Rosén (2004). However, we depart from this literature in the way training costs are allocated between workers and firms, and by incorporating richer job turnover dynamics based on on-the-job search and contract quality. In our model, workers can be separated from firms for two reasons: an exogenous separation shock that may lead workers to unemployment, and on-the-job search as workers look for new job offers while working. When employed workers receive a new job offer, they can choose to exert efforts to break their contract, incurring costs that depend on the economy's contract quality.

Motivated by our empirical results and the training literature that considers job turnover as the fundamental barrier to training investments, we focus on three main channels that vastly differ across stages of development and directly impact training: differences in job destruction rates (Donovan, Lu and Schoellman, 2020), differences in contract quality and institutions, which shape labor market dynamics and contract length (e.g. Hall and Jones, 1999; Acemoglu, Johnson and Robinson, 2005), and differences in self-employment shares (Gollin, 2002, 2008). Thus, we calibrate our model to a representative country at different income levels and perform several exercises aimed at answering the following three questions: (1) how much of the wage-growth differences across countries can be accounted for with on-the-job training; (2) why do developed economies invest more in training; and (3) what is the optimal training policy at different stages of development?

We first show that our model matches all of the cross-country differences in wage-growth for countries above \$10,000 of per-capita GDP. Nevertheless, it overpredicts the wage-growth for economies at the bottom of the world income distribution. When we include these countries, the model explains 55% of the wage-growth differences. Moreover, we find that the contribution from human capital in explaining workers wage growth is large for every economy and that it decreases with income. This happens because the high level of job destruction in the poorest economies prevents workers climbing up the job ladder. As income increases, fewer workers are separated from their jobs and become unemployed, which generates larger increases in wages through job-to-job transitions. Finally, we show that in our model 70% of the difference in wage-experience profiles is driven by on-the-job training. Moreover, focusing on productivity differences generated by this channel, we show that on-the-job training explains 10% to 15% of the income differences across countries in our quantitative model, which is a large share of the total differences stemming from life cycle human capital gains found in recent studies.

We then conduct a factor-decomposition of training to explore the evolving importance of the different channels at different stages of development and find three main results. First, most of the training gap between the poorest and richest economies is explained by differences in self-employment shares. This is driven by the high rates of self-employment prevalent in poor economies arising from the endogenous allocation of workers as a result of the wage-sector's high labor market frictions and lower relative wage-sector to self-employment productivity. Second, these labor market frictions remain key to explaining training investments as countries develop and self-employment shares fall. The mechanism driving this is the wage-sector's worker turnover. In particular, high job separation rates and low contract quality make worker turnover more likely, and thus depress the incentives to invest in training in low- and medium-income economies relative to richer economies. Third, when we decompose the importance of these labor market frictions along its two key components, we find that job destruction is the most important factor to explain the lack of training in poorer economies while frictions in job-to-job transitions are more important to explain the training differences between more developed economies.

In our framework, training investments are a negotiated outcome stemming from a joint decision of workers and firms. When deciding the optimal training level, workers and firms do not internalize the other party's and future firms' benefits from training, which generates inefficiencies in training investments. Motivated by the existence of these inefficiencies, we study the optimal training subsidy at different stages of development. We find that a training subsidy is a possible policy that could correct the distortions, and that these might be particularly desirable in poor countries where economic environments disincentivize training. We show that as self-employment and job separation increases or contract quality decreases, the optimal training subsidy rises to incentivize firms to invest in training.

Related Literature. This paper relates to several strands of the literature. First, our theory combines insights from two related strands of the literature studying on-the-job human capital accumulation. Our model builds on the theoretical literature on general training investments, first proposed by Becker (1964), and later developed by others such as Acemoglu (1997), Acemoglu and Pischke (1998) and Moen and Rosén (2004). Moreover, by embedding this firm-worker training investment dynamic into a search model, we relate to the literature that tries to disentangle the contributions of human capital and search dynamics on earnings (e.g. Bunzel et al., 1999; Rubinstein and Weiss, 2006; Barlevy, 2008; Yamaguchi, 2010; Burdett, Carrillo-Tudela and Coles, 2011; Bowlus and Liu, 2013; Bagger et al., 2014; Gregory, 2019). These papers differ from ours along several key dimensions. First, a large contingent of these papers assume that on-the-job human capital accumulation does not follow from an optimization problem facing tradeoffs between work and learning, and is simply an exogenous by-product of work.[1] Second, the focus of these papers contrasts sharply with the goal of our theory, which is to explain cross-country differences in training and income. Moreover, this literature analyzes how job search and human capital accumulation contributes to explaining workers' wage growth for specific developed economies. We also contribute to this literature by analyzing this decomposition for countries at all income levels.

By exploring the role of worker training in explaining differences in GDP per worker across countries, our paper relates to a large strand of the literature that measures the contribution of different factors in explaining cross-country income differences (e.g. Klenow and Rodriguez-Clare, 1997; Caselli, 2005; Hsieh and Klenow, 2010), and in particular to the ones focused on human capital.[2] Our paper focuses on one understudied source of cross-country human capital difference, namely on-the-job human capital accumulation. Thus, our work relates to the recent literature that highlights the potential importance of life-cycle human capital accumulation differences across countries (De la Croix, Doepke and Mokyr, 2018; Lagakos et al., 2018*b*,*a*; Islam et al., 2019). This literature, however, does not explain how these differences in on-the-job human capital accumulation patterns across countries emerge. Our paper attempts to fill this gap by delving into the processes and features giving rise to the low workers' skill acquisition prevalent in poor countries, and focuses its attention on employer-provided training.

[1]Wasmer (2006) and Flinn, Gemici and Laufer (2017) incorporate micro-founded human capital investment decisions, but they focus on studying the distinction between firm-specific and general training.

[2]These focused on explaining cross-country productivity differences by quantitatively measuring the role of educational attainment (e.g. Hall and Jones, 1999; Erosa, Koreshkova and Restuccia, 2010; Jones, 2014), measuring school quality (e.g. Hanushek and Woessmann, 2012; Schoellman, 2012, 2016), or differences in skill specialization in secondary and post-secondary curricula (Alon et al., 2017; Alon and Fershtman, 2019).

Third, our paper is related to the literature that explores the relationship between labor market dynamics and development. In particular, we incorporate insights from (1) the literature on cross-country job turnover differences (Donovan, Lu and Schoellman, 2020); (2) the vast literature on institutional quality differences across countries (reviewed in Acemoglu, Johnson and Robinson (2005)); and (3) the literature focusing on cross-country self-employment share differences (e.g. Gollin, 2002, 2008; Poschke, 2018). We contribute to this development literature by incorporating the interaction between these channels and firm-provided training.[3]

Finally, our focus on training and cross country analysis closely relates to two recent papers. The first is Doepke and Gaetani (2020) who study cross-country differences in on-the-job skill acquisition focusing on employment protection, which affects firms' and workers' incentives to invest in skills. The second is Engbom (2020) who studies differences in endogenous human capital formation in a search model, and focuses on how the costs of doing business affect human capital. Our work differs from theirs in the channels that we study, which include different labor market frictions and self-employment. More importantly, we focus on explaining the trend component of training with respect to per-capita GDP while they study different channels that vary across countries but may not directly explain the relationship between income and training. While they focus on developed economies, we provide evidence and quantitative analysis for countries at all stages of development. The closest paper in analyzing human capital differences at all stages of development is Manuelli and Seshadri (2014). They focus on individual worker decisions, abstracting from firms, and suggest that lower TFP in developing economies raises the cost of accumulating human capital and this lowers households' incentives to invest in human capital after schooling. In this paper, we offer a very different explanation by focusing on firm-provided training, which we show, also empirically, is a key component of adults' human capital

[3] Moreover, through the interaction between employment distribution across firms and training, this paper is related to the misallocation literature, which studies the productivity losses stemming from the large contingent of small unproductive firms in developing countries (e.g. Hsieh and Klenow, 2009; Restuccia and Rogerson, 2013; Bento and Restuccia, 2017; Poschke, 2018). Our paper focuses on documenting a new channel through which these generate productivity losses: lack of on-the-job training.

investments.

Chapter 1

Empirical Evidence on Cross-Country On-the-Job Training Provision

In this chapter, we start by describing the data sources and defining key concepts. Then, we proceed to documenting some facts about on-the-job human capital accumulation and the development process. We include further details in Appendix Section A.1.

1.1 Data Description

To document our cross-country facts, we rely on labor and firm surveys for more than 100 countries. For developing countries, we use the Enterprise Survey (WB-ES). For developed countries, on the other hand, we rely on the European Union Labor Force Survey (EU-LFS), the Adult Education Survey (EU-AES), and the Continuing Vocational Training (EU-CVT) enterprise survey. Our cross-country evidence encompasses developing and developed economies ranging from \$1,000 to \$60,000 of per-capita GDP.

The Enterprise Surveys (ES) from the World Bank are a collection of firm-level surveys of a representative sample of an economy's private manufacturing and service sectors covering approximately 136,000 firms over 140 low- and middle-income countries. The ES usually interviews owners and top managers of the establishments in the sample who can request assistance of accountants or human resources managers to answer certain questions. The ES has a set of country-specific questions according to each country's characteristics and a set of standardized questions that allow cross-country comparison. We rely on the two ES waves, between 2002 - 2005 and 2006 - 2017, for which they have standardized questions on worker training provisions, and use the second wave for the main specifications which provide individual weights.

For the EU enterprise data, we rely on the Continuing Vocational Training Survey (CVT). This survey provides information on enterprises' investment in continuing vocational training of their staff, providing information on participation, time spent, and the costs of such training. In our analysis, due to data availability, we rely on 3 of the 5 waves of CVTS conducted in 2005, 2010, and 2015, as these cover all EU member states and Norway.

For the European countries' worker level-data, we rely on data from the EU-LFS and EU-AES. The EU-LFS is a large household sample survey that provides data on labor participation, unemployment and job characteristics, socio-economic characteristics, and education and training of adults (aged 15+). The survey is conducted on all members of the EU and the 3 European Free Trade Association countries. Although data collection dates back to 1983 for some countries, and the series are generally available from 1992 (according to EU membership), we use time series ranging from 2009 to 2018 for all countries for consistency.

The EU-AES' official objective is to collect information on participation in education and learning activities including job-related training, among other things. Thus, this survey is conducted specially to understand the patterns of adults' education. The AES is one of the main data sources for EU lifelong learning statistics and it covers around 666,000 adults aged 25 - 64. These data were collected during 2007, 2011, and 2017 in 26, 27, and 28 EU Member states, respectively.

Finally, we rely on two secondary United States' data sources for empirical robustness checks, calibration and model validation. First, we provide historical evidence on firm-provided training provision as robustness based on the National Household Education Surveys Program which consist on data on educational activities. Furthermore, to calibrate the model to this country, which we use as benchmark in the quantitative analysis, we rely on the 1995 US Survey of Employer-provided Training (US-SEPT) which was conducted during personal visits to more than 1,000 private establishments.

1.2 Defining On-the-Job Training

We first carefully define training and its characteristics to ensure consistency across different data sources and to be able to provide meaningful economic interpretations through the lens of the model. We present more detailed definitions of training and other sources of human capital for comparison in Appendix Section A.2.

We define *"Training"* following the definition of "Non-formal education and Training" category from ISCED (2011)[1], which is any organized and structured learning activity outside the formal education system. This definition has two main components. First, it differentiates *"training"* from *"schooling"* as it is a learning activity that happens outside the formal education system. Therefore, training does not consist of programs as MBAs that may be a source of human capital for workers. Second, this activity must have certain degree of organization and structure, which differentiates *"training"* from other *"Informal Learning"* activities such as reading journals, visiting museums or learning through media in an unstructured or unplanned way.[2]

[1]The International Standard Classification of Education, adopted by the UNESCO, provides "uniform and internationally agreed definitions to facilitate comparisons of education systems across countries."

[2]*Informal Learning* is defined as a type of a learning activity that is not structured and that is more related to workers' self-investments. Some categories described in the data encompassing this category are: learning by reading printed material or using computers, learning through media (television, radio or videos), learning through guided tours in industrial sites or museums, and visiting learning centers such as libraries. These are self-directed and employers are not usually involved.

We decompose *Training* into *Formal Training* and *Informal Training* to have consistency across data sources. *Formal Training* is defined as training that has a structured and defined curriculum, and includes classroom work, seminars, workshops, among other activities planned in advance. Formal training activities are typically separated from the active workplace and show a high degree of organization by a trainer or institution. Further, this training is typically more general, not geared to specific tasks, machinery, or equipment specific to certain jobs or workers. *Informal Training* is less structured and more related to job-specific skills for workers. It also differs from formal training in that it is tailored to specific workers' needs and is connected to the active workplace. Thus, informal training tends to be more hands-on and task related. It encompasses guided on-the-job training, job rotation, exchanges, and other forms of learning through colleagues and training arising from participation in learning circles.

1.3 On-the-Job Training Facts

The wide variety of data sources allow us to analyze training patterns for more than 100 countries, and to describe in detail the key sources giving rise to adult human capital accumulation. In this section, we document two key facts about firm-provided training. First, we document that on-the-job firm-provided training increases with GDP per capita. We find that this pattern is partially driven by the large share of self-employed workers who do not receive employer-provided training in poor economies, but also we show that the share of firms offering training, the share of participants per firm, the amount of hours per participant and the training cost per participant increases with income. Second, we show that job-related firm-provided training is the main source of adults' education, and that this type of training explains the bulk of the differences in adults' human capital investments across countries.

Fact 1 ***A positive cross-country correlation between firm-provided training and income exists***

We first focus on formal training, which is available from and consistent across enterprise surveys for all 100 countries in our data to study the correlation between on-the-job training and cross-country income. Initial vocational training, employee orientation, apprenticeships and informal training are explicitly excluded. After analyzing the formal training measure, we will focus on broader forms of training for a smaller sample of countries for which data is available. We construct country-year measures of the share of employees who receive formal training with the following formula:

$$\%\textit{Trained Workers} = \frac{\textit{Firms' Trained Workers}}{\textit{All employees in firms}} \times (100 - \textit{Self Emp Share})$$

The WB-ES provides information about which firms provide training, along with the share of workers who were trained in those firms. We use these two measures to construct a country-year measure of the share of employees offered training. Since only firms are surveyed, we then adjust this measure by the share of self-employment for the main specification, assuming that self-employed workers do not receive training from employers.[3] In the EU-CVT, on the other hand, we have data on the share of workers who were trained in all the enterprises surveyed, which we also adjust by self-employment.[4]

Formal on-the-job training increases with development. In Figure 1.1, we show the results of our combined measure of on-the-job training and GDP per capita. We find that as

[3]We also construct a second measure where we adjust our measure of trained workers by the share of workers who are not self-employed and who do not work in agriculture, assuming that workers in the agricultural sector do not receive training. This follows from recent findings suggesting that the returns to experience (hence human capital accumulation) are much lower in agriculture than in the manufacturing and service sectors (Lagakos et al., 2018*b*; Islam et al., 2019). The results do not change significantly when we use either of these two measures (see Appendix Figure A.2)

[4]We restrict the sample from the WB-ES to 2005 - 2015 for comparability with EU-CVT. The WB-ES tends to overweight larger firms, which causes mean firm-based employment to be counterfactually large in some countries. Poschke (2018) shows the log mean employment is lower than 4 even for countries with more than 60,000 USD of GDP per worker for different data sources. Thus, we restrict our sample from the WB-ES to all countries with log mean employment lower than 4 to avoid countries largely overweighting big firms. We show the same pattern with the unrestricted sample in Appendix Figure A.1.

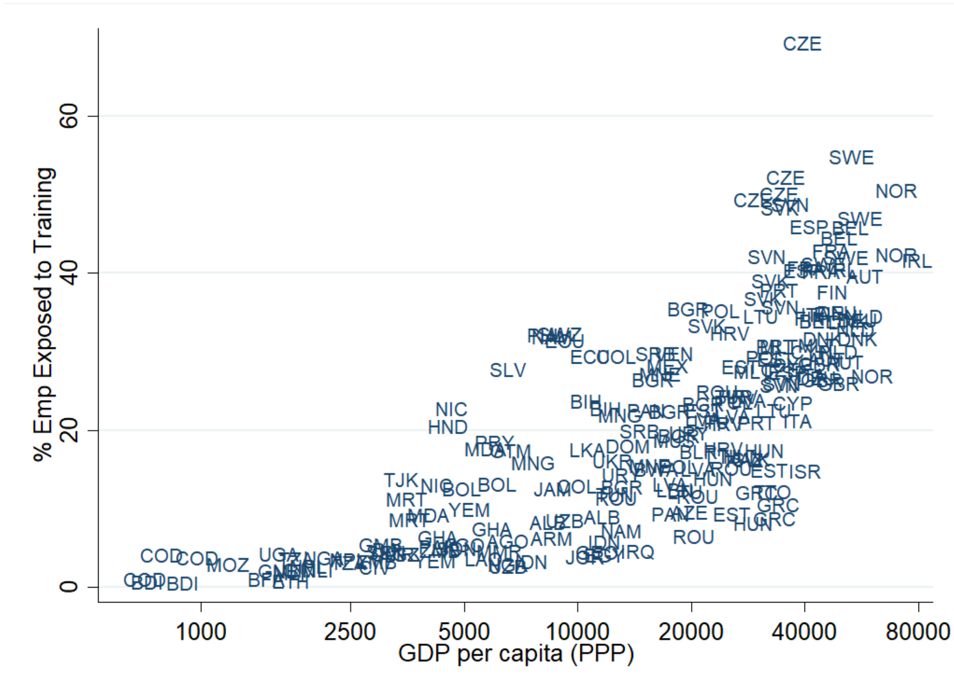
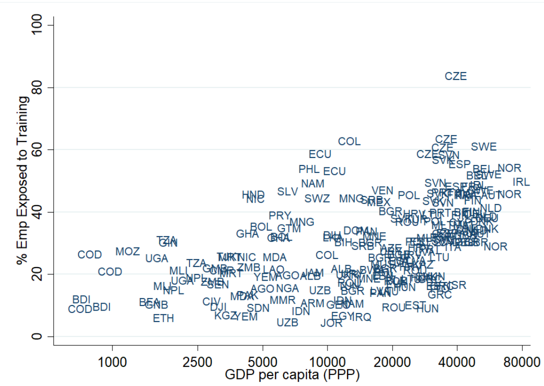


Figure 1.1: Share of Employment Formally Trained and Development

Note: The share of formally trained employment follows from adjusting the share of workers who were trained by firms by the share of self employment. Data on the share of employees trained inside firms comes from the World Bank Enterprise Survey (WB-ES) for all developing economies and from EU Continuing Vocational Training Survey (EU-CVT) for European economies. Both surveys contain data on whether firms provided formal training in the previous fiscal year, and the share of employees who participated. For the WB-ES we use the standardized wave with data from 2005 - 2017 for which we have firm weights. We restrict the sample from the WB-ES for the years between 2005 - 2015 to have the same years as the EU-CVT, and we restrict to countries with mean log employment in firms lower than 4. Data on GDP per capita and self-employment comes from the World Bank Indicators.

countries become more developed, on-the-job training increases substantially. In particular, for the poorest countries in our sample, with a per-capita GDP of about $1,000, only approximately 5% of employees are exposed to training. In contrast, this share rises to approximately 50% for the richest countries, with great variation in between. It is also noteworthy that the data from the WB-ES and the EU-CVT overlap for the income range common to both, denoting both harmony between the training definitions, and a consistent pattern between training and income in the two data sources.

Self-employment is a key mechanism driving low training in poor economies. We now show that the large share of self-employment prevalent in developing countries is key to explaining the low levels of on-the-job training in these settings. In panel (a) of Figure 1.2, we show that the share of workers who are offered training rises with income even when unadjusted for self-employment. However, the difference between poor and rich economies is much more compressed in this case, suggesting that the high share of self-employment exhibited in poor countries, and the strong correlation of this with income – evidenced in panel (b) – are a key factor driving low training in poor economies.



(a) % Trained workers in Firms

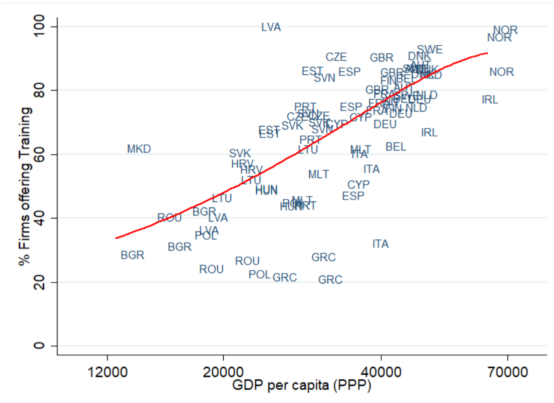
(b) Self Employment

Figure 1.2: Unadjusted On-the-Job Training Shares and Self-Employment

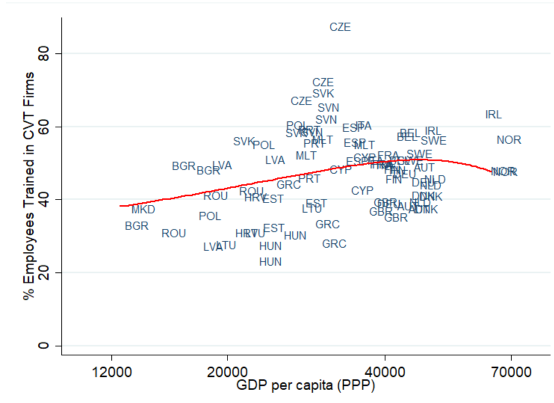
Note: This figure shows both margins from the formal training measure: in panel A we show workers who are trained by employer as a share of total workers in firms and panel B shows the share of workers who are self-employed. Data on the share of trained workers in the wage sector comes from the share of workers formally trained in the World Bank Enterprise Survey and the European Union Continuing Vocational Training Survey, which focuses on formal training and does not include initial vocational training or specific-worker targeted training. Self-employment data comes from the World Bank Indicators, which provide ILO estimates for each country-year.

Wage-sector training increases with development in every margin. We can analyze the relationship between training and income within the wage sector using enterprise survey data from European countries. Although we rely on fewer countries, the survey time frame allows us to have large income variation. We find that this positive correlation is prevalent among both the extensive and intensive margins. In Figure 1.3, we show that richer countries exhibit both a larger share of firms offering training (extensive margin), along with a larger share of trainees inside firms offering training, and a larger share of hours in training relative to total hours worked (intensive margin).[5] In addition, richer countries exhibit a larger cost of training, which potentially

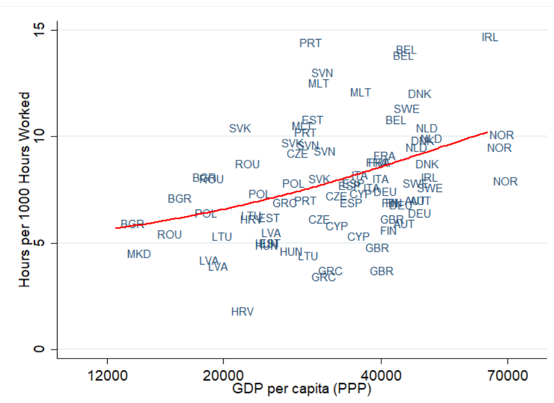
[5]In Appendix Figure A.4 we show that the share of hours in training relative to total hours worked increases with income, but total hours of training per participant remains constant. This is consistent with workers working fewer hours as income increases (Bick, Fuchs-Schuendeln and Lagakos, 2018). In Figure A.4 we plot the time spent in training for the EU-CVT and the EU-AES.



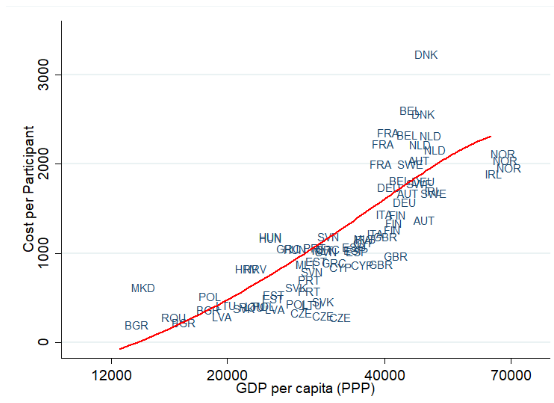
(a) % Firms Offering Training



(b) % of Participants within firm



(c) Hours per 1000 Hrs Worked



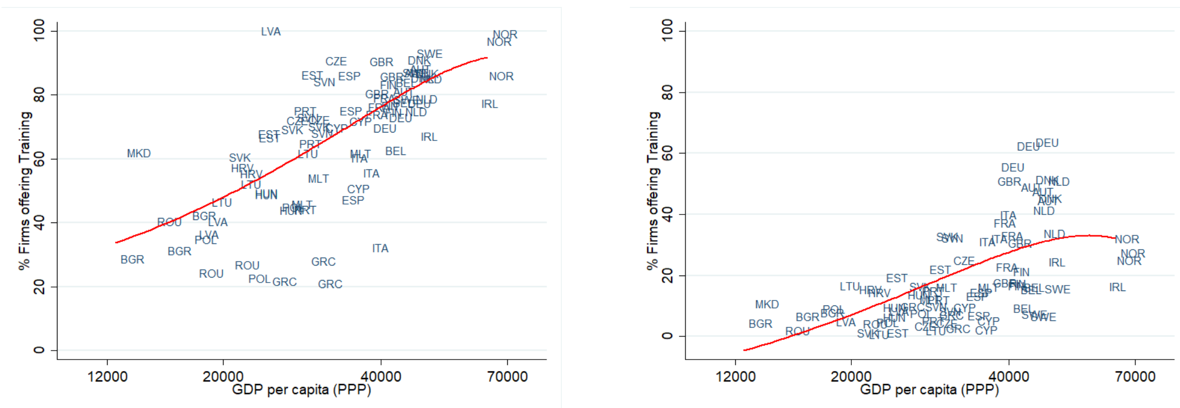
(d) Cost per Participant

Figure 1.3: OTJ Training Margins within the Wage Sector

Note: This figure shows all margins of training coming from the EU Continuing Vocational Survey. Panel a shows the share of firms that offered training, which is defined by firms that offer any type of continuing vocational training in the previous fiscal year. Panel b shows the share of participants within the firms who participated in training conditional on the firm offering training at all. Panel c shows the hours per 1000 hours worked by all employees in the firms (participants and non-participants of the training). Panel d shows the total cost of training per participant that includes direct and indirect training costs (wages of trainers and wage lost by not working during training).

proxies training quality.

Continuing and initial vocational training increases with development. Our formal training measure is based on continuing vocational training (it does not include worker orientation or initial training), which seems the most relevant margin to explain life cycle increases in productivity. Nevertheless, it could be the case that continuing vocational training is larger in developed economies, but the initial vocational training (IVT), that takes place when the worker starts the job, is larger in developing economies. Although we do not have measures of the share of workers who receive initial vocational training, we do have measures on the share of firms offering IVT and CVT, which are depicted in Figure 1.4. In both cases, there is a positive correlation with development. As countries become richer, firms invest more in both IVT and CVT, which rules out that there is a difference in the timing of human capital investment across countries. Interestingly, Germany has very high levels of IVT, which aligns with previous studies that analyzed German training programs.



(a) Firms offering CVT

(b) Firms offering IVT

Figure 1.4: Share of firms Offering CVT and IVT

Note: These data come from the EU Continuing Vocational Training. Panel a shows the share of firms that offer continuing vocational training (CVT) and panel b shows the share of firms offering initial vocational training (IVT). CVT is defined as all training for workers except for the initial training to show workers job-specific skills for the new job or to teach workers general knowledge about the firm as they enter to a new job, which is included on the initial vocational training. Data on GDP per capita comes from the World Bank Indicators.

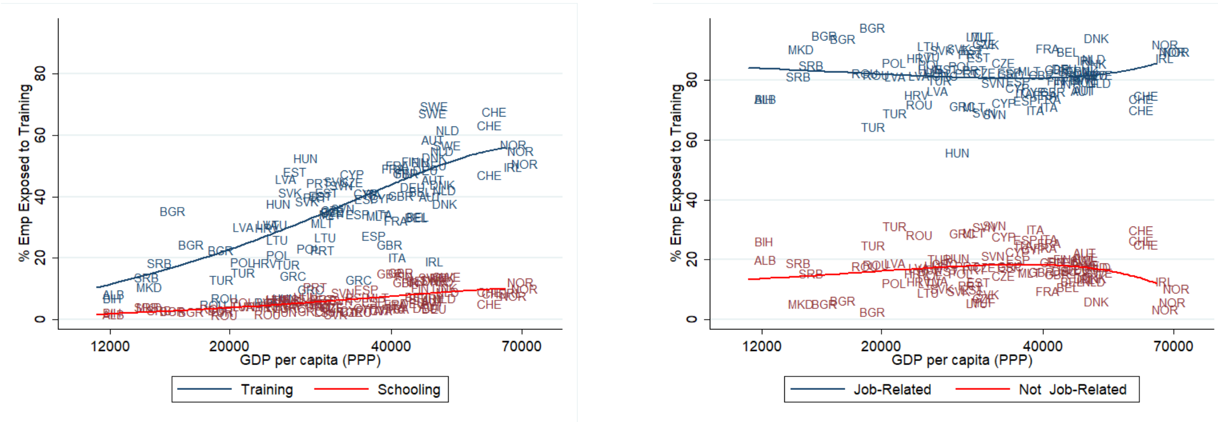
Fact 2 ***Firm-provided training is the main source of adults' education.***

Until this point, we have shown a strong correlation between on-the-job training and development using enterprise-level data. However, if on-the-job training is a small fraction of adults' human capital investments, this positive correlation will not be useful to explain cross-country human capital differences. Thus, we now turn our attention to labor force and worker surveys containing detailed information on workers' training activities and education, which allow us to quantify the role of on-the-job training relative to other human capital sources. In particular, we focus on data from the European Union Adult Education Survey and the Labor Force Survey, which contains information on the characteristics of all education and training investments in European countries. We document that the most important source of human capital investment for adults is job-related training financed by firms.[6]

Most of adults' education is job-related training. In Figure 1.5 we show how the proportion of workers exposed to different types of education varies with cross-country income. Panel (a) shows that the vast majority of adult education (around 90% of all adult education reported in the last year) is training, while less than 10% is schooling. Additionally, panel (b) shows that around 80% of workers who reported participating in some type of training (blue measures in panel a) claimed that this was job-related, and interestingly this share is uncorrelated with cross-country income. Moreover, Appendix Table A.1 shows the same pattern when looking at the share of adults who reported being involved in training in the labor force survey (EU-LFS). On average 84% of adults in European countries reported that the education they received had been job-related and only 16% mentioned personal or social reasons as the purpose of their training or education. This evidence suggests that job-related training is a primary source of adult learning and human capital accumulation.

Almost all of the job-related training is sponsored at least partially by firms. Figure 1.6 shows how the proportion of job-related training, financed at least partially by the firm or completely by the worker, varies across European countries. The graph shows that the vast majority of job-related training is sponsored by firms. In particular, less than 5% of workers for all countries receive some training that is directly related to their job and that is entirely self-financed. Moreover, the share of adults who fully self-financed their job-related education is constant as a function of per-capita GDP, which reflects the fact that the increase in job-related training with income is driven by firms offering more training, and not by workers themselves

[6]In Appendix Table A.1 and Table A.2, we further show that on-the-job training predominantly occurs during working hours, and its objective is to improve technical and job specific skills. Table A.1 shows that, on average, workers in the European Union were reportedly trained during paid working hours for 70% of the training length. In Table A.2, we first show the purpose of training reported by firms for the EU-CVT surveys in 2010 and 2015. These purposes range from general IT, management, and team working to problem solving and improve technical skills. Between 10% to 40% of firms reported having done some training on most of the categories, but the category with the highest level of reporting (around 70% of firms had training with this purpose) was technical and job-specific skills, which suggests a productivity-enhancing nature.



(a) Type of Education

(b) Job Relation of Training

Figure 1.5: Adult Human Capital Accumulation Characteristics

Note: These data comes from the EU Adult Education Survey. Panel (a) shows the difference in share of adults who participated in any type of educational activity. "Training" refers to our definition of informal + formal training, or the category of education defined by "Non-formal education and Training" from the International Standard Classification of Education 2011 (ISCED 2011) while "schooling" refers to "Formal education and training" according to the International Standard Classification of Education 2011 (ISCED 2011). Panel (b) refers to the share of training that was job related from all the training reported in panel (a) (blue line). Data on GDP per capita comes from the World Bank Indicators.

investing more in education. Doing some back-envelope accounting, our results show that 90% of all human capital investment is training, 80% of all training is job-related, and that almost 100% of this job-related training is financed by firms. This means around 72% of all human capital investments is at least partially provided and financed by firms. This striking result indicates that the most important source of adult human capital accumulation corresponds to firm investments in human capital, and not workers' self-investment outside the firm. Moreover, as robustness in Appendix Figure A.6, we provide United States' historical evidence from 1991 to 2005 and show job-related training provision accompanies economic growth in the time series as well.

These patterns imply that employer-provided training is a key determinant of on-the-job human capital accumulation, and that firms play a substantial role in adult human capital investments. This suggests that canonical human capital accumulation models *à la* Ben-Porath, which do not include firm-level decisions, provide an incomplete picture of workers' human capital accumulation after formal education or schooling concludes. This, in addition to our first fact showing that on-the-job training increases with development, suggests that understanding

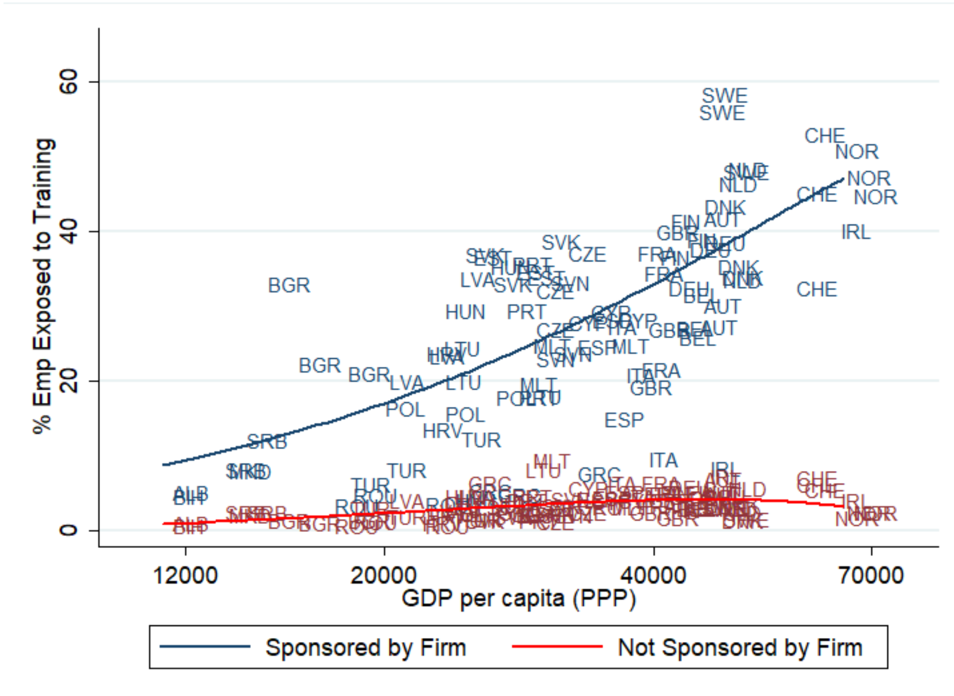


Figure 1.6: Training Financing

Note: The data on training financing come from the EU Adult Education Survey. The graph shows the difference in share of adults who participated in sponsored and non-sponsored job-related training relative to GDP per capita. Data on GDP per capita come from the World Bank Indicators.

firms' decisions to provide training is key to understand cross-country human capital and income differences.

1.4 Acknowledgements

Chapter 1 is part of the paper "Human Capital Accumulation and Development: The Role of On-the-Job Training" coauthored with Xiao Ma and Daniela Vidart. As of May 2021, this paper is being prepared for submission for publication. The dissertation author was a primary investigator of this material.

Chapter 2

Modeling Cross-Country Differences in Experience Wage Profiles and Human Capital Accumulation

2.1 Model

To explain the positive correlation between training and development and its link to firm-level decision-making, we focus on two main themes: self-employment, and job turnover. The self-employment theme is motivated by our empirical evidence, and particularly by the fact that the high prevalence of this type of work is a key mechanism driving low training in poor economies. The job turnover theme is rooted in the literature, specifically on the fundamental problem of financing training investments first identified by Becker (1964). Training investments are less likely to occur if the probability of losing the worker is high. We focus on two factors affecting turnover, which are especially salient and have been widely studied in the cross-country context: differences in contract quality, and labor market frictions. As Acemoglu, Johnson and Robinson (2005) stress *"Economic institutions matter for economic growth because they shape* *the incentives of key economic actors in society, in particular, they influence investments in physical and human capital and technology, and the organization of production..."*

Environment. The model economy is populated by a continuum of workers whose life spans 2 periods. Every period, the same number of workers who die are born, and we normalize the size of each generation's population to be 1. All workers are born ex-ante equal, but accumulate human capital through training at potentially different rates. Workers offer 1 unit of labor inelastically to the market every period. Their utility is assumed to be linear in consumption, and thus, they maximize the present value of consumption:

$$\max_{c^Y, c^O} c^Y + \frac{c^O}{1+\rho}$$

where superscripts Y and O denote young and old ages, and $\rho > 0$ governs time preference. In the steady state, $\rho = r$, therefore workers are indifferent between consuming in each period.

Consumption Good Production and Self-employment. The consumption good is a composite of goods from two different sectors: the traditional sector good Y_T and the modern sector good Y_M,

$$Y = (\gamma Y_T^{\sigma} + (1-\gamma) Y_M^{\sigma})^{\frac{1}{\sigma}}$$

Production in the traditional sector is characterized by a constant-returns-to-scale function:

$$Y_T = A_T N_T$$

where A_T is productivity and N_T is labor in that sector. This sector is related to self-employment and we assume training is not provided to workers. In our model, we normalize the price of the good produced by the modern sector to be 1. Therefore, the price of the good produced by the traditional sector is:

$$P_T = \frac{\gamma}{1-\gamma}\left(\frac{Y_T}{Y_M}\right)^{\sigma-1}.$$

Modern Sector. This sector is characterized by frictional labor markets. There is a unit measure of firms, which are heterogeneous in productivity $z \sim H(z)$ and produce a homogeneous good. Once workers and firms are matched, worker i's production in firm j is given by:

$$y_{i,j} = A_M z_j h_i.$$

where A_M is productivity in this sector, z_j is the firm-specific productivity and h_i is worker i's efficiency units of labor (human capital). This production function suggests human capital and technology are complementary.[1] By integrating this expression across all workers within firm j, we get total production of firm j:

$$y_j = \int_{i \varepsilon j} A_M z_j h_i di,$$

and aggregating production over all firms, we get production in the modern sector:

$$Y_M = A_M \int_{j \varepsilon J} z_j \int_{i \varepsilon j} h_i \, di \, dj.$$

Job Search and Matching. Firms post vacancies $v(z)$ at the start of each period, with a contract stipulating the wage rate $w(z)$ and working period — which we assume to be two periods for young workers and one period for old workers. The vacancies cost is defined by $c_v \frac{v^{1+\gamma_v}}{1+\gamma_v}$ and we require vacancy costs to be convex (i.e. $\gamma_v > 0$) such that firms with different productivity levels coexist. The total number of vacancies is then $V = \int v(z) dH(z)$.

There is a probability δ of exogenous destruction of workers' contracts in the beginning of the second period when they become old. These exogenously separated old workers enter the unemployment pool and look for a job full time jointly with all newborn workers. Moreover, a portion η of remaining old workers search on the job. Therefore, the amount of searchers is denoted by $\tilde{U} = (1 + \eta(1-\delta) + \delta)N_M$, where N_M is the share of each generation's workers in the

[1] This assumption is consistent with studies finding complementarities between technology and human capital (Acemoglu and Zilibotti, 2001; Porzio, 2017)

modern sector.

For analytical tractability, we assume the matching function is $M(\tilde{U},V) = \min\{\tilde{U},V\}$, and that c_v is small enough such that $V > \tilde{U}$, which ensures full employment in the equilibrium. As usual market tightness is defined by $\theta = \frac{V}{\tilde{U}}$.

Contract Enforceability and Worker's Optimal Separation Policy. If the remaining old workers who search while on the job get an outside offer, they can make efforts to break the contract with probability p, by incurring the costs of breaking the contract $c_p^{\gamma_p} \frac{p^{1+\gamma_p}}{1+\gamma_p}$ per efficiency unit. The costs of breaking the contract represent costs of being caught and filing lawsuits, with a lower c_p representing weaker institutional environments. We assume $\gamma_p > 0$ such that the marginal cost of breaking the contract increases with probability p.[2]

We first solve workers' optimal choice of leaving probability taking the level of training investment as given for two reasons. First, when the new offer arrives at the end of the period, training already occurred. Second, firms and workers need to internalize workers' probability of leaving the firm to decide the optimal level of training. Thus, they must choose training according to the optimal breaking contract efforts conditional on each new offer. In a firm with productivity z, a worker chooses the optimal leaving probability $p \in [0,1]$ when faced with an outside offer w' solving:

$$\max_{p \in [0,1]} (w' - w(z))p - c_p^{\gamma_p} \frac{p^{1+\gamma_p}}{1+\gamma_p}$$

We solve for $p(w(z), w')$ which is a piece-wise function,

$$p(w(z), w') = \begin{cases} 0 & \text{if } w' < w(z) \\ \frac{1}{c_p} (w' - w(z))^{\frac{1}{\gamma_p}} & \text{if } 0 < w' - w(z) < c_p^{\gamma_p} \\ 1 & \text{if } w' - w(z) > c_p^{\gamma_p} \end{cases}$$

[2]Consistent with previous literature (Acemoglu and Pischke, 1999), in our setting firms cannot break the contract as they pay for their share of training cost in the first period. Note that if firms have long-term reputation, the absence of contractual problems from firms breaking contracts is reasonable.

This result is intuitive. If the new wage offer is lower than the wage at the current firm, workers do not want to switch and the investment in breaking the contract is zero. On the other hand, if the new wage is large enough ($w' > w(z) + c_p^{\gamma_p}$), workers want to switch firms, and therefore, they will break the contract with a probability of 1. If the cost of breaking the contract increases, workers are less willing to switch and thus they invest fewer resources into breaking the contract. As expected, the worker's probability of leaving increases with the outside option w' and decreases with current wages $w(z)$.

Training Determination. A young worker has initial human capital $h^Y = 1$ (normalization) and can be trained for s efficiency units of time to enjoy an increase in the next-period efficiency units of labor:

$$h^O = h^Y + \zeta s^{\gamma_s}$$

where ζ is a constant, and $0 < \gamma_s < 1$ governs the diminishing returns of training. In each period, training is decided jointly by firms and workers and the cost of training is paid jointly by them when training occurs. There is a constant cost c_s per unit time of training, reflecting trainers' wages and material costs. In principle, training also reduces trainees' production time. As analytical properties will not be affected by training time costs, we omit them here, but we add them in the quantitative analysis as this is a key feature of the data.

