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

Xiao Ma — *Peking University*

Alejandro Nakab — *Universidad Torcuato Di Tella*

Daniela Vidart — *University of Connecticut*

This Version: February 2023; **First Version:** October 2019

Abstract

Workers in richer countries experience faster rates of wage growth over their lifetimes than workers in poorer countries. We offer an explanation for this pattern by showing that workers in richer economies receive more firm-provided training. Using cross-country enterprise and worker-level data, we document that the share of workers who receive firm-provided training increases with development, and that firm-provided training is a key determinant of workers' human capital. We then build a general equilibrium search model with firm-provided training investments. Our model suggests that firm-provided training accounts for 38% of cross-country wage growth differences and 12% of cross-country income differences.

Keywords: On-the-Job Training, Human Capital Accumulation, Lifecycle Wage Growth
JEL Codes: E24, J24, O11, O15, J63, J64, M53

*Email: xiaoma@phbs.pku.edu.cn, anakab@utdt.edu, and daniela.vidart@uconn.edu.

We are grateful to Titan Alon, Ruixue Jia, Munseob Lee, David Lagakos, Marc Muendler, Tommaso Porzio, Valerie Ramey, Natalia Ramondo, Todd Schoellman, and seminar participants at ITAM, Peking University, Shanghai University of Finance and Economics, University of California San Diego, University of Connecticut, Universidad de San Andres, Universidad de Montevideo, Universidad Torcuato Di Tella, Banco de Mexico, University of Colorado Boulder and NEUDC for helpful comments. Finally, we would like to thank Kevin Donovan, Will Lu, and Todd Schoellman for sharing data. 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. All potential errors are our own.

1 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., 2018b; Islam et al., 2019). Several factors can explain this pattern, including cross-country differences in human capital accumulation, labor market frictions, and long-term work contracts. These possible drivers differ in both their scope to explain cross-country differences and their policy implications. As such, 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 a key source of workers' human capital accumulation: firm-provided training. In order to assess the importance of firm-provided training in explaining cross-country wage growth differentials, we carefully measure workers' post-schooling human capital investments and explore how they differ across countries. Our results provide an explanation for why post-schooling human capital accumulation is greater for workers in more-developed economies, and thus why their lifetime wage growth is higher.

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 start by carefully reviewing the labor literature that explores the link between job-related on-the-job training, human capital accumulation, and wages. We show that on-the-job training is consistently found to have large and significant effects on workers' productivity and wages in a variety of settings and time periods, and that the returns to training do not vary systematically across countries. We then present novel cross-country evidence on on-the-job training and its importance for workers' human capital acquisition. We rely on enterprise surveys covering more than 400,000 firms across 102 countries and worker-level surveys containing detailed information on workers' training investments for more than 600,000 people across 26 countries. These surveys allow us to construct harmonized representative measures of on-the-job training provision across countries with PPP-adjusted GDP per capita ranging from \$1,000 to \$60,000 and thus spanning a broad range of development levels. Our definition of training encompasses any organized or sustained on-the-job learning activity occurring outside of the formal education system, and thus captures several important sources of workers' human capital acquisition such as participation in seminars or workshops, along with more task-related learning arising from coworker instruction. We document two novel facts.

First, we document that the share of workers who receive firm-provided training rises strongly with country-level GDP per capita. We show that a key margin explaining this positive correlation is poor countries' large share of self-employed workers who do not receive employer-provided training. However, we still find evidence of this positive correlation when we restrict our attention to firm employment. Richer countries have a larger share of firms offering training, along with a larger share of trainees within these firms 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.[1] Second, we show that job-related firm-provided training is the main source of on-the-job human capital accumulation for workers. We find that 72% of all reported adult education corresponds to job-related training, and that almost all of this training is financed by firms.

Taken jointly, the evidence on the impact of firm-provided training on wages and productivity found in the literature and our novel facts linking firm-provided training with development suggest that the systematic cross-country differences in on-the-job training investments may play a key role in explaining cross-country wage growth and income differences. In addition, this evidence suggests that firm-provided training, and therefore firms, play a substantial role in adults' human capital investments. 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 and its consequences on workers' wage growth, 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.[2] The wage sector, on the other hand, is characterized by labor market frictions and firm heterogeneity à la Burdett and Mortensen (1998). Firms rent physical capital, post vacancies and wages, and meet workers by random search. We incorporate general training investments that follow the theoretical framework 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-breaking costs. In our model, workers can be separated from firms for two reasons: an exogenous separation

[1]Moreover, we find that country-level differences in observables such as occupation and education composition among others can only explain a small portion of the positive correlation between training and GDP per capita.

[2]This follows from empirical evidence found by Lazear and Moore (1984), who show that self-employed workers exhibit mostly flat wage-experience profiles.

shock that may lead workers to unemployment, and on-the-job search as workers look for new offers while working. When employed workers receive a new job offer, they can choose to exert efforts to break their contract, incurring contract-breaking costs.

We calibrate our model to match representative economies at different income levels. We focus on three main channels that vary greatly across different stages of development to explain the training gap between poor and rich countries: self-employment shares, job turnover rates, and physical capital endowments. The emphasis on these channels stems both from our empirical findings and the literature. The focus on self-employment is motivated by our empirical evidence suggesting that the high prevalence of this type of work is a key driver of the low rates of firm-provided training investments in poor economies, and the evidence found by Gollin (2002, 2008), showing that self-employment shares are much higher in developing compared to developed countries across all sectors. The focus on job turnover and thus on labor market frictions is rooted in the fundamental problem of financing training investments highlighted in the training literature and first identified by Becker (1964). In this problem, firms will be less likely to provide training investments if the probability of losing the worker is higher. As documented by Donovan et al. (2020), job turnover rates are considerably higher in developing countries relative to developed countries.[3] This suggests that the low training investments prevalent in developing countries may be partly explained by job turnover dynamics and labor market frictions. Finally, we also focus on physical capital differences because developing countries exhibit lower physical capital endowments (Caselli, 2005) which could affect the returns to skills and shape the incentives for training due to capital-skill complementarities (Krusell et al., 2000).

We find that the model explains 58% of wage growth differences across all countries. We then decompose the wage growth predicted by our model across all income levels into training and job turnover components in order to quantify their relative importance. We find that training explains 62% of the cross-country differences in wage growth profiles predicted by our model. Thus, since our model captures 58% of the cross-country differences in returns to experience, firm-provided training accounts for about 36% of cross-country wage growth differences.

We then conduct two decomposition analyses in order to explore the importance of the different channels to explain the training gap at different stages of development. First,

[3]Donovan et al. (2020) also show that the negative cross-country correlation between job turnover rates and income cannot be explained by observables (such as occupation, industry, education and firm size distribution).

we perform a sectoral accounting analysis, and find that a third of the aggregate training gap between the poorest and richest economies is explained by differences in the share of the self-employment sector in aggregate employment. Second, we perform a factor decomposition analysis and find that labor market frictions constitute the main driver of the differences in training investments across countries, explaining around 80% of the training gap at all income levels. The higher job separation rates prevalent in low- and medium-income economies and stemming from job destruction and job-to-job transitions not only could lead to higher shares of self-employment, but also depress the incentives to invest in training in the wage sector. 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 explaining the lack of training in poorer economies, while frictions in job-to-job transitions are more important in explaining the training differences between more-developed economies. In addition, we show that differences in physical capital productivity and sectoral productivity levels jointly explain the remaining 20% of the training gap.

Finally, we show that on-the-job training explains 12% of the income differences across countries in our quantitative model. Thus, the contribution of on-the-job training to cross-country income differences is sizeable: Lagakos et al. (2018a) show that differences in experience-related human capital explain around 20% of the income differences across countries.[4]

Related Literature. 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). By embedding the firm-worker training decision into a search model, our work also relates to the literature that focuses on disentangling the contributions of human capital and search dynamics to earnings (e.g., Rubinstein and Weiss, 2006; Barlevy, 2008; Yamaguchi, 2010; Burdett et al., 2011; Bagger et al., 2014; Gregory, 2021). These papers differ from ours along several key dimensions. First, a large contingent of these papers assume that on-the-job human capital accumulation is an

[4]We also consider the robustness of our quantitative results to several model extensions, including incorporating learning-by-doing (LBD) for workers, among others. Our results are very robust across these different model specifications. The relative importance of firm-provided training fluctuates between 29% and 38% when explaining cross-country wage growth differences, and between 9% and 15% when explaining cross-country income differences.

exogenous by-product of work and does not follow from an optimization problem where workers face tradeoffs between work and learning.[5] 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. In particular, this literature analyzes how job search and human capital accumulation contribute to explaining workers' wage growth for specific developed economies. We contribute to this literature by extending this decomposition analysis for countries at all income levels.[6]