It is worth noting that we assume all training raises general human capital, so the benefits accrue even if the worker changes firms. Moreover, we assume that if s_W and s_F are optimal training levels from workers' and firms' perspective, respectively, training $s = \min\{s_W, s_F\}$. This assumption says the training level is determined by either party with lower affordability. It is a reasonable assumption: for instance, if firms bear all the training costs, workers may desire large training levels, yet firms would not like to pay for them. Thus, the optimal level of training for workers and firms is determined by Proposition 1:

Proposition 1 (Firms' and Workers' Optimal Training Levels) *In a firm with productivity* *level z, if μ_i is the proportion of training costs borne by group i (workers or firms) then:*

$$s_i(z) = \left(\frac{\zeta \gamma_s MR_i(z)}{(1+r)\mu_i c_s} \right)^{\frac{1}{1-\gamma_s}},$$

where, in a firm with productivity z, current wage w, new offers of wage w', a wage distribution from offers $F(w)$, and optimal investments to break contract $p(w,w')$ (denoted by p(w')), the marginal benefits of training for workers and firms are:

$$MR_W(z) = (1-\delta)\left(\underbrace{\left(1-\eta\int p(w')dF(w')\right)w}_{\textit{if stay in current firm}} + \underbrace{\eta\int p(w')w'dF(w')}_{\textit{if move to new firm}} - \underbrace{\eta\int c_p^{\gamma_p}\frac{p(w')^{1+\gamma_p}}{1+\gamma_p}dF(w')}_{\textit{cost of breaking contract}} \right) + \underbrace{\delta\int w'dF(w')}_{\textit{U back to a firm}}$$

$$MR_F(z) = \underbrace{(1-\delta)\left(1-\eta\int p(w')dF(w')\right)(A_M z - w)}_{\textit{future profits, from staying workers}}.$$

Proposition 1 explains how optimal training is decided with different divisions of training costs. As the share each group pays increases, the optimal level of training decreases. Moreover, taking the share of costs paid as given, workers' training levels depend on the expected wage flows if they stay in the firm or switch employers. On the other hand, firms choose the optimal level of training to maximize their net profits, which increase with firms' productivity and the probability of keeping the worker. One key difference between workers and firms is that firms cannot reap the gains from training after the trained worker leaves.

In this model, firms are willing to invest in general training. This departure from Becker (1964) is due to frictional labor markets, because of which firms are able to extract partial rents from training (Acemoglu and Pischke, 1999). We differ from the general training literature (e.g. Acemoglu and Pischke, 1999; Engbom, 2020) in that we assume the cost shares workers and firms pay are common across firms and we add a time cost of training when we take the model to the data. In Appendix Section B.8, we show the model's results for three different specifications on how the training costs are financed: firms paying all training costs, joint internal efficiency where workers and firms distribute the costs efficiently (they pay the share proportional to their income gains), and the one we presented in our paper. The evidence suggests that if we include training time costs, the common share is the right assumption as it is the only one that can jointly match the training pattern with respect to firm productivity and the aggregate training levels.[3]

Moreover, in Appendix Section B.8, we show the training patters in the aggregate and the optimal training decisions for firms and workers for the calibration that matches all labor market moments. We find that firm decisions define training investments as firms always want lower levels of training than workers. Thus, we now focus in understanding firm-level decisions.

Proposition 2 (Labor Market Frictions and Firms' Training) *In a firm with productivity level z, firms' optimal training $s_F(z)$:*

(1) increases with costs of breaking the contract c_p;

(2) decreases with on-the-job search probability η;

(3) decreases with exogenous separation rate δ.

Results in Proposition 2 indicate that a higher probability of job separation leads to lower training. This result provides the mechanisms for which better institutions and lower job destruction generates more training investments in our model.

Moreover, a wage increase has two opposing forces affecting training decisions. On the one hand, the incentives to invest in training decrease because firms capture a lower share of the surplus of the match but, on the other hand, the probability of keeping the worker increases, which generates higher training incentives. Interestingly, a wage increase in the aggregate economy do not impact the probability of keeping workers but the labor shares do decrease. In this case, training investments go down, which means that higher firm competition for workers translates to lower training investments.

[3]Training levels increase with firm productivity when the firm pays all the cost or when the share of the cost that the firm pays is constant. Nonetheless, the case of joint internal efficiency generates a decrease in training investments with firm productivity, which is counterfactual, as in the data, firms that are bigger and more productive invest more in training. Moreover, the model needs really implausible high levels of training productivity to match the levels of training observed in the data when firms pay the total cost of training. This suggests that the common share is the right assumption as it is the only one that can jointly match the training pattern with firm productivity and the aggregate levels of training.

Solving Firms' Problem. Firms choose wage $w(z)$, vacancies $v(z)$, and young workers' training $s(z)$ each period to maximize profits. Their value function can be written as:

$$
\begin{aligned}
J(z,l_{-1}^{O},w_{-1},X_{-1}) = \max_{\{w,v,s\}} & \underbrace{l_{-1}^{O}(1-\delta)\left(1-\eta\int p(w_{-1},w')dF(w')\right)(A_M z-w_{-1})}_{\text{profits from remaining workers}} \\
& +\underbrace{\frac{v}{\theta}\frac{1}{1+\eta(1-\delta)+\delta}(A_M z-w-\mu c_s s)}_{\text{profits from hiring young workers}}+\underbrace{\frac{v}{\theta}\frac{\eta(1-\delta)\int p(w',w)dF_{-1}(w')\bar{l}(w)+\delta\bar{l}}{1+\eta(1-\delta)+\delta}(A_M z-w)}_{\text{profits from hiring old workers}} \\
& -\underbrace{\frac{c_v v^{1+\gamma_v}}{1+\gamma_v}}_{\text{vacancy costs}}+\frac{J(z,l^{O},w,X)}{1+r} \\
& s.t. \quad l^{O}=\frac{v}{\theta}\frac{1}{1+\eta(1-\delta)+\delta}(1+\zeta s^{\gamma_s}), \quad X=\Gamma(X_{-1}), \quad w\geq b\bar{w}
\end{aligned}
$$

where we use the subscript -1 to denote pre-determined variables. l^O is the total supply of efficiency units by old workers before exogenous separations. $F_{-1}(w)$ is the wage distribution of job offers during the last period. The first term represents the net profits generated by all the workers who remain in the firm from the last period. The second term represents the profits (net of training costs) from hiring young workers. The third term represents the profits from poaching old workers who are willing to move to the current firm. The on-the-job movers have average efficiency units $\bar{l}(w)=1+\frac{\int \zeta p(w_{-1}(z),w)s_{-1}(z)^{\gamma_s}dF_{-1}(w_{-1}(z))}{\int p(w_{-1}(z),w)dF_{-1}(w_{-1}(z))}$, whereas the average efficiency units of unemployed old workers are $\bar{l}=1+\int \zeta s_{-1}(z)^{\gamma_s}dF_{-1}(w_{-1}(z))$. Note that $s(z)$ is determined according to Proposition 1, whereas $w(z)$ and $v(z)$ are determined according to FOCs. In particular, as shown by Burdett and Mortensen (1998), $w(z)$ is determined by a first-order differential equation, combined with the minimum wage $b\bar{w}$.

Intuitively, firms have incentives to increase wage offers to poach workers from other firms and to keep their own workers from being poached. Nevertheless, higher wages generate a higher labor share, which decreases profits. Thus, the wage distribution is determined by these two offsetting forces. Because hiring workers generates profits, firms want to post vacancies but will stop posting at some point, as costs of additional vacancies are increasing.

Equilibrium. In the steady-state equilibrium, labor and goods markets are clear for both traditional and modern sectors for each period. We abstract from workers' reshuffling between modern and traditional sectors, which needs tracking employment and training histories of each worker and is computationally intractable. Several assumptions can lead to immobility between sectors, including complete asset market, family networks, or large switching costs. Therefore, workers must also be indifferent in terms of expected utility between going to the traditional sector and the modern sector in the first period:

$$P_T A_T + \frac{P_T A_T}{1+\rho} = \int_z \left(w(z) - \mu_W c_s s(z) + \frac{1+\zeta \mu s(z)^{\gamma_s}}{1+\rho} MR_W(z) \right) dF(w(z))$$

The left-hand side of the equation represents the present discounted value of working as self-employed while the right-hand side shows the expected discounted labor income from working in the wage-sector for both periods. Finally, output per capita is defined as:

$$\frac{Y_i}{N_i} = A_T \left[\gamma (\Lambda_T)^{\sigma} + (1-\gamma) \left(\frac{A}{A_T} \Lambda_M \int z \left(1 + \frac{\zeta s(z)^{\gamma_s}}{2} \right) dF(z) \right)^{\sigma} \right]^{\frac{1}{\sigma}}$$

where Λ_i is the share of workers in sector i. Thus, output per capita depends on how workers are distributed between sectors and firms, sector-specific productivities and training investments in the economy.

Self-Employment Shares and Training. Proposition 1 and 2 provide the mechanisms for which better institutions and lower job destruction generate more training investments in our model. We now focus on how training is affected due to changes in self-employment shares. Every change that lowers the returns of working in the wage sector relative to the traditional sector generates a higher share of workers allocated in the self-employment sector and a decrease in training in the aggregate, conditional on the wage-sector investments. For instance, if δ increases, the expected return of working in the wage-sector decreases because it is more difficult for workers to move up the job ladder. This increases the economy's self-employment share and pushes aggregate human capital downward.

2.2 Quantitative Model Additions

In this section, we extend our two-period analytical model for quantitative analysis and take our model to the data. Thus, we add some features to closely replicate key aspects of the labor market and economic environment:

Workers. On the worker side, we assume workers live for T periods, since job search models are usually calibrated using high-frequency labor flows data. With this change in mind, an age-a worker's utility function can be written as:

$$\max_{\{c_\tau\}} \sum_{\tau \geq a}^{T} \left(\frac{1}{1+\rho}\right)^{\tau-a} c_\tau$$

We still normalize each generation's population to be 1, hence the total population in the economy is T. We assume human capital depreciates at rate d, and therefore evolves as $h' = (1-d)h + \zeta s^{\gamma_s}$. As in our analytical model, we abstract from workers' reshuffling between modern and traditional sectors, and thus, in the equilibrium, workers are indifferent between working in the wage sector and as self-employed in the first period.

Firms. Training costs are assumed to be proportional to the average wage $c_s \bar{w}$, while training per unit time also causes a δ_s decrease in efficiency units for production. With this assumption, firms' and aggregate training levels remain constant in response to proportional changes of A_T and A_M. As we aim to understand why developed countries invest more in training, we explicitly avoid capturing the training difference driven by the TFP level. With workers of different ages, firms can impose different training levels on different individuals. In particular, training intensities decrease with age.

Exogenous Job-to-Job Moves. We assume that the moving probability p has some lower bound $\underline{p} > 0$. This aims to capture that a portion of job-to-job flows are associated with wage losses (see Table 4 in Haltiwanger et al. (2018)). The economic intuition is that some job-to-job moves reflect idiosyncratic shocks related to family, health, or geographic reasons.

Labor Market. In the quantitative model, we use a widely-estimated matching function $M(\tilde{U}, V) = c_M \tilde{U}^{\Psi} V^{1-\Psi}$ for the modern sector. This matching function produces unemployment and meaningful elasticities of matches with regard to searchers $\tilde{U}$ and vacancies V. With $\theta = \frac{V}{\tilde{U}}$ denoted as market tightness as usual, we denote $q(\theta) = \frac{M}{V}$ as the vacancy filling rate and $\frac{M}{\tilde{U}} = q(\theta)\theta$ as the job finding rate.

Conditions for Simulations. To save space, we delegate the optimal conditions for the quantitative model to the Appendix Section B.1, as they provide the same intuition as in our analytical model. We show the optimal levels of training depending on firm productivity and age of workers, and the conditions for wages and vacancies for firms.

2.3 Model Parametrization

We proceed to calibrate the model in two steps. First, we calibrate the model to the United States as our baseline economy. For this, we draw on 12 moments describing labor market dynamics and training investments, which allow us to identify model parameters. Then, we perform a second calibration for a representative economy at 10 different income levels to understand how training investments change with development. For this purpose, we jointly recalibrate the parameters δ, c_p, A_M, and A_T to match self-employment, job destruction rate, job-to-job transition, and income levels for each representative economy.

2.3.1 Calibrating the Model to the United States

Preassigned parameters. We first directly set some parameters following the literature. We calibrate a model of quarterly frequency and set the quarterly discount rate ρ to 0.01 such that the annualized interest rate is 0.04. Each individual works for 40 years, and therefore, the whole lifetime is set to $T = 160$ quarters. The ratio of the lowest wage to the average wage is calibrated to be $b = 0.5$ following Hornstein, Krusell and Violante (2011) who compute the mean-min ratio of wages to be around 2 from the U.S. Census data. We choose the elasticity in the matching function to be $\psi = 0.7$, as estimated by Shimer (2005). We use $\frac{1}{1-\sigma} = 3$ for elasticity of substitution between the traditional and modern sectors in the aggregate production function as in Feng, Lagakos and Rauch (2018). Finally, we set the relative intensity of on-the-job search intensity to be 0.4 following Faberman et al. (2017) who found that the average amount of offers per month for employed people is around 40% of that for unemployed people in the U.S. data.

We calibrate two other parameters using other countries' data given that there is no estimate for the United States. First, to generate nontrivial wage dispersion, we need firms' hiring costs to be convex in the amount of vacancies. There are relatively few estimates on the convexity in vacancy costs γ_v. Dix-Carneiro et al. (2019) find γ_v ranges from 0.8 to 2.3 for Brazilian firms, whereas Blatter et al. (2016) found a relatively low convexity value (0.2) for Swiss firms. We use $\gamma_v = 1$ in our baseline calibration. Second, we calibrate losses in production hours per unit of training time to be $\delta_s = 0.7$, by taking the average from European countries' labor force surveys. Finally, we normalize the United States' aggregate productivity to be $A_M = A_T$ and unity.

Table 2.1: Pre-Assigned Parameters

Parameter	Model	Source
ρ - Discount rate	0.01	Annualized interest rate of 0.04
T - Number of periods	160	40 years of work
η - On-the-job search intensity	0.4	Faberman et al. (2017)
b - Ratio of lowest wage to average wage	0.5	Hornstein, Krusell and Violante (2011)
ψ - Elasticity of matching function to searchers	0.7	Shimer (2005)
$\frac{1}{1-\sigma}$ - Elasticity of substitution	3	Feng, Lagakos and Rauch (2018)
γ_v - Convexity of vacancy costs	1	Dix-Carneiro et al. (2019) 0.8-2.3
δ_s - losses in production hours per unit time	0.7	EU-LFS 2004 Training Module
A_T, A_M - Productivity in M and T sectors	1	Normalization

Parameters to estimate. The remaining parameters to estimate are: the constant in the matching function, c_M; training costs per time as a share of the average wage rate, c_s; the constant in vacancy costs, c_v; the constant in the function of leaving costs, c_p; the constant in training returns, ζ; the convexity in training returns, γ_s; the traditional-sector share in the aggregate production function, γ; the convexity in the function of leaving costs, γ_p; the shape parameter of Pareto productivity distribution, κ; exogenous separation rates, δ; lower bound of leaving probability, $\underline{p}$; and the share of training costs paid by firms, μ_F.

Targeted moments and fit. To calibrate those remaining parameters, we target the following moments listed in Table 2.2: average unemployment rates in 1994 - 2007, as computed by Hornstein, Krusell and Violante (2011); the ratio of the amount of vacancies to the amount of unemployed people, from FRED for 2000 - 2007 (data available after 2000); the share of self-employment in total employment for 1994 - 2007, from the World Bank; the Pareto parameter of firm employment distribution, as estimated by Axtell (2001); workers' average wage growth after job-to-job transitions, as computed by Haltiwanger et al. (2018); the ratio of training time in firms with 100-499 employees to that of firms with 50 - 99 employees; and the ratio of training costs to wage costs of training. We compute the last two moments using the 1995 Survey of Employer-Provided Training implemented by BLS, which has both employers' and employees' information. We add the percent wage growth of 20 and 40 years' experience, as estimated by Lagakos et al. (2018*b*), to calibrate training returns. Finally, we add three more moments—job-to-job and job-to-unemployment probabilities and training intensity—which we explain next.

For job transition dynamics, we rely on two moments—the share of employed people remaining in the same firm and the share of employed people remaining employed after a quarter—provided by Donovan, Lu and Schoellman (2020) for 40 countries of different development levels. As we are interested in how institutional quality affects job transition at different income levels, we first regress probabilities for countries available in Donovan, Lu and Schoellman (2020) on all institutional measures from the World Bank Worldwide Governance Indicators and then predict the probabilities for all countries using the coefficients. Finally, we regress these probabilities on GDP per capita to have the average probabilities at different income levels.[4] We use the predicted values for the United States in our baseline calibration. Although these predicted values are a little higher than actual U.S. values, we choose to use the predicted values to be consistent with our calibration in the second step for representative economies at different income levels. Appendix Section B.2 discusses alternative ways of computing job transition probabilities and compares the results.

Table 2.2: Moments in the Model vs Data

Moments	Data	Model
Panel A: Targeted Moments		
1. Moments: labor market		
1.1 Unemployment rate (%)	6.5	6.4
1.2 Ratio of #Vacancies to #Unemployed	0.55	0.61
1.3 Traditional sector employment share (%)	7.0	6.6
1.4 Pareto parameter of firm size distribution	1.06	1.1
1.5 % workers remaining in same firm after one quarter	0.94	0.94
1.6 % workers remaining employed after one quarter	0.97	0.97
1.7 Workers' av wage growth after job-to-job transition	0.13	0.14
1.8 % job-to-job transition from high to low wage firms	0.22	0.24
2. Moments: training intensity and value		
2.1 Average training intensity (% time)	2.2	2.2
2.2 Ratio of training intensity in firms with 100-499 employees to that with 50-99 employees	1.2	1.1
2.3 Ratio of training costs to wage costs of training	0.24	0.24
2.4 Percent wage increase of 20 years' experience (%)	80	80
2.5 Percent wage increase of 40 years' experience (%)	88	89

The table reports the targeted moments in the data and in the model.

[4] Alternatively, we construct a second measure, for which we directly regress the available probabilities in Donovan, Lu and Schoellman (2020) on GDP per capita and predict the probabilities at different income levels. The resulting fitted probabilities are very similar.

Finally, it is important to note that the available data do not provide a direct measure of overall firm-provided training for all countries. For instance, we do not have measures of informal training for most of our economies. Thus, we first measure the average hours of formal training per worker from the data.[5] We then impute overall training intensity for every economy, relying on two assumptions according to the US SEPT survey: the average worker spends two hours in informal training for each hour spent in formal training and there are 50% more workers participating in informal training than in formal training.[6]

Table 2.2 shows the model almost exactly matches all the moments related to training. Moreover, the model almost exactly or very closely matches all the moments reflecting labor market dynamics although it slightly overestimates the share of job-to-job transitions from high-to-low wage firms (0.22 in the data vs 0.24 in the model). Overall, the model does really well in matching the targeted moments.

Calibrated Parameters. We report the calibrated parameters in Table 2.3. Our parameters are reasonable compared with the literature. Our parameter γ_s can be interpreted as the diminishing returns of human capital investment (in terms of effective hours) in producing new human capital. Its calibrated value is in the ballpark of the estimates in the literature: for instance, Imai and Keane (2004) find this parameter to be 0.22, while Manuelli and Seshadri (2014) estimate this parameter to be 0.48. Moreover, training a young worker for the full quarter (480 working hours) increases her hourly wage by 5%, which lies in the range of empirical studies on U.S. training returns as reviewed by Leuven (2004) and Bassanini et al. (2005).[7] For more evidence on the model calibrated parameters and model dynamics, we illustrate in Appendix Section B.3 how the

[5] We multiply shares of workers exposed to formal training by hours spent in formal training per participant, which are predicted using the relationship between hours of formal training per participant and GDP per capita from the EU-CVT data.

[6] In the U.S., 60% of workers are receiving formal training and 90% are receiving informal training.

[7] For example, using the NLSY data, Veum (1995) finds that increasing one hour of formal training improves hourly wage by 0.01%. Also using NLSY data, Frazis and Loewenstein (2005) find that 60 hours of formal training increases wage by 3$\sim$5% — our calibration implies 2.7% wage growth for 60 hours of training in one quarter. The comparison with Veum (1995) and Frazis and Loewenstein (2005) is imperfect, because it is unclear whether training in their data happened within one quarter or in multiple quarters.

moments help identify the model's parameters by calculating the elasticity for moments to each parameter.

Table 2.3: Calibrated Parameters

Parameter	c_M	c_s	c_v	γ_s	γ	γ_p	κ	ζ	$\underline{p}$	μ_F	δ	c_p	d
Value	0.58	0.20	1.28	0.31	0.31	6.79	3.92	0.05	0.14	0.80	0.03	5.13	0.007

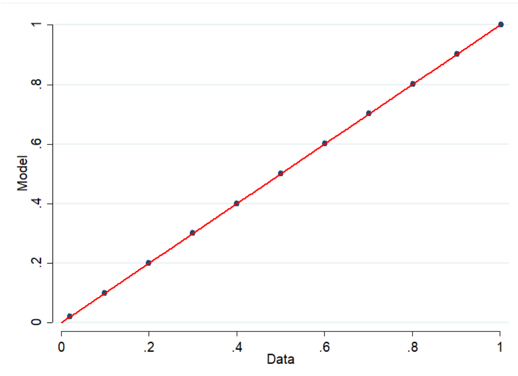
Note: The table lists the parameters that were determined using simulated method of moments and their values in the quantitative analysis. The estimated parameters are: the constant in the matching function, c_M; training costs per time as a share of the average wage rate, c_s; the constant in vacancy costs, c_v; the constant in the function of leaving costs, c_p; the constant in training returns, ζ; the convexity in training returns, γ_s; the traditional-sector share in the aggregate production function, γ; the convexity in the function of leaving costs, γ_p; the shape parameter of Pareto productivity distribution, κ; exogenous separation rates, δ; lower bound of leaving probability, $\underline{p}$; and the share of training costs paid by firms, μ_F.

2.3.2 Cross-Country Calibration

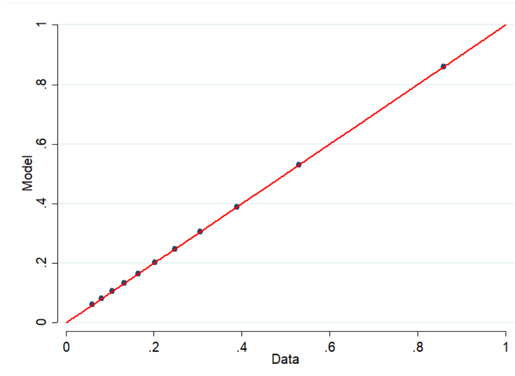
We focus our analysis on three main features that radically decrease with development: the share of workers who are self-employed, job-to-job transitions, and job destruction rates (Gollin, 2002, 2008; Donovan, Lu and Schoellman, 2020). In this section, we calibrate the model for a representative economy at 10 different income levels other than the United States' one. To do this, we keep the baseline calibrated parameters and re-calibrate δ, c_p, A_m, and A_T to match income levels, self-employment, the share of workers who stay in the same firm, and the share of workers who stay employed from quarter to quarter.

We first show how the model fits the targeted moments in Figure 2.1. On the x-axis we show the moments in the data, on the y-axis we show the moments in the model and we plot the 45-degree line. Overall, our model matches the targeted data moments well. We exactly match GDP per capita, self-employment shares, and the share of workers who remain employed from quarter to quarter. Moreover, the model slightly overestimates the share of workers who stay in the same firm from quarter to quarter (0.7 in the model and 0.66 in the data for the poorest economy).

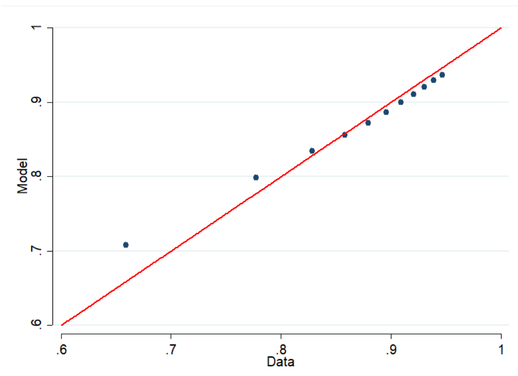
In Figure 2.2 we show the calibrated parameters given by the representative economy



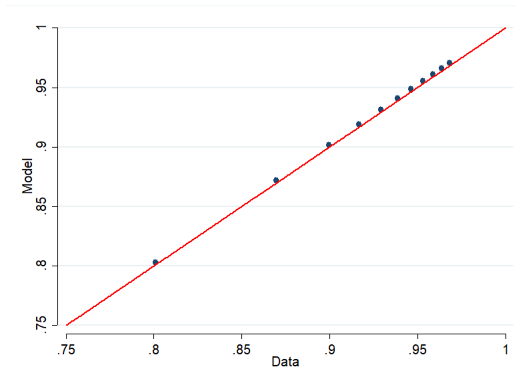
(a) GDP per capita



(b) Self employment share



(c) % workers in same firm

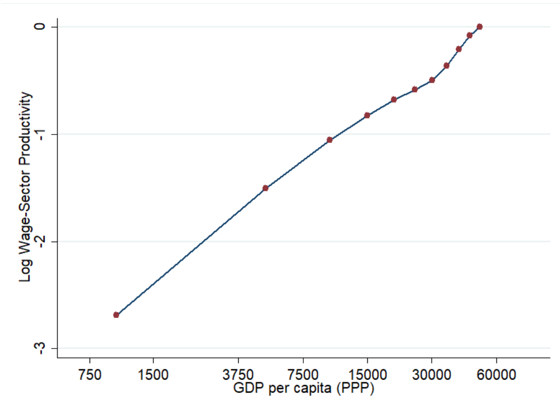


(d) % workers still employed

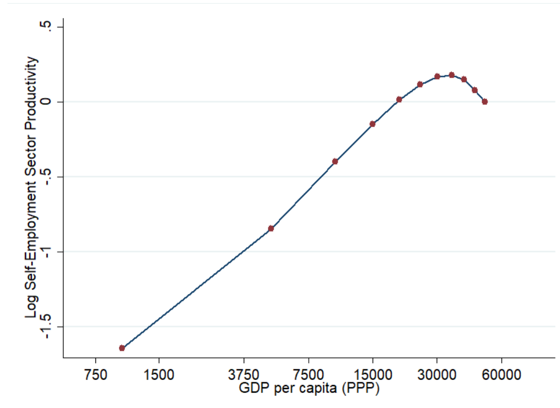
Figure 2.1: Cross Country Targeted Moments

Note: This figure shows the targeted moments in the model (vertical axis) and in the data (horizontal axis). We consider 10 representative economies at income levels of $2,000, $5,000, $10,000, $15,000, $20,000, $25,000, $30,000, $35,000, $40,000 and $45,000 for GDP per capita ($50,000 is the US level). Panel A shows GDP per capita (relative to the U.S.). Panel B shows self-employment. Panel C shows the share of workers who remain in the same firm after one quarter. Panel D shows the share of workers who are employed in any firm for two consecutive quarters.

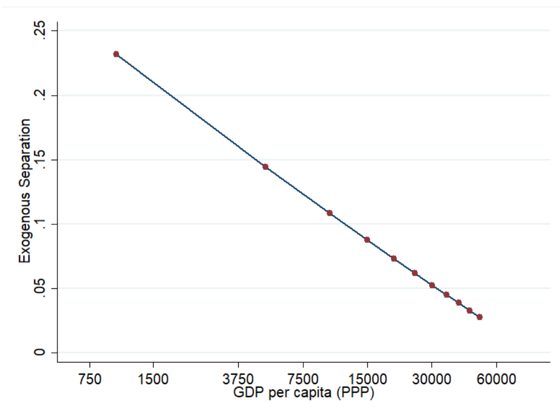
calibration. In Figure 2.2a we plot the level of productivity in the modern sector and in panel 2.2b we plot the level of productivity in the self-employment sector and, as expected, both productivity levels increase with GDP. The increase in the wage sector's productivity is faster than the self-employment sector's productivity increase, which partially shapes the decrease in self-employment with development. In Figure 2.2c and 2.2d we plot the parameters shaping the labor market dynamics. The job destruction rate δ decreases and the cost of breaking the contract c_p increases with income. These results align with larger turnover, more self-employment and worse institutional quality in developing economies.



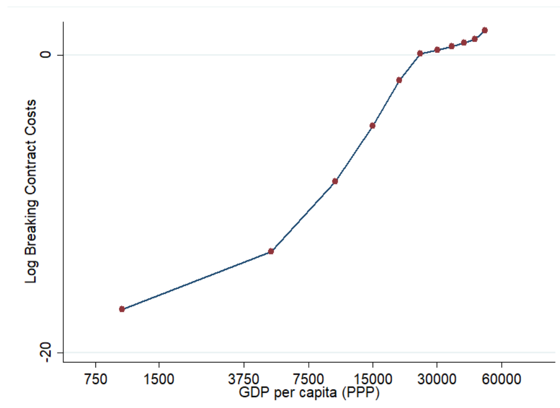
(a) Modern Sector Productivity



(b) Self-Employment Sector Productivity



(c) Exogenous Separation



(d) Breaking Contract Costs

Figure 2.2: Cross Country Calibrated Parameters

Note: This figure shows the calibrated parameters for each economy in the model as a function of Log(GDP per capita). We consider 10 representative economies at income levels of \$2,000, \$5,000, \$10,000, \$15,000, \$20,000, \$25,000, \$30,000, \$35,000, \$40,000 and \$45,000 for GDP per capita (\$50,000 is the US level). Panel A shows the wage sector productivity (A_M in the model). Panel B shows the relative productivity between the self-employment sector and the wage sector (A_T/A_M). Panel C shows the quarterly exogenous separation rate implied by the model (δ). Panel D shows the log of the breaking contract costs, c_p.

2.3.3 Non-targeted Moments and Model Validation.

In Table 2.4, we compare several untargeted moments in the model to the data for the U.S. and across countries. In Panel 1, we first focus on the within-country moments. Our model implies a similar aggregate labor share and average unemployment duration as in the BLS data for the period 1994-2007. Our model has slightly higher slope of the labor share on firm market shares as estimated by Autor et al. (2020), which captures how concentration affects wage returns relative to firm revenues. Moreover, our model implies that bigger and more productive firms invest more in training, which aligns with the patterns in the data.

Table 2.4: Non-Targeted Moments in the Model vs Data

Untargeted Moments	Data	Model
1. The US		
1.1 Employees' Labor Share (%)	55	63
1.2 Average unemployment duration (weeks)	17.5	15.3
1.3 Slope of labor share on firm market share	$[-2.37, -0.35]$	-0.69
2. Across countries		
2.1 Slope of labor shares on log GDPPC (adj 1)	-0.02	-0.03
2.2 Slope of labor shares on log GDPPC (adj 2)	0.02	0.05
2.3 Slope of std firm size on log GDPPC	0.18	0.10

The table reports some non-targeted moments in the data and in the model. The aggregate labor share and average unemployment duration are calculated using BLS data for the period 1994-2007. The slope of the labor share on firm market shares comes from Autor et al. (2020). The measures of the labor share and income to calculate moments 2.1 and 2.2 come from Gollin (2002). The first measure (adjustment 1) assumes the self-employment sector labor share is 1 while the second measure (adjustment 2) assumes that labor share in the self-employment sector is identical to its counterpart in the wage sector. The slope of the wage increase in 20 years of experience on the logarithm of per-capita GDP comes from Lagakos et al. (2018*a*). The slope of the standard deviation of employment with respect to income comes from Poschke (2018).

Panel 2 in Table 2.4 focuses on cross-country non-targeted moments. First, we compare the relationship between different measures of the labor share and income from Gollin (2002). The first measure (adjustment 1) assumes the self-employment sector labor share is 1 while the second measure (adjustment 2) assumes that labor share in the self-employment sector is identical to its counterpart in the wage sector. Furthermore, much of the literature shows how firms are smaller on average in developing economies and how this impacts productivity in the aggregate. We provide one informative moment of this distribution, which is the slope of the standard deviation of employment with respect to income, which is 0.1 in our model (slightly lower than 0.18, as found by Poschke (2018)).

Finally, we turn our attention to the main non-targeted moment we want to analyze. We plot the training intensity from the data and model as a function of GDP per capita in Figure 2.3. The model does well in matching both the elasticity of training with respect to per-capita GDP and also the levels. Training intensity in the data may be noisier, especially for middle-income countries for which the WB-ES may overweight bigger firms and where self-employment is low enough to increase the importance of the wage sector's training intensity. Thus, it is possible that for this reason the model slightly underestimates the training intensity for middle-income countries.

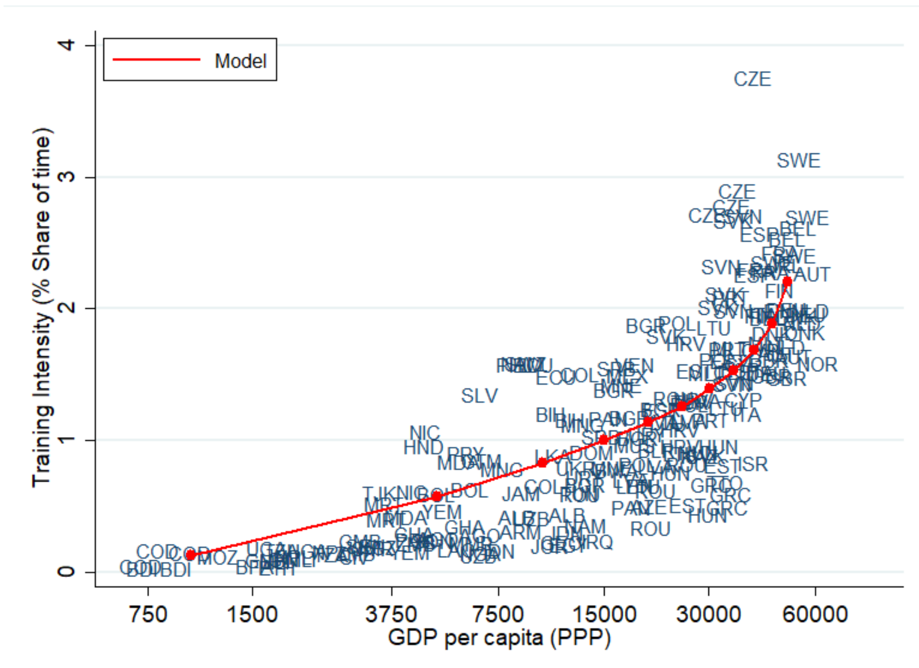


Figure 2.3: Training in Data and Model

Note: This graph shows the quadratic fit of the cross-country training intensity (measured in the share of time that an average worker spends in training) as a function of Los(GDP per capita). The green line represents the quadratic fit for the cross-country measure in the model and the blue line represents its counterpart in the data. The grey shadow represents the 95% confidence intervals.

These results suggest that our channels capture most of the difference in training across countries. In Appendix Table B.4 we show suggestive evidence on the correlations between training investments and job turnover measures from Donovan, Lu and Schoellman (2020), self-employment, firm size distribution, and institutional quality proxies. As we add each one of these explanatory variables, we show how the coefficient on GDP per capita decreases. Once we add the first principal component that includes all variables, we explain all the correlations between GDP per capita and training, which suggest that institutional quality, job separation, and self-employment capture most of the trend component of on-the-job training with respect to cross-country income.

These results do not imply that our channels explain all training differences across countries, but that we capture how training varies with income. From the model, it is clear that things that affect separation rates, the probability of hiring, firm-worker matching quality, or the vacancy costs will affect the contracts and training investments. In Appendix Section B.7 we rely on labor market institutions indexes constructed by Botero et al. (2004) to understand how the cost of firing workers, labor market institutions as the minimum wage and unemployment benefits correlate with our measure of training. We find that these measures increase the explanatory power over training but they do not account for part of the trend component between training and income. This is consistent with a result by Donovan, Lu and Schoellman (2020) that shows some labor market institutions are important determinants of cross-country variation in labor market flows but do not explain the trend relationship with respect to income.

As a robustness check, in Appendix Section B.4, we provide an alternative calibration in which we calibrate the model for 100 countries using country-specific levels of self-employment and training and let training productivity in the model vary across countries. We show that with this alternative calibration all our results hold.[8] In the following sections we aim to answer three main questions: (1) how much of the wage-growth differences across countries can be accounted for with on-the-job training; (2) why do developed economies invest more in training; and (3) what is the optimal training policy at different stages of development?

[8] By introducing country-specific training returns, we show the model matches all moments used in the baseline calibration and it also exactly matches training levels for all countries. As expected, we find that training productivity is mostly flat with income and thus there is not much unexplained by differences in training returns for the trend component of training levels.

2.4 Cross-Country Wage Growth and Income Differences

We first analyze how much each of the channels in our model contributes to explaining the difference in workers' wage growth between developed and developing economies.

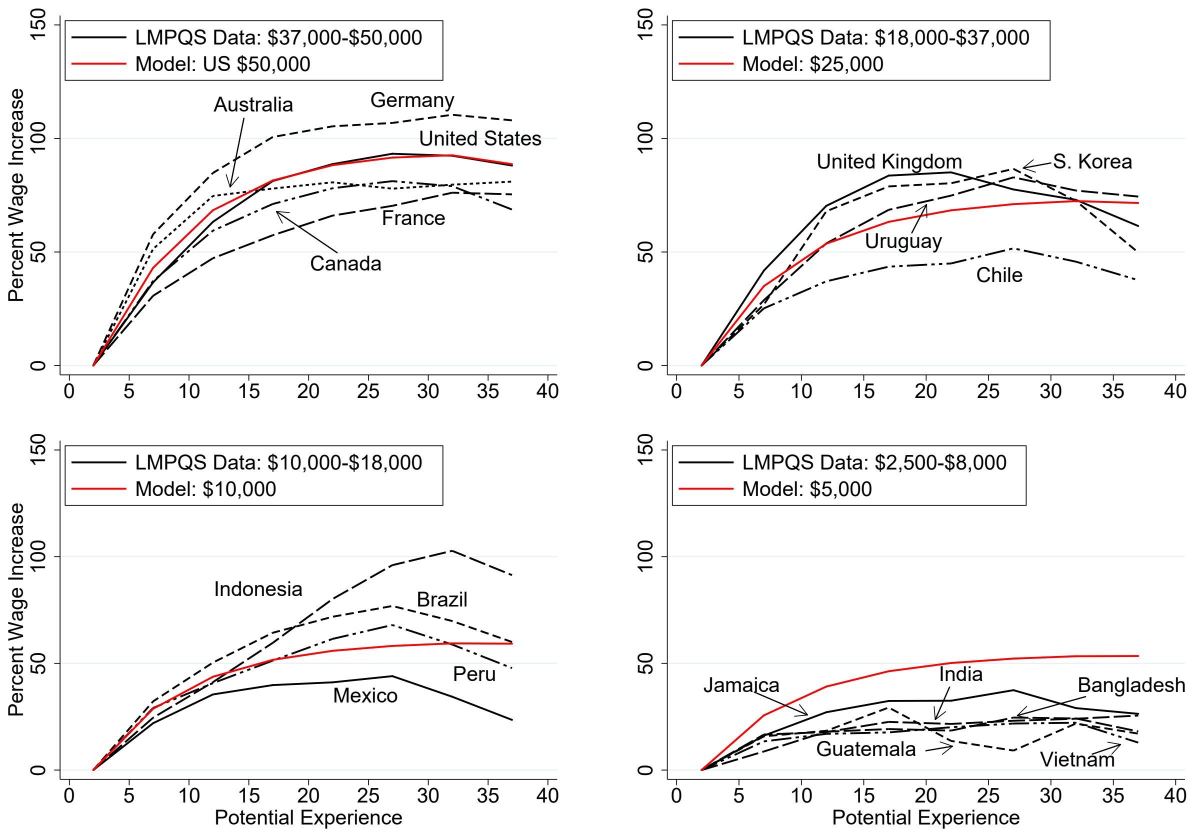


Figure 2.4: Cross-country Eperience-Wage Profiles: LMPQS vs Model

Note: This figure replicates Figure 2 from David Lagakos, Benjamin Moll, Tommaso Porzio, Nancy Qian and Todd Schoellman (2018*b*) (LMPQS) and adds the wage-experience profiles from our model in red. In the vertical axis we plot the percent increase in wages at each potential experience bin, and in the horizontal axis we plot potential experience in years. Panel a shows countries between $37,000 and $50,000 and the model outcome for the U.S calibrated economy ($50,000). Panel b shows countries between $18,000 and $37,000 and the model outcome for the calibrated economy at $25,000. Panel c shows countries between $10,000 and $18,000 and the model outcome for the calibrated economy at $10,000. Panel d shows countries between $2,500 and $8,000 and the model outcome for the calibrated economy at $5,000.

Figure 2.4 plots experience-wage profiles for 18 economies at all income levels from Lagakos et al. (2018*b*). Each panel shows the profiles from countries within a particular income range and the model's profile for an economy within that same range. Our model matches the profiles well at all income levels except for the ones at the bottom of the world income distribution. The calibrated economy at $5,000 has a steeper experience-wage profile than its counterparts in the data. This suggest that other factors that we do not include in our model may play an important role to explaining the low wage growth in these economies.

To analyze how much of the cross-country difference in returns to experience our model accounts for, we plot in Figure 2.5 the returns to 20 years of experience of LMPQS economies and the same measure from our model as a function of per-capita GDP. As expected, the model matches very well the wage growth for middle- and high-income countries and overestimates the wage growth for workers in the poorest economies. Regressing the returns on log(per-capita GDP) we find a slope of 0.26 in LMPQS and a slope of 0.14 in our model which implies we capture 55% of the differences in returns to experience. Nonetheless, the model captures all of the difference for the economies above \$10,000. Lagakos et al. (2018*b*) finds that occupation and schooling differences capture around 20% and 30% respectively of the differences in wage-growth across countries. This implies that our channel captures most of the cross-country differences in wage-growth that are not explained by these two other factors.