By exploring the role of workers' training in explaining differences in GDP per worker across countries, our paper relates to a large strand of the literature that measures the importance 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 studies focusing on human capital differences, such as those arising from disparities in educational attainment (e.g., Hall and Jones, 1999; Erosa et al., 2010; Jones, 2014) and school quality (e.g., Hanushek and Woessmann, 2012; Schoellman, 2012; Martellini et al., 2022). Our paper focuses on on-the-job human capital accumulation, an understudied source of cross-country human capital differences. Thus, our work relates to the recent literature that highlights the potential importance of differences in lifecycle human capital accumulation across countries (De la Croix et al., 2018; Lagakos et al., 2018a,b; Islam et al., 2019). This literature, however, does not explain how these cross-country differences in on-the-job human capital accumulation emerge. Our paper attempts to fill this gap by delving into the processes and features giving rise to the low skill acquisition prevalent among workers in poor countries by focusing on employer-provided training.

Our paper is also 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 et al., 2020); (2) the literature documenting cross-country capital intensity differences, as reviewed in Caselli (2005); and (3) the literature focusing on cross-country differences in self-employment

[5] Exceptions to this are Wasmer (2006) and Flinn et al. (2017), who incorporate micro-founded human capital investment decisions. However, they focus on studying the distinction between firm-specific and general training.

[6] Through our paper's focus on on-the-job training to explain cross-country wage growth differences, our paper is rooted in the classic labor economics literature that examines the impact of on-the-job training on workers' human capital accumulation and earnings. Numerous studies provide evidence on the effect of on-the-job training, and particularly the effect of firm-sponsored training on workers' outcomes in a variety of contexts. In Section 2.1 we provide a summary of this evidence, and show that job-related firm-sponsored training is consistently found to have large and significant effects on workers' human capital acquisition and wages, and that the returns to training do not vary systematically across countries.

shares (e.g., Gollin, 2002, 2008; Poschke, 2018, 2019). We contribute to this development literature by incorporating the interaction between these channels and firm-provided training.[7]

Our paper also closely relates to recent papers focusing on a cross-country analysis of training. The first of these papers is Doepke and Gaetani (2020), who focus on the effect of employment protections on firms' and workers' incentives to invest in skills in order to study cross-country differences in on-the-job skill acquisition. The second paper is Engbom (2021), who studies how the costs of doing business affect human capital formation using a search model featuring endogenous human capital investments. Our work differs from both of these papers by focusing on different channels to explain on-the-job training differences, which include different labor market frictions, physical capital endowments, and self-employment. More importantly, while they focus on developed economies, we provide empirical and quantitative evidence for countries at all stages of development and focus on explaining the trend component of training with respect to per-capita GDP.[8]

The paper is organized as follows. Section 2 introduces our data and empirical findings. Section 3 presents the theory, and in Section 4 we calibrate a quantitative version of the model. Section 5 shows evidence of the drivers of wage growth differences across countries, the factor decomposition of training, income accounting results, and robustness of these results to model extensions. We conclude in Section 6.

[7]Moreover, through the interaction between firm employment distribution and training, this paper relates to the misallocation literature, which studies the productivity losses stemming from the extensive existence 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 causing productivity losses: the lack of on-the-job training.

[8]Finally, by analyzing human capital differences at all stages of development, our paper relates to two other papers. First, our paper relates to Manuelli and Seshadri (2014), who find that the lower TFP levels prevalent in developing economies raise the costs of accumulating human capital, thus lowering households' incentives to invest in human capital after schooling. In contrast with our theory, Manuelli and Seshadri (2014) focus on worker-level decisions on human capital while abstracting from firm-level decisions. Second, our paper relates to Guner and Ruggieri (2022), who document how inequality of labor earnings varies with GDP per capita, and build a search model with heterogeneous workers and firms and on-the-job training investments to interpret their findings. Their focus on the dynamics of earnings inequality differs from ours: we focus on how firm-provided training varies with development, and quantitatively link these training differences to cross-country wage growth and income gaps.

2 Empirical Evidence on On-the-Job Training

In this section, we start by reviewing the labor literature that explores the link between firm-sponsored on-the-job training, human capital accumulation, and wages. We then describe the data sources and carefully define on-the-job training in our data. Finally, we document facts about on-the-job human capital accumulation and the development process.