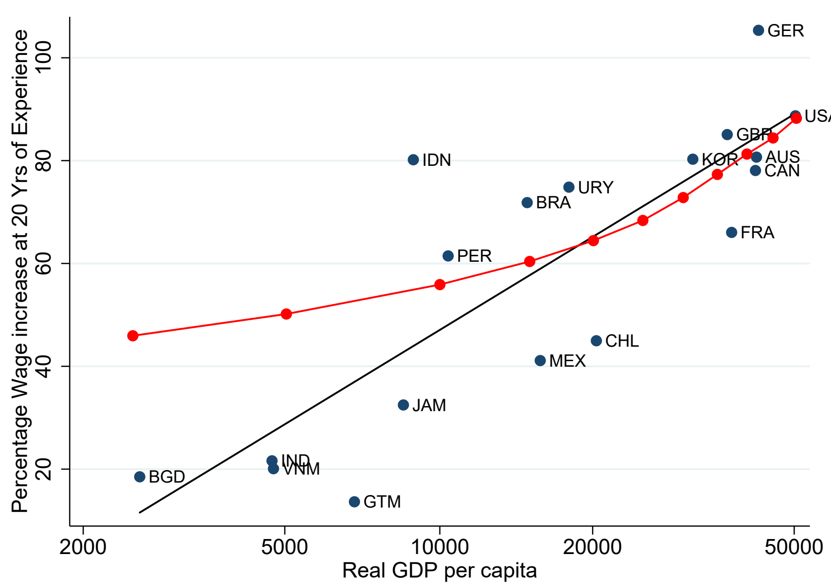


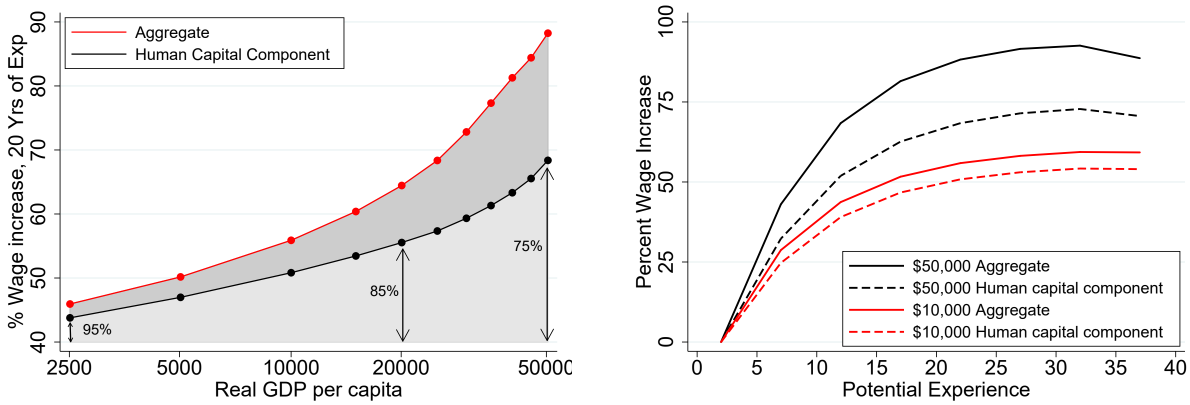
Figure 2.5: Cross-country Returns to 20 Years of Experience: LMPQS vs Model

Note: This figure replicates Figure 3 from David Lagakos, Benjamin Moll, Tommaso Porzio, Nancy Qian and Todd Schoellman (2018*b*) (LMPQS) and adds the returns to experience from our model in red. The slope in the LMPQS data is 26 while the slope of regressing the model's returns on log per-capita GDP is 14.

There is a growing literature that tries to disentangle how human capital accumulation and job-search contributes to explaining workers' earnings dynamics (e.g. Bunzel et al., 1999; Rubinstein and Weiss, 2006; Barlevy, 2008; Yamaguchi, 2010; Burdett, Carrillo-Tudela and Coles, 2011; Bowlus and Liu, 2013; Bagger et al., 2014; Gregory, 2019). These papers analyze how job search and human capital accumulation contribute to explaining workers' wage growth only for developed economies. We continue by calculating how much of the wage growth is driven by firm-training and labor market dynamics for countries at all income levels. In Figure 2.6a, we decompose the effects on the human capital and the job-turnover components of wage-growth as a function of income. We find that the contribution of human capital is large for every economy and that it decreases with income. This happens because the high level of job destruction in the poorest economies prevents workers climbing up the job ladder. As income increases, fewer workers are separated from their jobs and become unemployed, which generates larger increases in wages through job-to-job transitions. Moreover, Figure 2.6b shows the entire wage-experience profile decomposition for two economies at \$50,000 and \$10,000, which shows the same pattern. More importantly, we find that the human capital channel accounts for 75% of the difference in workers' wage growth between these economies.

We now focus on productivity differences driven by this channel. Using our calibrated representative economies, we simulate the model with different assumptions on training investments and plot the resulting per-capita GDP from each model in Figure 2.7. In orange we plot the original model, in blue we plot the model with no training, and in red we plot a case where all economies have the same training investments as the poorest economy.[9] When there is no training output is the lowest. Output increases when we add the poorest economy level of training to the model with no training, and increases even more when we endogenice training, which reflects the fact that training boosts productivity in the aggregate. The heterogenous increase in output with respect to income suggest that adding training improves output more in developed economies

[9]For this last case we use the training level for each firm and age-type worker, and we assume that all economies have that exact same training pattern in each firm and worker age type.



(a) Decomposition and Income (b) Profile Decomposition

Figure 2.6: Cross-country Eperience-Wage Profiles Composition

Note: This figure replicates Figure 2 from David Lagakos, Benjamin Moll, Tommaso Porzio, Nancy Qian and Todd Schoellman (2018*b*) (LMPQS) and adds the wage-experience profiles from our model in red. In the vertical axis we plot the percent increase in wages at each potential experience bin, and in the horizontal axis we plot potential experience in years. Panel a shows countries between $37,000 and $50,000 and the model outcome for the U.S calibrated economy ($50,000). Panel b shows countries between $18,000 and $37,000 and the model outcome for the calibrated economy at $25,000. Panel c shows countries between $10,000 and $18,000 and the model outcome for the calibrated economy at $10,000. Panel d shows countries between $2,500 and $8,000 and the model outcome for the calibrated economy at $5,000.

than in developing economies.

Thus we ask: what is the share of the income differences across countries explained directly by training in our model? To answer this question, we plot the difference between the log(per-capita GDP) in the full model and its counterpart from the two models with no training and with the poorest economy training. We plot the fitted values in the secondary axis of Figure 2.7. This difference represents the percentage increase in output from the model with no training (or low training) relative to the full model. The slope of the percentage increase in output from the model with no training to the full model on log(per-capita GDP) provides the share of the income differences explained directly by training in our model. Our quantitative model suggests that on-the-job training explains 15% of the income differences across countries, which is sizeable.[10] Moreover, doing the same exercise but using the model with the poorest economy's training level for all economies, the model generates an income difference coming from training of 10%, which

[10]Lagakos et al. (2018*a*) shows that adding returns to experience helps explain around 20% of the income differences across countries, and thus, our result suggests that firm training is one of the main sources of workers' human capital accumulation post-schooling.

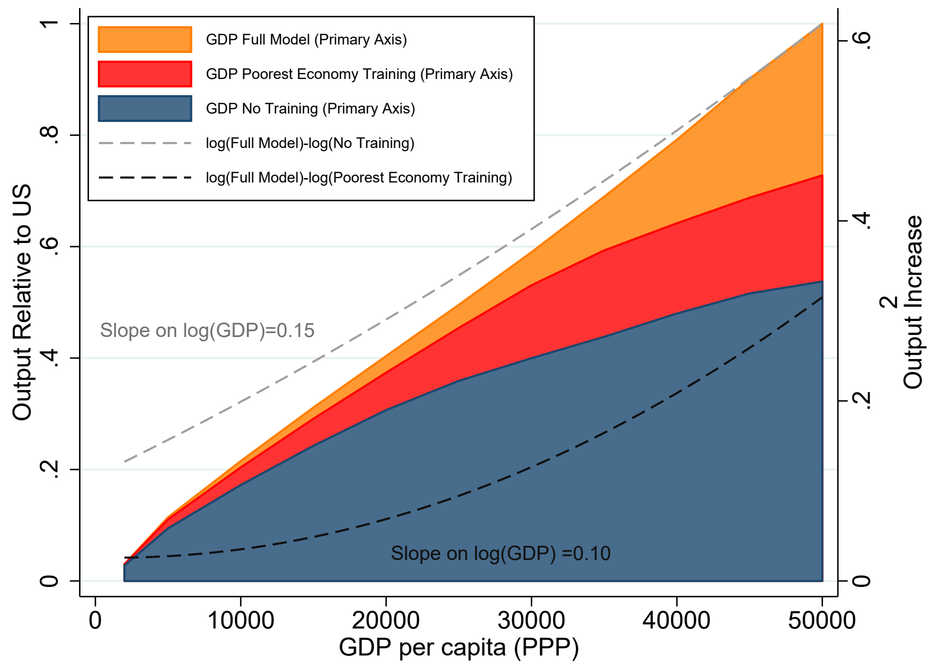


Figure 2.7: Income Increase due to Training

Note: This figure shows the percentage increase in output from training calculated as the log change in output from the model shutting down training (increasing c_s to an extremely large value) to the full model as a function of GDP per capita. Each observation comes from using the calibrated version of the model for each country. Data from GDP per capita comes from the World Bank Indicators. The slope of 0.16 represents the share of the increase in GDP per capita explained by training in the model.

reflects a lower bound on how training explains income differences in the model.

2.5 Training Decomposition

In this section, we analyze how much of the cross-country training differences each of our channels account for. We aim to understand what drives the lack of training in developing economies and the role each channel plays at different stages of development.

From the results on training intensity and self-employment shares, we first analyze how the sectorial allocation explains training differences in the aggregate. Denoting S as the training investment in the modern sector and T as the wage-sector employment share, the difference in training can be simply decomposed as:

$$log(S_{us}T_{us}/S_{base}T_{base}) = log(S_{us}T_{us}/S_{us}T_{base}) + log(S_{us}T_{base}/S_{base}T_{base})$$

The first term reflects the training increase due to the change in the self-employment share and the second term represents the increase in training in the modern sector, conditional on the sectoral allocation. We plot these two components in Figure 2.8. We find that the role of self-employment in explaining training differences decreases substantially with development. It goes from 60% in the poorest economies to 10% in the richest ones. This result is very intuitive; aggregate human capital will be low if there is a small proportion of workers in the modern sector, even if training levels are large in that sector. This result offers strong policy implications as this channel amplifies the productivity increases from reallocating workers away from self-employment studied in the structural transformation and migration literature.

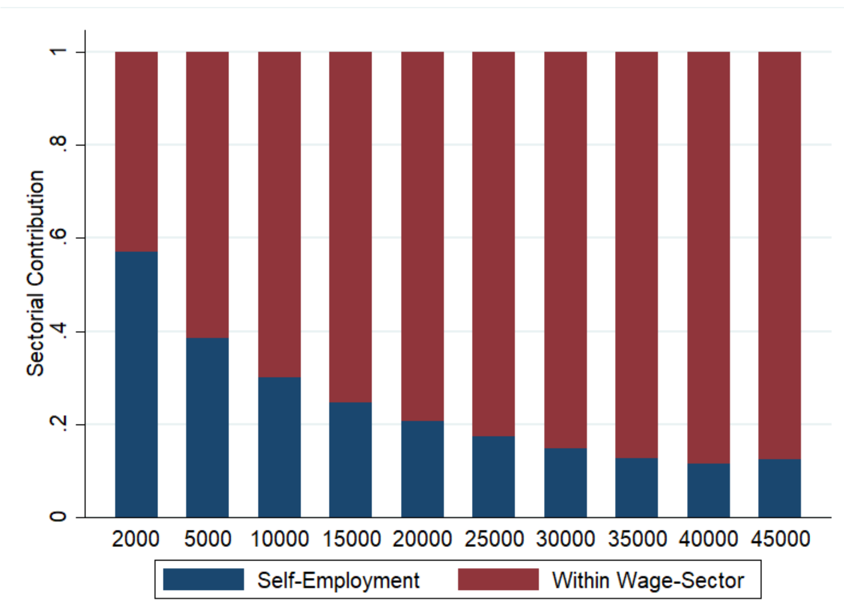


Figure 2.8: Training Decomposition by Sectoral Component

Note: This figure shows how (1) changing the self-employment share while keeping training in the wage sector fixed, and (2) changing the wage sector training level while keeping the self-employment share fixed contribute to explaining the difference in training between each economy and the U.S.

There are 4 parameters that vary across country: δ, which shapes job destruction; c_p, which shapes job-to-job transitions; and A_T and A_M, which shapes income and self-employment shares. We proceed to do a factor-decomposition driven by these parameters. We explain the procedure through one example. Let's assume we want to capture how each parameter contributes to explaining the training gap between the United States and the poorest economy. We first start from the U.S. calibration and change c_p to its poorest economy counterpart and we measure the change in training. Then, we start from the U.S. calibration, but with the poorest economy's δ and change c_p to the poorest economy value again and compute the new change in training. We repeat this exercise for all possible combinations to calculate the average change in training coming from c_p. Then, we do the same simulations to calculate the average change in training from the other parameters.[11] After these calculations we have the result on how each parameter contributes to explaining the training gap between the U.S. and the poorest economy. For our factor-decomposition, we repeat this procedure comparing the U.S. with the economy at each income level and plot the results in Figure 2.9.

The importance of job destruction, reflected in the exogenous separation, tends to decrease with income while the importance of the labor market friction coming from the cost of breaking contracts increases with income. On the other hand, the contribution from the change in productivity is roughly constant (around 20%). The first important result suggests that labor market frictions explain 80% of the differences in training across countries. The second take away is that, for poor economies, the most important channel to explain the lack of training is job destruction, but as income increases the difference in training comes largely from frictions in job-to-job transitions.

To dig deeper into the mechanisms driving these results, in the Appendix Figure B.5, we

[11] Note that there are 6 ways of doing these changes (from the baseline, we might first change c_p, then δ and finally A_T and A_M, or we might first include δ and then the other ones in another order, etc.). These different ways of calculating the change in training may provide different results depending on the complementarity between the mechanisms. Therefore, we simulate the model following all possible orders and calculate the average increase in training coming from each parameter in all simulations.

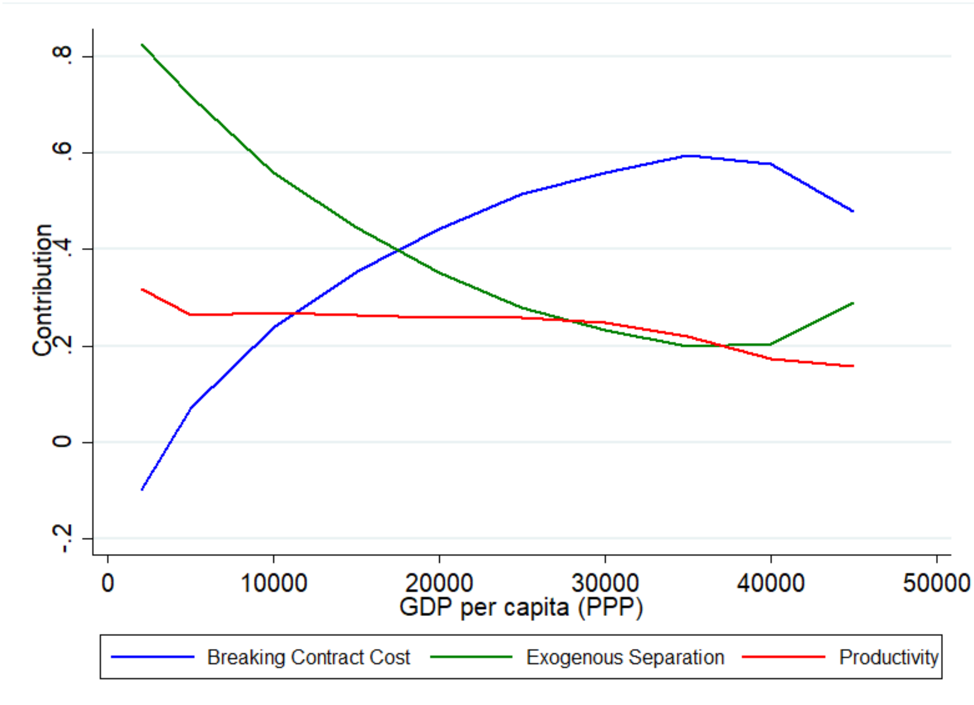


Figure 2.9: Share of Training Gap Covered by each Parameter Change

Note: This figure shows the contribution of each channel to explaining the training gap between the economies at each income level and the U.S. The green line presents the contribution from changing δ; the red line presents the contribution from changing A_t and A_M simultaneously; and the blue line presents the contribution from c_p.

decompose each channel into the wage sector and self-employment contributions.[12] The effect coming from changing sector-specific productivities only impacts the share of workers in each sector but does not generate changes in training levels within the wage sector (relative training cost and revenue are invariant across different income levels).[13] The effect coming from c_p and δ are driven mostly by the change in self-employment for poor economies and by changes in firm-level investments in training for richer economies.

Moreover, to understand the model's dynamics, we go one step further and decompose the increase in training within the wage sector into the partial equilibrium and the general

[12]We decompose the effects in the following way: $log\left(\frac{S_iT_i}{S_{base}T_{base}}\right) = log\left(\frac{S_iT_{base}}{S_{base}T_{base}}\right) + log\left(\frac{T_i}{T_{base}}\right)$ where S represents the wage-sector training level and T represents the share of workers in the wage sector, both by changing parameter i.

[13]The wage sector changes are not exactly zero because we solve the model numerically and the convergence may not be perfect in every case and some small noise can be reflected in the results.

equilibrium effects. The partial equilibrium effect represents firms' training level in the new parameter scenario while keeping the wage and employment distributions fixed. The general equilibrium effect represents the wage sector's additional increase in training (letting the wage and employment distribution change). We plot the decomposition from the change in the exogenous job destruction and the job-to-job transition friction in Appendix Figure B.6.

In partial equilibrium we always observe the expected signs. As c_p increases and δ decreases, the probability of keeping workers goes up, increasing training investments. Nevertheless, the effects from general equilibrium may be negative. Lower separation to unemployment generate fewer unemployed workers in the pool of searchers, and thus, firms must post higher wages to attract workers. Moreover, lower probability of job destruction means workers move to more productive firms faster because they fall to unemployment in fewer occasions. In this case, higher wages pull training investments down while the employment distribution shift towards more productive firms pushes training up. The negative GE effect observed reflects the fact that the wage effect predominates.

Moreover, higher costs to break contracts through higher c_p means firms keep workers longer and it is more difficult to poach workers from other firms. These effects push wages downward and this encourages firms to invest more in training as they capture a higher share of the surplus. Although fewer job-to-job switches shifts the employment distribution towards less productive firms (which would generate a decrease in training) the wage effect predominates.

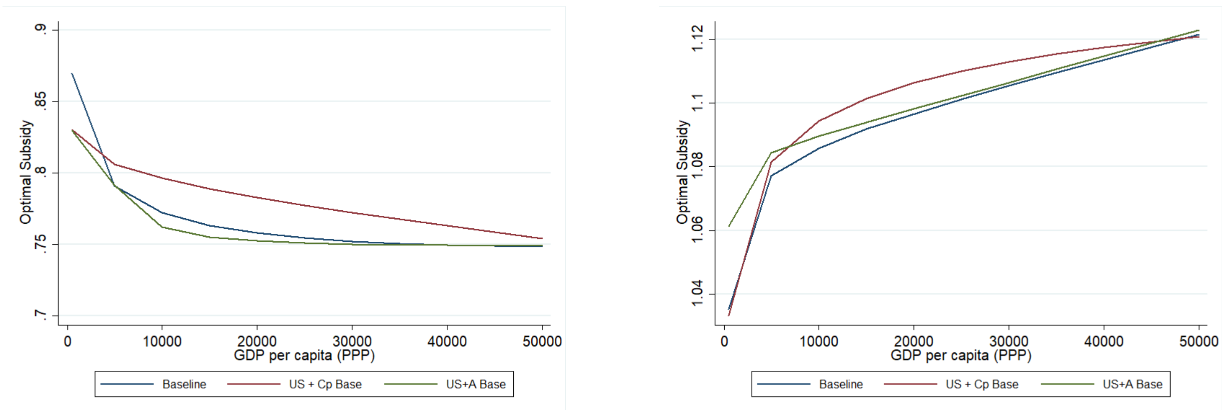
There are 3 main takeaways from the training decomposition. First, most of the training gap between the poorest and richest economies is explained by differences in self-employment shares. This is driven by the lack of training in the self-employment sector, and the high rates of self-employment prevalent in poor economies arising from the endogenous reallocation of workers from the wage sector to the self-employment sector as a result of high labor market frictions in the former sector. Second, these labor market frictions remain key to explaining training investments as countries develop, and self-employment shares fall. The mechanism driving this is worker turnover in the wage sector. In particular, high job separation rates and low contract quality make worker turnover more likely, and thus, depress the incentives to invest in training in medium-income economies relative to richer economies. Collectively, these two facts indicate that labor market frictions are key to explaining training differences across countries of different levels of development. Third, when we decompose the importance of these labor market frictions along its two key components, we find that job destruction is the most important factor to explain the lack of training in poorer economies while frictions in job-to-job transitions are more important to explain the training differences between more developed economies.

2.6 Cross-Country Optimal Policy

In this section, we discuss the inefficiencies present in our framework, show empirical evidence on the existence of policies aimed at increasing training around the world and conduct an optimal policy analysis.

Our model has two sources of inefficiencies in training investments. In Appendix Section B.8 we show the training patters for firms, workers, and the aggregate, and show that the firm decision is the relevant margin as firms always want lower levels of training in the calibration that matches all labor market dynamics. Thus, firms do not internalize the full benefits from training investments for workers. Second, even if there is a joint decision and the optimal investment level is the same for both firms and workers, another source of inefficiency still exists. Workers and firms never internalize the benefits from training for future employers if separation occurs (Acemoglu, 1997). Therefore, higher output in future employers and the full benefits for workers are not internalized, which means there is space for policies to improve welfare.

In Appendix Section B.6 we do an extensive description of government subsidy policies around the world and show that training subsidies are indeed very common. We review government policies to incentivize employer-provided training from 36 countries from all continents and income levels and 21 U.S. states. There are a wide variety of training policies or subsidies ranging from 5% - 10% of training costs in Mozambique, or \$20 - \$100 in India, to 120% of training costs in Austria, or \$5,415 in New Zealand per participant. For, U.S. states, on the other hand, the most common policy is a reimbursement for about 75% of the training cost. The large set of government programs described suggests governments all around the world consider training an important channel to improve productivity, and that training investments are probably inefficient.



(a) Optimal Subsidy Rate

(b) Net Output Increase

Figure 2.10: Optimal Training Subsidy and Development

Note: This figure shows the results from simulating the model adding a subsidy to the total training costs (direct costs of training and workers' wages for full training length), financed with lump-sum taxes. Panel (a) shows the optimal subsidy at each income level, and panel (b) shows the increase in net output driven by the policy. The blue line presents the results from the policy at each income level with the calibrated parameters of that economy. The red and green lines present the results from the optimal subsidy using the U.S. baseline calibration and only changing one parameter at a time to the value from each income level representative economy.

Motivated by the existence of training investment inefficiencies and by the empirical evidence showing training subsidies are very common, we perform an optimal policy analysis. We simulate the model with the baseline calibration, adding a subsidy to the total training costs, which includes the direct costs of training and the workers' wages for the full training length. We assume the training program is financed with lump-sum taxes to agents and we analyze how output (net of training subsidy costs and vacancy costs) changes for different rates. In Figure 2.10 we plot the optimal subsidy rate and the net output increase for the optimal policy at different income levels. The blue line represents the optimal policy for each representative economy with all its calibrated parameters. The U.S. benchmark case shows that net output increases by 12%

when it is maximized with a subsidy of 75%. Moreover, as countries become less developed, the optimal subsidy must be larger, although the increase in output is lower.

The other lines in Figure 2.10 represent the U.S. baseline calibration changing just one parameter at a time to the calibrated one for each income level. For instance, the red line at \$40,000 reflects the baseline calibration for the United States with the c_p from the \$40,000-calibrated economy. We do this exercise to show the fact that the decrease in the optimal subsidy rate with income comes partially from all our channels. Thus, we show that subsidies should be larger in developing economies when self-employment or job destruction is higher or when contract quality is lower.

2.7 Conclusion

Human capital accumulation plays a key role in economic growth and development. While recent research has highlighted the importance of on-the-job human capital accumulation in explaining cross-country income differences, how workers accumulate human capital on-the-job at different stages of development is still underexplored. We study one key source of on-the-job human capital accumulation, namely firm-provided training. We exploit rich data sources to show that firm-provided training increases with development and that this happens in every margin of training. Moreover, we find that firm-provided training is the most important source of human capital investments in workers' careers. Then, we build a GE search model with firm heterogeneity and training investments that help us identify the mechanisms mediating these facts.

Our results have strong policy implications. A simple sectoral decomposition suggests that self-employment is key in explaining the lack of on-the-job training in the poorest economies. Thus, our channel amplifies the productivity increases from reallocating workers away from self-employment studied in the structural transformation and migration literature. Furthermore, by digging deeper into the channels generating the differences in training, we find that the high levels of job destruction is the most important factor preventing training investments in poor economies, while frictions in job-to-job transitions are more important in explaining training differences between developed economies. These imply that to increase productivity and improve net output, policies to improve match quality between firms and workers may be desirable in developing economies while policies to generate better contracts may be more beneficial in richer countries. Finally, our model predicts considerable inefficiencies in human capital investments and sizeable aggregate gains from training subsidies to firms. More importantly, training subsidies to firms may be particularly beneficial in poor countries where economic environments disincentivize training.

The importance of on-the-job training could be even larger if there are complementarities with other sources of human capital, such us schooling or co-worker spillovers. A fruitful area for future research is to study how different human capital accumulation sources interact with each other and what the implications of these interactions are to conduct more efficient public policy for countries at different stages of development.

2.8 Acknowledgements

Chapter 2 is part of the paper "Human Capital Accumulation and Development: The Role of On-the-Job Training" coauthored with Xiao Ma and Daniela Vidart. As of May 2021, this paper is being prepared for submission for publication. The dissertation author was a primary investigator of this material.

Part II

Firm Dynamics, Export Status and Wage Profiles

Chapter 3

Learning By Exporting and Wage Profiles: New Evidence From Brazil

3.1 Introduction

It is well-known that exporters are more productive than non-exporters. This could be driven by self-selection of the best firms into exporting activities, or by productivity improvements after exporting. In particular, Atkin, Khandelwal and Osman (2017) find that exporting improves firms' technical efficiency using a randomized experiment, and De Loecker (2007) shows that firms' productivity gains after exporting may increase when firms export to high-income countries.[1] Whereas these existing studies mostly focus on firm-level outcomes, exporting may also impact workers. It is well-documented that workers earn higher wages in exporters than in non-exporters (Bernard and Jensen, 1995). However, despite much attention to firm-level differences in life-cycle wage growth in recent studies (Herkenhoff et al., 2018; Jarosch, Oberfield

[1]For more evidence on the comparison of productivity levels between exporters and non-exporters, see also Bernard and Jensen (1999), Aw, Chung and Roberts (2000); Van Biesebroeck (2005), Lileeva and Trefler (2010), and Aw, Roberts and Xu (2011).

and Rossi-Hansberg, 2018; Gregory, 2019),[2] little is known about how firms' exporting activities shape workers' life-cycle wage dynamics.

In this paper, we fill this gap. We rely on Brazilian linked employer-employee administrative data and customs records between 1994–2010, and assemble a long panel with detailed job information and firms' exporting activities. We document that workers' experience-wage profiles are steeper in exporters than in non-exporters. Aside from self-selection of better firms into exporting, we show that workers' experience-wage profiles are steeper when firms export to industrialized destinations. We discuss possible explanations and propose that this result is likely driven by faster human capital accumulation of workers in firms that export to advanced economies. To support our preferred hypothesis, we use the Enterprise Surveys and document that exporters are more likely to train workers than non-exporters, especially when they adopt foreign technology.

We begin our analysis by applying the standard approach to estimate workers' experience-wage profiles (Mincer, 1974), regressing log hourly wage on dummies of experience bins, schooling, time effects, and individual effects. The well-known challenge is that experience is collinear with time and individual fixed effects, making it impossible to separately identify experience and time effects. To solve this problem, we apply Heckman, Lochner and Taber (1998) approach (HLT) , following Lagakos et al. (2018*b*). The centerpiece of this approach is to assume no experience effects in the final working years, in order to isolate time effects from returns to experience. Applying this approach and controlling for industry composition, we show that after 20 years of experience, workers wage growth is 78% in non-exporters and 96% in exporters, indicating a sizeable difference of 18 percentage points in life-cycle wage growth between exporters and non-exporters.

To understand what drives the difference in experience-wage profiles between exporting and non-exporting firms, we further construct firm-year-level experience-wage profiles based on

[2] Herkenhoff et al. (2018) and Jarosch, Oberfield and Rossi-Hansberg (2018) study the effects of exposure to coworkers, and Gregory (2019) explores the impact of firm-specific human capital accumulation.

the HLT method's assumption that old workers' wage growth purely comes from time effects. We obtain two main results. First, productivity proxies and firm fixed effects explain most of the differences in experience-wage profiles between exporters and non-exporters, hinting that exporters essentially provide higher returns to experience. Second, after controlling for firm size, labor composition, and firm fixed effects, workers life-cycle wage growth is higher in firms exporting to industrialized destinations than in non-exporters, whereas firms exporting to non-industrialized destinations do not enjoy similar increases. We also find that this increase in returns to experience materializes immediately following firms' entry into industrialized destinations, yet it does not show up before entry.

It is possible that destination-specific returns still originate from firms' selection into exporting, as firms may have workforce improvements prior to exporting. To lessen this concern, we conduct an event study using the 1999 currency devaluation episode, which led to a quasi-experimental surge in Brazilian firms' exporting activities. We focus on non-exporters prior to the devaluation shock. We find that firms exporting to industrialized destinations after this devaluation experienced a large jump in their experience-wage profiles after exporting, whereas firms exporting to non-industrialized destinations did not.

We discuss four possible explanations for our destination-specific effect: (1) selection of firms into different export destinations; (2) differential changes in labor composition; (3) job search and screening; and (4) human capital accumulation. Although we cannot entirely rule out other hypotheses, we construct a set of robustness checks and show that faster human capital accumulation when exposed to advanced destinations is the most likely hypothesis.

Anecdotal evidence, based on interviews to leading exporters in Latin America,[3] supports our hypothesis of human capital accumulation. These interviews show that exports to different types of markets imply very different hurdles. As Artopoulos, Friel and Hallak (2010) note,

[3]The interviews were conducted by the Inter-American Development Bank under the project "The Emergence of New Successful Export Activities in Latin America," aiming to provide the experience of some leading exporters in Latin American countries. These studies include exporters in Brazil (Rocha et al., 2008), Argentina (Artopoulos, Friel and Hallak, 2010), Chile (Agosin and Bravo-Ortega, 2009), and Uruguay (Snoeck et al., 2009).

"successfully entering markets in developed economies with differentiated products requires potential exporters to make substantial efforts to upgrade the physical characteristics of their products and to make their marketing practices more sophisticated" (p. 6). With sophisticated technology and demanding customers, firms exporting to advanced destinations often need to invest in the capability of the workforce, in conjunction with specific training institutes or through on-the-job training provided by the firm.

We go beyond those exporters' experiences and provide direct evidence on the relationship between exporting, human capital accumulation and technology adoption, using the World Bank Enterprise Surveys for more than 100 countries. We find that exporters are more likely to offer on-the-job training than non-exporters, after controlling for firm size, industry, country, and year fixed effects. Therefore, human capital accumulation seems to drive at least a portion of the steeper experience-wage profiles in exporters. We also find that exporters which adopt foreign technology are more involved in training workers than exporters who do not. This indicates that destination-specific effects may originate from advanced knowledge that enhances human capital, in line with anecdotal evidence from interviews.

Our analysis of exporting and life-cycle wage growth has important aggregate implications. Through the lens of our empirical results, trade liberalization affects workers' life-cycle earnings growth by reallocating labor toward better firms and exposing workers to advanced destinations. This dynamic effect on workers' earnings, if overlooked, would lead researchers to underestimate the impact of trade liberalization on workers' welfare and income inequality. Moreover, the interaction between life-cycle wage growth and advanced destinations suggests that trade may disproportionately benefit workers' human capital in poor countries, providing support to export promotion policies in those economies.

This paper relates to several strands of the literature. We directly contribute to the large literature on learning by exporting. Recent research shows that through acquiring new knowledge from exporting, firms could improve their technical efficiency (Aw, Chung and Roberts, 2000;

De Loecker, 2013; Atkin, Khandelwal and Osman, 2017) or understanding of export demand (Albornoz et al., 2012; Morales, Sheu and Zahler, 2019). Few studies explore how workers may also acquire knowledge from exporting. Exceptions are Mion and Opromolla (2014) and Muendler and Rauch (2018) who find that employees' previous experience in exporting firms is valuable for their new employers' choices of export markets. In contrast with these studies, we look into how exporting activities affect workers' life-cycle wage growth within the firm. Our results indicate that exporting may enhance workers' human capital, especially with exposure to advanced export destinations.

Second, we make contact with research on trade and workers' earnings. Much empirical work finds wage differences between exporters and non-exporters but abstracts from experience effects (e.g., Bernard and Jensen, 1995).[4] Our evidence shows that the exporter wage premium increases with workers' experience and relies on export destinations. A few recent studies explore how trade openness affects wage growth. Our paper relates to Dix-Carneiro (2014) who estimates industry-specific returns to experience in Brazil to study welfare gains of trade liberalization. Our results imply that between-firm labor reallocation and interacting with export destinations amplify the effects of returns to experience on gains from trade. Our paper also relates to Fajgelbaum (2019) who quantitatively finds higher wage growth in exporters, due to wage renegotiations and increased job surplus after exporting. Our evidence shows that human capital accumulation may also induce higher wage growth in exporters, especially when exposed to advanced destinations.[5]

Third, we relate to research on life-cycle wage profiles. The literature has shown that workers' life-cycle wage growth is heterogeneous across firms, due to factors such as job search

[4]The literature finds that the exporter wage premium is composed of differences in labor composition and wage premia for workers with identical characteristics, including Schank, Schnabel and Wagner (2007), Frias, Kaplan and Verhoogen (2009), and Krishna, Poole and Senses (2014). These existing studies abstract from workers' experience effects.

[5]Fajgelbaum (2019) abstracts from human capital, and the effects of exporting on wage growth rely on wage renegotiations and export revenue. However, we find that export revenue cannot explain our destination-specific effects, suggesting that other factors also matter. Another difference is that Fajgelbaum (2019) abstracts from workers' age, and therefore wage growth may reflect common trends which do not exactly correspond to life-cycle wage growth studied in this paper.

(Bagger et al., 2014), coworkers (Herkenhoff et al., 2018; Jarosch, Oberfield and Rossi-Hansberg, 2018), and firm-specific learning (Gregory, 2019).[6] To our knowledge, our study is the first to empirically study the role of firms' exporting activities. Recent studies highlight the importance of life-cycle wage growth in accounting for cross-country income differences (Lagakos et al., 2018*b*). Our results imply that incentivizing exporting in poor countries may reduce the cross-country income gap.

Finally, we connect with the literature on international knowledge diffusion. Many studies use macro aggregates, such as TFP and R&D, and empirically link trade with knowledge diffusion (e.g., Coe and Helpman, 1995), as reviewed in Wolfgang Keller (2004). Our results highlight that workers human capital accumulation may reflect trade-induced knowledge flows. Recent theoretical papers also explore the relation between trade-induced knowledge diffusion and firm productivity growth (e.g., Alvarez, Buera and Lucas, 2013; Perla, Tonetti and Waugh, 2015; Sampson, 2016; Buera and Oberfield, 2020).[7] Alvarez, Buera and Lucas (2013) and Buera and Oberfield (2020) show that interacting with sellers from more productive countries induces larger knowledge diffusion in domestic markets, whereas our results imply that knowledge diffusion may also originate from exporting to more productive destinations.

This paper is organized as follows. Section 3.2 describes our findings on experience-wage profiles for exporters and non-exporters, and highlights the interaction between wage profiles and export destinations. Section 3.3 exploits the Brazilian currency crisis to address the endogeneity issue of exporting. Section 3.4 discusses possible explanations for the destination-specific effect. Section 3.5 provides evidence on training and foreign technology adoption for exporters, using firm-level data from more than 100 countries. Section 3.6 concludes.

[6]Besides firm-level factors, Islam et al. (2019) show that a lot of factors, such as sectors, occupations, and Internet penetration, are associated with returns to experience.

[7]See Lind and Ramondo (2019) for a review.

3.2 Experience-Wage Profiles and Exporting

In this section, we present a set of stylized facts on how exporting affects experience-wage profiles in Brazil. Section 3.2.1 describes the data, and Section 3.2.2 shows the cross-sectional pattern of experience-wage profiles. Section 3.2.3 discusses the identification challenges and provides our method to apply Mincer regressions to formally estimate experience-wage profiles. Sections 3.2.4 to 3.2.6 report our main findings on differences in experience-wage profiles between exporters and non-exporters, and highlight that experience-wage profiles are steeper when firms export to industrialized destinations.

3.2.1 Data

Our analysis focuses on Brazil, which constitutes a good case study for several reasons. First, Brazil has great data availability and quality, as this subsection shows. Second, the Brazilian case is typical of developing countries, especially in Latin America, and thus our analysis is relevant for policy making. Third, Brazilian exporters sell to a wide range of destinations, allowing the exploration of how export destinations shape experience-wage profiles. For example, in 2010, Brazil's exports were not only directed to high-income countries (10% of total exports sold to the U.S., 25% to Europe, and 4% to Japan), but also to middle-income and low-income countries (23% to Latin America, 15% to China, and 10% to Middle East and Africa). Appendix C.1 describes details of the Brazilian economy, export trends, export products, and destination markets over our sample period.

We rely on the RAIS (Relao Anual de Informaes Sociais of Brazilian labor ministry MTE) database with comprehensive linked employer-employee information in Brazil between 1994–2010. It provides a complete depiction of workers employed in the Brazilian formal sector, because firms are mandated (by law) to annually provide workers' information to RAIS

(Menezes-Filho, Muendler and Ramey, 2008).[8] Each datapoint represents a worker-firm-year observation, containing worker ID, firm ID, and workers information on schooling, age, hourly wage, occupations, and other demographic indicators. These data provide a great laboratory to study returns to experience in the Brazilian formal sector.

One limitation of the data is the absence of information about the informal sector. Therefore, appropriate caution is necessary to interpret our empirical findings from RAIS. Appendix C.2 discusses the characteristics of the Brazilian informal sector and shows that including informal workers in the sample may strengthen our empirical results.[9]

Because we are mainly interested in the interaction between experience-wage profiles and exporting, we restrict our empirical analysis to manufacturing industries, which are tradable and extensively studied in the firm literature. In addition, we focus on full-time male workers aged between 18–65 and employed in firms with the number of employees (including females and part-time workers) larger than 10.[10] If a worker has multiple records in a year, we select the record with the highest hourly wage (Dix-Carneiro, 2014). Under these restrictions, we obtain a sample of 71,748,105 observations between 1994–2010, including 16,629,730 unique worker IDs and 228,890 unique firm IDs.

We use firm IDs to merge the RAIS data with Brazilian customs declarations for merchandise exports collected at SECEX (Secretaria de Comrcio Exterior) for the years 1994–2010, following Aguayo-Tellez, Muendler and Poole (2010). We define a firm as an exporter in a given

[8]The ministry of labor estimates that above 90% of formally employed workers in Brazil were covered by RAIS throughout the 1990s. The data collection is typically concluded by March following the year of observation (Menezes-Filho, Muendler and Ramey, 2008). One benefit of this data is that the reports are substantially accurate. This accuracy stems from the fact that workers' public wage supplements rely on the RAIS information, which encourages workers to check if information is reported correctly by their employers.

[9]Another important limitation is the possible inconsistency in correctly reporting the workers ID number (PIS). Firms may choose to fire and rehire a worker several times throughout any given year to allow the worker to withdraw unemployment benefits multiple times in a single year. This phenomenon may lead companies to incorrectly or repeatedly report a workers ID.

[10]The restrictions on full-time male workers follow Lagakos et al. (2018*b*), due to large changes in female labor participation rate over time. According to the World Bank's estimates for those aged 15+ in Brazil, female labor force participation rate increased from 45% in 1994 to 54% in 2010, whereas male labor force participation rate changed from 81% to 77%. The restriction on firm size aims to avoid the issues of self-employment.

year if the firm has at least one export transaction in that year. Moreover, we divide destinations into industrialized and non-industrialized regions based on the classification provided in Appendix C.3. The SECEX data for the years 1997–2000 also provide firm-level export quantity and value (U.S.$) by 8-digit HS products and destinations. This allows us to measure the structure of export destinations in more detail for these years, which will be used for robustness checks. We discuss more details of the data in Appendix Section C.3.

Table 3.1 describes characterizations of the RAIS database, based on worker-firm-year observations. On average, workers in exporters are slightly older and more educated, earn higher hourly wages, and tend to work in cognitive occupations, relative to non-exporters. Moreover, exporters are much larger in terms of employment size than non-exporters. These pieces of evidence are consistent with the exporter premium typically found in the literature (e.g., Bernard et al., 2003; Verhoogen, 2008). Finally, 49% of workers in the sample stay in exporters, and therefore exporting activities are nontrivial in our sample.

Panel B of Table 3.1 also characterizes dynamic features of the database. Several features stand out. First, the average duration per worker-firm link is 2.78 years, with an average worker working for 1.55 firms in the database. The low duration of the average worker in the sample ($2.78 \times 1.55 = 4.31$ years) is driven by workers switching into industries other than manufacturing or the informal sector, and by young workers entering the workforce in later periods of the sample. Second, we find that 33% of jobs are destroyed after one year—either by workers' switching to another firm or exiting the database entirely. Finally, we also compute firms' transition matrix and find that firms' statuses are stable, with 84% of exporters remaining exporters after one year.[11]

[11]The exit rates for manufacturing firms include the probability of them becoming nonmanufacturing firms or having employment size less than 10. In non-exporters' exit rates (0.11), 0.07 is due to becoming nonmanufacturing firms or having employment size less than 10. In exporters' exit rates (0.05), 0.02 is due to becoming nonmanufacturing firms or having employment size less than 10.

Table 3.1: Sample Statistics

Panel A: Cross-Sectional Characteristics

Observations (72 million)	Non-exporter Mean	Non-exporter S.D.	Exporter Mean	Exporter S.D.
workers' characteristics:				
age	31.97	9.72	32.78	9.39
schooling	8.23	3.46	9.06	3.77
log(hourly wage), Brazilian Real$	0.67	0.77	1.20	0.96
cognitive occupations (1 if yes)	0.20	0.40	0.25	0.43
share of workers in the sample	0.51	–	0.49	–
firms' characteristics:				
log(employment)	4.51	1.56	7.05	1.71
by destinations:				
industrialized regions	–	–	0.06	–
non-industrialized regions	–	–	0.19	–
both types of regions	–	–	0.75	–
log(exports per worker), U.S.$	–	–	8.09	2.31

Panel B: Dynamic Characteristics

Observations (72 million)	Mean	S.D.		Mean
duration of worker-firm links (years)	2.78	2.76		
num of firms per worker	1.55	0.96		
by worker: probability (t to t+1)				
same firm	0.67	–		
different firm	0.06	–		
exit	0.27	–		
by firm: probability (t to t+1)				
non-exporter: to non-exporter	0.84		exporter: to non-exporter	0.11
to exporter	0.05		to exporter	0.84
exit	0.11		exit	0.05

Note: Because Brazil experienced large inflation during the sample period, we adjust log(hourly wage) for inflation using 1994 as the baseline year. The inflation data are drawn from Penn World Table 9.0 (Feenstra, Inklaar and Timmer, 2015). Cognitive occupations refer to professionals, technicians, and other white-collar workers. Firm employment size is computed based on all workers within the firm in the raw sample (including female and part-time workers) to reflect actual firm size. The export value data are only available in 1997–2000, and hence log(exports per worker) are based on these four years. In computing firms' switching probabilities, we weight switching statuses between years t and $t+1$ by firm employment size at year t. This aims to be consistent with workers' and firms' statistics, which are computed based on firm-worker-year observations.

3.2.2 A First Glance at Experience-Wage Profiles

Using the raw data, we first show differences in experience-wage profiles between exporters and non-exporters in the cross section. We measure workers' potential experience as years elapsed since finishing schooling (min{age-18,age-6-educ}). In each year, we obtain experience-wage profiles by computing the average log hourly wage for workers in each 5-year experience bin $x \in X = \{1\text{–}5, 6\text{–}10, ..., 36\text{–}40\}$, separately for workers observed in exporting and non-exporting firms. Because we are interested in life-cycle wage growth, we normalize the value of the first experience bin (1–5 years of experience) to be 0 for each experience-wage profile. Finally, we average profiles across years to obtain experience-wage profiles for exporters and non-exporters, respectively.

Table 3.2: Average Log Wage of Workers with 36–40 Yrs of Exp Relative to 1–5

	(1)	(2)	(3)	(4)	(5)
	Panel A: Aggregate profiles				
				Rel. to non-exporters' first bin	
	all	1994–2000	2001–2010	first bin	40 years of exp
Exporter	0.74	0.67	0.79	0.29	1.04
Non-Exp	0.49	0.48	0.51	0	0.50
Difference	0.25	0.19	0.28	0.29	0.54
	Panel B: Aggregate profiles by education level				
	illiterate	primary	middle school	high school	college
Exporter	0.22	0.69	0.84	1.29	1.43
Non-Exp	0.18	0.46	0.55	0.82	1.08
Difference	0.04	0.23	0.29	0.47	0.35
	Panel C: Aggregate profiles by occupation				
	professionals	technical	other white-collar	Skilled blue-collar	unskilled blue-collar
Exporter	1.10	0.99	0.52	0.57	0.23
Non-Exp	0.85	0.71	0.34	0.44	0.16
Difference	0.25	0.28	0.18	0.13	0.07
	Panel D: Aggregate profiles by firm size				
	10-50	50-100	100-500	500-1000	1000+
Exporter	0.55	0.61	0.69	0.77	0.81
Non-Exp	0.43	0.50	0.59	0.58	0.47
Difference	0.12	0.11	0.10	0.19	0.34

Note: This table reports the average log wage for workers with 36–40 years of experience relative to 1–5 years of experience (normalization). In each year, we obtain experience-wage profiles by computing the average log hourly wage for workers in each 5–year experience bin, separately for workers observed in exporters and non-exporters. We normalize the value of the first experience bin (1–5 years of experience) to be 0 for each experience-wage profile. Finally, we average profiles across years to obtain experience-wage profiles for exporters and non-exporters, respectively. Columns (4)–(5) of Panel A use the average log wage of workers with 1–5 years of experience in non-exporters as normalization.

In Table 3.2, we report the average log wage for workers with 36–40 years of experience relative to 1–5 years of experience (normalization). Column (1) in Panel A shows that, in exporters (non-exporters), the average log wage of workers with 36–40 years of experience is 0.74 (0.49) higher than workers with 1–5 years of experience.[12] This pattern holds in different time periods (Columns (2)–(3)). More notably, it is not caused by lower starting wages of workers in exporters. In the last two columns of Panel A, we recompute the average log wage of each experience bin relative to workers with 1–5 years of experience in non-exporters for any given year. We find that workers with 1–5 years of experience already have higher wages in exporters than in non-exporters. This gap grows larger as workers' experience increases.

In light of potential composition effects (exporters are larger and have better workforce), in Panels B to D of Table 3.2, we recompute the result in Column (1) of Panel A within the same workers' education levels, occupations, or firm size categories. Consistent with recent papers (Islam et al., 2019; Lagakos et al., 2018*b*), we find that the experience-wage profile is steeper for workers with higher education levels (Panel B), in cognitive occupations (Panel C),[13] and in larger firms (Panel D). Moreover, we find that within all of these categories, workers have higher life-cycle wage growth in exporters than in non-exporters.

There are many identification problems with this first-pass attempt: for example, workers observed in exporters in a given year may have previously accumulated working experience in non-exporters in their earlier career. Nonetheless, the preliminary evidence from the raw data indicates that workers in exporters may have steeper experience-wage profiles than workers in non-exporters. With this suggestive pattern in mind, we proceed to formally estimate experience-wage profiles.

[12] Our results are comparable to Lagakos et al. (2018*b*) who use Brazilian Population Census and document that the percent wage increase of 36–40 years of experience relative to 1–5 years of experience is around 60% (see Figure 1 in Lagakos et al. (2018*b*)).

[13] Cognitive occupations refer to professionals, technicians, and other white-collar workers.

3.2.3 Experience-Wage Profiles: Estimation Method

We estimate experience-wage profiles by Mincer regressions, following the labor literature (e.g., Deaton, 1997; Lagakos et al., 2018*b*). We restrict our sample to workers in the same firm for two consecutive years, as there may be imperfect portability of human capital across firms and wage gains/losses related to job separations. We estimate the following regression:

$$\Delta \log(w_{it}) = \sum_{x \in X} \phi_s^x D_{it}^x + \beta_s \Delta e_{it} + (\gamma_{st} - \gamma_{st-1}) + \varepsilon_{it}, \tag{3.1}$$

where i and t represent individuals and years respectively. The subscript s is the level of aggregation for estimating experience effects (e.g., industries, exporters and non-exporters), which will be specified in later implementation. $\Delta \log(w_{it})$ denotes log hourly wage growth from $t-1$ to t for an individual i within the same firm. By using a difference in log hourly wages within the same firm across two periods, we control for individual and firm fixed effects that affect wage levels, as in the employer-employee literature (e.g., Abowd, Kramarz and Margolis, 1999; Card, Heining and Kline, 2013).[14]

D_{it}^x is a dummy variable that takes the value 1 if a worker's potential experience (min{age-18,age-6-educ}) is in group $x \in X = \{1\text{–}5, 6\text{–}10, ...\}$ at time t. The parameter ϕ_s^x measures wage growth for one year of experience accumulated in the experience group x. By avoiding a specific parametric function of experience effects, we allow returns of experience to nonparametrically differ across different stages of the life cycle.[15] We also control for changes in schooling, Δe_{it}, in all our regressions. In addition, γ_{st} represents time effects on wage levels at time t (e.g., TFP, price levels).

Estimating Equation (3.1) faces the well-known collinearity problem regarding experience, individual effects, and time effects in the labor literature (Deaton, 1997). This is easily seen as

[14] Our setting also captures match-specific fixed effects affecting workers' wage levels.

[15] The nonparametric approach of modelling experience effects is commonly used (e.g., Lagakos et al., 2018*b*). Given the large sample size of our data, we choose this approach that allows more precision. Another common way to model experience effects is to assume a quadratic functional form (e.g., de la Roca and Puga, 2017).

$\sum_x D_{it}^x = 1$ is perfectly correlated with the constant $(\gamma_{st} - \gamma_{st-1})$ for each aggregation level s and time t.[16] Intuitively, wage growth over time can be induced by experience or better aggregate economic conditions (e.g., TFP growth). Therefore, to disentangle returns to experience from aggregate trends, we must impose more structure into the model. First, we decompose time effects into trend and cyclical components:

$$\gamma_{st} = g_s t + e_{st}, \tag{3.2}$$

where g_s denotes linear time trends. Specially, we restrict cyclical components to average zero over the time period $\sum_t e_{st} = 0$ and to be orthogonal to the time trend $\sum_t e_{st} t = 0$. These two restrictions resolve the collinearity problem in Equation (3.1) and are also made in Deaton (1997) and Aguiar and Hurst (2013) in estimating life-cycle profiles.

To pin down the wage trend g_s, we adopt the HLT method in Lagakos et al. (2018*b*). The method draws on the basic prediction of a large number of theories of life-cycle wage growth that there are little experience effects in the final working years.[17] Implementing the HLT approach requires assumptions on two parameters: the number of years with no experience effects, and the depreciation rate. Following Lagakos et al. (2018*b*), we consider 10 years at the end of the working life (31–40 years of experience) with no experience effects and a 0% depreciation rate. We conduct our estimation of Equation (3.1) by iterating on g_s until individuals have no experience effects in the last 10 years of their working life.

3.2.4 Experience-Wage Profiles and Export Status

We first apply Equation (3.1) to estimate experience-wage profiles separately for manufacturing workers in exporters and non-exporters between 1994–2010. Figure 3.1 presents the

[16] In other words, the current year and a person's entering year and initial experience pin down their potential experience. The person's entering year and initial experience are captured by individual effects.

[17] See Lagakos et al. (2018*b*) for a detailed description of the method and Rubinstein and Weiss (2006) for a review of theories about life-cycle wage growth.

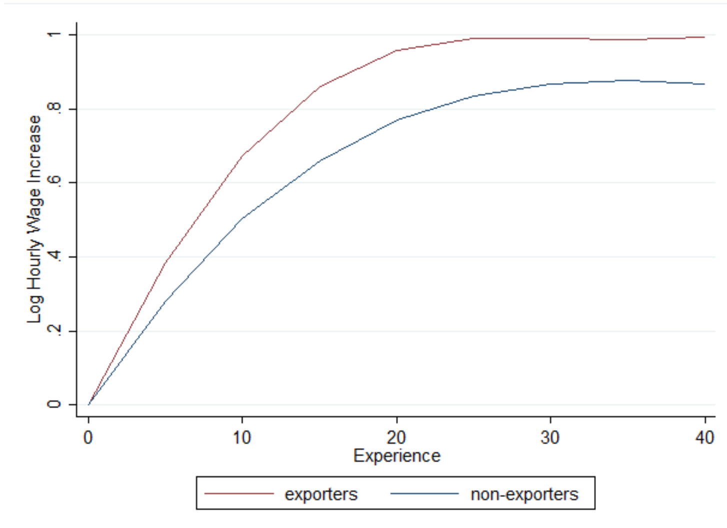


Figure 3.1: Returns to Experience in Exporters and Non-exporters

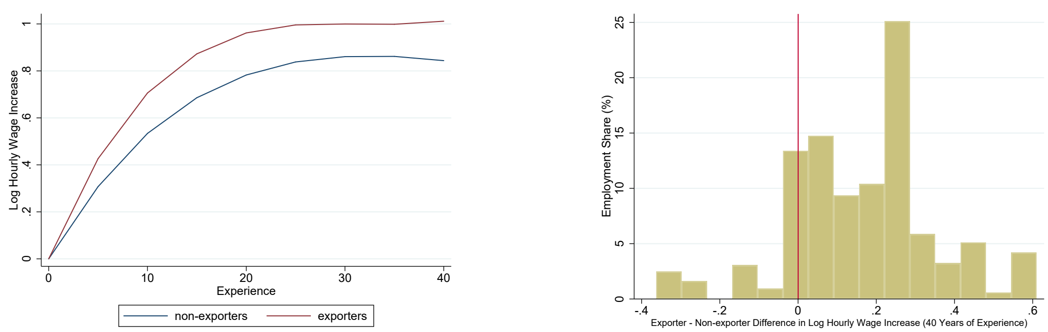
Note: This graph presents experience-wage profiles for exporters and non-exporters, by estimating Equation (3.1) separately for manufacturing workers in exporters and non-exporters between 1994–2010.

log wage growth with regard to potential experience, for a hypothetical person working for 40 years from the beginning of their career. Consistent with the cross-sectional evidence, we find that workers in exporters have a larger life-cycle wage growth: after 40 years of experience, their wage growth is 13 percentage points higher than workers in non-exporters.

Different reasons can explain this difference in experience-wage profiles between exporters and non-exporters. First, an important driver of the result could be industry composition. This is motivated by two well-established results in the literature: (1) different industries have different returns to experience (e.g., Islam et al., 2019); (2) trade induces industry specialization and labor reallocation, possibly driven by comparative advantage (e.g., Costinot, Donaldson and Komunjer, 2012) or home market effects (e.g., Head and Ries, 2001). Therefore, if exporters are more concentrated in industries with higher returns to experience than non-exporters, exporters will also have steeper experience-wage profiles.

In Appendix Section C.4.1, we examine in detail the role of industry composition in driving the difference of experience-wage profiles between exporters and non-exporters. We find a large degree of heterogeneity in returns to experience across industries, indicating that trade-induced labor reallocation could potentially have a large impact on the aggregate returns to experience. However, for Brazil, exporters are more concentrated in industries with lower returns to experience than non-exporters, and therefore industry composition cannot explain the aggregate difference in returns to experience between exporters and non-exporters.[18]

As industry composition cannot explain our results, the difference in returns to experience between exporters and non-exporters must be driven by firm-level differences within industries. To explore this, we estimate Equation (3.1) separately for workers within exporters and non-exporters, for each 3-digit industry. For precision, we focus on industries with more than 0.1% of total employment and require at least 10 workers in each year-experience-bin (separately for exporters and non-exporters). This leaves us with 78 industries with estimated experience-wage profiles for both exporters and non-exporters, and these industries represent 96% of manufacturing employment in the sample.



(a) Within-industry Experience-wage Profile

(b) Distribution of Differences

Figure 3.2: Log Hourly Wage Increase by Exporters and Non-exporters

Note: This figure presents the results from estimating Equation (3.1), separately for workers within exporters and non-exporters in each 3-digit industry between 1994–2010. Panel (a) is the (employment-weighted) within-industry experience-wage profiles for workers in exporters and non-exporters, where the weight reflects industry-level employment. Panel (b) is the cross-industry distribution of within-industry differences in returns to 40 years of experience between exporters and non-exporters.

Figure 3.2a plots the (employment-weighted) within-industry experience-wage profiles for workers in exporters and non-exporters. To avoid effects of industry composition, we apply

[18]We estimate experience-wage profiles separately for workers in each 3-digit manufacturing industry. We find that after 40 years of experience, workers' wage growth would be 2 percentage points lower in exporters than in non-exporters because of the difference in employment distributions across industries. This pattern is consistent with Brazil's comparative advantage in low-tech products (see Table 3.3 in Bonelli and Pinheiro (2008)) and that returns to experience may increase with technology levels (see Table 5 in Islam et al. (2019)).

identical weights (total industry-level employment) to construct profiles for exporters and non-exporters. For a hypothetical person working for 40 years from the beginning of their career, the life-cycle wage growth is 16 percentage points higher in exporters than in non-exporters, and 71% of this difference is achieved within the first 5 years of experience.

Figure 3.2b shows the cross-industry distribution of within-industry differences in returns to 40 years of experience between exporters and non-exporters. We find that experience-wage profiles are steeper in exporters than in non-exporters for 85% of industries, which account for 89% of manufacturing employment in the sample.

Therefore, within-industry factors drive the difference in experience-wage profiles between exporters and non-exporters. We documented in Table 3.1 that exporters are larger and have larger shares of cognitive and educated workers. The pattern in Figure 3.2a could partly reflect workforce composition and selection of firms into exporting. Moreover, additional benefits from exporting may occur due to increased revenues or interactions with destination markets. Thus, in the following subsection, we investigate how the difference in life-cycle wage growth between exporters and non-exporters is driven by differences in firms' characteristics, export status, and the interaction with different destination markets.

3.2.5 Firm-level Wage Profiles and Export Destinations

This subsection aims to understand plausible drivers of the differences in returns to experience between exporters and non-exporters. To make progress, we construct firm-year-level returns to experience in each experience bin as follows:

$$\phi^x_{\omega,t} = \frac{\sum_{i\in\omega} D^x_{it}\Delta\log(w_{it})}{\sum_{i\in\omega} D^x_{it}} - \frac{1}{2}\left(\frac{\sum_{i\in\omega} D^{31-35}_{it}\Delta\log(w_{it})}{\sum_{i\in\omega} D^{31-35}_{it}} + \frac{\sum_{i\in\omega} D^{36-40}_{it}\Delta\log(w_{it})}{\sum_{i\in\omega} D^{36-40}_{it}}\right). \tag{3.3}$$

$\frac{\sum_{i\in\omega} D^x_{it}\Delta\log(w_{it})}{\sum_{i\in\omega} D^x_{it}}$ represents the average individual-level log hourly wage growth between year $t-1$ and t, for workers in firm ω in both periods and in experience bin $x \in X = \{1\text{–}5,...,36\text{–}40\}$. The second term reflects the average of log wage growth for workers within firm ω and in the last two experience bins. This term aims to capture firm-specific wage trends, based on the same assumption of the HLT approach that there are no experience effects in the last 10 years of the working life.[19]

By applying Equation (3.3), we not only control for firm, individual, and match-specific fixed effects that affect workers' wage levels, but also capture time-variant conditions (e.g., TFP growth, and supply and demand shocks of products) that alter wages for all workers within the firm. For instance, if the firm raises all workers' wage by the same proportion due to increased revenue or upgraded technology after exporting, this effect will not show up in Equation (3.3). However, if the wage growth is relatively higher for young workers than old workers, this relative difference in wage growth is interpreted as reflecting returns to experience. Section 3.4 discusses possible causes for this difference and connects the empirical results with existing theory.

In Table 3.3, we regress firm-year-level returns to 20 years of experience on firm characteristics. The dependent variable corresponds to $5 \times \sum_{x\in\{1-5,\ldots,16-20\}} \phi^{x}_{\omega,t}$. The variable refers to the hypothetical life-cycle wage growth of a worker staying in firm ω for 20 years from the beginning of their career, with returns to experience fixed at time t. This variable provides a measure of time-variant firm-level returns to experience. We choose to report returns to 20 years of experience, because many firms do not have workers in all experience bins. This choice is also motivated by Figure 3.2a showing that workers have little returns to experience after 20 years of experience.[20] To lessen the effect of extreme values, we truncate the sample by dropping the highest and lowest 1% of the dependent variable.

In Column (1) of Table 3.3, the independent variables are exporter dummies by destinations and a set of industry and year fixed effects. The baseline group is non-exporters. We find

[19] If only one term of $\frac{\sum_{i\in\omega} D^{31-35}_{it}\Delta\log(w_{it})}{\sum_{i\in\omega} D^{31-35}_{it}}$ and $\frac{\sum_{i\in\omega} D^{36-40}_{it}\Delta\log(w_{it})}{\sum_{i\in\omega} D^{36-40}_{it}}$ exists, we use the existing one to construct firm-specific wage trends.

[20] This result is partly due to our use of the HLT approach, which assumes zero experience effects in the 31–40 years of experience. Our results are robust to using other ranges of potential experience to construct dependent variables.

Table 3.3: Firm-year-level Log Hourly Wage Increase (20 Years of Experience)

	(1)	(2)	(3)	(4)
Exporter, non-industrialized dests	0.212***	0.101***	-0.007	-0.020
	(0.014)	(0.014)	(0.023)	(0.023)
Exporter, industrialized dests	0.225***	0.114***	0.083**	0.071**
	(0.025)	(0.025)	(0.036)	(0.036)
Exporter, both types of dests	0.315***	0.097***	0.070***	0.046*
	(0.012)	(0.014)	(0.027)	(0.027)
Log(firm employment)		0.103***		0.085***
		(0.004)		(0.014)
Share of high-school grads		0.238***		0.045
		(0.019)		(0.042)
Share of cognitive occupations		0.310***		0.172***
		(0.027)		(0.057)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	Yes
Obs	361,850	361,850	361,850	361,850
R-squared	0.014	0.017	0.299	0.299

Note: This table presents estimates from regressions of firm-year-level log hourly wage increase after 20 years of experience on firm characteristics for the period 1994–2010. The baseline group is non-exporters. The shares of high-school graduates and cognitive workers in the workforce are computed based on our restricted sample, from which we obtained our estimates of firm-year-level experience-wage profiles. Cognitive occupations refer to professionals, technicians, and other white-collar workers. Firm employment size is computed based on all workers within the firm in the raw sample (including female and part-time workers) to reflect actual firm size. Notably, our results are quantitatively very similar if we use our restricted sample (full-time male workers) to compute firm employment size. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

that after 20 years of experience, workers' wage increase is 21–31 percentage points higher in exporters than in non-exporters. These numbers are comparable in magnitude to the within-industry difference found from Mincer regressions (Figure 3.2a)—18 percentage points after 20 years of experience.

In Column (2), we further control for the shares of high-school graduates and cognitive workers in the firm's workforce. This allows us to capture labor composition effects, because cognitive and more educated workers have steeper experience-wage profiles (Islam et al., 2019; Lagakos et al., 2018*b*). In addition, we control for firm employment size, which proxies for a firm's productivity level as productive firms hire more (Hopenhayn, 1992). As expected, experience-wage profiles are higher in larger firms, or firms with more cognitive and educated workers. However, after including these controls, the resulting exporters' premium in returns to experience is almost halved. By taking coefficients of these controls in Column (2) of Table 3.3 and differences in controls between exporters and non-exporters shown in Table 3.4,[21] we find that firm size is the most important factor in explaining the drop in exporters' premium between Columns (1) and (2). For exporters exporting to both industrialized and non-industrialized destinations, firm size explains 76% of the difference ($0.315 - 0.097$), the share of high-school graduates explains 14%, and the share of cognitive workers explains 10%.

The large effect of firm size suggests the importance of firm productivity in affecting firm-level returns to experience. However, firm size may not entirely reflect firm productivity, if labor markets are frictional (Meghir, Narita and Robin, 2015) or productivity partly reflects product quality (Lentz and Mortensen, 2008). Considering this, we further control for firm fixed effects in Columns (3) and (4), capturing time-invariant unobserved firm productivity levels and other characteristics. By introducing firm fixed effects, we are using firms that switch export status to identify exporters' premium in returns to experience. Surprisingly, exporting to non-industrialized destinations now leads to insignificant changes in returns to experience, whereas exporting to industrialized destinations results in statistically significant and positive gains. Consistently, exporting to both types of destinations has positive (yet lower) gains than solely exporting to industrialized destinations. We find similar results when we add back other controls in Column (4), but nonetheless these controls are less important in affecting exporters' premium in the presence of firm fixed effects.

In Appendix Table C.4, we exploit firm-level data on export value by destinations, which are available for the 1997–2000 period, for robustness. We measure a firm's exposure to industrialized destinations by a continuous variable: the share of exports to industrialized destinations in its total exports. We regress firm-year-level returns to 20 years of experience on an exporter dummy, the share of firms' exports to industrialized destinations, and identical controls as in Table

[21] We use the first four rows of Table 3.4.

Table 3.4: Difference in Firm Characteristics (Relative to Non-exporters)

	Log(emp)	Share of high-school grads	Share of cogn occs
without Firm FE:			
Exporter, non-industrialized dests	0.818 (0.006)	0.058 (0.001)	0.042 (0.001)
Exporter, industrialized dests	0.800 (0.010)	0.062 (0.002)	0.046 (0.002)
Exporter, both dests	1.610 (0.006)	0.126 (0.001)	0.069 (0.001)
with Firm FE:			
Exporter, non-industrialized dests	0.143 (0.004)	0.008 (0.001)	0.004 (0.001)
Exporter, industrialized dests	0.136 (0.007)	0.013 (0.002)	0.002 (0.001)
Exporter, both types of dests	0.264 (0.006)	0.021 (0.001)	-0.001 (0.001)

Note: This table presents coefficients on exporter dummies, from regressions of firm-year-level characteristics on exporter dummies by destinations and a set of year and industry fixed effects. The baseline group is non-exporters. Cognitive occupations refer to professionals, technicians, and other white-collar workers. Firm employment size is computed based on all workers within the firm in the raw sample (including female and part-time workers) to reflect actual firm size. Robust standard errors are in parentheses.

3.3. We also control for export value per worker, as destination-specific effects may originate from increased revenue due to exporting. We find that after controlling for firm fixed effects, labor composition, and firm size, exporter status and export value do not affect returns to experience, whereas higher shares of exports to industrialized destinations significantly increase returns to experience. Appendix Table C.5 finds similar results, using export-weighted GDP per capita across destinations as a measure of exposure to industrialized destinations.

Before providing a detailed review of plausible causes for destination-specific effects in Section 3.4, we show more supportive evidence for the existence of these effects.

3.2.6 Dynamics of Exporting to Industrialized Destinations

In this subsection, we construct an event study on the dynamics of experience-wage profiles. We study whether changes in returns to experience—due to exporting to industrialized destinations—materialize immediately when firms start exporting. In particular, we perform the following regression:

$$y_{\omega,t} = \sum_{\tau=-4}^{\tau=-2} \beta_\tau 1\{industrial\}_{\omega,t^*+\tau} + \sum_{\tau=0}^{\tau=4} \beta_\tau 1\{industrial\}_{\omega,t^*+\tau} + \beta_{pre} \sum_{\tau\leq-5} 1\{industrial\}_{\omega,t^*+\tau}$$
$$+ \beta_{post} \sum_{\tau\geq5} 1\{industrial\}_{\omega,t^*+\tau} + \mathbf{X}'_{\omega,t}\mathbf{b} + \theta_\omega + \psi_{j(\omega,t)} + \delta_t + \varepsilon_{\omega,t}. \tag{3.4}$$

As before, the dependent variable is firm-year-level returns to 20 years of experience: $y_{\omega,t} = 5 \times \sum_{x\in\{1-5,\ldots,16-20\}} \phi^x_{\omega,t}$. In the regression, we control for firm fixed effects θ_ω, industry fixed effects $\psi_{j(\omega,t)}$, and year fixed effects δ_t. Firm-level controls $\mathbf{X}_{\omega,t}$ include the shares of high-school graduates and cognitive workers, firm size, and a dummy variable indicating whether the firm is exporting to a non-industrialized destination.

The β_τ parameters of primary interest are coefficients on indicators for time periods relative to the firm's first export entry into industrialized destinations ($\tau = 0$).[22] We exclude an indicator for the period immediately before the firm's export entry into industrialized markets, and hence the parameters represent changes in returns to experience relative to that period. The coefficients are identified by firms starting as non-exporters or exporters only to non-industrialized destinations and then turning to export to industrialized destinations in our sample period. For the β_τ parameters after entry, we also require that firms remain exporting to industrialized destinations, and therefore β_τ (for $\tau > 0$) is interpreted as changes in returns to experience for a firm that still exports to industrialized destinations in τ periods after first entry. We are aware that this regression could potentially suffer from selection bias, as those better-performing firms may choose to start exporting (Fajgelbaum, 2019). Still, it is a good exercise to understand the dynamics of experience-wage profiles before and after firms export to industrialized destinations.

Figure 3.3 presents the results from estimating Equation (3.4). After first entry into industrialized destinations, a significantly positive jump occurs in firms' experience-wage profiles,

[22] We focus on firms that do not start as exporters to industrialized destinations when they make first appearance in the sample, but experience entry into industrialized destinations later.

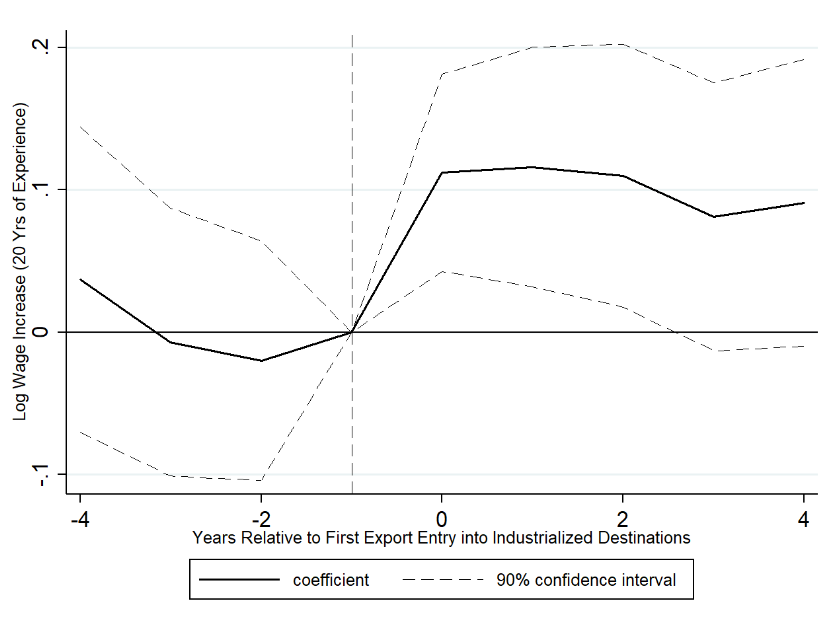
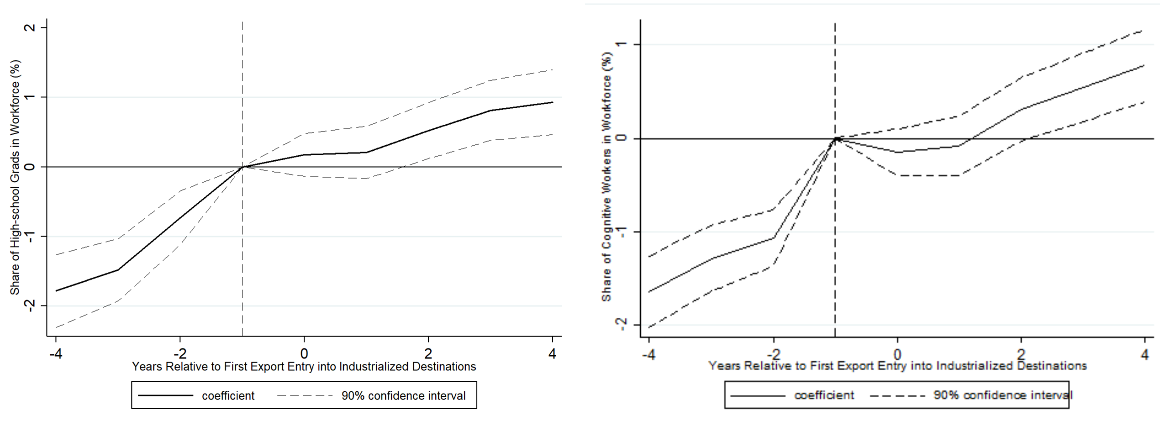


Figure 3.3: Dynamics of Firms' First Entry Into Industrialized Destinations (Survivors)

Note: The figure shows the β_τ parameters from estimating Equation (3.4). The dependent variable is firm-year-level returns to 20 years of experience. The regression controls for firm fixed effects, industry fixed effects, year fixed effects, the shares of high-school graduates and cognitive workers in the workforce, firm size, and a dummy variable indicating whether the firm is exporting to a non-industrialized destination. To estimate the β_τ parameters after entry, we require that firms remain exporting to industrialized destinations.

whereas experience-wage profiles do not significantly shift before firms' export entry. In addition, the increase in returns to experience stays roughly constant after entry, indicating that exporting to industrialized destinations is associated with persistent higher life-cycle wage growth. Appendix Figure C.7 shows the results from the same regression, but we do not enforce a requirement that firms remain exporting to industrialized destinations after entry to identify $\beta_\tau, \tau > 0$. We find in that case that the gains in experience-wage profiles tend to decline several years after firms' entry into industrialized destinations, as some firms gradually quit exporting to industrialized destinations. Appendix Figure C.8 estimates the β_τ parameters for the firm's first export entry into non-industrialized destinations at time $t = t^*$ ($\tau = 0$). We find no statistically significant change in returns to experience before and after firms export to non-industrialized destinations.

Finally, in Figure 3.4, we assign the shares of high-school graduates and cognitive workers as dependent variables in Equation (3.4) to analyze the dynamics of labor composition around firms entry into industrialized destinations. We find that firms gradually improve their labor



(a) Share of High-school Grads (%)

(b) Share of Cognitive Occupations (%)

Figure 3.4: Dynamics of Firms' First Entry Into Industrialized Destinations

Note: The figure shows the β_τ parameters from estimating Equation (3.4). The dependent variable is the share of high-school graduates in the workforce in Panel (a) and the share of cognitive workers in the workforce in Panel (b). All regressions control for firm fixed effects, industry fixed effects, year fixed effects, and a dummy variable indicating whether the firm is exporting to a non-industrialized destination. To estimating the β_τ parameters after entry, we require that firms remain exporting to industrialized destinations.

composition even before exporting, and this gradual improvement continues after export entry. This suggests that export entry may be endogenous. Nevertheless, their labor composition does not significantly change immediately after export entry, whereas we observe the immediate jump in life-cycle wage growth after firms' entry into industrialized destinations. This suggests that the jump in returns to experience is more likely driven by changes in export status rather than changes in labor composition. In the following section, we rely on a quasi-experiment to address the endogeneity problem in the exporting decision.

3.3 Case Study: Brazilian Currency Crisis

From our previous analysis, it is possible that the estimated destination-specific returns to experience may still reflect firms' selection into exporting, as exporting firms may experience improvements prior to exporting. To corroborate our argument that the destination-specific effects are shaped by exporting activities, this section describes an event study using the 1999 currency devaluation, which led to a quasi-experimental surge in exporting activities.

In January and February 1999, Brazil experienced a massive devaluation of its domestic currency, with the Brazilian Real per U.S. dollar increasing from 1.20 in December 1998 to 1.93 in February 1999, a 60% devaluation within two months.[23] The abrupt currency devaluation was detrimental to the economy in many ways, but nonetheless it improved Brazilian firms' competitiveness in the global market and induced more firms to export. In Figure 3.5b, we show that the probability of firms exporting strongly increased after 1999 (relative to year 1998, after controlling for firm fixed effects and industry fixed effects), while there was no effect in the year prior to the large devaluation episode and a small increase in the previous years. Similarly, Verhoogen (2008) finds that the Mexican peso crisis in 1994 led to more firms' entry into exporting.

We exploit this large devaluation episode and apply a difference-in-difference approach to analyze how exporting affects experience-wage profiles due to exogenous shifts (from individual firms' perspective) in exporting opportunities. We perform the following regression:

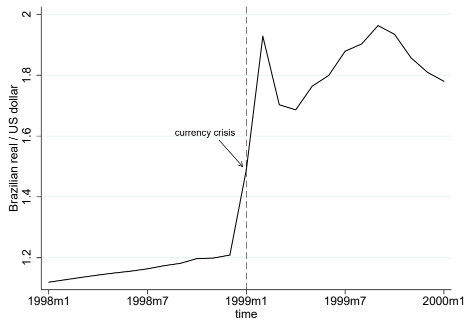
$$y_{\omega,t} = \sum_{d \in D} \beta_d \times 1\{d\}_t \times 1\{post\_1999\} + 1\{post\_1999\} + \theta_\omega + \mathbf{X}'_{\omega,t}\mathbf{b} + \psi_{j(\omega,t)} + \delta_t + \varepsilon_{\omega,t}. \quad (3.5)$$

The dependent variable is still firm-year-level returns to 20 years of experience: $y_{\omega,t} = 5 \times \sum_{x \in \{1-5,\ldots,16-20\}} \phi^x_{\omega,t}$. We control for firm fixed effects θ_ω, industry fixed effects $\psi_{j(\omega,t)}$, and year fixed effects δ_t. Firm-level controls $\mathbf{X}_{\omega,t}$ include the share of high-school graduates, the share of workers in cognitive occupations, and firm size. β_d captures changes in experience-wage profiles if firms started to export to destination d after the devaluation episode.[24] We estimate this regression on the set of non-exporters[25] before the Brazilian currency crisis.

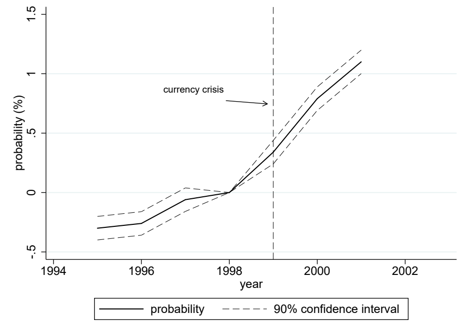
[23]The devaluation came as a surprise, and many factors may have led to this crisis. Many economists believed that the crisis had roots in the financial turmoil following the Asian financial crisis and fundamental problems of the Brazilian economy (such as budget and current account deficits). For a thorough discussion of the Brazilian currency crisis, see https://www.nber.org/crisis/brazil_report.html.

[24]The set of destinations is denoted as D.

[25]Specifically, we focus on firms that ran business (for at least one year), yet did not export during the 1996–1998 period.



(a) Brazil Currency Crisis



(b) Exporting Probability

Figure 3.5: Brazil Currency Crisis and Exporting Probability

Note: Panel (a) presents the monthly Brazilian nominal exchange rates (per U.S. dollar), which are drawn from https://fxtop.com/. Panel (b) presents the probability of a firm exporting in each year. To obtain the probability, we regress the dummy variable of the export status (1, if the firm exports, and otherwise 0) on firm fixed effects, industry fixed effects, and year fixed effects. We plot the coefficients on year effects relative to 1998 (the baseline year) in Panel (b).

In this difference-in-difference design, we impose two implicit assumptions for identification: (1) most changes in firms' export status after 1999 were due to improved competitiveness with currency devaluation; (2) this currency devaluation affected returns to experience through changes in exporting activities, but was uncorrelated with other factors that shift returns to experience. These assumptions are more likely to be true within a narrow time frame of the currency crisis; therefore, we estimate Equation (3.5) using the observations within 1–3 years around the episode year, 1999.

Table 3.5 presents the results. Regardless of the chosen time frame, the results show that firms which started exporting to industrialized destinations after currency devaluation saw increases in experience-wage profiles. On the other hand, the coefficients for firms exporting to non-industrialized destinations after the devaluation are not significant.

Moreover, in Appendix Tables C.8 and C.7, we assign the shares of cognitive workers and high-school graduates in the workforce as dependent variables. We show that within a year around the shock (between 1998 and 2000), no significant change occurred in the labor composition for firms exporting to industrialized destinations, whereas an improvement occurred in the labor composition (in terms of the share of high-school graduates) for firms that started exporting

Table 3.5: Firm-year-level Log Hourly Wage Increase (20 Years of Experience)

time	(1) 1998-2000	(2) 1997-2001	(3) 1996-2002
1{export to industrialized dests}	0.422*	0.392**	0.277**
×1{post_1999}	(0.236)	(0.164)	(0.124)
1{export to non-industrialized dests}	-0.076	-0.048	-0.106
× 1{post_1999}	(0.156)	(0.107)	(0.085)
1{export to both types of dests}	0.387	0.111	-0.047
×1{post_1999}	(0.377)	(0.220)	(0.159)
Year, industry and firm FE	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes
Obs	37,267	61,390	85,266
R-squared	0.563	0.446	0.382

Note: This table presents estimates from Equation (3.5). The dependent variable is firm-year-level log hourly wage increase after 20 years of experience. The regression includes firm, industry, and year fixed effects. Firm-level controls include the shares of high-school graduates and cognitive occupations, and firm size. Cognitive occupations refer to professionals, technicians, and other white-collar workers. Firm employment size is computed based on all workers within the firm in the raw sample (including female and part-time workers) to reflect actual firm size. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

to non-industrialized destinations. Therefore, consistent with our explanation in the previous subsection, improvement of labor composition is unlikely to explain increases in experience-wage profiles for exporters to industrialized destinations.

We are aware of two possible caveats to our approach. First, firms that started to export after currency devaluation were "marginal exporters," in the sense that they were close to export thresholds. Therefore, if our identification assumptions hold, our estimated effects actually capture "local average treatment effects." Second, although this experiment addresses the endogeneity of exporting, selection of destinations may still occur, as firms may choose different destinations after the devaluation due to time-variant factors (not captured by firm fixed effects), such as product quality (Manova and Zhang, 2012). We will discuss selection of destinations in the next section and show that it is unlikely to cause destination-specific returns to experience.

3.4 Connecting Destination-Specific Results with Theory

In this section, we briefly discuss four plausible explanations for our finding on the interaction between returns to experience and different destinations: (1) selection of firms into different export destinations; (2) differential changes in labor composition; (3) job search and screening; and (4) human capital accumulation. We propose that faster human capital accumulation from exposure to advanced countries is the most likely explanation. Appendix Section C.5 provides the detailed procedure and results of robustness checks.

Selection of Firms into Different Export Destinations. Our first hypothesis states that firms exporting to industrialized destinations are better than other exporters due to factors not captured by firm fixed effects, or they enjoy more favorable linkages with destinations. We argue that this is unlikely to explain our destination-specific effects. First, as Table 3.4 shows, firms exporting to both types of destinations appear to be the most productive even after controlling for firm fixed effects, as they are the biggest and have the largest shares of high-school graduates. Nonetheless, they do not enjoy the largest increase in experience-wage profiles after exporting. Second, our results in Column (4) of Table 3.3 remain similar after controlling for unit prices of exports,[26] as a proxy for product quality (Manova and Zhang, 2012). Finally, our results in Column (4) of Table 3.3 remain unchanged, after controlling for industry-year fixed effects or gravity variables (e.g., bilateral distance, sharing a language). Therefore, industry-year-level common shocks or bilateral linkages of destinations with Brazil cannot capture destination-specific returns to experience.

Differential Changes in Labor Composition. The second plausible hypothesis states that changes in labor composition after exporting favor exporters to industrialized destinations. We argue that this may not be the case for the following reasons. First, as Table 3.4 shows,

[26]The firm-level export value and quantity are available by destinations and 8-digit HS products in 1997–2000. We take an export-weighted average of unit prices across destinations and HS products to construct firm-year-level unit prices of exports. Given the heterogeneity in values of HS products, we experimented with first normalizing the unit price by the unit price of the same HS product exported from Brazil to the U.S.. The results remain very similar under this normalization.

firms exporting to both types of destinations have the best workforce among all firms, and their workforce become more educated after exporting. Changes in labor composition may favor firms exporting to both types of destinations, but nonetheless, firms exporting to industrialized destinations perceive the largest increase in returns to experience.

There could still be unobserved workers' characteristics not captured by education levels and occupations. We undertake two sets of robustness checks. First, we construct a proxy for workers' unobserved ability, using the residual wage of their first appearance in the sample, after removing year and age effects. Controlling for the average ability of the workforce does not change our results in Column (4) of Table 3.3. In addition, when we compute firm-year-level profiles in year t, we use workers employed within the same firm in both years $t-1$ and t. If current workers are unaware of whether firms would change export status in one year, we could compare profiles for firm-year-level observations in years $t-1$ and t with a switch in export status between those years. We rerun our regression in Column (4) of Table 3.3 with these observations around switches and find similar results.

Job Search and Screening. Our third hypothesis states that the observed destination-specific effects are due to job search and screening. Though we focused on workers staying in the same firm to construct firm-level wage profiles, workers' wage growth may still result from changes in job surplus and wage renegotiations in the presence of on-the-job search and firms' monopsony power.[27] Alternatively, workers' wage growth may originate from screening when information frictions occur (Jovanovic, 1979).[28] Moreover, given initial uncertainty about workers' abilities, exporters may offer back-loaded wage contracts.

[27]For example, in a calibrated model with wage bargaining like Cahuc, Postel-Vinay and Robin (2006), Fajgelbaum (2019) shows that workers in potential exporters experience faster wage growth due to wage renegotiations and larger job surplus after exporting. Our destination-specific results may thus arise due to larger surplus from exporting to industrialized destinations. Acemoglu and Pischke (1998) argue that firms monopsony power on workers ability information affects firms wage determination. Through the lens of their model, our results may arise if firms exporting to industrialized destinations have the least monopsony power and therefore design the steepest experience-wage profiles to avoid poaching from other firms.

[28]In particular, larger job surplus after exporting may interact with screening based on workers' abilities (Helpman et al., 2017) or match-specific quality (Donovan, Lu and Schoellman, 2020), leading to different patterns of worker turnover and our observed experience-wage profiles.