2.1 Review of Literature on the Effects of Firm-Provided On-the-Job Training

Our focus on firm-provided training is rooted in the numerous studies which document large and persistent impacts of on-the-job training (and particularly firm-sponsored job-related training) on wages in a variety of settings and countries. In this section, we perform a review of the labor literature that explores the link between on-the-job training, human capital accumulation, and wages. We focus specifically on job-related on-the-job training given the two facts we present below.

We summarize the evidence found by 86 studies in Table A.1. The studies considered document overwhelmingly positive and often significant effects of work-related training on wages, productivity, and other variables such as promotions, job continuity, and skill content of tasks. In addition, there is evidence suggesting that the effect of the work-related firm-provided training consistently exceeds that of other types of training interventions, and particularly those in which the employer is not actively involved. In Figure A.1 we plot the returns to training found by these studies against the GDP per capita in the year and country where the study was conducted.[9] We find that there are no systematic differences in the returns to training across countries. Appendix A includes further details on this literature review.

2.2 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 World Bank Enterprise Survey (WB-ES). For developed countries, on the other hand, we rely on the European Union Continuing Vocational Training (EU-CVT) enterprise survey, the Labor Force Survey (EU-LFS),

[9] In order to increase the number of studies we can compare, the studies we focus on examine the effects of the incidence of on-the-job training on wages.

and the Adult Education Survey (EU-AES). Our cross-country evidence encompasses economies with per-capita GDP levels ranging from $1,000 to $60,000.

The WB-ES is a collection of firm-level surveys of a representative sample of an economy's private manufacturing and service sectors covering approximately 136,000 firms across 140 low- and middle-income countries. The ES usually consists of interviews with establishments' owners and top managers, who can request assistance of their firms' accountants or human resources managers to answer certain questions. The ES has a set of country-specific questions reflecting each country's characteristics and a set of standardized questions that enable cross-country comparison. We rely on the two ES waves, between 2002 and 2005 and between 2006 and 2017, which have standardized questions on workers' training. We use the second wave (which provides individual weights) for the main analysis. We rely on the EU-CVT for enterprise data in the EU. This survey provides information on enterprises' investments in continuing vocational training of their staff, providing information on participation, time spent, and the costs of such training. Due to data availability, our analysis relies on three of the five waves of the EU-CVT conducted in 2005, 2010, and 2015, which cover all EU member states and Norway.

We rely on data from the EU-LFS and the EU-AES for worker-level data in Europe. The EU-LFS is a large household survey that provides data on labor force participation, unemployment, job characteristics, socioeconomic characteristics, and education and training of adults (ages 15+). The survey is conducted in all of the EU member countries and the three European Free Trade Association countries. Although the data collection dates back to 1983 for some countries, the data series are generally available from 1992 according to the EU membership. We use the data ranging from 2009 to 2018 for all countries to ensure consistency. Finally, the EU-AES collects information on participation in education and learning activities including job-related training, among others. Thus, this survey is conducted with the specific objective of understanding adult education patterns. The AES is one of the main data sources for the EU lifelong learning statistics and it covers around 666,000 adults ages 25–64. These data were collected during 2007, 2011, and 2017 in 26, 27, and 28 EU member states, respectively. Appendix B includes further details on the data sources used.

2.3 Defining On-the-Job Training

Before turning our attention to the cross-country evidence on 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 define *training* following the definition of "Non-formal Education and Training" category from ISCED (2011),[10] stating that training is any organized and structured learning activity outside the formal education system. Our definition encompasses two broad types of training: *formal training* and *informal training*. *Formal training* has a structured and defined curriculum and includes classroom work, seminars, and workshops, among others. Formal training activities are typically separated from the active workplace and show a high degree of organization by a trainer or an institution. Furthermore, this type of training is typically more general and not geared toward tasks, machinery, or equipment specific to certain jobs or workers.[11] *Informal training* involves task-related learning connected to the active workplace and often arising from coworker instruction. It encompasses guided on-the-job training, job rotation, exchanges, and other forms of learning arising from participation in learning circles. Appendix C presents these definitions in detail.

Our definition of training has two main features. First, it differentiates *training* from schooling. Therefore, *training* does not encompass programs such as MBAs that may be a source of human capital for some workers. Second, training activities must have a certain degree of organization and structure, which differentiates *training* from learning-by-doing and *informal learning* activities such as reading journals, visiting museums, or learning through media in an unstructured or unplanned way.[12] Thus, our definition of training encompasses all organized and structured on-the-job learning activities occurring outside of the formal education system, and thus captures several important sources of workers' human capital acquisition involving firms such as participation in seminars or workshops, along with more task-related learning arising from coworker instruction.