We cannot entirely rule out numerous stories in the literature, but nonetheless we provide several checks to show that job search and screening are unlikely to explain destination-specific effects. First, as shown in Section 3.2.5, export value per worker does not affect returns to experience, indicating that changes in job surplus may not trigger destination-specific shifts in returns to experience.[29] Second, we find that exporting to industrialized destinations leads to higher returns to experience in more manual or less skill-intensive industries, where workers may have lower bargaining power. Third, as workers' tenure can be used as a proxy for firms' monopsony power, we control for workers' average tenure, which does not change our results in Column (4) of Table 3.3. Finally, as shown in Section 3.2.6 and 3.3, the jump in experience-wage profiles happens immediately after entry into industrialized destinations. If exporters offer back-loaded wage contracts, we must expect an initial decline in experience-wage profiles after switching to exporting.

Human Capital Accumulation and Knowledge Diffusion. There is a long tradition, starting with Becker (1964), using experience-wage profiles to implicitly measure human capital accumulation (e.g., Caselli, 2005; Manuelli and Seshadri, 2014). Clearly, one potential way to interpret our destination-specific results is through human capital theory. In addition, the literature argues that knowledge diffusion is central to human capital accumulation (e.g., Lucas and Moll, 2014), and that trade transmits knowledge across borders (e.g., Buera and Oberfield, 2020).

Our destination-specific returns to experience are consistent with faster human capital accumulation due to exposure to advanced countries. First, workers enjoy steeper experience-wage profiles if firms export to industrialized destinations, in line with larger knowledge diffusion from trading with more advanced destinations (Alvarez, Buera and Lucas, 2013; Buera and Oberfield, 2020). Moreover, increases in returns to experience from industrialized destinations

[29] Even if we control for export value per worker, job surplus may still be higher if firms exporting to industrialized destinations enjoy higher markups than other firms. There is not much evidence on it. If any, Keller and Yeaple (2020) find that the markups of U.S. multinationals' affiliates decline with the GDP per capita of the host country. De Loecker and Eeckhout (2020) estimate the aggregate markup across countries, and there is no clear relationship between markups and countries' development levels.

are larger in industries with smaller shares of high-skill and cognitive workers. This is compatible with the theory that the least productive agents typically enjoy the largest gains in human capital from knowledge diffusion (Lucas and Moll, 2014). Third, increases in returns to experience due to industrialized destinations are larger in industries with more differentiated goods, which might be associated with larger benefits for knowledge adoption.

Therefore, although we cannot entirely rule out other hypotheses, we propose that human capital accumulation due to knowledge diffusion is the most likely hypothesis to explain destination-specific returns to experience.

3.5 Exporting, Training and Technology Adoption

The Inter-American Development Bank has conducted a series of case studies on leading exporters in Brazil and other Latin American countries.[30] One consistent finding is that exporting to industrialized destinations usually requires the adoption of more sophisticated production technology, which often induces these exporters to invest in the capability of the workforce by providing training.[31] This finding supports our preferred hypothesis and indicates that the adoption of advanced technology may be the driver of faster human capital accumulation in firms that export to industrialized destinations.

In this section, we go beyond the anecdotal evidence and employ the World Bank Enterprise Surveys to provide direct evidence on workers' human capital accumulation in non-exporters and exporters. The enterprise survey (ES) is a firm-level survey of a representative sample of

[30] Apart from Brazil (Rocha et al., 2008), studies for other Latin American countries also exist, including Argentina (Artopoulos, Friel and Hallak, 2010), Chile (Agosin and Bravo-Ortega, 2009) and Uruguay (Snoeck et al., 2009).

[31] A good example is Artefama, the largest exporter of wood furniture in Brazil in 2006. To export to Europe and the U.S., this company imported production machinery from Italy, invested in new equipment to dry wood, and adopted electronic control mechanisms, since the domestic markets did not require the same standards as export markets. In the meantime, the company offered a special in-house two-year training program to its workers, which inadvertently benefited other firms. As Rocha et al. (2008) describe, "Artefama, although unwillingly, supplied skilled workers to fulfill the needs of other firms in the region, and stimulated the appearance of new entrepreneurs among their own employees" (p. 59).

an economy's private manufacturing and service industries,[32] covering around 100 countries (mostly low- and middle-income). These surveys provide two standardized waves conducted in 2002–2005 and 2006–2017 and cover a variety of topics such as firms' financial information, business environment, infrastructure, technology adoption, and on-the-job training. Owners and top managers usually answer the ES. This survey includes 1200–1800 interviews in large economies, 360 in medium-sized economies, and 150 in small economies. Finally, firms with fewer than 5 employees are usually omitted, and firms with 100% government/state ownership are not eligible to participate.

The ES provides a set of standardized questions that allow for cross-country comparison. We rely on those standardized questions. Appendix Section C.6 provides the details of the questions we use. Questions on training investments, exports, and R&D investments are recorded based on firms' activities in the last fiscal year, and the question on foreign technology adoption refers to the firm's technology in the current year. In all regressions of this section, we include country, year, and industry fixed effects. We also control for firm size, labor share, managerial experience in the industry, and the share of high-school graduates in the workforce. These control variables are computed using firms' information in the last fiscal year. Appendix Table C.11 presents the summary statistics of the variables we use in the ES. Consistent with the exporter premium found in the literature (Bernard et al., 2003), exporters have larger employment size, more experienced managers, and higher shares of educated workers in the workforce than non-exporters.

The ES asks each firm if it provides formal on-the-job training to its permanent workers. Therefore, we investigate if exporters provide more on-the-job training than non-exporters, which is direct evidence of differential human capital investments between exporters and non-exporters.

[32]The ES interviews formal firms in manufacturing and service industries (ISIC codes 15–37, 45, 50–52, 55, 60–64, and 72, ISIC Rev. 3.1). This survey has two types of questionnaires, one for manufacturing firms and one for service firms, which have questions in common for some topics and specific questions for others. The ES uses a stratified random sampling, which means that firms are grouped according to firm size, industry, and region, and a random sample within those groups is representative of that strata. In some particular surveys for some countries, the ES includes informal firms and/or firms with fewer than 5 employees.

Table 3.6: Exporting and On-the-job Training

	(1)	(2)	(3)	(4)	(5)
	Panel A: Training and Exporting				
Exporter	0.16***	0.06***	0.06***	0.06***	0.07***
	(0.011)	(0.011)	(0.012)	(0.012)	(0.019)
Obs	109,698	107,568	86,226	83,202	44,242
R-squared	0.136	0.190	0.202	0.202	0.254
	Panel B: Training, Exporting and Technology				
Non-Exporter # Foreign Tech	0.18***	0.13***	0.12***	0.12***	0.11***
	(0.017)	(0.018)	(0.020)	(0.020)	(0.032)
Exporter # No Foreign Tech	0.16***	0.07***	0.06***	0.05***	0.06***
	(0.013)	(0.013)	(0.015)	(0.015)	(0.018)
Exporter # Foreign Tech	0.29***	0.15***	0.14***	0.15***	0.14***
	(0.020)	(0.021)	(0.026)	(0.026)	(0.028)
Obs	79,184	78,394	63,631	61,881	26,731
R-squared	0.149	0.192	0.202	0.202	0.281
Log(Emp)	No	Yes	Yes	Yes	Yes
Labor share	No	No	Yes	Yes	Yes
Managerial experience in sector	No	No	No	Yes	Yes
% High school grads	No	No	No	No	Yes

Note: This table presents estimates from regressing a dummy variable that takes the value 1 if the firm offers formal on-the-job training, on export status. The baseline groups are non-exporters for Panel A and non-exporters with no foreign technology adoption for Panel B. Exporters are defined as firms with positive sales to foreign markets. We control for country, year, and industry fixed effects in all regressions. Firm-level control variables are log (employment), the ratio of labor costs to total sales, the share of high-school graduates in the workforce, and managers' years of experience in the operating sector. We use the second standardized wave of the ES with the provided weights. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

With this in mind, we regress a dummy variable representing if the firm offers on-the-job training on export status. Panel A in Table 3.6 shows the results. Exporters are 6–16 percentage points more likely to invest in on-the-job training compared to non-exporters, under different sets of control variables, suggesting that more opportunities exist for human capital accumulation in exporters than in non-exporters.[33]

We explore whether exporters with access to better technology are more willing to

[33]The ES also provides information on whether firms indirectly export. Indirect exporters are firms that do not export but are selling goods to another firm which then exports the same goods, and they are not counted as exporters in our regressions. We find that indirect exporters also have a larger probability of training their workers than non-exporters, but their probability of training is lower than direct exporters. This evidence suggests that both production of exported goods and direct contact with destination markets may benefit workers' human capital, though we cannot rule out selection effects.

invest in human capital. In Panel B of Table 3.6, we regress our dummy variable of on-the-job training on interaction terms between export status and a dummy variable that equals 1 if the firm adopts foreign technology. Being an exporter is associated with larger human capital investments. Conditional on export status, if firms adopt foreign technology, the probability of investing in on-the-job training also increases. These results are consistent with studies finding complementarities between technology and human capital (Acemoglu and Zilibotti, 2001; Porzio, 2017). In Appendix Table C.12, we show that exporters who invest in R&D are more likely to train workers than exporters who do not. This also supports that exporters with access to better technology are more willing to invest in human capital. In Appendix Table C.13, we use triple interactions between foreign technology adoption, R&D investments, and export status, and find that all three are associated with larger human capital investments conditional on the other two.

Finally, in Appendix Tables C.14, we replicate these results using the Enterprise Surveys for Brazil and show that all results hold. Therefore, the evidence from the ES supports our main hypothesis in Section 3.4 that higher returns to experience observed in Brazilian exporters correspond at least partially to faster human capital accumulation, and that exposure to advanced countries may induce faster human capital accumulation.

3.6 Conclusion

Using Brazilian employer-employee and customs data, this study documents that workers' life-cycle wage growth is faster in exporters than in non-exporters. Apart from selection of firms with higher returns to experience into exporting, we find that workers enjoy steeper experience-wage profiles when firms export to industrialized destinations. We discuss several plausible explanations for the destination-specific effects on experience-wage profiles. We propose that faster human capital accumulation when exposed to advanced destinations is the most likely explanation. Using the Enterprise Surveys for more than 100 countries, we further corroborate this hypothesis by showing that exporters which adopt foreign technology are more involved in training workers than exporters which do not.

Understanding the effects of trade on workers' wages is important because of its implications for aggregate welfare and inequality. We view our study as one of the first steps to empirically understanding the effects of trade on workers' life-cycle wage growth, complementing recent efforts using structural models to study trade and wage growth (Fajgelbaum, 2019; Guner, Ruggieri and Tybout, 2019). Our results also raise the possibility that workers' human capital accumulation may interact with destination markets. A fruitful area for future study is how this interaction impacts the effects of globalization on workers' income levels and inequality in countries with different development levels.

3.7 Acknowledgements

Chapter 3 is currently being prepared for submission for publication and is coauthored with Xiao Ma. The dissertation author was a primary investigator of this material. I would like to thank Xiao in particular for all the invaluable discussions we had over these years. I learned a lot from our interactions.

Appendix A

Appendix of Chapter 1

A.1 Data Sources

We rely on enterprise and workers' surveys for developed and developing economies. We rely on the World Bank Enterprise Survey for developing economies and we rely on European Union data for developed economies. We use the EU Labor Force Survey (EU-LFS), the Adult Education Survey (EU-AES) and the Continuing Vocational Training Survey (EU-CVT). Moreover, we use other data sources to test some implications from our quantitative model. With this purpose, we rely on the Chinese Industrial Census (an administrative dataset for all manufacturing firms in China) to test the correlation between firms' features and training investments. Finally, we rely on the World Bank Worldwide Governance Indicators to have measures of institutional quality, data from Botero et al. (2004) to proxy labor market indicators, and on data from Donovan, Lu and Schoellman (2020) to have measures of job destruction and job-to-job transitions to test cross-country correlations.

A.2 Detailed Definitions on Educational Sources

We first carefully define different types of on-the-job human capital acquisition to ensure consistency across different sources. We separate the sources of skill acquisition into four categories that allow for data comparability, while also having simple economic interpretations that can be mapped onto our model. The categories rank from the most structured and planned in advance (schooling) to the least structured (informal learning, which is not structured at all). For expositional purposes, and since we focus on firm-sponsored investments, we also consider a secondary distinguishing quality within each source, which is the financing source for the educational investment (firm vs. worker sponsored).

Schooling: Formal education and training according to the International Standard Classification of Education 2011 (ISCED 2011) is defined as: "education that is institutionalized, intentional and planned through public organizations and recognized private bodies and in their totality constitute the formal education system of a country. Formal education programs are thus recognized as such by the relevant national education authorities or equivalent authorities, e.g. any other institution in cooperation with the national or sub-national education authorities. Formal education consists mostly of initial education. Vocational education, special needs education and some parts of adult education are often recognized as being part of the formal education system."

Training: Non-formal education and training is defined as "any organized and sustained learning activities outside the formal education system. Non-formal education is an addition, alternative and/or complement to formal education. Non-formal education may therefore take place both within and outside educational institutions and cater to people of all ages. Depending on national contexts, it may cover educational programs to impart adult literacy, life-skills, work-skills, and general culture. Note that within non-formal education we can have formal training or informal training depending on its level of organization."

We rely on definitions for *formal training* and *informal training* from the CVT survey manuals. Continuing vocational training (*formal training*) refers to education or training activities that are planned in advance, organized, or supported with the specific goal of learning and financed in total or at least partially by the enterprise. These activities aim to generate "the acquisition of new competences or the development and improvement of existing ones" for firms' employees. Persons employed holding an apprenticeship or training contract should not be considered for CVT. Random learning and initial vocational training are explicitly excluded and measured separately. These courses are typically separated from the active workplace (e.g., the classroom or training institution), show a high degree of organization by a trainer, and its content is designed for a group of learners (e.g., a curriculum exists).

As defined by the CVT survey, "Other forms of CVT" that we refer to as *informal training*, have the purpose of learning and are typically connected to the active work and the active workplace, but they can also include participation (instruction) in conferences, trade fairs, etc. These are often characterized by self-organization by the individual learner or by a group of learners and is often tailored according to the workers' needs. The following types of "other forms of CVT" are identified in the survey:

1. Guided-on-the job training: "It is characterised by planned periods of training, instruction or practical experience in the workplace using the normal tools of work, either at the immediate place of work or in the work situation. The training is organised (or initiated) by the employer. A tutor or instructor is present. It is an individual-based activity, i.e. it takes place in small groups only (up to five participants)."

2. Job rotation, exchanges, secondments or study visits: "Job rotation within the enterprise and exchanges with other enterprises as well as secondments and study visits are other forms of CVT only if these measures are planned in advance with the primary intention of developing the skills of the workers involved. Transfers of workers from one job to another which are not part of a planned developmental programme should be excluded."

3. Learning or quality circles: "Learning circles are groups of persons employed who come together on a regular basis with the primary aim of learning more about the requirements of the work organisation, work procedures and workplaces. Quality circles are working groups, having the objective of solving production and workplace-based problems through discussion. They are counted as other forms of CVT only if the primary aim of the persons employed who participate is learning."

4. Self-directed learning/e-learning Self-directed learning/e-learning: "individual engages in a planned learning initiative where he or she manages the settings of the learning initiative/activity in terms of time schedule and location. Self-directed learning means planned individual learning activities using one or more learning media. Learning can take place in private, public or job-related settings. Self-directed learning might be arranged using open and distance learning methods, video/audio tapes, correspondence, computer based methods (including internet, e-learning) or by means of a Learning Resources Centre. It has to be part of a planned initiative. Simply surfing the internet in an unstructured way should be excluded. Self-directed learning in connection with CVT courses should not be included here."

5. Participation in conferences, workshops, trade fairs and lectures: "Participation (instruction received) in conferences, workshops, trade fairs and lectures are considered as training actions only when they are planned in advance and if the primary intention of the person employed for participating is training/learning"

Initial vocational training is defined as a formal education program (or a component of it) where working time alternates between periods of education and training at the work-place and in educational institutions or training centers. This program consists of learning activities for workers initializing a job.

Informal learning: It is defined as "intentional learning which is less organized and less structured than the previous types. It may include for example learning events (activities) that occur in the family, in the workplace, and in the daily life of every person, on a self-directed, family-directed or socially directed basis. Categories used for informal training are: learning from peers, colleagues; learning by using printed material, learning by using computers, learning through media (television, radio or videos); learning through guided tours as museums; learning by visiting learning centers as libraries."

A.3 Empirical Results: Additional Tables and Graphs

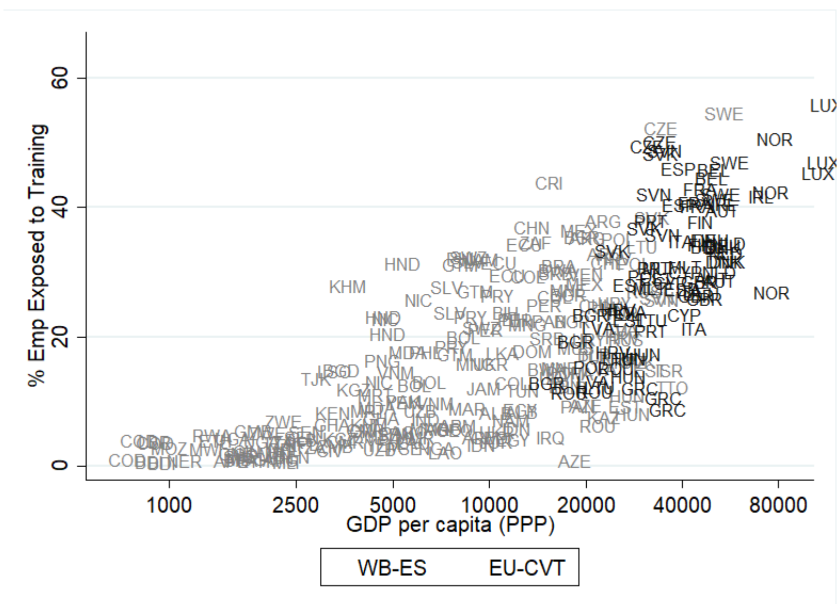
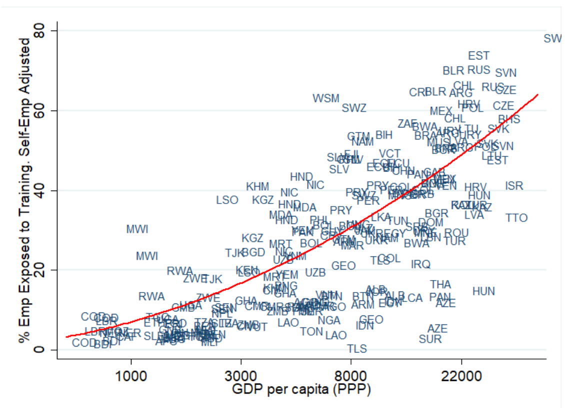


Figure A.1: Share of Employment Formally Trained and Development (Full Sample)

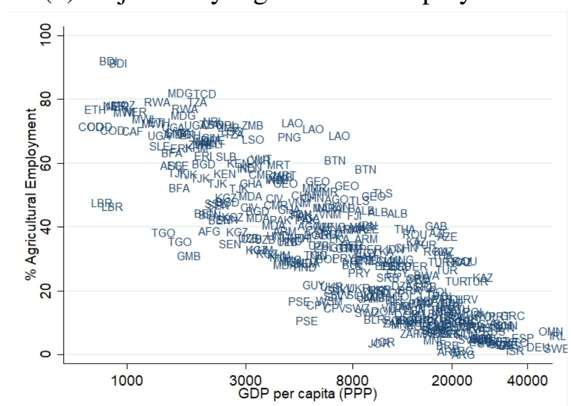
Note: The share of employment formally trained comes from adjusting the share of workers who were trained by firms by the share of self-employment. Data on share of employees trained inside the firms comes from the World Bank Enterprise Survey for all developing economies and from EU Continuing Vocational Training Survey for European countries. Both surveys ask if the firm provided formal training in the previous fiscal year and the share of employees who participated. For the World Bank Enterprise Survey we use the standardized wave with data from 2005-2017 for which we have firm weights and we plot all countries with no restrictions. Data on GDP per capita and self-employment comes from the World Bank Indicators.



(a) Adjusted by Self Employment

(b) Adjusted by Agricultural Employment

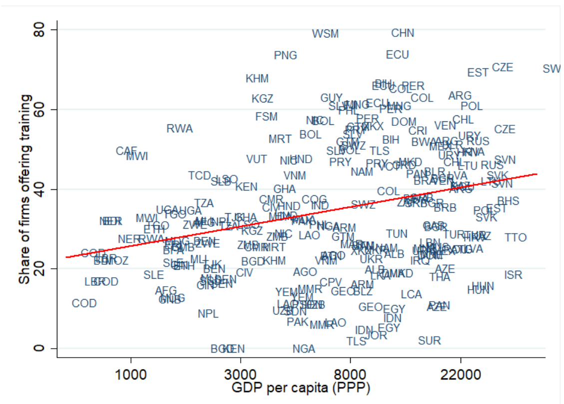
(c) Share of Self-Employed Workers



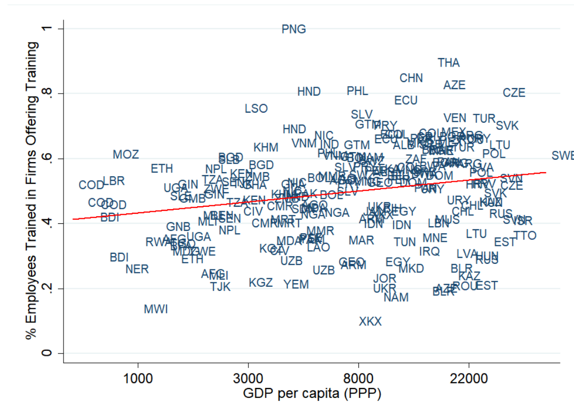
(d) Share of Agricultural Workers

Figure A.2: Share of Workers Exposed to Firms that offer Training

Note: The share of employment trained in Panel a come from adjusting the share of workers who were trained in the wage-sector by the share of self-employment plotted in Panel c. The share of employment trained in Panel b comes from adjusting the share of workers in the wage-sector by the share of agricultural workers plotted in Panel d. Data on share of employees trained inside the firms comes from the World Bank Enterprise Survey for all developing economies. Data on self-employment and the share of agricultural workers come from the World Bank Indicators. Data on GDP per capita comes from the PWT.

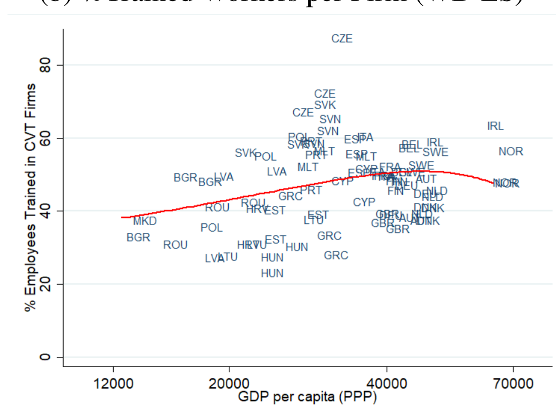


(a) %Firms Offering Training (WB-ES)



(b) %Trained Workers per Firm (WB-ES)

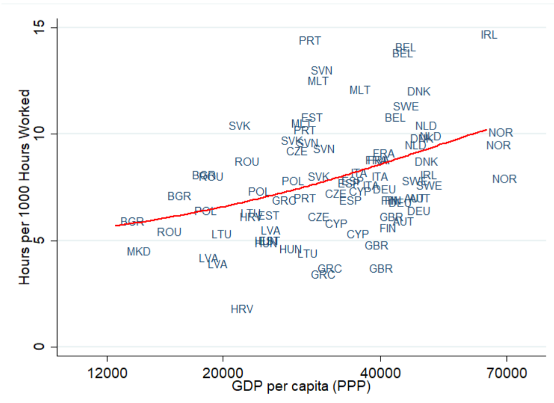
(c) %Firms Offering Training (EU-CVT)



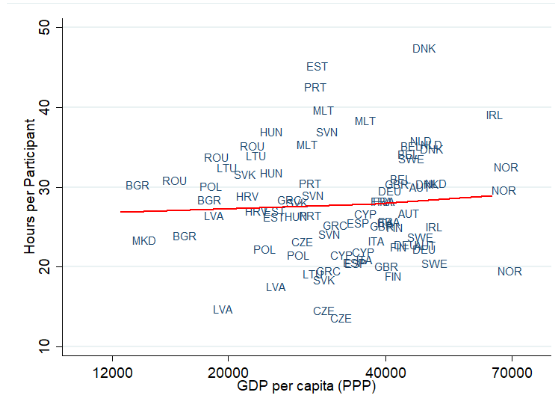
(d) %Trained Workers per Firm (EU-CVT)

Figure A.3: Intensive and Extensive Margin

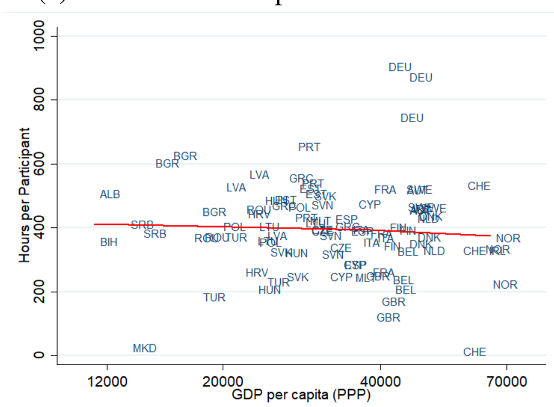
Note: These figures show the share of firms offering training and the share of workers trained for the World Bank Enterprise Survey and the European Union Continuing Vocational Training Survey. Panel a shows the share of firms and Panel b shows the share of participants per firm in the manufacturing and service sector weighted by the WB-ES-provided weights. For the World Bank Enterprise Survey, we use the standardized wave with data from 2005-2017 for which we have firm weights and we plot all countries with no restrictions. Panel c and d show the counterparts from the EU-CVT provided by the publicly available results (trng_cvt_01s and trng_cvt_12s). Data on GDP per capita comes from the Penn World Table.



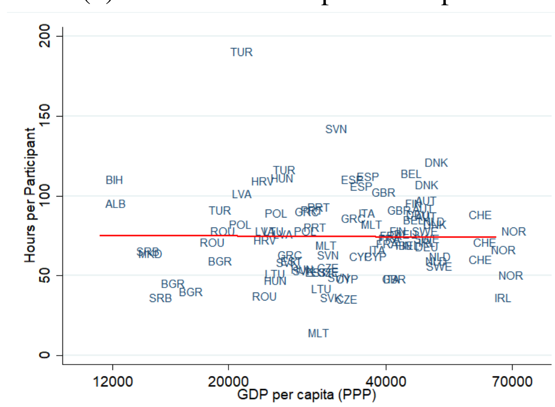
(a) EU-CVT: Hours per 1000 Hrs Worked



(b) EU-CVT: Hours per Participant



(c) EU-AES: Formal Education



(d) EU-AES: Training

Figure A.4: Time Spent in Education

Note: These figures show the time spent in training for different data sources. Panel a represents the hours per 1000hrs worked for participants and Panel b represents the hours spent in training per participant coming from the publicly available EU-CVT data (trng_cvt_22s and trng_cvt_25s). Panel c represents the hours spent in formal education, which we call schooling, and Panel d represents the hours spent in training per participant coming from the EU-AES publicly available data (trng_aes_147). Data on per capita GDP comes from the Penn World Table.

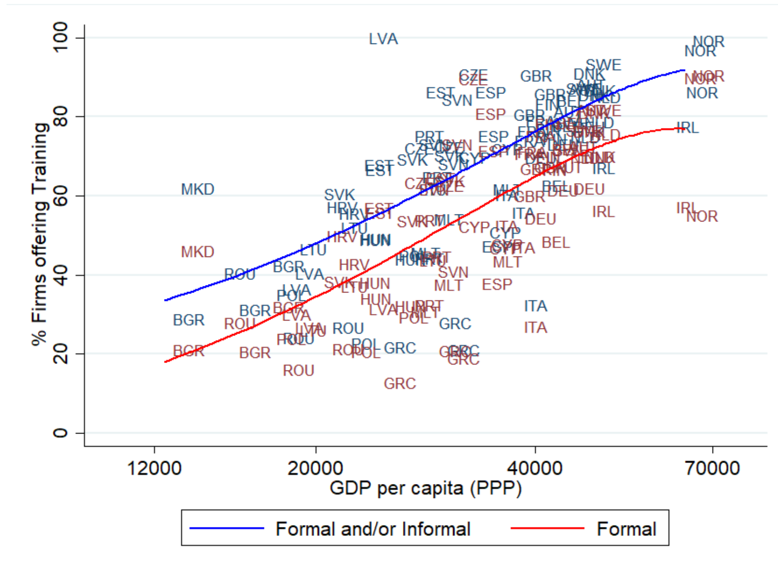


Figure A.5: Share of Firms Offering Formal and Informal Training

Note: This figure shows the share of firms offering formal training in red and the share of firms offering formal and informal training in blue as a function of per-capita GDP. Data comes from the EU- CVT survey and formal training is defined as CVT and informal training as "other forms of CVT" as defined in the data description section. The data comes from the EU-CVT publicly available data (trng_cvt_01s). Data on GDP per capita comes from the Penn World Table.

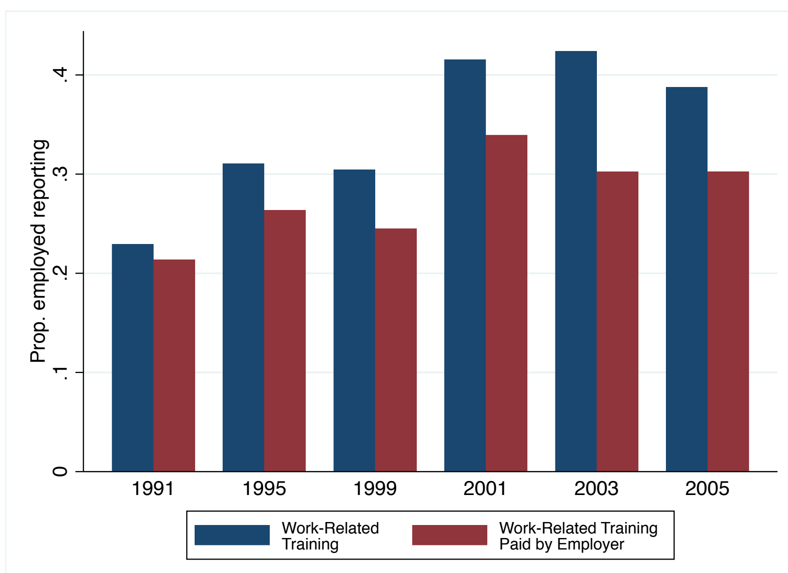
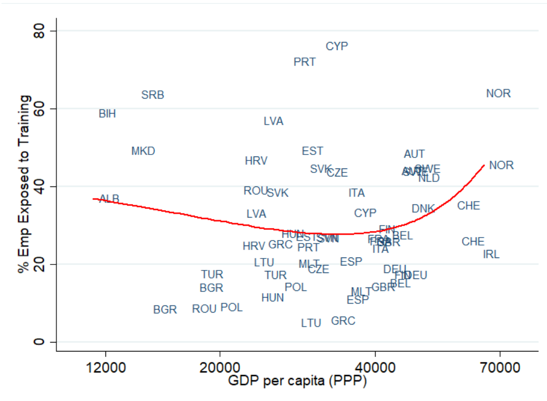
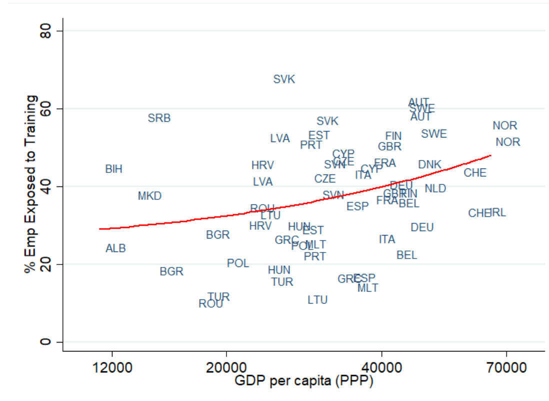


Figure A.6: Share of Workers Reporting Training by Year in the US

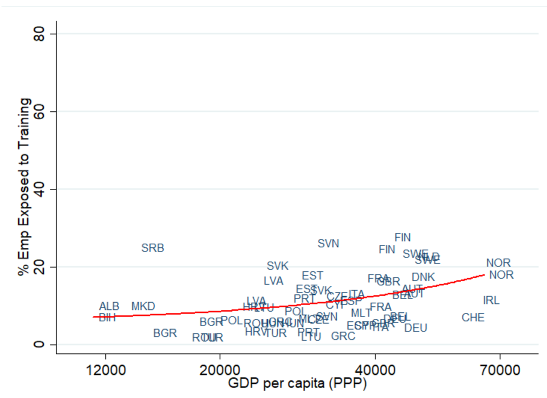
Note: This figure shows U.S. workers' participation rate in work-related training (training, workshops, seminars, courses, or classes for work related reasons in the past 12 months) and work-related training sponsored by employer (training paid at least partially by employer). We use all years with data on these variables and exclude the 2016 survey from the analysis presented here due to definitional changes. Data comes from the National Household Education Survey (NHES).



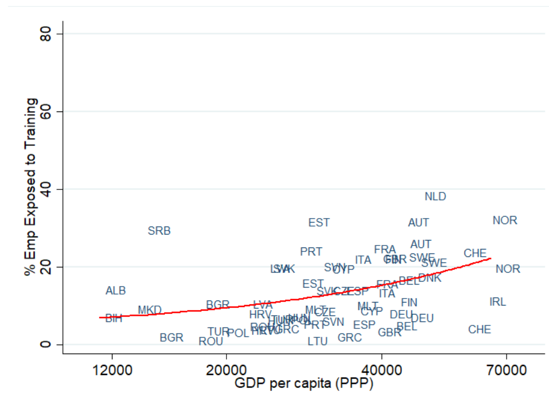
(a) Through Peers



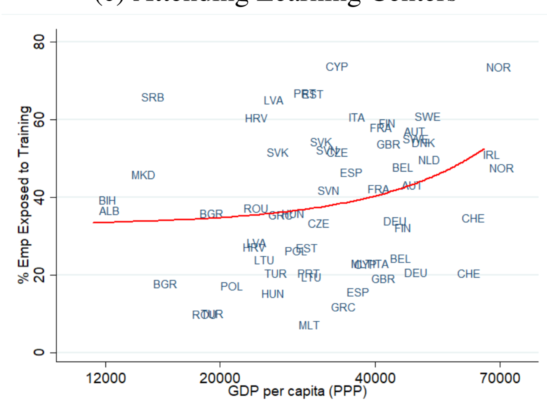
(b) Using Printed Material



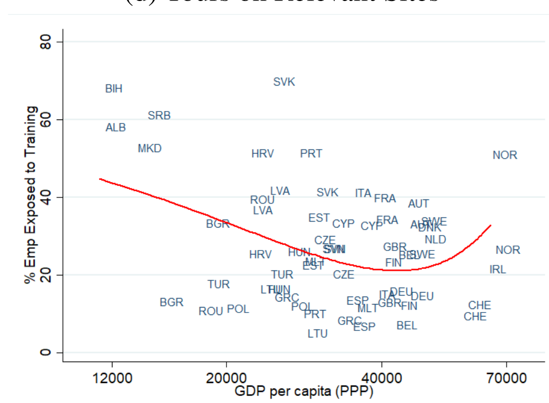
(c) Attending Learning Centers



(d) Tours on Relevant Sites



(e) Using Computer



(f) Using Media

Figure A.7: Informal Learning AES

Note: These figures show participation rate in informal learning that includes learning through peers (Panel a), using printed material (Panel b), attending to learning centers (Panel c), tours on learning sites (Panel d), using computers (Panel e) and using media (Panel f). Data are publicly available from the EU-AES (trng_aes_202). Data on per-capita GDP comes from the Penn World Table.

Table A.1: European Union Labor Force Survey (EU-LFS)

	Hours	During Working Hours		Reason	
	Employed Population	During paid hours	Outside paid hours	Job related	Personal Social
European Union - 25 (2004-2006)	66	69.3	30.7	83.9	15.9
Germany	74			90.8	9.1
France	85	87.4	12.6	93.3	6.7
United Kingdom	35	70.3	29.7	79.2	20.8
Italy	58	56.5	43.5	83.5	15.1
Spain	102	38	62	61.7	38.3
Poland	40	59.4	40.6	91.3	8.7
Romania	80	34	66	80.3	19.7
Netherlands	76	54.8	45.2	86.1	13.9
Belgium	69	68.7	31.3	82.5	17.5
Greece	80	40.4	59.6	72.7	27.3

This data comes from Eurostat, past series, LFS ad Hoc Module 2003 (trng_nfe6 for reason, trng_nfe7 for working hours and trng_nfe15 for hours). We show the outcomes from the most populated European countries ranked by population size.

Table A.2: Training Purpose (EU-CVT)

	Average By firm Size in 2010				Average By firm Size in 2015			
	All	10-49	50-249	250+	All	10-49	50-249	250+
General IT	27.3	23.7	34.5	54.7	12.8	12	15.2	15.8
Professional IT	16.9	14.5	21	37.5	10.2	9.8	11.6	11.2
Management	32	26.2	43.7	74.3	23.4	19.9	30.9	49.2
Team working	32.5	29	38.3	61.6	19.6	19.1	20.5	22.6
Customer handling	38.5	35.4	44.1	62.7	25.6	25	26.5	31.3
Problem solving	30.1	28.5	31.2	50	13.5	13.3	14.1	13.8
Office administration	26.9	24.3	32.3	45.1	13.4	13.6	14	8
Foreign language	15.3	11	24	46.9	7.9	5.9	13.2	17.5
Technical or job-specific	69	67.2	73.2	81.2	64.6	63.1	68.5	71.9
Oral or written communication	14.7	12.7	16.9	36.5	3.5	3.3	4	4.4
Numeracy and/or literacy	7	6.7	6.5	14.7	1.2	1.3	1.2	1.1
Other skills and competences	11	11.2	10.4	10.3	19.9	20.3	18	19.8

Note: This table shows the main skills targeted by CVT courses by type of skill and size class. This represents the share of enterprises providing CVT courses (publicly available data trng_cvt_29s). The table shows the share for all firms and for the firm size categories 10 - 49, 50 - 249, and 250+. A particular course may cover more than one category.

Appendix B

Appendix of Chapter 2

B.1 Quantitative Model: Conditions for Simulations

Workers' Expected Utility. With linear utility and $\rho = r$, we assume workers do not save and spend all their income each period. Thus, for a worker of age a, their utility comes from (future) income flows produced by workers' current human capital, and potential income flows from human capital accumulation. For a worker of age a in a firm with productivity z, we denote $J_{c,a}(z)$ as the expected value of income flows per efficiency unit of current human capital, and $J_{h,a}(z)$ as the expected value of income flows from human capital accumulation. With little abuse of notation, we use $J_{c,a}(u)$ and $J_{h,a}(u)$ for unemployed workers.

First, note that in the last period of workers' lifetime ($a = T$), workers have no incentive to accumulate human capital. Thus, we can obtain

$$J_{c,T}(z) = w(z);\ J_{h,T}(z) = 0;\ J_{c,T}(u) = 0;\ J_{h,T}(u) = 0.$$

For younger workers ($a < T$), we can obtain their values by backward induction.

$$
\begin{aligned}
J_{c,a}(z) &= w(z) + \frac{1-d}{1+r}\delta\left[\theta q(\theta)\int J_{c,a+1}(z)dF(w(z)) + (1-\theta q(\theta))J_{c,a+1}(u)\right] \\
&+ \frac{1-d}{1+r}(1-\delta)\left[J_{c,a+1}(z) + \eta\theta q(\theta)\int p_{a+1}(z,z')(J_{c,a+1}(z') - J_{c,a+1}(z)) - c_p^{\gamma_p}\frac{p_{a+1}(z,z')^{1+\gamma_p}}{1+\gamma_p}dF(w(z'))\right] \\
J_{h,a}(z) &= -\mu_W\left(c_s\bar{w} + \delta_s A_M z\right) + \frac{\zeta s_a(z)^{\gamma_s}}{1+r}\delta\left(\theta q(\theta)\int J_{c,a+1}(z)dF(w(z)) + (1-\theta q(\theta))J_{c,a+1}(u)\right) \\
&+ \frac{\zeta s_a(z)^{\gamma_s}}{1+r}(1-\delta)\left[J_{c,a+1}(z) + \eta\theta q(\theta)\int p_{a+1}(z,z')(J_{c,a+1}(z') - J_{c,a+1}(z)) - c_p^{\gamma_p}\frac{p_{a+1}(z,z')^{1+\gamma_p}}{1+\gamma_p}dF(w(z'))\right] \\
&+ \frac{\delta}{1+r}\left[\theta q(\theta)\int J_{h,a+1}(z)dF(w(z)) + (1-\theta q(\theta))J_{h,a+1}(u)\right] \\
&+ \frac{1-\delta}{1+r}\left[J_{h,a+1}(z) + \eta\theta q(\theta)\int p_{a+1}(z,z')(J_{h,a+1}(z') - J_{h,a+1}(z))dF(w(z'))\right] \\
J_{c,a}(u) &= \frac{1-d}{1+r}\left[\theta q(\theta)\int J_{c,a+1}(z)dF(w(z)) + (1-\theta q(\theta))J_{c,a+1}(u)\right] \\
J_{h,a}(u) &= \frac{1}{1+r}\left[\theta q(\theta)\int J_{h,a+1}(z)dF(w(z)) + (1-\theta q(\theta))J_{h,a+1}(u)\right]
\end{aligned}
$$

$p_a(z,z')$ is the leaving probability conditional on getting an offer from a firm with productivity z', obtained by evaluating leaving probability for an average worker of age a in firm z:[1]

$$
\max_{p\in[\underline{p},1]} \left[(J_{c,a}(z') - J_{c,a}(z))\bar{h}_a(z) + J_{h,a}(z') - J_{h,a}(z)\right]p - c_p^{\gamma_p}\frac{p^{1+\gamma_p}}{1+\gamma_p}\bar{h}_a(z)
$$

$\bar{h}_a(z)$ is the average human capital of age a workers in firm z, which will be derived soon.

Employment Distribution. Let N_m be the amount of workers who enter the modern sector at each generation. Then, in the beginning of each period, the amount of searchers in the modern sector is:

$$
\tilde{U} = \sum_{a=1}^{T}(u_a + (1-u_a)\eta)N_m
$$

which is the sum of the unemployed and on-the-job searchers across different age groups. The

[1]For computational tractability, we do not use different values of leaving probability for individual workers of age a in firm z. Because costs of leaving increase with human capital and income flows from current human capital are larger than benefits from future human capital accumulation in most cases of our simulation (except for early ages when workers have little human capital), this simplification is also reasonable.

unemployed population (before job search and matching) for the youngest cohort is $u_1 = N_M$ and proceeds as $u_{a+1} = \delta N_M + (1-\theta q(\theta))(1-\delta)u_a \;\forall\; 1 \leq a \leq T-1$.

We define the measure of employment $m_a(z)$ for workers of age a in firms with productivity z. Hence, the employment distribution across firms for the youngest cohort is simply $m_1(z) = \frac{\theta q(\theta) f(w(z))w'(z)}{g(z)} u_1$ after search and matching processes. For older cohorts, their measure of employment proceeds as

$$m_{a+1}(z) = \underbrace{(1-\delta)\left[1-\eta\theta q(\theta)\int p_{a+1}(z,z')dF(w(z'))\right]m_a(z)}_{\text{stayers}}$$
$$+\underbrace{u_{a+1}\frac{\theta q(\theta)f(w(z))w'(z)}{g(z)}}_{\text{hires from unemployed}} + \underbrace{(1-\delta)\eta\frac{\theta q(\theta)f(w(z))w'(z)}{g(z)}\int m_a(y)p_{a+1}(y,z)dG(y)}_{\text{hires from job-to-job moves}}$$

Training. Firms' optimal training is determined by:

$$\mu_F\left(\delta_s A_M z + c_s \bar{w}\right) = \zeta\gamma_s s_{F,a}(z)^{\gamma_s-1}(A_M z - w(z))\Psi(z,1,a)$$

where $\Psi(z,t,a) = \sum_{\tau=t}^{T-a}(1-d)^{\tau-1}(1-\delta)^{\tau}\prod_{k=1}^{\tau}\left(\frac{1-\eta\theta q(\theta)\int p_{a+k}(z,z')dF(w(z'))}{1+r}\right)$. And workers' optimal training is determined by:

$$\mu_W\left(\delta_s A_M z + c_s \bar{w}\right) = \zeta\gamma_s s_{W,a}(z)^{\gamma_s-1}\frac{J_{c,a}(z)-w(z)}{1-d}$$

where $\frac{J_{c,a}(z)-w(z)}{1-d}$ is workers' return for an extra efficiency unit of human capital in the next period. The optimal training is $s_a(z) = \min(s_{F,a}(z), s_{W,a}(z))$. In comparison with our analytical model, the optimal training now depends on the present value of all future returns, adjusted for the depreciation rate of training and workers' separation rates (for firms). Notably, the optimal training decreases with workers' age, as training young workers produces longer-lasting returns than training old workers. Also note that training does not depend on workers' training and employment histories, which enables us to track the dynamics of average human capital for a firm's labor force.

Evolution of Human Capital. Specifically, define $\overline{h}_a(z)$ as the average human capital of age-a workers in firms with productivity z. The human capital of the youngest cohort is $\overline{h}_1(z)=1$. We could obtain the dynamics of human capital as:

$$
\begin{aligned}
\overline{h}_{a+1}(z) = & \underbrace{\frac{m_a(z)}{m_{a+1}(z)}(1-\delta)\left[1-\eta\theta q(\theta)\int p_{a+1}(z,z')dF(w(z'))\right](\overline{h}_a(z)(1-d)+\zeta s_a(z)^{\gamma_s})}_{\text{stayers}} \\
& + \underbrace{\frac{\theta q(\theta) f(w(z))w'(z)}{m_{a+1}(z)g(z)}}_{\text{new meets/employment}}\left[\underbrace{\eta(1-\delta)\int p_{a+1}(y,z)(\overline{h}_a(y)(1-d)+\zeta s_a(y)^{\gamma_s})m_a(y)dG(y)}_{\text{meet on-the-job searchers}}\right] \\
& + \underbrace{\frac{\theta q(\theta) f(w(z))w'(z)}{m_{a+1}(z)g(z)}}_{\text{new meets/employment}} \underbrace{u_{a+1}\overline{h}^u_{a+1}}_{\text{meet unemployed}}
\end{aligned}
$$

where $\overline{h}^u_{a+1} = \frac{(1-\theta q(\theta))u_a\overline{h}^u_a(1-d)+\delta\int(\overline{h}_a(z)(1-d)+\zeta s_a(z)^{\gamma_s})m_a(z)dG(z)}{\delta N_M+(1-\theta q(\theta))(1-\delta)u_a}$ refers to the average human capital of unemployed people with $\overline{h}^u_1=1$.

Vacancies and Wage Determination. We now focus on the conditions for vacancies and wages. The condition for firms' optimal level of vacancies and wages is given by:

$$
\begin{aligned}
c_v v(z)^{\gamma_v} = & \underbrace{\sum_{a=1}^{T}\frac{q(\theta)(A_M z-w(z))}{\sum_a u_a+\eta(N_M-u_a)}\left[\eta(1-\delta)\int p_a(y,z)\overline{h}^s_a(y)(1-d)m_{a-1}(y)dG(y)+u_a\overline{h}^u_a\right]\frac{\Psi(\phi,0,a)}{(1-d)^{-1}}}_{\text{benefits from new hires' human capital}} \\
& + \sum_{a=1}^{T-1}\frac{q(\theta)[\eta(1-\delta)\int p(y,z)m_{a-1}(y)dG(y)+u_a]}{\sum_a u_a+\eta(N_M-u_a)} \\
& \underbrace{\times\sum_{t=0}^{T-a}D(z,t,a)\left[\zeta s_{a+t}(z)^{\gamma_s}(A_M z-w(z))\Psi(z,1,a+t)-\mu_F c_s(z)s_a(z)\right].}_{\text{benefits from training new hires}}
\end{aligned}
$$

We define $D(z,t,a)=\prod_{k=1}^{t}\left(\frac{1-\eta(1-\delta)\theta q(\theta)\int p_{a+k}(z,z')dF(w(z'))-\delta}{1+r}\right)$ with $D(z,0,a)=1$. $\bar{h}^s_a(y)=$

$\overline{h}_{a-1}(y)(1-d)+\zeta s_{a-1}(y)^{\gamma_s}$, and $c_s(z)=\delta_s A_M z + c_s\bar{w}$.

The differential equation of wages can be obtained by totally differentiating the above equation with regard to $w(z)$, as firms choose wages to maximize the value of each vacancy.

$$
\begin{aligned}
&\sum_{a=1}^{T}\frac{q(\theta)}{\sum_a u_a+\eta(N_M-u_a)}\left[\eta(1-\delta)\int p_a(y,z)\overline{h}_a^s(y)(1-d)m_{a-1}(y)dG(y)+u_a\overline{h}_a^u\right]\frac{\Psi(\phi,0,a)}{(1-d)^{-1}}\\
&=\sum_{a=1}^{T}\frac{q(\theta)(A_Mz-w(z))}{\sum_a u_a+\eta(N_M-u_a)}\frac{\partial\left[\eta(1-\delta)\int p_a(y,z)\overline{h}_a^s(y)(1-d)m_{a-1}(y)dG(y)+u_a\overline{h}_a^u\right]\frac{\Psi(\phi,0,a)}{(1-d)^{-1}}}{\partial w(z)}\\
&+\sum_{a=1}^{T-1}\frac{q(\theta)}{\sum_a u_a+\eta(N_M-u_a)}\times\\
&\frac{\partial[\eta(1-\delta)\int p(y,z)m_{a-1}(y)dG(y)+u_a]\sum_{t=0}^{T-a}D(z,t,a)\left[\zeta s_{a+t}(z)^{\gamma_s}(A_Mz-w(z))\Psi-\mu_F c_s(z)s_a(z))\right]}{\partial w(z)}
\end{aligned}
$$

Note that this is a differential equation with regard to wage $w(z)$. To solve this, we can multiply each side by $w'(z)$. With this transformation, the right-hand side becomes the derivative with regard to productivity z, and thus, we can numerically evaluate $w'(z)$. Combined with the lowest wage $b\bar{w}$, we can iterate the wage structure $w(z)$ until convergence.

B.2 Calibrating Labor Market Dynamics

In this study we use three moments that are key to performing our counterfactuals. In the baseline calibration, we add the job-to-job and job-to-unemployment probabilities. In this subsection, we explain in detail how we measure these moments and provide alternative calibrations using different measures and their implications for our results.

For job turnover dynamics, we use two measures in our calibration: the share of employed people remaining in the same firm and the share of employed people remaining employed after a quarter. We rely on data from Donovan, Lu and Schoellman (2020), which provide these two probabilities for many countries. Their study is the first, and only, in providing the relationship between these probabilities and development. Nevertheless, the countries in their sample do not match the countries and years in our sample in most cases and, therefore, we must build predicted measures using their data. Moreover, as the purpose of this study is to provide comparisons across countries, we must have consistent measures for all countries.

Our main measure uses the variation on institutional quality across countries, which shapes labor market dynamics, particularly job turnover. If contracts are better and easily enforced, job turnover will be lower. Thus, we first regress Donovan, Lu and Schoellman (2020) probabilities in all institutional measures from the World Bank Worldwide Governance Indicators and predict the probabilities for all countries in our sample using these variables (imposing an upper bound of 1 given that we are predicting probabilities). Figure B.1 shows these two predicted probabilities in grey. This measure has two issues that are the high noise in the predicted value and that these two probabilities are really close together for some developed countries. This second issue is more relevant, as it implies there are almost no job-to-job transitions, which is counterfactual. Thus, we construct a smooth probability measure by predicting our previously built measure with GDP per capita. We plot these new predicted values with respect to per-capita GDP in red in Figure B.1.

We construct a second measure to use as robustness for our main specification. We directly predict the probabilities from Donovan, Lu and Schoellman (2020) with per-capita GDP and plot the outcome in blue in Figure B.1. We show that the probability of a worker staying in the same firm is the virtually the same for both measures. For the probability of a worker staying employed after a quarter, our main measure is a little higher than the one just predicted with GDP for poorer economies, which imply that there is even higher job-turnover in developing economies. Finally, it is worth noting that our results do not change substantially by adding either one of these two measures.

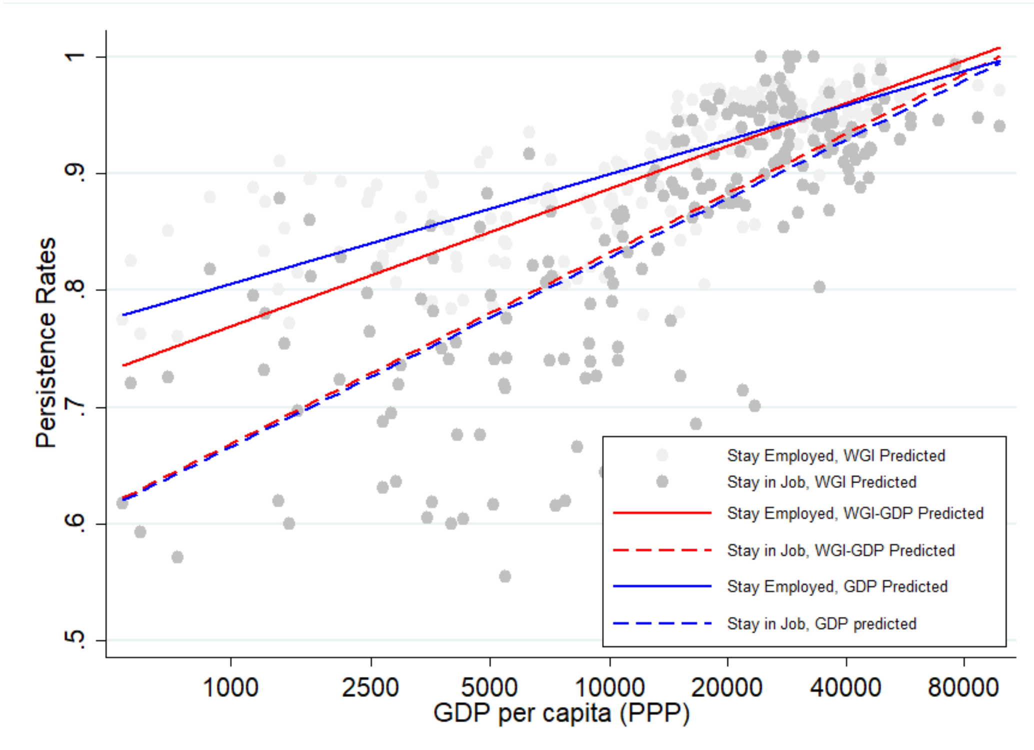


Figure B.1: Job Transition Probabilities

Note: This figure shows the three measures for the share of employed people remaining in the same firm and the share of employed people remaining employed after a quarter. The first measures in the grey scatter plot represents the result of first regressing Donovan, Lu and Schoellman (2020) probabilities in all institutional measures from the World Bank Worldwide Governance Indicators and then predicting the probabilities for all countries in our sample. The predicted red lines are the linear prediction of the first measure with respect to per capita GDP and the blue lines are the linear prediction of Donovan, Lu and Schoellman (2020) probabilities with respect to per capita GDP directly. Data on per-capita GDP comes from the Penn World Table.

B.3 Identification of Model Parameters

We now illustrate how the moments help identify parameters. We calculate the elasticity for moments to each parameter and provide the results in Table B.1. First, we describe the parameters closely related to labor market outcomes. For the constant in the vacancy cost function c_v, the most sensitive moment is the ratio of vacancies to unemployment. Similarly, c_m affects the economy's matching efficiency, and the moments that identify this parameter are both the vacancy-to-unemployment ratio and the unemployment rate. As expected, the share of workers who switch jobs due to an idiosyncratic shock, $\underline{p}$, is identified through the wage growth from job-to-job switches and the share of workers who switch from high-to-low paying firms. The traditional sector share in production, γ, has the main impact in the self-employment share. Lastly, a larger shape parameter of the Pareto productivity distribution, κ implies fewer productive firms, which reduce the wage sector's relative return to the self employment sector and the average wage growth after job-to-job transitions.

Table B.1: Elasticities of Targeted Moments to Parameters

	Labor Market Dynamics					Training Dynamics				Frictions		
	c_v	c_m	$\underline{p}$	γ	κ	γ_s	c_s	ζ	μ_F	γ_p	c_p	δ
Unemp Rate	0.2	**-1.4**	-0.1	-0.3	-0.2	0.1	0.0	-0.1	0.0	0.0	-0.1	**0.8**
Vacancies/Unemp	**-0.7**	**1.0**	0.1	0.1	0.2	**-0.6**	0.0	**0.5**	-0.2	0.1	0.3	**-0.8**
Self-Emp Share	0.1	**-0.5**	0.4	**4.1**	**2.1**	1.1	0.1	**-1.5**	0.4	0.2	**0.5**	**0.6**
Pareto Parameter	0.3	0.0	0.4	0.0	**0.6**	0.0	0.0	0.0	0.0	0.2	**0.6**	0.3
% wkr leaving Firm	-0.1	**0.5**	0.2	0.0	-0.1	-0.1	0.0	0.1	0.0	-0.1	-0.2	0.4
% wkr J-to-U	0.0	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.8**
Av wage growth J-J	0.2	-0.2	**-1.2**	0.0	**-1.5**	0.0	0.0	0.1	0.0	**-0.5**	**-0.9**	0.2
% J-J high-low	-0.1	0.2	**0.5**	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.3	-0.2
Training Intensity	0.0	**-0.6**	-0.2	-0.3	0.0	**0.6**	-0.2	**0.6**	**-1.2**	0.1	0.2	-0.4
Trng large-small	-0.1	-0.1	0.0	0.0	-0.2	0.1	0.0	0.0	0.0	0.0	-0.1	0.0
Direct/wage cost	0.0	-0.1	0.1	0.0	0.4	-0.1	**1.0**	0.0	0.0	0.1	0.1	0.1
% wage inc 40 yrs	0.0	-0.1	-0.2	0.0	-0.2	**-1.5**	-0.1	**2.0**	**-0.7**	0.0	0.0	**-0.5**

The table reports the elasticity for moments to each parameter where we highlight in bold the elasticities greater than 0.5 in absolute values. The elasticities are measured by calculating the percent increase in each moment after a 1% change around the calibrated parameter value keeping the rest parameters fixed.

Second, we focus on the parameters directly related to training. The parameter c_s pins down the importance of direct training costs and is identified by the ratio of direct costs to wage costs of training. The parameter ζ determines how training translates to efficiency units and has a large impact on training intensity, the wage increase after 40 years, and the self-employment share because higher training returns make the wage sector more attractive.[2] Finally, training intensity decreases with μ_F — the share of the training cost firms pay. This indicates that optimal training levels are mostly determined by firm choices (as they are lower than workers' choices), which indicates the presence of inefficient training levels. We will discuss optimal policies to reduce training inefficiency in Section 8.

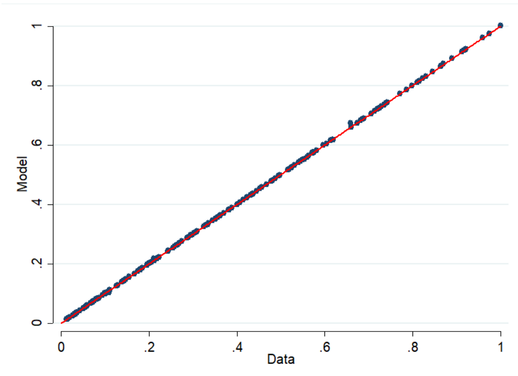
[2]Moreover, γ_s that defines the convexity on the training function also has the biggest impact on the wage growth and self employment share through the impact on training intensity. The signs are more complicated to analyze due to the training function choice because an increase of γ_s increases the marginal returns ($\zeta\gamma_s s^{\gamma_s-1}$) but reduces the overall training returns (ζs^{γ_s}) for $s<1$.

We now focus on the main parameters that mediate our channels. The breaking contract cost friction is composed of two parameters. The convexity in the cost, γ_p, has the biggest impact on the average wage growth from job-to-job transitions, as a higher γ_p makes it more costly to increase the leaving probability in response to higher wage offers. Moreover, c_p has the greatest impact on wage growth in job-to-job transitions for the same reason, but it also impacts labor market outcomes such as market tightness, self-employment share, and the Pareto parameter more strongly. It also has a positive impact on training intensity, in line with our analytical model. Finally, the share of workers who are exogenously separated, δ, increases the unemployment rate and the job-to-unemployment rate, while reducing market tightness (due to more unemployed people). Moreover, it has a relatively strong negative effect on training intensity, in line with our analytical model as well.

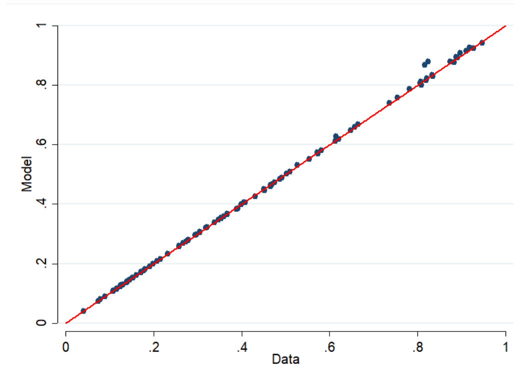
B.4 Alternative Cross-Country Calibration

In this section, we calibrate the model to all 100 countries for which we have data on training. The difference with the main specification is that we use the observed variation in self-employment and also match training intensity directly. Therefore, in this calibration we target 5 different moments: (1) Real GDP per capita, (2) traditional sector employment share, (3) exogenous separation, (4) difference in endogenous separation from job-to-job transitions and (5) training intensity. We keep the calibrated parameters from the U.S. baseline calibration except for δ, c_p, A_m, A_T and differently from the baseline calibration we also let ζ vary across countries.

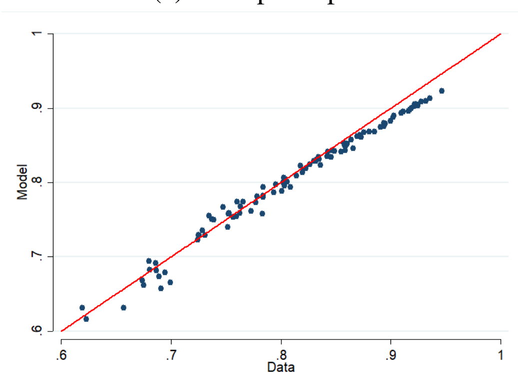
We first show how the model fits the targeted moments in Figure B.2. On the x-axis we show the moments in the data, on the y-axis we show the moments in the model and we plot the 45-degree line for the targeted moments except training. Overall, our model matches the targeted moments well. In Figure B.3a we plot training intensity from the data and the model as a function of GDP per capita, and in Figure B.3b we plot the training intensity in the model (y-axis) and data (x-axis). When we let ζ change across countries, the model exactly matches training intensity for every country.



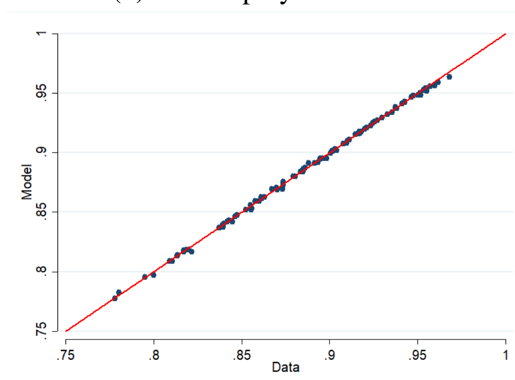
(a) GDP per capita



(b) Self employment share



(c) % workers in same firm

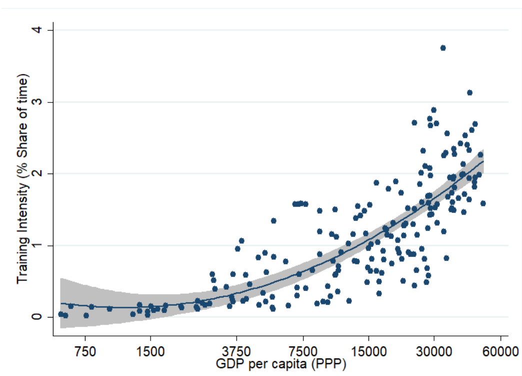


(d) % workers still employed

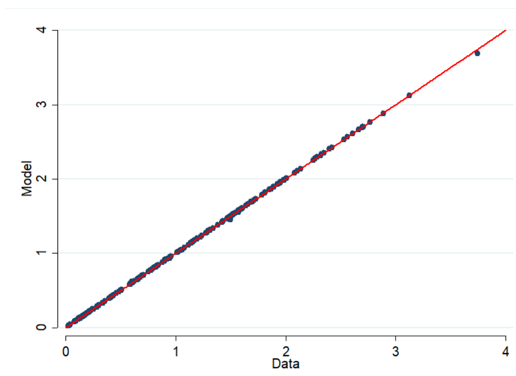
Figure B.2: Cross Country Targeted Moments

Note: This figure shows the targeted moments in the model (vertical axis) and in the data (horizontal axis). Panel A shows GDP per capita. Panel B shows self employment. Panel C shows the share of workers who remain in the same firm after one quarter. Panel D shows the share of workers who are employed in any firm for two consecutive quarters.

Finally, we show the calibrated parameters given by the cross-country calibration in Figure B.4. We get the same elasticities and patterns from the main specification. Interestingly, we add the dynamics on the training productivity and show that this parameter is mostly flat with respect to GDP per capita. This result reinforces our conclusion that most training differences come from factors captured in our parameters.



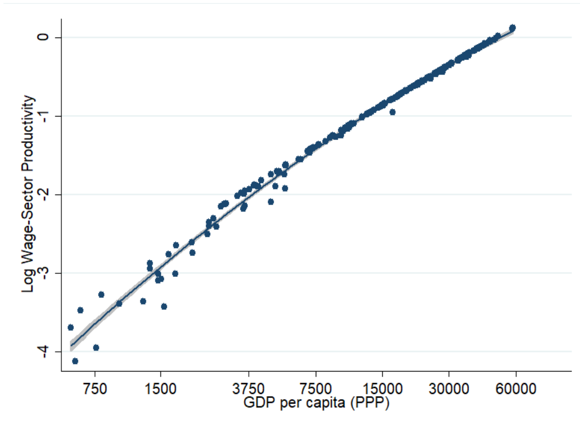
(a) Training Intensity



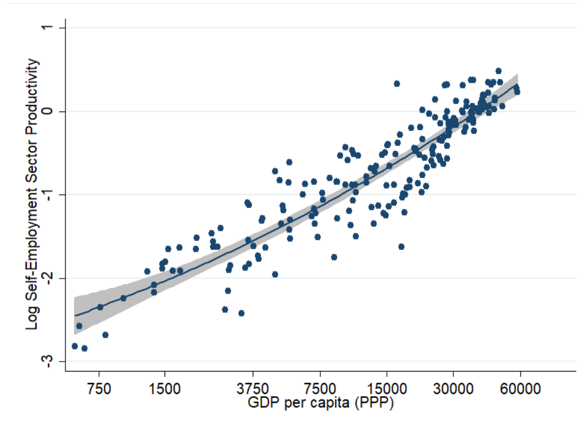
(b) Training Data vs Model

Figure B.3: Training in Data and Model

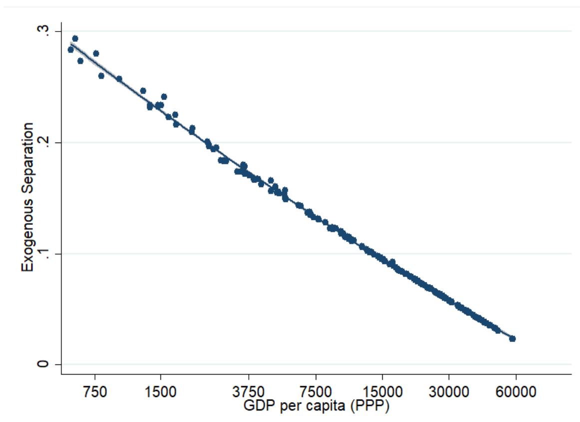
Note: This graph shows the quadratic fit of the cross-country training intensity (measured in the share of time that an average worker spends in training) as a function of Los(GDP per capita). The green line represents the quadratic fit for the cross-country measure in the model and the blue line represents its counterpart in the data. The grey shadow represents the 95% confidence intervals.



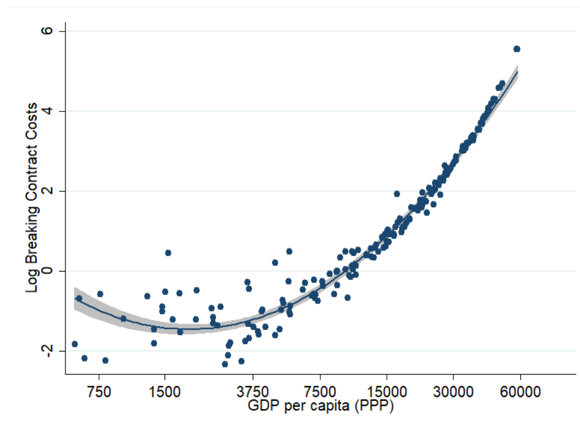
(a) Modern Sector Productivity



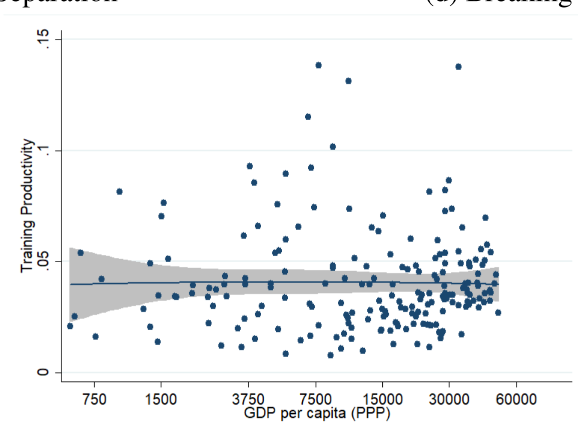
(b) Self-Employment Sector Productivity



(c) Exogenous Separation



(d) Breaking Contract Costs

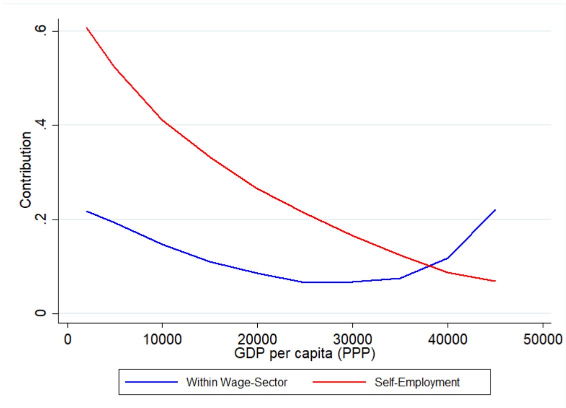


(e) Productivity of training

Figure B.4: Cross Country Calibrated Parameters

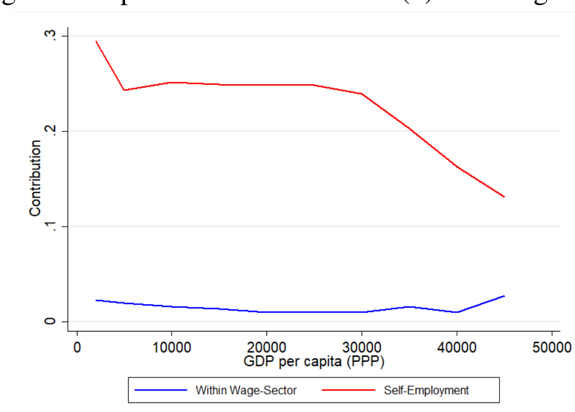
Note: This figure shows the calibrated parameters for each economy in the model as a function of Log(GDP per capita).Panel A shows the wage sector productivity (A_M in the model). Panel B shows the relative productivity between the self-employment sector and the wage sector (A_T/A_M). Panel C shows the quarterly exogenous separation rate implied by the model (δ). Panel D shows the log of the breaking contract costs ($\log(c_p * A_M)$).

B.5 Training Decomposition



(a) Exogenous Separation

(b) Breaking Contract Costs



(c) Productivity

Figure B.5: Training decomposition change by parameter and sector

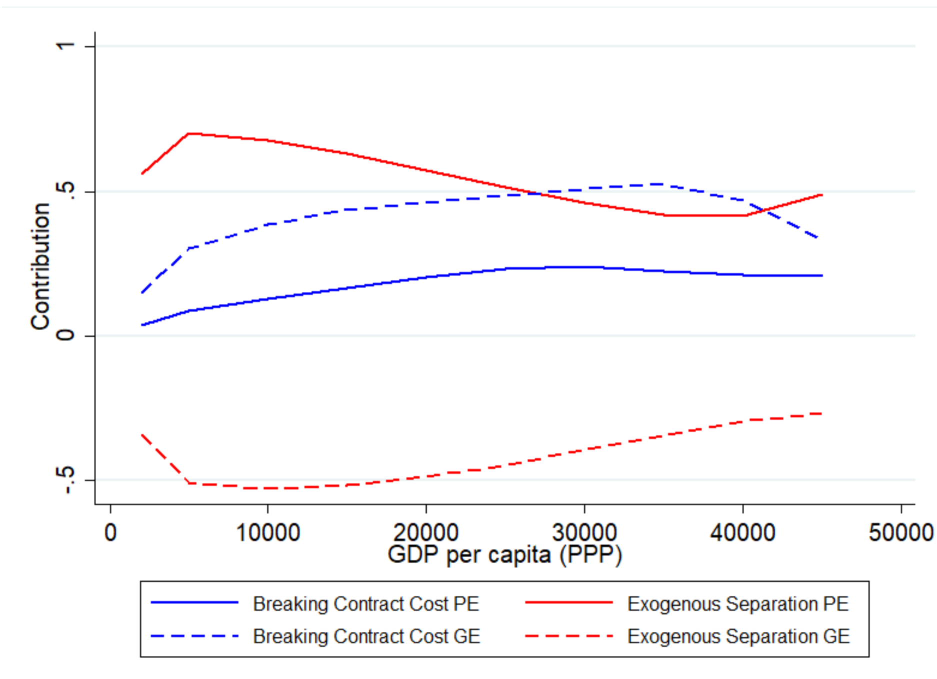


Figure B.6: Partial and General Equilibrium from Parameter Changes

B.6 Training Investment Inefficiency and Subsidies

It is clear that training investments are usually not efficient due to worker turnover and incomplete contracts. In this section, we do an extensive description of government subsidy policies and show that training subsidies are indeed very common. Figure B.7 presents the countries for which we found data on training incentives. We review government policies from countries in all continents (we provide examples for 36 countries from all income levels) and we also show examples for 21 U.S. States for which there is data available on government policies to incentivize employer-provided training. We present the survey for the cross-country policy examples in Table B.2 and the examples for the U.S.' States in Table B.3.

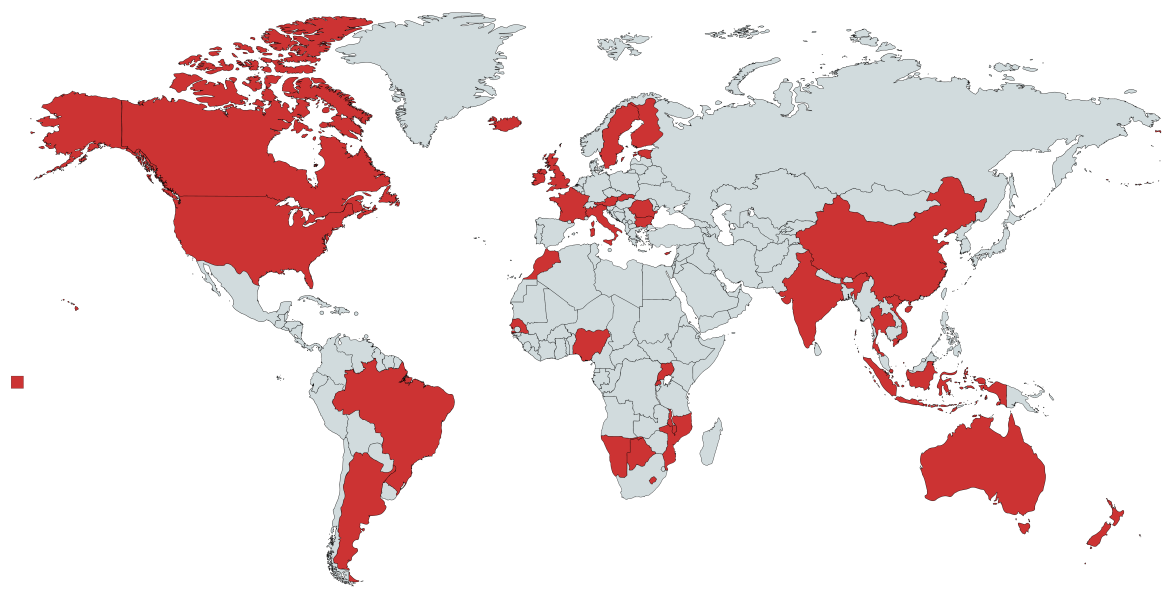


Figure B.7: Examples for Government Training Incentives

Table B.2: Training Subsidies Across Countries

Country	Year	Subsidy or Incentive to employer
Argentina	2000 - present	30% of training costs are tax deductible
Australia	2019 - present	50% or $2,200 of training costs granted (or to employee)
Austria	2002 - 2015	120% of training costs tax deductible
Botswana	1985 - present	200% training costs tax deductible
Brazil	2012 - present	No set maximum of tax deductible training expenses
Bulgaria	2007 - present	80-90% of training costs granted
Canada	2014 - present	50-83% of training costs granted
China	2015 - present	8% of total payroll may be deducted from taxable income
Cyprus	1979 - present	60-80% of training costs granted
Czech Republic	1992 - present	100% of training costs tax deductible
Estonia	2012 - present	100% of training costs tax deductible
France	2005 - present	Wages of trainees are paid
Finland	2014 - present	50% of employee's average wage is tax deductible
Iceland	1998 - present	100% of training courses reimbursed
India	2016 - present	Funding of $20-100 for IVT (apprenticeship)
Indonesia	2019 - present	200% of learning costs tax deductible for corporate taxpayers
Ireland	1999 - present	Maximum not specified, but depends on project and industry
Italy	2017 - present	50% of training costs of tax deductible
Lesotho	1980 - present	50% of wage bill reimbursed
Malawi	1999 - present	20-50% of training costs reimbursed
Mauritius	2003 - present	75% of training costs reimbursed
Morocco	2014 - 2020	20% of training costs reimbursed (large projects)
Mozambique	2002 - present	5-10% of taxable income may be deducted
Namibia	1995 - 2020	75% of training costs reimbursed
New Zealand	1983 - present	Funding of $5,415 (or to trainee)
Nigeria	1971 - present	Reimbursement of 50% of payroll tax paid
Romania	2000 - present	No set maximum of tax deductible training expenses
Rwanda	2014 - present	70% of training costs granted
Senegal	2014 - present	80-90% of training costs granted
Singapore	2016 - present	90% of training course fees reimbursed
Slovakia	2003 - present	100% of training tuition tax deductible
Sweden	1996 - present	Training costs are tax deductible (no specified maximum)
Thailand	2002 - present	200% of training costs tax deductible
Uganda	1997 - present	100% of training costs tax deductible
United Kingdom	2017 - present	95% of IVT (apprenticeship) costs paid
Vietnam	2019 - present	100% of training costs subsidized for female-owned enterprises

Table B.3: Training Subsidies within United States

Country	Year	Subsidy or Incentive to employer
Alabama	2014 - present	75% of training costs reimbursed
Arizona	2015 - 2020	50-75% of training costs reimbursed
Colorado	2018 - present	60% of training costs reimbursed
Florida	1993 - present	50-75% of training costs reimbursed
Georgia	1994 - present	50% of training costs tax deductible
Hawaii	1991 - present	50% tuition costs reimbursed
Illinois	1992 - present	50% of training costs reimbursed
Kentucky	1984 - present	50% of training costs reimbursed
Maryland	1989 - present	50% of training costs reimbursed
Massachusetts	2008 - present	50% of training costs reimbursed
Mississippi	2013 - present	50% of training costs reimbursed
Montana	2005 - present	Funding of $5,000 for training
Nebraska	2005 - present	Funding of $800-4,000 for training
New Hampshire	2007 - present	50% of training costs reimbursed
New Jersey	1992 - present	50% of training costs reimbursed
New Mexico	1972 - present	50-75% of training costs reimbursed
Pennsylvania	1999 - present	Funding of $600-1,200 per trainee
Rhode Island	2006 - present	50% of training costs reimbursed
Washington	1983 - present	50% of training costs reimbursed
Wisconsin	2012 - present	50% of training costs reimbursed
Wyoming	1997 - present	Funding of $1,000 per trainee

B.7 Our Channels in the Data

In Table B.4 we show suggestive evidence on the correlations between training investments and job turnover measures from Donovan, Lu and Schoellman (2020), self-employment, firm size distribution, and institutional quality proxies in the data. It shows the results of regressing the share of employment exposed to training on GDP per capita, the share of employment in small firms (to account for the composition effect not captured by job separation), the probability of staying in the same job, and the first principal component of all 5 institutional measures from the World Bank Worldwide Governance Indicators.[3] The semi-elasticity of GDP per capita with respect to our training measure is 8.69. As we add each one of the explanatory variables, we show how the coefficient on GDP per capita decreases. Once we add the first principal component that includes all the other variables, we explain all the correlation between GDP per capita and training, which suggests that institutional quality, job separation, and self-employment captures most of the pattern described.

Labor Market Institutions. From the model, it is clear that things affecting separation rates, the probability of hiring, or the vacancy costs will affect contracts and training investments. It is intuitive to think that higher unemployment benefits or higher firing costs will generate lower levels of training. These could be potential mechanisms to explain why more developed economies invest more in training, and thus, we test this hypothesis in the data. We rely on the labor market institutional indexes constructed by Botero et al. (2004) to understand how the cost of firing workers and labor market institutions (such as the minimum wage and unemployment

[3]We use data from the PWT and the World Bank Indicators for GDP per capita, self-employment, and capital stock. For institutional quality we rely on the World Bank Worldwide Governance Indicators, which provides indexes on country-specific institutional characteristics. The characteristics provided by the WGI are: "Voice and Accountability," "Political Stability," "Government Effectiveness," "Regulatory Quality," "Rule of Law," and "Control of Corruption." Moreover, we use data on separation rates estimated using the results provided by Donovan, Lu and Schoellman (2020). Due to a mismatch between their sample and our sample, we are not able to relate these two measures directly. Nevertheless, we can conduct a 2 step estimation process. We first regress the probability of staying in the same job on all the institutional variables, which gives us an R-squared of almost 80%. Then, we predict the probability of staying in the same job for all countries using the institutional indicators which gives us predicted separation rates for most of the country-years in our sample.

Table B.4: Share of Workers Exposed to Training

(a) WB-ES and EU-CVT

	(1)	(2)	(3)	(4)	(5)	(6)
GDP per Capita	8.69***	4.69**	5.77***	7.43***	5.98***	1.32
	(0.61)	(1.89)	(1.14)	(0.70)	(0.70)	(1.32)
Log per Capita K		3.58**				
		(1.69)				
self-employed			-0.14***			
			(0.047)			
Prob Same Job				20.2***		
				(7.58)		
1st comp Institutions					1.87***	
					(0.44)	
1st comp All						5.61***
						(1.04)
Constant	-57.8***	-57.0***	-25.3*	-63.5***	-33.4***	11.6
	(6.75)	(6.68)	(12.9)	(7.38)	(7.52)	(12.8)
Year FE	YES	YES	YES	YES	YES	YES
Observations	211	211	211	211	211	211
R^2	0.626	0.635	0.640	0.637	0.651	0.663

(b) WB-ES

	(1)	(2)	(3)	(4)	(5)	(6)
GDP per Capita	8.09***	3.54***	7.63***	7.17***	6.63***	2.87***
	(0.64)	(1.02)	(0.62)	(0.66)	(0.70)	(0.83)
self-employed		-0.22***				
		(0.041)				
% Emp in Small Firms			-25.1***			
			(3.91)			
Prob Same Job				18.2**		
				(7.40)		
1st comp Institutions					1.33***	
					(0.50)	
1st comp All						-5.03***
						(0.75)
Constant	-52.5***	-1.98	-43.3***	-58.4***	-37.5***	-7.02
	(5.33)	(10.6)	(5.34)	(6.41)	(6.44)	(7.33)
Year FE	YES	YES	YES	YES	YES	YES
Observations	194	194	194	194	194	194
R^2	0.517	0.570	0.567	0.532	0.534	0.600

benefits) correlate with our measure of training. We regress our measure of training from the ES and CVT on GDP per capita and each index separately, year and country fixed effects and show the results in Table B.5. Training increases as the legally mandated notice period to fire workers increases, meaning that as the firing costs increase, turnover rates decrease, and agents stay longer in their jobs. In our sample, the amount of severance payment does not seem to be significant to explain training on-the-job. Moreover, rows 4, 5, and 6 have different measures on the strength of unemployment benefits in different countries. As unemployment benefits increase, training investment decreases. This shows that when the workers' outside option is better, they are harder to retain workers and training investments decrease. This same pattern is observed when countries have meaningful minimum wages and outside options are higher. Nevertheless, although all these measures increase the explanatory power over training on-the-job, they do not account for part of the explanatory power of GDP per capita. These results reflect the fact that, although important, these measures as unemployment benefits and labor market characteristics, which are not included in our model (i.e., differences in minimum wages, laws to protect workers,or firing costs) are not the key elements to explain the positive correlation between training and income. This result is consistent with Donovan, Lu and Schoellman (2020), who find that labor market institutions are an important determinant of cross-country variation in labor market flows (job separation, destruction, and job-to-job transitions) but that they do not explain the trend relationship between development and labor market flows.