2.4 Facts on On-the-Job Training

The wide variety of data sources allows us to analyze training patterns for 102 countries in the main analysis and describe the key sources giving rise to adults' human capital

[10]The International Standard Classification of Education (ISCED) provides "uniform and internationally agreed definitions to facilitate comparisons of education systems across countries."

[11]Initial vocational training, employee orientation, and apprenticeships are excluded from formal training. We focus on continuing vocational training since it appears to be more relevant for explaining lifecycle increases in productivity. Nevertheless, in Figure D.5 we show that initial vocational training is also positively correlated with development and thus rule out the possibility that our results stem from a difference in the timing of human capital investments across countries.

[12]*Informal learning* is defined as a type of a learning activity that is not structured and is more related to workers' self-investments. Please see Appendix C for details.

accumulation. We now document two key facts about firm-provided training.

Fact 1 ***There exists a positive cross-country correlation between firm-provided training and income.***

In order to study the cross-country correlation between on-the-job training and income, we first focus on formal training since the data for this is available and consistent across the enterprise surveys we use (the WB-ES and the EU-CVT), yielding results for 102 countries at all levels of development. We construct country-year measures of the share of employees who receive formal training using the formula:

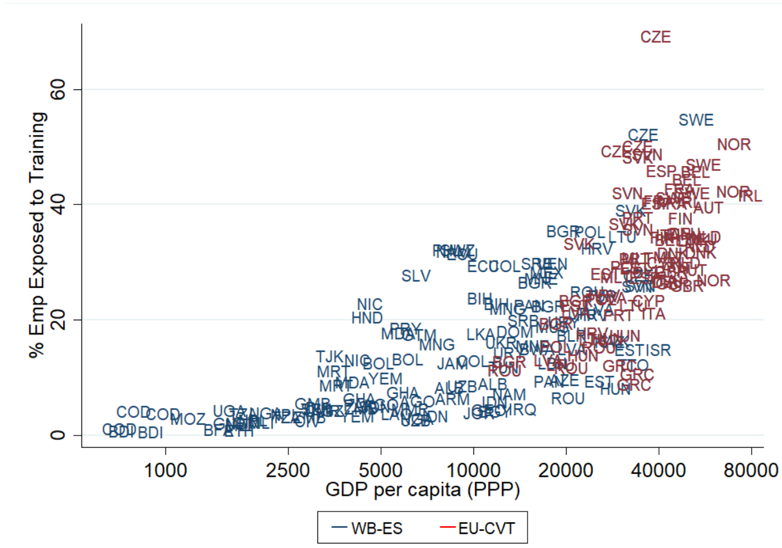
$$\%\,Trained\ \ Workers = \frac{Firms'\ Trained\ Workers}{All\ employees\ in\ firms} \times (100 - Self.\ \ Emp.\ \ Share).$$

The WB-ES and the EU-CVT provide information about which firms provide training, along with the share of workers who receive training in those firms. We use these two measures to construct the country-year measure of the share of employees exposed to 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.[13]

Formal on-the-job training increases with development. In Figure 1, we show the results of our combined measure of formal on-the-job training and GDP per capita. We find that as countries become more developed, formal 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 workers 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 data from 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.

[13]We restrict the sample from the WB-ES to 2005–2015 for comparability with the EU-CVT. We use firm weights provided by the data. The WB-ES tends to overweight larger firms, which causes mean firm-based employment to be counterfactually large in some countries. Poschke (2018) shows that log mean employment is lower than 4 even for countries with more than \$60,000 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 large firms. We show that the pattern documented is robust to performing the analysis on the unrestricted sample in Figure D.1.

Figure 1: Share of Formally Trained Employment and Development



Notes: The share of formally trained employment follows from adjusting the share of workers who receive training from firms by self-employment shares. Data on the share of employees trained within firms come from the WB-ES for all developing economies and from the EU-CVT for European economies. Both surveys contain data on whether firms provided formal training in the last year and the share of employees who participated. Data on GDP per capita and self-employment come from the Penn World Tables and World Bank Indicators, respectively.

Self-employment is a key driver of low levels of 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 countries. In Panel (a) of Figure 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 more compressed in this case, suggesting that the high share of self-employment exhibited in poor countries is a key factor driving low training levels in poor economies, as evidenced by Panel (b).

In Figure D.2, we further show that this positive correlation between training and income for workers employed by firms is prevalent at both the extensive and intensive margins.[14] In particular, richer countries exhibit both a larger share of firms offering training (extensive margin), and larger shares of trainees and more hours in training relative to total hours worked within these firms (intensive margin). In addition, richer countries exhibit a higher cost of training per participant, which potentially proxies for training quality.

[14]We rely on enterprise survey data from European countries to conduct this analysis. Although this analysis encompasses fewer countries, the relatively wide survey time frame and country coverage allow for sizeable income variation. Furthermore, in Figure D.4, we show the prevalence of similar patterns for developing economies using the WB-ES data: the share of firms offering training and the share of participants per firm increase with development.

Figure 2: Unadjusted On-the-Job Training Shares and Self-Employment

(a) % Trained workers in Firms **(b)** Self-employment

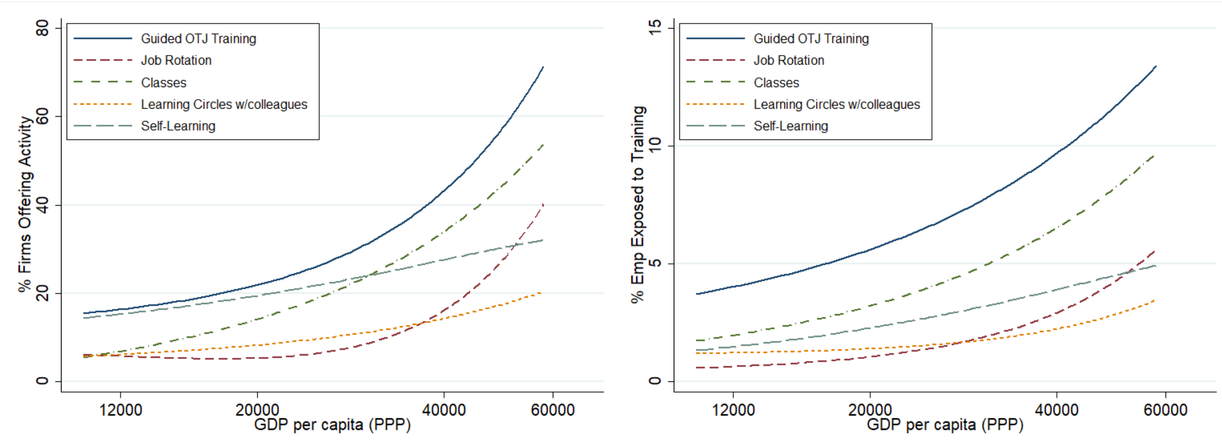
Notes: Panel (a) shows the share of workers who are trained by their employers out of total workers in firms, and Panel (b) shows the share of workers who are self-employed. Data on formal training come from the WB-ES for all developing economies and from the EU-CVT for European economies.

Informal on-the-job training increases with development. For our previous results, we focused on formal training in order to cover countries in all stages of development. However, using the EU-CVT, we are able to show evidence on the relationship between income and informal training, which is typically connected to the active workplace and is often tailored according to the learner's individual needs. This is important because more-developed countries could be providing more formal training at the expense of informal training. For all EU countries in 2005 and 2010, for which we have detailed data, we construct measures of the share of employees trained and the share of firms offering five different types of training: guided on-the-job training; job rotation and exchanges; participation in conferences, workshops, trade fairs, and lectures; participation in learning or quality circles; and self-directed learning. In Panels (a) and (b) of Figure 3, we plot the quadratic fit of the training measures with respect to GDP per capita for the share of firms that offer each one of these activities and the share of workers who participate, respectively. We find that all informal training activities increase with development.[15]

In Appendix E, we then show that the positive cross-country correlation between training and income showed in this section can only be partially explained by observables,

[15]It might be possible that due to a lack of resources, firms in poor countries do not offer training and workers replace this human capital source with informal learning. However, this does not seem to be the case. Figure D.3 provides measures of all types of informal learning in the AES survey (e.g., learning from family and friends and by using printed material or media, among others), and we show these measures have at best weakly positive correlations with development.