Table B.5: Training and Labor Market frictions (Botero et al 2004)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Log(GDP pc)	8.41***	8.50***	18.9***	17.8***	20.0***	7.66***
	(1.45)	(1.50)	(4.48)	(4.55)	(4.76)	(1.36)
Legally mandated notice period	0.67*					
	(0.35)					
Legally mandated sev payment		-0.059				
		(0.22)				
Months of contributions for U.B.			12.0			
			(12.2)			
% monthly salary deducted for U.B.				-14.1**		
				(6.94)		
Waiting period for U.B.					-32.7***	
Minimum Wage Index						-10.5**
						(4.04)
Constant	-42.7***	-39.6***	-150***	-117**	-121***	-22.4*
	(13.2)	(13.3)	(44.2)	(47.1)	(44.7)	(13.4)
Observations	183	183	132	132	132	184
R^2	0.421	0.412	0.389	0.395	0.430	0.440
Log(GDP pc) restricted sample	8.42***	8.42***	18.8***	18.8***	18.8***	8.36***
	(1.48)	(1.48)	(4.52)	(4.52)	(4.52)	(1.43)
Observations	183	183	132	132	132	184
R^2	0.412	0.412	0.384	0.384	0.384	0.415

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

B.8 Cost Shares

In this section, we provide a cost share analysis. In our model, worker and firm choices of training depend on the marginal revenue and the cost shares as shown. Note that we could make different assumptions on what the cost shares are, and thus, we could have different training patterns.

Proposition 3 (Joint Internal Efficiency) *In a firm with productivity level* ϕ *and wage* $w(\phi)$ *if*

$$\mu(\phi)^* = \frac{MR_W(\phi)}{MR_W(\phi) + MR_F(\phi)}$$

then

$$s^*(\phi) = \left(\frac{\zeta e^{\alpha} \gamma_s (MR_W(\phi) + MR_F(\phi))}{(1+r)c_s} \right)^{\frac{1}{1-\gamma_s}}$$

which maximizes the joint surplus from training

$$\max_s \frac{\zeta e^{\alpha} s^{\gamma_s}}{1+r} (MR_W(\phi) + MR_F(\phi)) - c_s s$$

Proposition 3 suggests there is a unique division of training costs that maximizes the joint surplus of firms and workers from training. However, there is still under-investment in training because of the incomplete contract (Acemoglu, 1997) — workers and firms cannot internalize the benefits of training for future employers if separation occurs.

We first show that the marginal return from training increases with firm productivity for both workers and firms, and that this increase is faster for firms. On the one hand, as firms become more productive, the probability of losing the worker is lower, which means firms will enjoy higher revenue from workers for longer. That dynamic, jointly with the increase in training returns with firm productivity, generates the increase in firms' training marginal returns depicted in Figure B.8.b. Moreover, workers have larger expected revenue from training as they will capture the increase in human capital if separation occurs, thus having larger marginal revenue than firms as shown in Figure B.8.a for every firm.

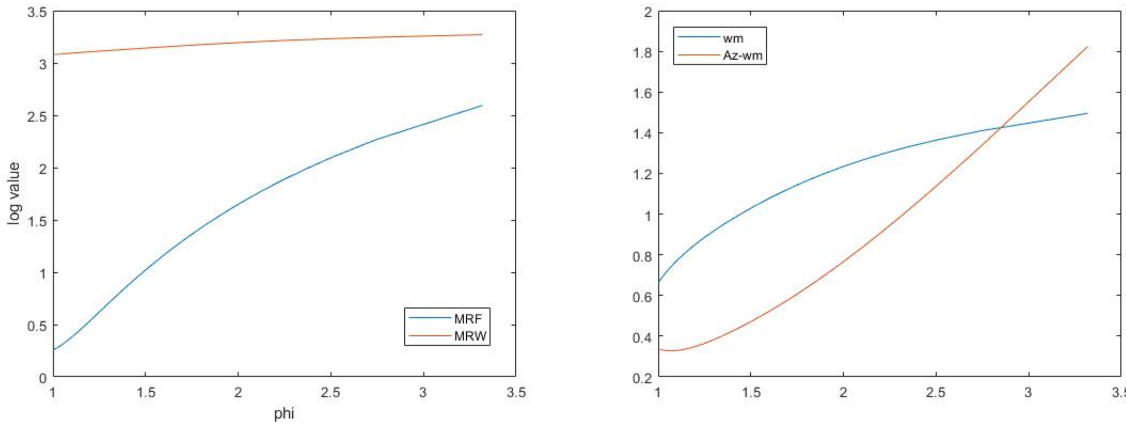


Figure B.8: Marginal Returns from Training

To think about human capital investments, we must also consider the investment costs, which are also increasing due to the opportunity cost. Note that although there is a constant direct cost –C_s– workers lose 70% of production time while being trained. Figure B.9 shows the optimal training levels workers and firms would choose when the firm pays for all training costs (Figure B.9.a in blue), when the worker pays all the costs (Figure B.9.a in orange) and when each one pays for the share they capture from the investment (joint internal efficiency case in Figure B.9.b). Workers have larger probabilities of leaving to better firms when their firms are small and unproductive, and thus, the difference between firms and workers' marginal revenues from training are the largest at the bottom of the firm productivity distribution and workers will want to invest in training more than firms. When firms become more productive they are willing to invest more in training as the increase in revenue is larger than the increase in costs, but the opposite is true for workers. In the joint internal efficiency case, as the ratio MR_F/MR_W increases, firms will start paying a higher share of the training cost which decreases the training investment as firms want lower levels of training than workers.

Figure B.10 shows the optimal level of training for three relevant cases (joint internal efficiency, calibrated shares and firm paying all the costs). When firms pay all the cost or when the

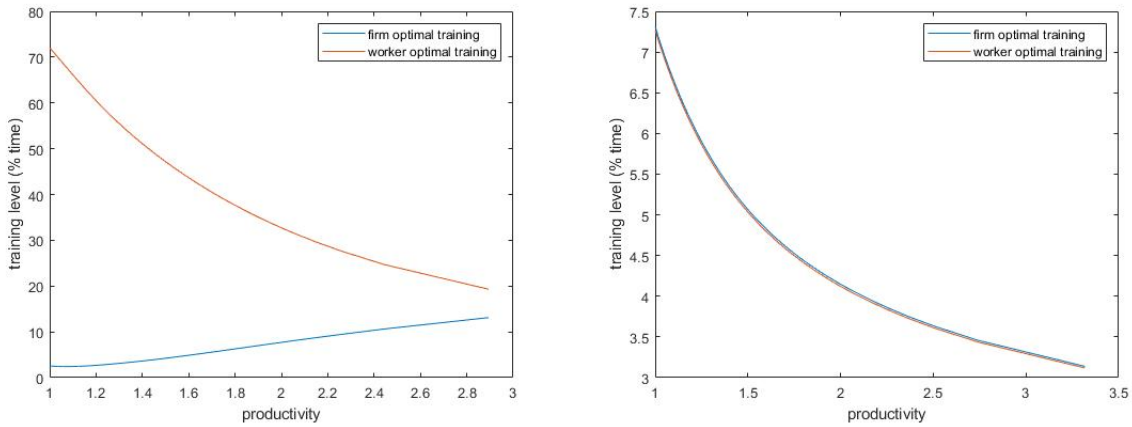


Figure B.9: Workers and Firms Optimal Training Levels

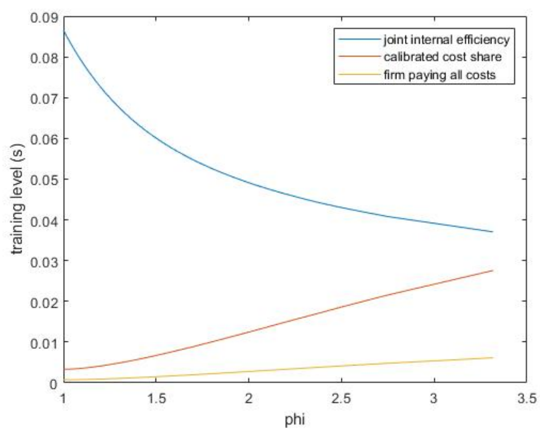


Figure B.10: Training and Cost Shares

share of the cost firms pay is constant, training levels increase with productivity. Nonetheless, the case of joint internal efficiency is different because training investments decrease with productivity, which is counterfactual. This evidence suggests that more productive firms do not seem to finance substantially larger shares of training costs than smaller firms, which, if true, would generate a decrease in training with productivity.

Appendix C

Appendix of Chapter 3

C.1 Brazilian Economic Background

Up to the 1990s, Brazil was a relatively closed economy to international trade. In the 1990s, with the economic liberalization, reductions in import tariffs, and the Mercosur Agreements, Brazil began opening to international trade. After 1999, exports started to increase substantially due to changes in the exchange rate regime and the large devaluation episode. This process sped up after 2002, with a new depreciation episode and an improvement of international agricultural prices. Table C.1 shows the trends of exports for manufacturing goods, agricultural goods, and fuel over our sample period. It is clear that there was a sharp increase in exports after 2000, and that manufacturing goods represent a large share of Brazil's exports.

Moreover, Rocha et al. (2008) explain how Brazil's exports are highly diversified across a variety of products. Apart from agricultural goods, Brazil intensively exports chemical products, pharmaceutical products, aircrafts, automobiles, and home appliances. In 2004, there were more than 10,000 different 8-digit HS products exported by more than 15,000 firms.

Table C.1 presents the share of Brazil's exports to each destination. In the 1990s, thanks to the Mercosur agreement, there was an increase in the share of exports destined to Latin American

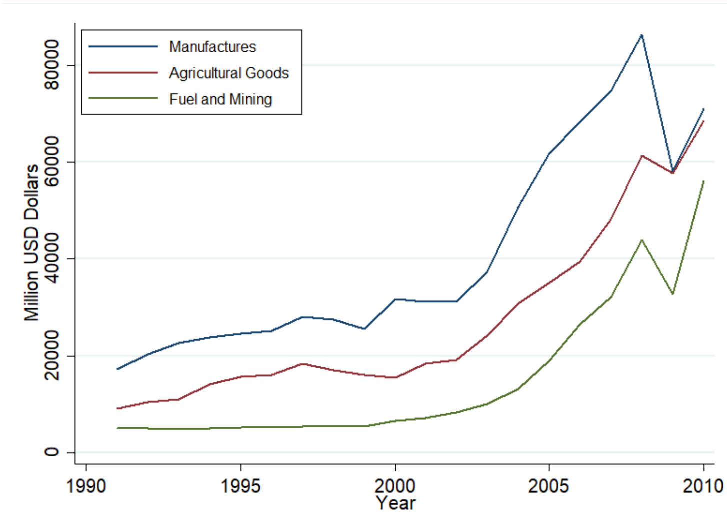


Figure C.1: Brazil's Exports in 1990–2010

Note: The data come from the WTO. This graph shows the value of exports in millions of dollars for manufacturing goods, agricultural goods, and fuels and mining products in the period 1990–2010.

countries, in particular Argentina to which its share increased from 2% in 1990 to 11% in 2000. While the U.S. was one of the biggest markets for Brazilian exporters in 1990 with 25% of total exports, this share decreased to 10% in 2010. Moreover, between 1990 and 2010, there was an increase in the share of exports going to East Asia and the Pacific, mostly explained by the increase in exports going to China (1% in 1990 to 15% in 2010). The most important takeaway from these shares is that Brazil exports to a wide variety of destinations with around half of total exports going to richer economies and half going to other developing economies.

Table C.2 presents the share of total exports, the value, and the revealed comparative advantage index for main products Brazil exported in the years 2010 and 1990. 22% of Brazil's exports in 1990 and 42% in 2010 were raw materials. This means that around 80% (60%) of its exports were manufactured goods in 1990 (2010). Moreover, although the share of raw materials in total exports increased in this period, it is worth noting that the export value of manufactured products also substantially increased.

Table C.1: Share of Exports (%) by Trading Partners

	2010	2000	1990
By Region			
Europe & Central Asia	25.63	30.78	31.93
East Asia & Pacific	25.11	10.93	15.34
Latin America & Caribbean	23.26	24.99	11.67
United States	9.64	24.29	24.62
Middle East & North Africa	7.33	3.35	0
Sub-Saharan Africa	2.49	1.52	1.91
By Country (Top 15)			
China	15.25	1.97	1.22
United States	9.64	24.29	24.62
Argentina	9.17	11.32	2.05
Netherlands	5.07	5.07	7.94
Germany	4.03	4.58	5.69
Japan	3.54	4.49	7.48
United Kingdom	2.3	2.72	3.01
Chile	2.11	2.26	1.54
Italy	2.1	3.89	5.14
Russian Federation	2.06	0.77	0
Spain	1.93	1.83	2.24
Venezuela	1.91	1.37	0.85
Korea, Rep.	1.86	1.05	1.73
Mexico	1.84	3.11	1.61
France	1.79	3.25	2.87

Note: This table presents the share of exports to each destination market. The data are collected from the WITS (the World Integrated Trade Solution). The countries and Regions are ranked by the share of exports in 2010.

Table C.2: Exports by Products

	Product Share (%)		Value (U.S.$ Mill)		RCAI	
	2010	1990	2010	1990	2010	1990
By Type						
Raw materials	41.93	21.37	84671	6713	2.93	1.84
Intermediate goods	27.29	39.01	55109	12252	1.28	1.75
Consumer goods	14.62	20.81	29517	6537	0.44	0.56
Capital goods	14.27	15.45	28822	4854	0.42	0.35
By Product						
Minerals	15.63	8.93	31557	2804	10.79	10.26
Food Products	13.4	16.83	27056	5287	4.21	4.46
Vegetable	10.88	9.02	21961	2831	3.81	2.61
Fuels	9.83	2.17	19843	682	0.61	0.03
Transportation	8.55	7.32	17272	2299	0.88	0.35
Mach and Elec	8.03	11.17	16216	3509	0.28	0.32
Metals	7.14	17.17	14412	5393	0.9	2.89
Animal	6.7	2.07	13526	650	3.46	0.8
Chemicals	5.06	4.89	10221	1535	0.57	0.62
Wood	4.33	5.28	8740	1659	2.11	0.95
Miscellaneous	2.98	2.43	6023	762	0.33	0.17
Plastic or Rubber	2.65	2.56	5341	804	0.57	0.5
Stone and Glass	1.96	1.37	3954	431	0.36	0.56
Textiles and Clothing	1.12	3.97	2265	1248	0.28	0.67
Hides and Skins	0.92	1.03	1865	323	1.5	1.62
Footwear	0.82	3.78	1653	1188	1.07	1.95

Note: This table presents the share of exports in Columns 1–2, the value of exports in Columns 3–4, and the revealed comparative advantage indices in Columns 5–6 for the years 2010 and 1990. The data are collected from WITS (World Integrated Trade Solution). The products and products types are ranked by the share of exports in 2010.

C.2 Brazilian Economic Background and Informality

One possible drawback of the analysis is that we focus on the formal sector. Therefore, it is important to discuss the economic and political background of the Brazilian informal labor market in recent decades. The 90s was a period of instability for the Brazilian economy. Brazil opened up to international trade, with the Mercosur Agreements signed in 1991 and 1994. However, the 90s started with another major recession that led to high unemployment. Under these circumstances, the share of unregistered employees in total employees grew by 2 percentage points from 1990 to 2003. The 2000s were, in some sense, the opposite of what the 90s were. In the 2000's, the inflation was finally tamed, and the economy was considerably more open due to those policies adopted in the 90s. After 2002, an economic expansion took place with a rapid increase in GDP, improvements in social-economic indicators, and a considerable decrease in the amount of unemployment and unregistered workers. For an extensive review of policies and the background about the informal sector in Brazil, see Dix-Carneiro et al. (2019).

Figure C.2 shows unregistered workers as a share of total employees. The informality rate sharply declined in recent decades, from around 33% in the 1990s to 23% in the 2010's. Besides employees, Brazilian employment also includes self-employed workers, employers, and unpaid workers, and these three types of employment may not appear in the RAIS (except for employers who receive a wage). Figure C.3 shows the share of self-employed workers, employers and unpaid workers in Brazilian total employment. These three types of employment represented 30–40% of Brazilian employment in the 90's and 2000's.

We obtain Brazilian Population Census from IPUMS International to compare experience-wage profiles for Brazilian wage workers and self-employed workers. We estimate experience-wage profiles by applying the HLT method. Differing from the Mincer regressions estimated in Section 3.2.3, because we cannot identify individuals in Brazilian Population Census, we regress log hourly wage on a set of experience dummies, schooling, cohort effects, and year effects. We

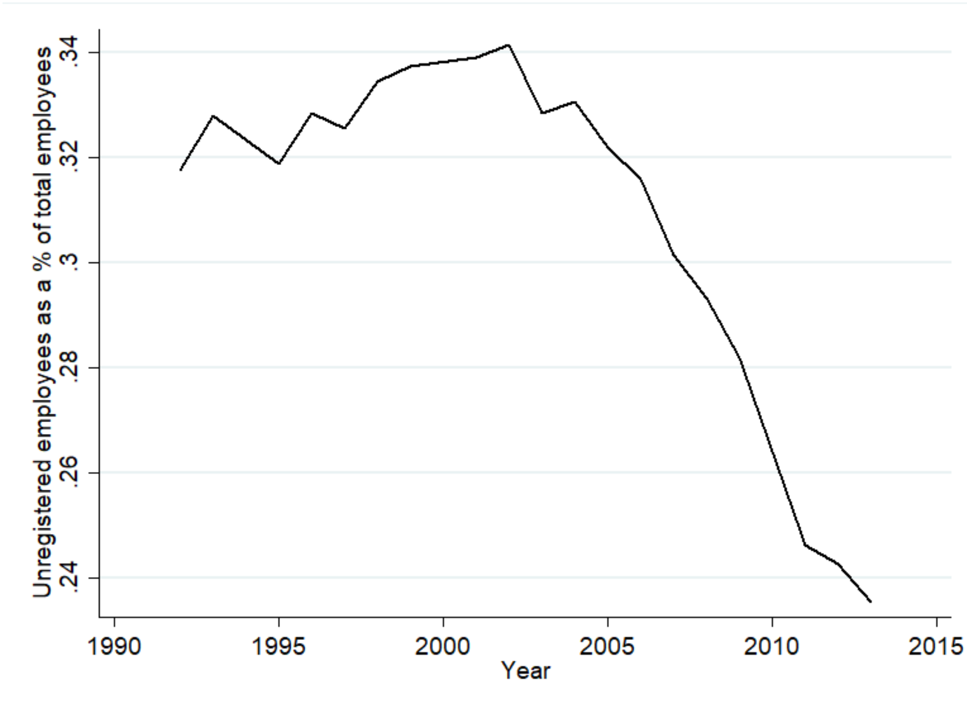


Figure C.2: Share of Unregistered Employees in Total Employees

Note: The figure shows the share of unregistered employees in total employees. The data come from the PNAD censuses.

do not enforce a first difference of log wage across years, as we are not able to identify individuals. Our regression is identical as in Lagakos et al. (2018*b*), with 10 years of no experience effects at the end of the working life and 0% depreciation rate. As shown in Figure C.5, we find that wage workers have steeper profiles than self-employed workers.

Moreover, for two years (2000 and 2010), we have information on the contract status of wage workers. We split the sample into wage workers with formal contracts and with no formal contracts. Because the data are only available for two years, we are not able to apply the HLT method. As some reference, we draw the experience-wage profiles in the cross section, following the process in Section 3.2.2. Figure C.5 plots both profiles and shows that formal workers have stepper experience-wage profiles. Dix-Carneiro et al. (2019) show that Brazilian informal workers tend to be mostly allocated in non-tradable sectors and within tradable sectors, most of workers are formally employed. Moreover, they show that the transition between formality and informality is relatively low. Therefore, given our focus on tradable industries, informality should not be a

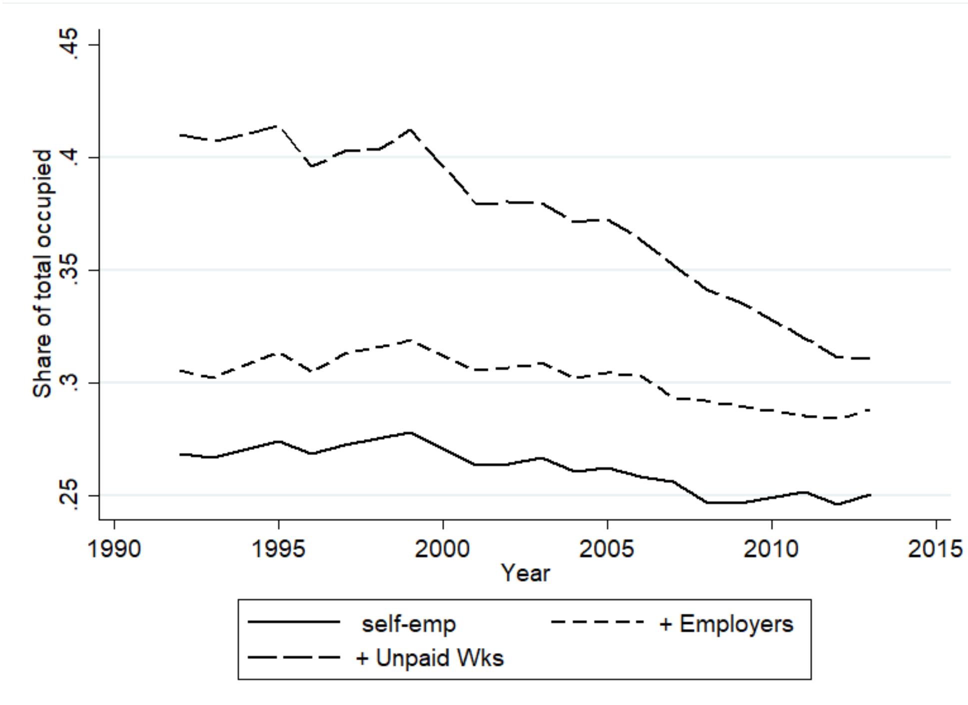


Figure C.3: Share of Non-employees Occupied Population

Note: The data come from the PNAD censuses. The share of self-employed people represents the ratio of the amount of self-employed workers to total occupied population. The share "+ employers" is the share of self-employed and employers in total occupied population. The share "+ Unpaid" is the share of self-employed, employers, and unpaid workers in total occupied population.

big issue. Nevertheless, even considering informal workers, because exporters are mostly formal firms, it is likely that non-exporters hire informal workers more intensively than exporters. By missing informal workers, we may underestimate the difference in experience-wage profiles between exporters and non-exporters in our main results.

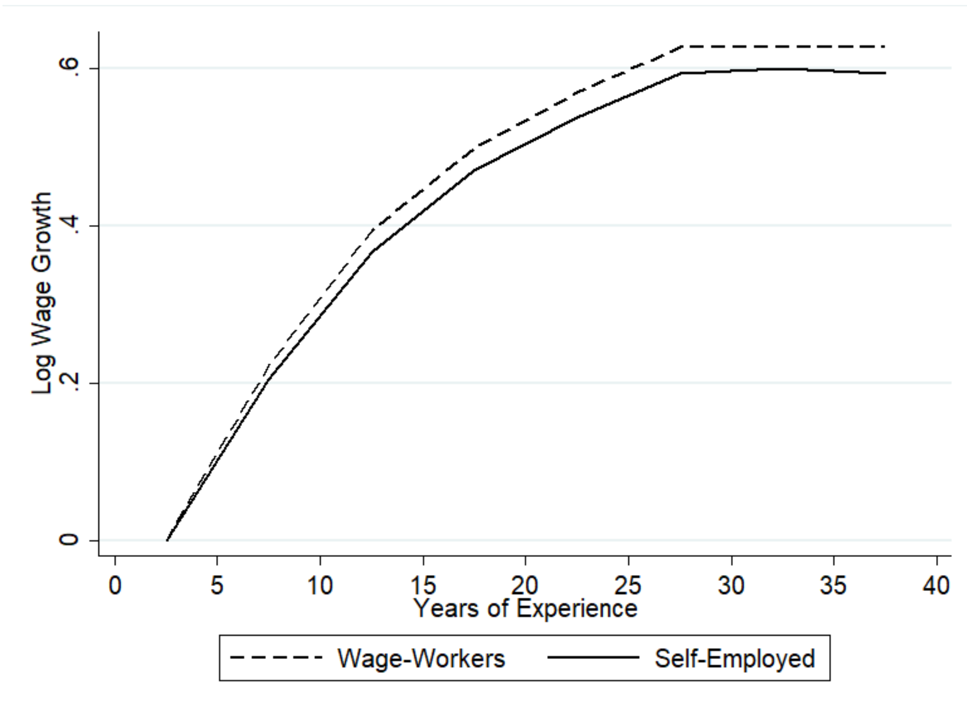


Figure C.4: Experience-wage Profiles for Wage Workers and Self-Employed

Note: The figure shows experience-wage profiles separately for male wage workers and male self-employed workers, derived from the HLT method (identical regression as in Lagakos et al. (2018*b*)). In applying the HLT method, we assume 10 years of no experience effects at the end of the working life (31–40 years of potential experience) and a 0% depreciation rate. We rely on Brazilian Census data available in IPUMS for the years 1991, 2000, and 2010.

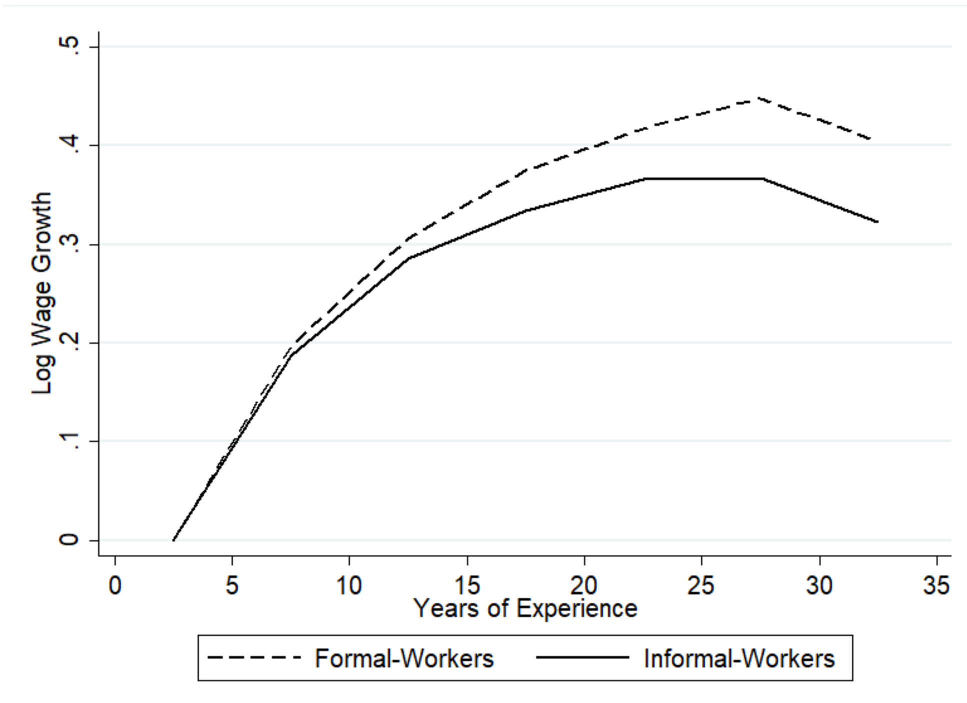


Figure C.5: Experience-wage Profiles: Workers With and Without Formal Contracts

Note: The figure shows experience-wage profiles separately for male wage workers with and without formal contracts. In each year, we obtain experience-wage profiles by computing the average of log hourly wage for workers in each 5–year experience bin, separately for workers with and without formal contracts. We normalize the value of the first experience bin (1–5 years of experience) to be 0 for each experience-wage profile. Finally, we average profiles across years to obtain experience-wage profiles for workers with and without formal contracts, respectively. We rely on Brazilian Census data available in IPUMS for the years 2000 and 2010.

C.3 Description of the RAIS and Customs Data

We use Brazilian employer-employee data named RAIS (Relacao Anual de Informacoes Sociais). Establishments receive 14-digit unique permanent tax codes (CNPJ), from which we can identify firms by the first 8 digits of the code (Muendler, Rauch and Tocoian, 2012). For this study, we focus on firms and aggregate establishments into the affiliated firms. Firms are mandated by law to annually provide workers' information to RAIS, and therefore the data contain annual information on all workers employed in the Brazilian formal sector. The data are available from 1986. Nonetheless, the detailed data on age and hours worked are only available after 1994, and these two variables are important for us to accurately measure experience-wage profiles.

The occupation classification in RAIS is based on the CBO (Classificao Brasileira de Ocupaes), which has more than 350 categories and can be aggregated to 5 broad occupations (professionals, technical workers, other white-collar workers, skilled blue-collar workers, and unskilled blue-collar workers). The industry classification is based on the CNAE (Classificao Nacional de Atividade Econmica), which has 564 5-digit industries. Although there are available data on agriculture and services, we only focus on manufacturing industries, as manufacturing firms are tradable and extensively studied in the literature. The data contain monthly average wage and wages of December, which are measured by multiples of the contemporaneous minimum wage. We follow Menezes-Filho, Muendler and Ramey (2008) to transform these earnings into the Brazilian Real and deflate them to the August 1994 price level. For the cases with more than one observations per worker-year, we keep the observation with the highest hourly wage (Dix-Carneiro, 2014). Most workers are employed only at one firm in a year, and the average number of observations per worker-year is roughly 1.1.

We use firm IDs to merge the RAIS data with Brazilian customs declarations for merchandise exports collected at SECEX (Secretaria de Comrcio Exterior) for the years 19942010, following Aguayo-Tellez, Muendler and Poole (2010). Thus, we use RAIS merged with customs data for the 1994–2010 period. From Brazilian customs declaration, we have data on destination markets for all firms. We split destination markets into industrialized and non-industrialized destinations. We classify the following countries into each group:

- Industrialized destinations: US, EU Countries, Canada, Hong Kong, South Korea, Australia, Israel, Japan, New Zealand.
- Non-industrialized destinations: All the rest of the countries that are not included in the industrialized group; mainly include South American, Central American and African Countries, Russia, and China.

For customs records, we have data on export value and quantity by 8-digit HS products and destinations for the years 1997–2000. We use these additional data to provide some robustness checks as discussed in the main text.

C.4 Empirical Analysis: Additional Results

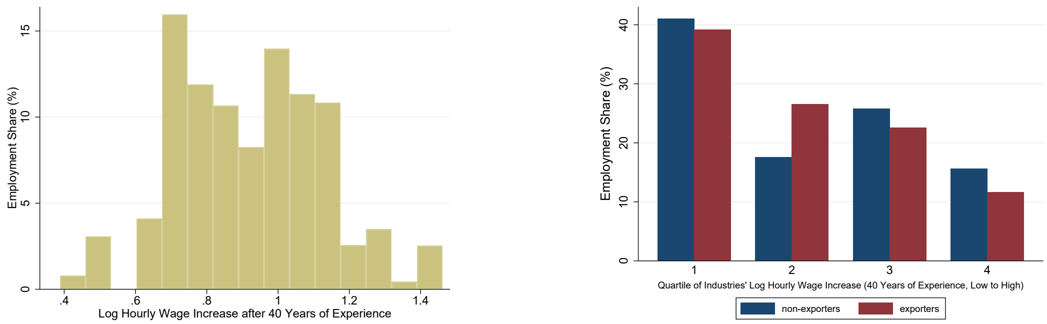
C.4.1 Between-industry Heterogeneity in Returns to Experience

The difference in the aggregate experience-wage profiles between exporters and non-exporters could be explained by different reasons. One important driver of the result could be industry composition. This is motivated by two well-established results in the literature: (1) different industries have different returns to experience (e.g., Dix-Carneiro, 2014; Islam et al., 2019); (2) trade induces industry specialization and labor reallocation, possibly driven by comparative advantage (e.g., Costinot, Donaldson and Komunjer, 2012) or home market effects (e.g., Head and Ries, 2001). Therefore, if exporters are more concentrated in industries with higher returns to experience than non-exporters, on average they will also have steeper experience-wage profiles.

We first examine the role of industry composition in driving the difference of experience-wage profiles between exporters and non-exporters. We perform regression Equation (3.1) separately for workers in each 3-digit manufacturing industry between 1994–2010. For precision, we focus on industries with more than 0.1% of total employment and at least 10 workers in each year-experience-bin. We obtain estimation results for 91 industries (99% of manufacturing employment in the sample).[1] Figure C.6a illustrates the cross-industry distribution of wage growth for a hypothetical worker with 40 years of experience in the same industry, which is computed as $5 \times (\phi_s^{1-5} + ... + \phi_s^{36-40})$. It is clear that there is a large degree of heterogeneity in returns to experience across industries.

Figure C.6b presents industry-level employment distributions in 1994–2010, for exporters and non-exporters respectively. We rank industries by returns to 40 years of experience, and for ease of description, we further split industries into 4 quartiles based on returns to experience. We find that more than 65% of workers in exporters are employed in industries with lower returns to

[1] The estimation does not work for some industries with few observations.



(a) Distribution of Log Hourly Wage Increase Across Industries

(b) Distribution of Employment by Exporters and Non-exporters

Figure C.6: Returns to Experience and Industry Heterogeneity

Note: This graph presents the results from estimating Equation (3.1), separately for workers in each 3-digit manufacturing industry between 1994–2010. Panel (a) is the cross-industry distribution of returns to 40 years of experience. Panel (b) presents the employment distribution of workers in exporters and non-exporters across industries ordered by different quartiles of returns to 40 years of experience.

experience than the median, compared to around 57% for non-exporters.

These findings have two main implications. First, trade changes workers' allocation across industries with heterogeneous returns to experience, as similarly found by Dix-Carneiro (2014). This force could generate gains or losses in workers' earnings growth, depending on each country's specialization pattern. For countries with comparative advantage in industries with higher returns to experience, trade openness can lead to higher earnings growth. On the other hand, for other countries such as Brazil, trade openness can generate lower earnings growth by allocating workers toward industries with lower returns to experience.

Second, in Brazil, industry composition cannot explain the aggregate difference in returns to experience between exporters and non-exporters. On the contrary, using industry-specific returns to experience and different employment distributions across industries for exporters and non-exporters, we find that after 40 years of experience, workers' wage increase would be 2 percentage points lower in exporters than in non-exporters due to industry composition.

In Table C.3, we explore what causes cross-industry heterogeneity in experience-wage profiles by regressing profiles on industry characteristics. We find that industries enjoy steeper experience-wage profiles, if they hire larger shares of high-school and cognitive workers. However, even controlling for education levels and occupations, there is still a large degree of cross-industry heterogeneity in experience-wage profiles unexplained.

Table C.3: Log Hourly Wage Increase (40 Years of Experience)

	(1)	(2)	(3)	(4)	(5)	(6)
log(industry employment)	-0.064**					-0.015
	(0.028)					(0.026)
Share of high-school grads		0.960***				0.696*
		(0.181)				(0.417)
Share of cognitive occupations			1.292***			0.588
			(0.278)			(0.391)
Share of employment in exporters				-0.086		-0.267*
				(0.161)		(0.139)
Differentiated industry					0.127**	-0.003
					(0.051)	(0.071)
Obs	91	91	91	91	91	91
R-squared	0.071	0.298	0.328	0.006	0.066	0.395
Mean (dep var)	0.914	0.914	0.914	0.914	0.914	0.914
S.D. (ind var)	0.888	0.135	0.106	0.212	0.480	–

Note: This table presents estimates from regressions of industry-level log hourly wage increase after 40 years of experience on industry characteristics, weighted by the number of each industry's observations in the restricted sample used to estimate profiles. An industry is defined as differentiated if its share of differentiated goods (based on 4-digit SITC goods exported by this industry) lies above the median of the share of differentiated goods across industries, according to the classification provided by Rauch (1999). The shares of high-school graduates, cognitive occupations, and exporters' employment in the workforce are computed based on our restricted sample, from which we obtained our estimates of industry-level experience-wage profiles. Cognitive occupations refer to professionals, technicians, and other white-collar workers. Industry-level employment is the average of the number of all types of workers in the raw sample (including female and part-time workers) between 1994 and 2010—which reflects actual industry size and is consistent with our treatment of firm employment. It is worth noting that our results are quantitatively very similar if we instead use full-time male workers in our restricted sample to compute industry-level employment size. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

C.4.2 Robustness checks on Destination-specific Effects

Table C.4: Firm-year-level Log Hourly Wage Increase (20 Yrs of Experience)

	(1)	(2)	(3)	(4)
Exporter	0.204***	0.112*	0.004	-0.003
	(0.026)	(0.059)	(0.066)	(0.066)
Exporter × Share of exports	0.087**	0.082*	0.186*	0.184*
to industrialized destinations	(0.044)	(0.045)	(0.099)	(0.099)
Exporter ×		-0.009		0.006
Log (export value per worker)		(0.008)		(0.019)
Log(firm employment)		0.094***		0.063
		(0.010)		(0.014)
Share of high-school grads		0.285***		0.307*
		(0.061)		(0.042)
Share of cognitive occupations		0.285***		-0.065
		(0.072)		(0.276)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	Yes
Obs	77,071	77,071	77,071	77,071
R-squared	0.011	0.014	0.487	0.487

Note: This table presents estimates from regressions of firm-year-level log hourly wage increase after 20 years of experience on firm characteristics (Firms 1997–2000). The baseline group is non-exporters. The shares of high-school graduates and cognitive occupations in the workforce are computed based on our restricted sample, from which we obtained our estimates of firm-year-level experience-wage profiles. Cognitive occupations refer to professionals, technicians, and other white-collar workers. Firm employment size is computed based on all workers within the firm in the raw sample (including female and part-time workers) to reflect actual firm size. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

Table C.5: Firm-year-level Log Hourly Wage Increase (20 Years of Experience)

	(1)	(2)	(3)	(4)
Exporter	0.210***	0.131**	0.032	-0.007
	(0.023)	(0.060)	(0.061)	(0.116)
Exporter × Log(GDP per capita) in destination	0.073***	0.065**	0.111*	0.107*
	(0.027)	(0.028)	(0.059)	(0.059)
Exporter × Log (export value per worker)		-0.010		0.006
		(0.008)		(0.019)
Log(firm employment)		0.094***		0.063
		(0.010)		(0.053)
Share of high-school grads		0.284***		0.306*
		(0.061)		(0.186)
Share of cognitive occupations		0.285***		-0.065
		(0.072)		(0.276)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	Yes
Obs	77,071	77,071	77,071	77,071
R-squared	0.011	0.014	0.487	0.487

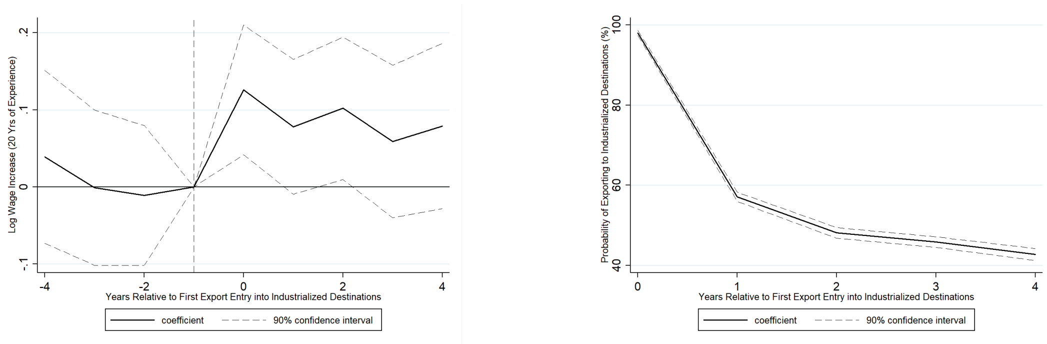
Note: This table presents estimates from regressions of firm-year-level log hourly wage increase after 20 years of experience on firm characteristics (Firms 1997–2000). The baseline group is non-exporters. We draw log real GDP per capita (2011 U.S.$) for each country in 2000 from Penn World Table 9.0 (Feenstra, Inklaar and Timmer, 2015) and compute a firm-year-level export-weighted log GDP per capita across destinations, normalized by log GDP per capita in Brazil. The shares of high-school graduates and cognitive occupations in the workforce are computed based on our restricted sample, from which we obtained our estimates of firm-year-level experience-wage profiles. Cognitive occupations refer to professionals, technicians, and other white-collar workers. Firm employment size is computed based on all workers within the firm in the raw sample (including female and part-time workers) to reflect actual firm size. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

Table C.6: Robustness of Column (4) Table 3.3 (Baseline: Exporter, Non-Ind Dests)

	Non-exporters		Exporter, industrial dests		Exporter, both dests	
baseline results	0.020 (0.023)		0.091** (0.039)		0.066*** (0.026)	
(1) add unit price of exports (1997–2000)	0.023 (0.066)		0.214** (0.104)		0.203*** (0.076)	
(2) add industry-year fixed effects	0.013 (0.023)		0.088** (0.039)		0.063** (0.026)	
(3) add gravity controls	-0.128 (0.177)		0.094** (0.047)		0.068*** (0.030)	
(4) add labor ability	0.024 (0.023)		0.091** (0.039)		0.068*** (0.026)	
(5) only switching periods	-0.018 (0.035)		0.151** (0.062)		0.070 (0.069)	
(6) add average tenure	0.016 (0.023)		0.089** (0.039)		0.068*** (0.026)	
(7) add differences in tenure between young and old	0.020 (0.023)		0.087** (0.039)		0.064** (0.026)	
By industry characteristics:	more	less	more	less	more	less
(8) more/less manual	0.049* (0.028)	-0.046 (0.043)	0.137*** (0.047)	-0.017 (0.070)	0.084*** (0.031)	0.016 (0.047)
(9) more/less skill-intensive	0.016 (0.039)	0.025 (0.029)	0.053 (0.070)	0.107** (0.047)	0.004 (0.043)	0.096*** (0.033)
(10) more/less differentiated	0.059* (0.035)	-0.017 (0.031)	0.157** (0.065)	0.050 (0.049)	0.088** (0.040)	0.048 (0.035)

Note: This table presents robustness checks of Column (4) of Table 3.3. All regressions control for the shares of high-school graduates and cognitive occupations in the workforce and firm size, as well as year and industry fixed effects. We use exporters to non-industrialized destinations as the baseline group, because they have the lowest returns to experience in the baseline results. Cognitive occupations refer to professionals, technicians, and other white-collar workers. Firm employment size is computed based on all workers within the firm in the raw sample (including female and part-time workers) to reflect actual firm size. For each firm, unit prices of exports are observed for each 8-digit HS and destination, and we take an average across destinations and HS products to obtain firm-year-level unit price of exports. Gravity controls refer to the average of log distance and bilateral cultural characteristics (with Brazil) across all of a firm's destinations, which are drawn from GeoDist database in CEPII (Mayer and Zignago, 2011). Old workers refer to workers in experience bins of 31–40 years, whereas young workers refer to workers in experience bins of 1–20 years. An industry is defined as more differentiated if its share of differentiated goods (based on 4-digit SITC goods exported by this industry) lies above the median of the industry-level share of differentiated goods across all manufacturing industries, according to the classification provided by Rauch (1999). An industry is defined as more skill-intensive (manual) if its share of high-school (manual) workers averaged across firms lies above the median of industry-level averages across all manufacturing industries. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

C.4.3 Dynamics of Experience-Wage Profiles



(a) Changes in Experience-wage Profiles (b) Survival Rate in Industrial Markets

Figure C.7: Dynamics of Firms' First Entry Into Industrialized Destinations

Note: The figure shows the β_τ parameters from estimating Equation (3.4). The dependent variable is firm-year-level returns to 20 years of experience in Panel (a) and a dummy variable that equals 1 if the firm exports to industrialized destinations in Panel (b). All regressions control for firm fixed effects, industry fixed effects, year fixed effects, the shares of high-school graduates and cognitive workers, firm size, and a dummy variable indicating whether the firm is exporting to a non-industrialized destination. To estimate the β_τ parameters, we do not enforce a requirement that firms remain exporting to industrialized destinations.