Figure 3: Informal Training



(a) % Firms Offering Activities **(b)** % Workers Offered Activities

Notes: This figure shows five types of informal training: planned training through guided on-the-job training; planned training through job rotation; planned training through participation in conferences, workshops, trade fairs, and lectures; planned training through participation in learning or quality circles; and planned training through self-directed learning/e-learning. Data on informal training come from the EU-CVT.

namely occupation, industry, education level of workers, and firm size. To do this, we decompose the trend in on-the-job training using shift-share accounting analysis. We find that these factors can only jointly explain 21% of the increase in training with income, which implies these observables drive only a small fraction of our results.

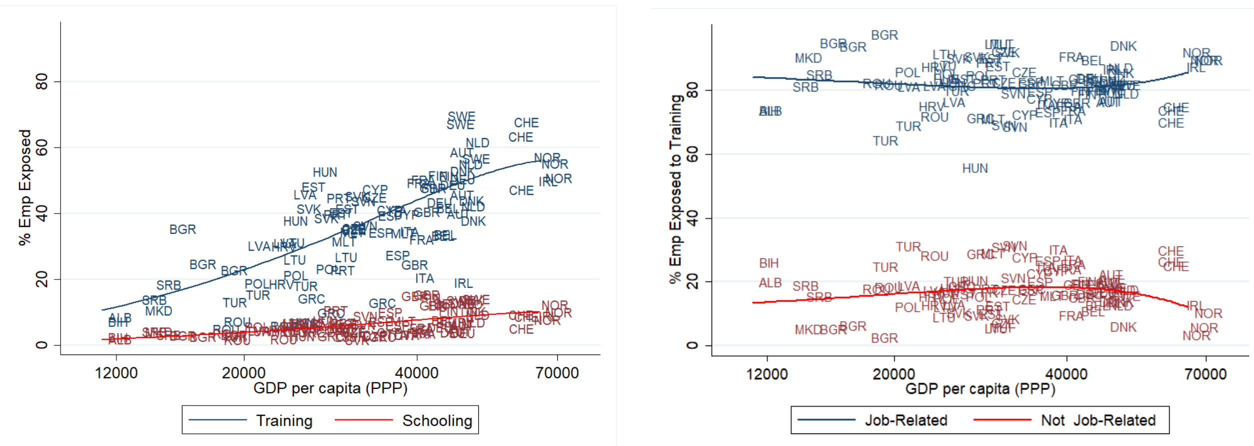
Fact 2 ***Firm-provided training is the main source of adult education.***

Previously we found a positive correlation between on-the-job training and development. However, although our measure of training encompasses several key sources of worker learning, it excludes others, namely schooling, learning-by-doing, and informal self-learning. We proceed by showing that our measure of on-the-job training is a predominant source of adults' human capital investments, and focuses on building general skills that improve workers' productivities, creating a scope for this positive correlation to explain cross-country human capital and income differences. With this purpose, we turn our attention to labor force and worker-level surveys containing detailed information on workers' training activities and education, which allows us to quantify the role of on-the-job training relative to other human capital sources of workers' learning.

Most of adult education is job-related training. We use data from the EU to show that on-the-job training is a predominant source of adults' human capital investments. First, we focus on data from the EU-AES and the EU-LFS, which collect information on the characteristics of all education and training investments in European countries. In

Figure 4 we show how the proportion of workers exposed to different types of education varies with cross-country income using data from the EU-AES. Panel (a) shows that the vast majority of adult education (around 90% of all adult education reported in the past year) is training, while less than 10% is schooling. Additionally, Panel (b) shows that around 80% of workers who report participating in some type of training claim that this is job-related, and interestingly, this share is uncorrelated with cross-country income.

Figure 4: Characteristics of Adults' Human Capital Accumulation



(a) Type of Education

(b) Job Relation of Training

Notes: Panel (a) shows the difference in the share of adults who participate in any type of educational activity. "Training" refers to our definition of informal + formal training, corresponding to "Non-formal Education and Training" from the International Standard Classification of Education 2011 (ISCED 2011). "Schooling" refers to "Formal Education and Training" according to the ISCED 2011. Panel (b) presents the share of job-related training in all the training reported in Panel (a) (blue line). Data come from the EU-AES.