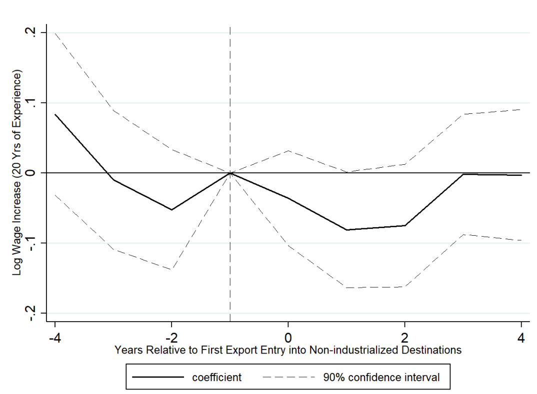


Figure C.8: Dynamics of Firms' First Entry Into Non-industrialized Destinations (Survivors)

Note: The figure shows the β_τ parameters from estimating Equation (3.4), except for that the β_τ parameters are coefficients on indicators for time periods relative to the firm's first export entry into non-industrialized destinations. The dependent variable is firm-year-level returns to 20 years of experience. The regression controls for firm fixed effects, industry fixed effects, year fixed effects, the shares of high-school graduates and cognitive workers in the workforce, firm size, and a dummy variable indicating whether the firm is exporting to an industrialized destination. To estimate the β_τ parameters after entry, we require that firms remain exporting to non-industrialized destinations.

C.4.4 Case Study: 1999 Devaluation Episode

Table C.7: Dependent Variable: Share of Cognitive Workers (Percentage Points)

time	(1) 1998-2000	(2) 1997-2001	(3) 1996-2002
1{export to industrialized dests}	0.481	1.492***	1.651***
×1{post_1999}	(0.630)	(0.549)	(0.479)
1{export to non-industrialized dests}	1.793***	1.604***	1.848***
× 1{post_1999}	(0.470)	(0.340)	(0.296)
1{export to both types of dests}	0.751	1.649***	2.429***
×1{post_1999}	(1.035)	(0.613)	(0.515)
Year, industry and firm FE	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes
Obs	37,267	61,390	85,266
R-squared	0.940	0.905	0.873
Average	21.12	21.44	21.74

Note: This table presents estimates from Equation (3.5). The dependent variable is the share of high-school graduates in the workforce, in terms of percentage points (%). The regression includes firm, industry, and year fixed effects. The last row shows the average share of high-school grads in the workforce (%) during the period. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

Table C.8: Share of High-school Grads, (Percentage Points)

time	(1) 1998-2000	(2) 1997-2001	(3) 1996-2002
1{export to industrialized dests}	0.181	0.254	0.410
×1{post_1999}	(0.387)	(0.388)	(0.320)
1{export to non-industrialized dests}	0.364	0.106	0.154
× 1{post_1999}	(0.301)	(0.215)	(0.196)
1{export to both types of dests}	-0.245	0.226	0.221
×1{post_1999}	(0.804)	(0.481)	(0.366)
Year, industry and firm FE	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes
Obs	37,267	61,390	85,266
R-squared	0.966	0.943	0.925
Average	17.24	17.39	17.63

Note: This table presents estimates from Equation (3.5). The dependent variable is the share of cognitive occupations in the workforce, in terms of percentage points (%). The regression includes firm, industry, and year fixed effects. Cognitive occupations refer to professionals, technicians, and other white-collar workers. The last row shows the average share of cognitive workers in the workforce (%) during the period. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

C.5 Detailed Discussions on Mechanisms

In this section, we discuss four plausible explanations for our finding on the interaction between returns to experience and different destinations: (1) selection of firms into different export destinations; (2) differential changes in labor composition; (3) job search and screening; and (4) human capital accumulation. We present detailed evidence and discussions for each hypothesis.

C.5.1 Selection of Firms into Different Export Destinations

Our first hypothesis is that firms exporting to industrialized destinations are better than other exporters due to factors not captured by firm fixed effects, or they enjoy more favorable linkages with destinations, which leads to higher returns to experience. We argue that this is unlikely to explain our findings.

First, as Table 3.4 shows, firms exporting to both types of destinations appear to be the most productive even after controlling for firm fixed effects, as they are the biggest and have the largest shares of high-school workers. Nevertheless, it is the firms exporting to industrialized destinations that enjoy the largest increase in experience-wage profiles after switching to exporting. This suggests that the destination-specific results we find may not be simply explained by better firms' selection into exporting to industrialized destinations.

Second, we exploit available data on export value and quantity in 1997–2000 to construct firm-year-level unit prices of exports as a proxy for product quality (Manova and Zhang, 2012).[2] We replicate the regression in Column (4) of Table 3.3 for the years 1997–2000 and control for unit prices of exports. We still find that exporting to industrialized destinations increases returns

[2]The firm-level export value and quantity are available by destinations and 8-digit HS products in 1997–2000. We take an export-weighted average of unit prices across destinations and HS products to construct firm-year-level unit prices of exports. Given the heterogeneity in values of HS products, we experimented with first normalizing the unit price by the unit price of the same HS product exported from Brazil to the U.S.. The results remain very similar under this normalization.

to experience, as shown in Row (1) of Appendix Table C.6. In Rows (2)–(3) of Appendix Table C.6, we also show that our results in Column (4) of Table 3.3 are robust to controlling for industry-year fixed effects or gravity variables (bilateral distance and sharing cultural characteristics). Therefore, industry-year-level common shocks or bilateral linkages of destinations with Brazil cannot capture destination-specific returns to experience.

C.5.2 Differential Changes in Labor Composition

The second plausible hypothesis is changes in labor composition after exporting. As shown in Table 3.4, firms exporting to both types of destinations have the largest shares of high-school graduates and cognitive workers in the workforce among all firms. Their workforce tends to become even more educated after switching to exporting, as shown by the coefficients in Table 3.4 after controlling for firm fixed effects. Therefore, it seems that changes in labor composition favor firms exporting to both types of destinations, but nonetheless firms exporting to industrialized destinations perceive the largest increase in returns to experience.

Although we control for education and occupations of the workforce in our regressions, it is possible that there are unobserved workers' characteristics, leading to higher returns to experience in firms exporting to industrialized destinations. We undertake two sets of robustness checks regarding this possibility. First, for each worker, we construct a proxy for their unobserved ability by using the residual of their log wage when she makes first appearance in the sample, after removing year and age effects. We can then obtain a measure of average ability of the workforce for each firm-year observation. We rerun the regression of Column (4) in Table 3.3 and control for this ability measure. Our destination-specific effects remain unchanged, as shown in Row (4) of Table C.6.[3]

In addition, when we compute firm-year-level experience-wage profiles in year t, we use workers employed within the same firm in both years $t-1$ and t. If current workers are unaware

[3] As expected, we find that firm-year-level experience-wage profiles increase with this ability measure.

of whether firms would change export status in one year, we could compare experience-wage profiles for firm-year-level observations in years $t-1$ and t with a switch in export status between years $t-1$ and t. We rerun the regression of Column (4) in Table 3.3 with these observations around switches and still find similar results, as shown in Row (5) in Table C.6.

C.5.3 Job Search and Screening

Our third hypothesis is that the observed destination-specific effects are due to job search and screening. Although we focus on workers staying in the same firm in the empirical analysis, workers' wage growth may still result from wage renegotiations due to job search. For example, in a calibrated model with wage bargaining like Cahuc, Postel-Vinay and Robin (2006), Fajgelbaum (2019) shows that workers in potential exporters experience faster wage growth due to wage renegotiations and larger job surplus after exporting. Our destination-specific results may thus arise due to larger surplus from exporting to industrialized destinations. Acemoglu and Pischke (1998) argue that firms monopsony power on workers ability information affects firms wage determination. Through the lens of their model, our results may arise if firms exporting to industrialized destinations have the least monopsony power and therefore design the steepest experience-wage profiles to avoid poaching from other firms.

Alternatively, workers' wage growth may originate from screening in the presence of information frictions (Jovanovic, 1979). In particular, larger job surplus after exporting could interact with screening based on workers' abilities (Helpman et al., 2017) or match-specific quality (Donovan, Lu and Schoellman, 2020), leading to different patterns of worker turnover and our observed experience-wage profiles. Moreover, given initial uncertainty about workers' abilities, exporters may offer back-loaded wage contracts that lead to steeper wage profiles.

We cannot entirely rule out all these plausible stories, but nonetheless we provide several robustness checks to show that job search and screening are unlikely to explain destination-specific effects. First, as we discussed in Section 3.2.5, we do not find that export value per worker affects returns to experience (Appendix Tables C.4–C.5). Therefore, changes in job surplus after exporting are unlikely to trigger destination-specific shifts in returns to experience.[4] Second, we divide the sample into sub-samples based on the industry-level shares of manual workers and high-school graduates in the workforce. We perform our regression in Column (4) of Table 3.3 on the sub-samples. The results show that exporting to industrialized destinations leads to higher returns to experience in more manual or less skill-intensive industries (Rows (8)–(9) of Table C.6), where workers may have lower bargaining power. Third, as firms' monopsony power can be measured by the length of workers' tenure, we add the average tenure (current firm) as well as differences in the average tenure (current firm) between old and young workers into our regression in Column (4) of Table 3.3. Our results remain quantitatively similar, as shown in Rows (6)–(7) of Table C.6. Finally, as shown in Section 3.2.6 and 3.3, the jump in experience-wage profiles happens immediately after entry into industrialized destinations. If exporters offer back-loaded wage contracts, we should expect an initial decline in experience-wage profiles after switching to exporting.

C.5.4 Human Capital Accumulation and Knowledge Diffusion

There is a long tradition, starting with Becker (1964), using experience-wage profiles to implicitly measure human capital accumulation (e.g., Caselli, 2005; Manuelli and Seshadri, 2014). Clearly, one potential way to interpret our destination-specific results is through human capital theory. In addition, the literature argues that knowledge diffusion is central to human capital accumulation (e.g., Lucas and Moll, 2014), and trade transmits knowledge across borders (e.g., Buera and Oberfield, 2020).

Our findings on destination-specific returns to experience are consistent with faster human

[4]Even if we control for export value per worker, job surplus could still be higher if firms exporting to industrialized destinations enjoy higher markups than other firms. There is not much evidence on it. If any, Keller and Yeaple (2020) find that the markups of U.S. multinationals' affiliates decline with the GDP per capita of the host country. De Loecker and Eeckhout (2020) estimate the aggregate markup across countries, and there is no clear relationship between markups and countries' development levels.

capital accumulation due to exposure to advanced countries. First, we find that firms enjoy steeper experience-wage profiles if they export to industrialized destinations. This is consistent with advanced knowledge from trading with industrialized destinations (Alvarez, Buera and Lucas, 2013; Buera and Oberfield, 2020). Moreover, in Rows (8)–(9) of Table C.6, we find that increases in experience-wage profiles due to industrialized destinations are larger in industries with smaller shares of high-skill and cognitive workers. This evidence is compatible with the knowledge diffusion literature, which typically predicts that the least productive agents experience the largest gains in human capital from knowledge diffusion (Lucas and Moll, 2014). Third, in Row (10) of Table C.6, we also find that increases in experience-wage profiles due to industrialized destinations are larger in industries that produce more differentiated goods, which might be associated with larger benefits for knowledge adoption.

Therefore, we propose that human capital accumulation due to knowledge diffusion is likely to explain destination-specific returns to experience, although we cannot entirely rule out other hypotheses. In the main text, we show that this hypothesis is also backed up by anecdotal evidence on exporters' experience and direct evidence on human capital investments, foreign technology adoption, and exporting.

C.6 Exporting, Training and Technology adoption

The ES is conducted by private contractors on behalf of the World Bank, and confidentiality is never compromised according to the ES unit. The ES is usually answered by owners and top managers with the assistance of accountants or human resources managers. Typically, the ES conducts 1200–1800 interviews in large economies, 360 in medium-sized economies, and 150 in small ones for manufacturing and service sectors. The ES interviews formal firms with more than 5 employees, although in some cases, it may include informal firms and/or firms with fewer than 5 employees. Firms with 100% government/state ownership are not eligible to participate in this survey. According to the WB-ES unit, there is a stratified random sampling. The strata for the ES are firm size, business sector and geographic region within a country, and random samples are selected within each strata. It over-samples large firms, but we rely on firm-level specifications and control for firm size in every regression, and regressions are also weighted by sample weights.

The ES has two types of questionnaires, one for manufacturing firms and one for service firms, which have identical questions for some topics and also specific questions. We rely mostly on the second standardized wave covering countries between 2006–2017. Although the first wave has similar questions, the available data do not provide weights which are needed to obtain more reliable estimates. Nevertheless, for the case of Brazil, we perform the regression about R&D investments using data in 2003, because 2003 is the only year for which there are available data on this variable for this country. In Tables C.9 and C.10, we provide the list of countries and years with available data in the first and second standardized waves.

To define exporter and indirect exporters, we use question D.3: ”In the last fiscal year, what percent of this establishments sales were: a. National sales, b. Indirect exports [sold domestically to third party that exports products], c. Direct exports?”. We rely on the following three questions to construct our dummy variables of on-the-job training, R&D investments and foreign Technology adoption. For training, the question L.10 is: “Over fiscal year [insert last complete fiscal year], did this establishment have formal training programs for its permanent, full-time employees?” For R&D investments, the question H.8 is ”In the last fiscal year, did the establishment invest in R&D?”. For foreign technology adoption, the question E.6 is ”Does this establishment at present use technology licensed from a foreign-owned company?”. Table C.11 presents the summary statistics of the variables in our sample.

Table C.9: Countries in the Enterprise Survey

Country	Year(s)	Country	Year(s)
Afghanistan	2008, 2014	Dem.Rep.Congo	2006, 2010, 2013
Albania	2007, 2013	Ecuador	2006, 2010, 2017
Algeria	2003	Egypt	2013, 2016
Angola	2006, 2010	El Salvador	2006, 2010, 2016
A.and Barbuda	2010	Eritrea	2009
Argentina	2006, 2010, 2017	Estonia	2005, 2009, 2013
Armenia	2005, 2009, 2013	Eswatini	2006, 2016
Azerbaijan	2005, 2009, 2013	Ethiopia	2011, 2015
Bahamas	2010	Fiji	2009
Bangladesh	2007, 2013	FYR Macedonia	2005, 2009, 2013
Barbados	2010	Gabon	2009
Belarus	2005, 2008, 2013	Gambia	2006, 2018
Belize	2010	Georgia	2005, 2008, 2013
Benin	2009, 2016	Germany	2005
Bhutan	2009, 2015	Ghana	2007, 2013
Bolivia	2006, 2010, 2017	Greece	2005
Bos. and Her.	2005, 2009, 2013	Grenada	2010
Botswana	2006, 2010	Guatemala	2006, 2010, 2017
Brazil	2003, 2009	Guinea	2006, 2016
Bulgaria	2005, 2007, 2009, 2013	Guinea-Bissau	2006
Burkina Faso	2006, 2009	Guyana	2010
Burundi	2006, 2014	Honduras	2006, 2010, 2016
Cambodia	2013, 2016	Hungary	2005, 2009, 2013
Cameroon	2006, 2009, 2016	India	2006, 2014
Cape Verde	2006, 2009	Indonesia	2009, 2015
Cen. Af. Rep.	2011	Iraq	2011
Chad	2009, 2018	Ireland	2005
Chile	2006, 2010	Israel	2013
China	2012	Ivory Coast	2009, 2106
Colombia	2006, 2010, 2017	Jamaica	2005, 2010
Congo	2009	Jordan	2006, 2013
Costa Rica	2005, 2010	Kazakhstan	2005, 2009, 2013
Croatia	2005, 2007, 2013	Kenya	2003, 2007, 2013
Czech Republic	2005, 2009, 2013	Kosovo	2009, 2013
Djibuti	2013	Kyrgystan	2005, 2009, 2013
Dominica	2010	Laos	2006, 2009, 2009, 2012
Dom. Republic	2005, 2010, 2016	Latvia	2005, 2009, 2013

Table C.10: Countries in the Enterprise Survey

Country	Year(s)	Country	Year(s)
Lebanon	2006, 2013	Serbia	2009, 2013
Lesotho	2009, 2016	Ser. and Mon.	2005
Liberia	2009, 2017	Sierra Leone	2009, 2017
Lithuania	2005, 2009, 2013	Slovakia	2005, 2009, 2013
Madagascar	2005, 2009, 2013	Slovenia	2005, 2009, 2013
Malawi	2005, 2009, 2014	Solomon Islands	2015
Malaysia	2015	South Africa	2007
Mali	2007, 2010, 2016	South Korea	2005
Mauritania	2006, 2014	South Sudan	2014
Mauritius	2005, 2009	Spain	2005
Mexico	2006, 2010	Sri Lanka	2004, 2011
Micronesia	2009	St. K. and Nevis	2010
Moldova	2005,'09,'13	Sudan	2014
Mongolia	2004, 2009, 2013	Suriname	2010
Montenegro	2009, 2013	Swaziland	2006
Morocco	2004, 2013	Sweden	2014
Mozambique	2007	Syria	2003
Myanmar	2014, 2016	Tajikistan	2005, 2008, 2013
Namibia	2006, 2014	Tanzania	2006, 2013
Nepal	2009, 2013	Thailand	2004, 2016
Nicaragua	2006, 2010, 2016	Timor-Leste	2009, 2015
Niger	2009, 2017	Togo	2009, 2016
Nigeria	2007, 2014	Tonga	2009
Oman	2003	Tri. and Tob.	2010
Pakistan	2007, 2013	Tunisia	2013
Panama	2006, 2010	Turkey	2005, 2008, 2013
P. New Guinea	2015	Uganda	2006, 2013
Paraguay	2006, 2010, 2017	Ukraine	2005, 2008, 2013
Peru	2006, 2010, 2017	Uruguay	2006, 2010, 2017
Philippines	2003, 2009, 2015	Uzbekistan	2005, 2008, 2013
Poland	2005, 2009, 2013	Vanuatu	2009
Portugal	2005	Venezuela	2006, 2010
Romania	2005, 2009, 2013	Vietnam	2005,
Russia	2005, 2009, 2012	W.B. and Gaza	2006, 2013
Rwanda	2006, 2011	Yemen	2010, 2013
Samoa	2009	Zambia	2007, 2013
Senegal	2007, 2014	Zimbabwe	2011, 2016

Table C.11: Sample Statistics

	Non-exporter		Exporter	
	Mean	S.D.	Mean	S.D.
Main variables				
On-the-job training	0.34	0.47	0.56	0.49
Foreign technology	0.11	0.32	0.25	0.43
R&D investments	0.19	0.39	0.41	0.49
Controls				
Log(employment)	3.08	1.22	4.33	1.43
Labor share (%)	21.89	19.77	19.59	18.47
Managerial years of experience in sector	16.43	10.22	19.25	10.53
Share of high-school grads (%)	61.47	35.90	66.91	33.22

Note: This table presents the summary statistics from the second standardized wave of the World Bank Enterprise Surveys, covering the period 2006–2017. This table shows the mean and the standard deviation of variables we use in the paper, computed across all firms from all countries and years. On-the-job Training, foreign technology, and R&D investments are dummy variables that equal 1 if firms perform the corresponding activity and 0 otherwise. We restrict the sample to firms with labor shares lower than 200% to avoid extreme values.

Table C.12: Exporting, On-the-Job Training and R&D Investments

	(1)	(2)	(3)	(4)	(5)
Non-Exporter # R&D Investment	0.29***	0.24***	0.25***	0.25***	0.22***
	(0.012)	(0.013)	(0.014)	(0.014)	(0.016)
Exporter # No R&D Investment	0.12***	0.06***	0.06***	0.06***	0.07***
	(0.015)	(0.014)	(0.015)	(0.015)	(0.025)
Exporter # R&D Investment	0.36***	0.25***	0.24***	0.24***	0.25***
	(0.023)	(0.022)	(0.024)	(0.024)	(0.029)
Obs	81,094	79,906	63,258	60,951	38,588
R-squared	0.191	0.231	0.248	0.249	0.269
Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes	Yes
Log(Emp)	No	Yes	Yes	Yes	Yes
Labor share	No	No	Yes	Yes	Yes
Managerial experience in sector	No	No	No	Yes	Yes
% High school grads	No	No	No	No	Yes

Note: This table presents estimates from regressing a dummy variable that takes the value 1 if the firm offers formal on-the-job training, on export status interacted with a dummy variable reflecting if the firm invests in R&D. The baseline group is non-exporters with no R&D investments. We control for country, year, and industry fixed effects in all regressions. Firm-level control variables are log (employment), the ratio of labor costs to total sales, the share of high-school graduates in the workforce, and managers' years of experience in the operating sector. We use the second standardized wave of the ES with the provided weights. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

Table C.13: On-the-job Training, Technology Adoption, R&D Investments and Exporting

VARIABLES	(1)	(2)	(3)	(4)	(5)
0b.exporter#0b.use_tech_foreign#1.rd_investment	0.29***	0.25***	0.26***	0.26***	0.21***
	(0.017)	(0.019)	(0.024)	(0.023)	(0.028)
0b.exporter#1.use_tech_foreign#0b.rd_investment	0.17***	0.13***	0.12***	0.12***	0.06
	(0.022)	(0.021)	(0.026)	(0.027)	(0.052)
0b.exporter#1.use_tech_foreign#1.rd_investment	0.42***	0.35***	0.33***	0.33***	0.32***
	(0.028)	(0.031)	(0.034)	(0.034)	(0.035)
1.exporter#0b.use_tech_foreign#0b.rd_investment	0.12***	0.06***	0.04***	0.05***	0.04*
	(0.017)	(0.015)	(0.017)	(0.017)	(0.023)
1.exporter#0b.use_tech_foreign#1.rd_investment	0.36***	0.25***	0.24***	0.23***	0.26***
	(0.030)	(0.027)	(0.031)	(0.031)	(0.029)
1.exporter#1.use_tech_foreign#0b.rd_investment	0.27***	0.17***	0.17***	0.18***	0.13***
	(0.031)	(0.029)	(0.030)	(0.031)	(0.037)
1.exporter#1.use_tech_foreign#1.rd_investment	0.46***	0.33***	0.32***	0.34***	0.26***
	(0.026)	(0.025)	(0.028)	(0.028)	(0.035)
Obs	61,220	60,586	48,249	46,840	25,549
R-squared	0.205	0.236	0.254	0.257	0.309
Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes	Yes
Log(Emp)	No	Yes	Yes	Yes	Yes
Labor Share	No	No	Yes	Yes	Yes
Managerial experience in Sector	No	No	No	Yes	Yes
% High School Grads	No	No	No	No	Yes

Note: This table presents estimates from regressing a dummy variable that takes the value 1 if the firm offers training on interaction terms between dummies of export status, foreign technology adoption, and R&D investments. Exporters are defined as firms with positive sales to foreign markets. The baseline group is non-exporters with no R&D investments and foreign technology. We control for country, year, and industry fixed effects in all regressions. Firm-level control variables are log (employment), the ratio of labor costs to total sales, the share of high-school graduates in the workforce, and managers' years of experience in the operating sector. We use the second standardized wave of the ES with the provided weights. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

Table C.14: On-the-job Training , Export Status, and Technology (Brazil)

	(1)	(2)	(3)	(4)
	Panel A: Training, Exporting, and R&D			
0b.exporter#1.rd_investment	0.17***	0.13***	0.12***	0.11***
	(0.026)	(0.026)	(0.026)	(0.026)
1.exporter#0b.rd_investment	0.26***	0.13***	0.12***	0.10**
	(0.042)	(0.043)	(0.044)	(0.044)
1.exporter#1.rd_investment	0.32***	0.15***	0.15***	0.13***
	(0.033)	(0.036)	(0.037)	(0.037)
Obs	1,574	1,560	1,518	1,517
R-squared	0.099	0.156	0.165	0.176
	Panel B: Training, Exporting, and Technology			
0b.exporter#1.use_tech_foreign	0.24***	0.19**	0.13	0.13
	(0.074)	(0.075)	(0.087)	(0.088)
1.exporter#0b.use_tech_foreign	0.37***	0.15**	0.16**	0.15**
	(0.060)	(0.073)	(0.077)	(0.078)
1.exporter#1.use_tech_foreign	0.60***	0.33***	0.36***	0.35***
	(0.053)	(0.064)	(0.067)	(0.068)
Obs	1,304	1,282	1,087	1,056
R-squared	0.244	0.289	0.313	0.311
Log(Emp)	No	Yes	Yes	Yes
Labor share	No	No	Yes	Yes
Managerial experience †	No	No	No	Yes

Note: This table presents estimates from regressing a dummy variable that takes the value 1 if the firm offers formal on-the-job training and 0 otherwise, on export status interacted with a dummy variable reflecting if the firm invests in R&D (Panel A) or adopts foreign technology (Panel B). Exporters are defined as firms with positive sales to foreign markets. The baseline groups are non-exporters with no R&D investments (Panel A) or no foreign technology adoption (Panel B). We control for industry fixed effects in all regressions. Firm-level control variables are log (employment), the ratio of labor costs to total sales, and managers' years of experience in the operating sector. We use data in 2009 for Panel B to be able to use the provided weights and data in 2003 for Panel A due to the lack of data on R&D investments for Brazil in 2009. †: For the regressions about R&D investments, we use the highest education level of top manager due to the lack of data on managerial experience. Robust standard errors are in parentheses. Significance levels: * 10%, ** 5%, *** 1%.

Bibliography

Abowd, John, Francis Kramarz, and David Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67(2).

Acemoglu, Daron. 1997. "Training and Innovation in an Imperfect Labour Market." *Review of Economic Studies*, 64(3).

Acemoglu, Daron, and Fabrizio Zilibotti. 2001. "Productivity Differences." *Quarterly Journal of Economics*, 116(2).

Acemoglu, Daron, and Jorn-Steffen Pischke. 1998. "Why Do Firms Train? Theory and Evidence." *The Quarterly Journal of Economics*, 113(1).

Acemoglu, Daron, and Jrn-Steffen Pischke. 1999. "Beyond Becker: Training in Imperfect Labour Markets." *The Economic Journal*, 109.

Acemoglu, Daron, Simon Johnson, and James A Robinson. 2005. "Institutions as a fundamental cause of long-run growth." *Handbook of economic growth*, 1: 385–472.

Agosin, Manuel R, and Claudio Bravo-Ortega. 2009. "The Emergence of New Successful Export Activities in Latin America: The Case of Chile." *IDB Working Paper*.

Aguayo-Tellez, Ernesto, Marc-Andreas Muendler, and Jennifer P. Poole. 2010. "Globalization and Formal-sector Migration in Brazil." *World Development*, 38(6).

Aguiar, Mark, and Erik Hurst. 2013. "Deconstructing Life Cycle Expenditure." *Journal of Political Economy*, 121(3).

Albornoz, Facundo, Hctor F.Calvo Pardo, Gregory Corcos, and Emanuel Ornelas. 2012. "Sequential Exporting." *Journal of International Economics*, 88(1).

Alon, Titan, and Daniel Fershtman. 2019. "Gradual Specialization and the Allocation of Talent."

Alon, Titan 2017. "Earning more by doing less: Human capital specialization and the college wage premium." *Unpublished Manuscript, Northwestern University.*

Alvarez, Fernando, Francisco Buera, and Robert Lucas. 2013. "Idea Flows, Economic Growth, and Trade." *NBER Working Paper*, 19667.

Artopoulos, Alejandro, Daniel Friel, and Juan Carlos Hallak. 2010. “Challenges of Exporting Differentiated Products to Developed Countries: The Case of SME-dominated Sectors in a Semi-industrialized Country.” *IDB Working Paper*.

Atkin, David, Amit K Khandelwal, and Adam Osman. 2017. “Exporting and Firm Performance: Evidence from a Randomized Experiment.” *The Quarterly Journal of Economics*, 132(2).

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen. 2020. “The fall of the labor share and the rise of superstar firms.” *The Quarterly Journal of Economics*, 135(2): 645–709.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu. 2011. “R&D Investment, Exporting, and Productivity Dynamics.” *American Economic Review*, 101(4).

Aw, Bee Yan, Sukkyun Chung, and Mark Roberts. 2000. “Productivity and Turnover in the Export Market: Micro-level Evidence from the Republic of Korea and Taiwan (China).” *World Bank Economic Review*, 14.

Axtell, Robert L. 2001. “Zipf Distribution of U.S. Firm Sizes.” *Science*, 293.

Bagger, Jesper, François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. 2014. “Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics.” *American Economic Review*, 104(6): 1551–96.

Barlevy, Gadi. 2008. “Identification of search models using record statistics.” *The Review of Economic Studies*, 75(1): 29–64.

Bassanini, Andrea, Alison Booth, Giorgio Brunello, Maria De Paola, and Edwin Leuven. 2005. “Workplace Training in Europe.” *IZA Working Paper*.

Becker, Gary. 1964. *Human Capital.* The University of Chicago Press.

Bento, Pedro, and Diego Restuccia. 2017. “Misallocation , Establishment Size , and Productivity Misallocation, Establishment Size , and Productivity.” *American Economic Journal: Macroeconomics*, 9(3): 267–303.

Bernard, A., J. Eaton, J. B. Jensen, and S. Kortum. 2003. “Plants and Productivity in International Trade.” *American Economic Review*, 93(4).

Bernard, Andrew, and Bradford Jensen. 1995. “Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987 .” *Brookings Papers: Microeconomics*.

Bernard, Andrew B, and J Bradford Jensen. 1999. “Exceptional Exporter Performance: Cause, Effect, or Both?” *Journal of international economics*, 47(1).

Bick, Alexander, Nicola Fuchs-Schuendeln, and David Lagakos. 2018. "How do Hours Worked Vary With Income? Cross-Country Evidence and Implications." *American Economic Review*, 108(1).

Blatter, Marc, Samuel Muehlemann, Samuel Schenker, and Stefan C. Wolter. 2016. "Hiring Costs of Skilled Workers and the Supply of Firm-Provided Training." *Oxford Economic Papers*.

Bonelli, Regis, and Armando Castelar Pinheiro. 2008. "New Export Activities in Brazil: Comparative Advantage, Policy or Self-Discovery? ." *IADB Research Network Working Paper R-551*.

Botero, Juan C, Simeon Djankov, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. 2004. "The regulation of labor." *The Quarterly Journal of Economics*, 119(4): 1339–1382.

Bowlus, Audra J, and Huju Liu. 2013. "The contributions of search and human capital to earnings growth over the life cycle." *European Economic Review*, 64: 305–331.

Buera, Francisco J, and Ezra Oberfield. 2020. "The Global Diffusion of Ideas." *Econometrica*, 88(1).

Bunzel, Henning, Bent J Christensen, Nicholas M Kiefer, Lars Korsholm, et al. 1999. *Equilibrium search with human capital accumulation.* Vol. 99, Centre for Labour Market and Social Research.

Burdett, Kenneth, and Dale T. Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review*, 39(2).

Burdett, Kenneth, Carlos Carrillo-Tudela, and Melvyn G Coles. 2011. "Human capital accumulation and labor market equilibrium." *International Economic Review*, 52(3): 657–677.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean Marc Robin. 2006. "Wage Bargaining with On-the-job Search: Theory and Evidence." *Econometrica*, 74.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics*, 128(3).

Caselli, Francesco. 2005. "Accounting for cross-country income differences." *Handbook of economic growth*, 1: 679–741.

Coe, David T, and Elhanan Helpman. 1995. "International R&D spillovers." *European economic review*, 39(5).

Costinot, Arnaud, Dave Donaldson, and Ivana Komunjer. 2012. "What Goods do Countries Trade? A Quantitative Exploration of Ricardo's Ideas." *Review of Economic Studies*, 79(2).

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* Washington D.C:The World Bank. An optional note.

De la Croix, David, Matthias Doepke, and Joel Mokyr. 2018. "Clans, guilds, and markets: Apprenticeship institutions and growth in the preindustrial economy." *The Quarterly Journal of Economics*, 133(1): 1–70.

de la Roca, Jorge, and Diego Puga. 2017. "Learning by Working in Big Cities." *Review of Economic Studies*, 84(1): 106–142.

De Loecker, Jan. 2007. "Do Exports Generate Higher Productivity? Evidence from Slovenia." *Journal of International Economics*, 73.

De Loecker, Jan. 2013. "Detecting Learning by Exporting." *American Economic Journal: Microeconomics*, 5(3).

De Loecker, Jan, and Jan Eeckhout. 2020. "Global Market Power." *Working Paper*.

Dix-Carneiro, Rafael. 2014. "Trade Liberalization and Labor Market Dynamics." *Econometrica*, 82(3).

Dix-Carneiro, Rafael, Pinelopi Goldberg, Costas Meghir, and Gabriel Ulyssea. 2019. "Trade and Informality in the Presence of Labor Market Frictions and Regulations." *Working Paper*.

Doepke, Matthias, and Ruben Gaetani. 2020. "Why Didn't the College Premium Rise Everywhere? Employment Protection and On-the-Job Investment in Skills." National Bureau of Economic Research.

Donovan, Kevin, Jianyu Lu, and Todd Schoellman. 2020. "Labor Market Flows and Development." *Working Paper*, , (4).

Engbom, Niklas. 2020. "Labor Market Fluidity and Human Capital Accumulation."

Erosa, Andres, Tatyana Koreshkova, and Diego Restuccia. 2010. "How important is human capital? A quantitative theory assessment of world income inequality." *The Review of Economic Studies*, 77(4): 1421–1449.

Faberman, R. Jason, Andreas I Mueller, Ayegl ahin, and Giorgio Topa. 2017. "Job Search Behavior among the Employed and Non-Employed." *NBER Working Paper*.

Fajgelbaum, Pablo D. 2019. "Labor Market Frictions, Firm Growth, and International Trade." *NBER Working Paper No. 19492*.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review*, 105.

Feng, Ying, David Lagakos, and James Rauch. 2018. "Unemployment and Development." *NBER Working Paper*.

Flinn, Christopher, Ahu Gemici, and Steven Laufer. 2017. "Search, matching and training." *Review of Economic Dynamics*, 25: 260–297.

Frazis, Harley, and Mark A. Loewenstein. 2005. "Reexamining the Returns to Training: Functional Form, Magnitude, and Interpretation." *The Journal of Human Resources*, 40(2).

Frias, Judith A., David S. Kaplan, and Eric A. Verhoogen. 2009. "Exports and Wage Premia: Evidence from Mexican Employer-Employee Data." *Working Paper*.

Gollin, Douglas. 2002. "Getting income shares right." *Journal of political Economy*, 110(2): 458–474.

Gollin, Douglas. 2008. "Nobody's business but my own: Self-employment and small enterprise in economic development." *Journal of Monetary Economics*, 55(2): 219–233.

Gregory, Victoria. 2019. "Firms as learning environments: Implications for earnings dynamics and job search." *Unpublished Manuscript*.

Guner, Nezih, Alessandro Ruggieri, and James Tybout. 2019. "Training, Offshoring, and the Job Ladder." *Working Paper*.

Hall, Robert E, and Charles I Jones. 1999. "Why do some countries produce so much more output per worker than others?" *The quarterly journal of economics*, 114(1): 83–116.

Haltiwanger, John C, Henry R Hyatt, Lisa B Kahn, and Erika McEntarfer. 2018. "Cyclical job ladders by firm size and firm wage." *American Economic Journal: Macroeconomics*, 10(2): 52–85.

Hanushek, Eric A, and Ludger Woessmann. 2012. "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation." *Journal of economic growth*, 17(4): 267–321.

Head, Keith, and John Ries. 2001. "Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of US-Canada Trade." *American Economic Review*.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics*, 1(1): 1–58.

Helpman, Elhanan, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen J. Redding. 2017. "Trade and Inequality: From Theory to Estimation." *Review of Economic Studies*, 84(1).

Herkenhoff, Kyle, Jeremy Lise, Guido Menzio, and Gordon M Phillips. 2018. "Production and Learning in Teams." *Working Paper*.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, 60(5).

Hornstein, Andreas, Per Krusell, and Giovanni L. Violante. 2011. "Frictional Wage Dispersion in Search Models: A Quantitative Assessment." *American Economic Review*.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4).

Hsieh, Chang-Tai, and Peter J Klenow. 2010. "Development accounting." *American Economic Journal: Macroeconomics*, 2(1): 207–23.

Imai, Susumu, and Michael Keane. 2004. "Intertemporal Labor Supply and Human Capital Accumulation." *International Economic Review*, 45.

Islam, Asif, Remi Jedwab, Paul Romer, and Daniel Pereira. 2019. "Returns to Experience and the Misallocation of Labor."

Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg. 2018. "Learning from Coworkers." *Working Paper*.

Jones, Benjamin F. 2014. "The human capital stock: a generalized approach." *American Economic Review*, 104(11): 3752–77.

Jovanovic, Boyan. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy*, 87(5).

Keller, Wolfgang. 2004. "International Technology Diffusion." *Journal of Economic Literature*, 42(3).

Keller, Wolfgang, and Stephen R. Yeaple. 2020. "Multinationals, Markets, and Markups." *Working Paper*.

Klenow, Peter J, and Andres Rodriguez-Clare. 1997. "The neoclassical revival in growth economics: Has it gone too far?" *NBER macroeconomics annual*, 12: 73–103.

Krishna, Pravin, Jennifer P. Poole, and Mine Zeynep Senses. 2014. "Wage Effects of Trade Reform with Endogenous Worker Mobility." *Journal of International Economics*, 93(2).

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2018*a*. "Life-cycle human capital accumulation across countries: lessons from US Immigrants." *Journal of Human Capital*, 12(2): 305–342.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2018*b*. "Life cycle wage growth across countries." *Journal of Political Economy*, 126(2): 797–849.

Lentz, Rasmus, and Dale Mortensen. 2008. "An Empirical Model of Growth Through Product Innovation." *Econometrica*, 76.

Leuven, Edwin. 2004. "A review of the wage returns to private sector training." *Working Paper*.

Lileeva, Alla, and Daniel Trefler. 2010. "Improved Access to Foreign Markets Raises Plant-level ProductivityFor Some Plants." *The Quarterly Journal of Economics*, 125(3).

Lind, Nelson, and Natalia Ramondo. 2019. "The Economics of Innovation, Knowledge Diffusion, and Globalization." *Oxford Research Encyclopedia of Economics and Finance*.

Lucas, Robert E., and Benjamin Moll. 2014. "Knowledge Growth and the Allocation of Time." *Journal of Political Economy*, 122(1).

Manova, Kalina, and Zhiwei Zhang. 2012. "Export Prices across Firms and Destinations." *Quarterly Journal of Economics*, 127.

Manuelli, Rodolfo E, and Ananth Seshadri. 2014. "Human capital and the wealth of nations." *American economic review*, 104(9): 2736–62.

Mayer, Thierry, and Soledad Zignago. 2011. "Notes on CEPIIs distances measures: The GeoDist database ." *CEPII Working Paper 2011-25*.

Meghir, Costas, Renata Narita, and Jean-Marc Robin. 2015. "Wages and Informality in Developing Countries." *American Economic Review*, 105(4).

Menezes-Filho, Naercio Aquino, Marc-Andreas Muendler, and Garey Ramey. 2008. "The Structure of Worker Compensation in Brazil, with a Comparison to France and the United States." *The Review of Economics and Statistics*, 90(2): 324–346.

Mincer, Jacob A. 1974. "Schooling, Experience, and Earnings." *NBER Working Paper*.

Mion, Giordano, and Luca David Opromolla. 2014. "Managers' Mobility, Trade Performance, and Wages." *Journal of International Economics*, 94(1).

Moen, Espen, and Asa Rosén. 2004. "Does Poaching Distort Training?" *The Review of Economic Studies*, 71(4).

Morales, Eduardo, Gloria Sheu, and Andrés Zahler. 2019. "Extended Gravity." *The Review of Economic Studies*, 86(6).

Mortensen, Dale, and Christopher Pissarides. 1994. "Job creation and job destruction in the theory of unemployment." *Review of Economic Studies*, 61(3).

Muendler, Marc-Andreas, and James E. Rauch. 2018. "Do Employee Spinoffs Learn Markets from Their Parents? Evidence from International Trade." *European Economic Review*, 105(5).

Muendler, Marc-Andreas, James E. Rauch, and Oana Tocoian. 2012. "Employee Spinoffs and Other Entrants: Stylized Facts from Brazil." *International Journal of Industrial Organization*, 30(5).

Perla, Jesse, Christopher Tonetti, and Michael E Waugh. 2015. "Equilibrium Technology Diffusion, Trade, and Growth." *NBER Working Paper No. 20881*.

Pissarides, Christopher. 2000. *Equilibrium Unemployment Theory (2nd ed.).* MIT Press.

Porzio, Tommaso. 2017. "Cross-Country Differences in the Optimal Allocation of Talent and Technology." *Job Market Paper*.

Poschke, Markus. 2018. "The firm size distribution across countries and skill-biased change in entrepreneurial technology." *American Economic Journal: Macroeconomics*, 10(3): 1–41.

Rauch, James. 1999. "Networks Versus Markets in International Trade." *Journal of International Economics*, 48.

Restuccia, Diego, and Richard Rogerson. 2013. "Misallocation and productivity." *Review of Economic Dynamics*, 16(1): 1–10.

Rocha, Angela Da, Alexandre Darze, Beatriz Kury, and Joana Monteiro. 2008. "The Emergence of New and Successful Export Activities in Brazil: Four Case Studies from the Manufacturing and the Agricultural Sector." *IDB Working Paper*.

Rubinstein, Yona, and Yoram Weiss. 2006. "Post schooling wage growth: Investment, search and learning." *Handbook of the Economics of Education*, 1: 1–67.

Sampson, Thomas. 2016. "Dynamic Selection: An Idea Flows Theory of Entry, Trade and Growth." *Quarterly Journal of Economics*, 131(1).

Schank, Thorsten, Claus Schnabel, and Joachim Wagner. 2007. "Do Exporters Really Pay Higher Wages? First Evidence from German linked Employer-employee Data." *Journal of International Economics*, 72(1).

Schoellman, Todd. 2012. "Education quality and development accounting." *The Review of Economic Studies*, 79(1): 388–417.

Schoellman, Todd. 2016. "Early childhood human capital and development." *American Economic Journal: Macroeconomics*, 8(3): 145–74.

Shimer, R. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review*, 95(1).

Snoeck, Michele, Lucia Pittaluga, Hector Pastori, Rosario Domingo, and Carlos Casacuberta. 2009. "The Emergence of Successful Export Activities in Uruguay: Four Case Studies." *IDB Working Paper*.

Van Biesebroeck, Johannes. 2005. "Exporting Raises Productivity in Sub-Saharan African Manufacturing Plants." *Journal of International Economics*, 67.

Verhoogen, Eric. 2008. "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector." *Quarterly Journal of Economics*, 123(2).

Veum, Jonathan R. 1995. " Sources of Training and Their Impact on Wages." *Industrial and Labor Relations Review*, 48(4).

Wasmer, Etienne. 2006. "General versus specific skills in labor markets with search frictions and firing costs." *American Economic Review*, 96(3): 811–831.

Yamaguchi, Shintaro. 2010. "Job search, bargaining, and wage dynamics." *Journal of Labor Economics*, 28(3): 595–631.