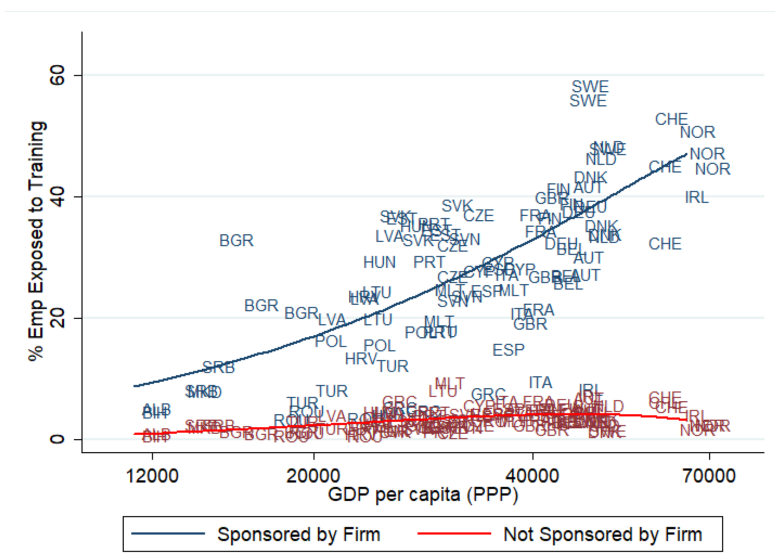
On average, 84% of adults in European countries report that the education they receive is job-related and only 16% mention personal or social reasons as the purpose of their training or education. This evidence suggests that job-related training is a primary source of adults' learning and human capital accumulation.[16] In addition, in Table D.1 we tabulate the share of firms providing CVT courses by type of skill targeted using data from the EU-CVT. This table shows that our measure of training focuses on building general skills such as management, customer handling, or technical skills, and

[16]In order to further show the importance of on-the-job training relative to other sources of worker learning, we also exploit our German BIBB data, which contains information on the sources from which workers report having learned the skills needed for their jobs (see Appendix B for further description of this data). Using this data, we build three measures of workers' skill acquisition: training, self-learning, and learning-by-doing. We find that during the period 1986-1999, about 62% of workers report training as their primary source of on-the-job skill acquisition, while these values are only about 32% and 4% for learning-by-doing and self-learning, respectively. These values are very stable across waves, and suggest that training is a key source of on-the-job learning for workers, and is significantly more important than other sources of learning.

that there are no systematic differences in the skills targeted by our training measure across countries.

Almost all of the job-related training is sponsored at least partially by firms. Figure 5 shows how training financing varies across European countries. We look at how the proportion of job-related training financed at least partially by the firm or completely by the worker varies with income. The graph shows that the vast majority of job-related training is sponsored by firms: less than 5% of workers in all countries receive some training directly related to their job that is entirely self-financed. Moreover, the share of adults who fully self-finance their job-related education is constant as a function of per-capita GDP, which reflects the fact that the correlation between job-related training and income is driven by firms offering more training, and not by workers themselves investing more in education. Doing some back-of-the-envelope calculation, our results show that 90% of all reported adult educational investments correspond to training, and 80% of all training is job-related and firm-sponsored. This means around 72% of all human capital investments are at least partially provided and financed by firms.

Figure 5: Training Financing



Notes: The graph shows the difference in the shares of adults who participate in firm-sponsored and non-firm-sponsored job-related training relative to GDP per capita. Data come from the EU-CVT.

Taken jointly, the evidence shown in this section suggests that firm-provided training is an important determinant of on-the-job human capital accumulation for workers, and that the systematic cross-country differences in on-the-job training investments we document may play a key role in explaining cross-country wage growth and income differences.

In addition, this evidence suggests that firm-provided training, and therefore firms, play a substantial role in adults' human capital investments. Thus, canonical 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.

3 Model

To shed light on the mechanisms giving rise to the positive correlation between training and development and its consequences for workers' wage growth, 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 (or traditional) sector and a wage (or modern) 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.

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

$$\max_{\{c^Y, c^O, k^Y\}} \; c^Y + \frac{c^O}{1+\rho} \qquad s.t. \quad Pc^Y = w^Y - \frac{k^Y}{\chi}, \quad Pc^O = \left(\frac{1-\delta_k}{\chi} + R\right) k^Y + w^O,$$

where superscripts Y and O denote young and old ages, and $\rho > 0$ governs time preference.[17] We treat the wage sector good as the numeraire, and P is the price of the consumption good. Young workers can invest in physical capital to save for the next period. While the Kaldor facts suggest a constant capital-to-output ratio, the recent Penn World Table shows that the capital-to-output ratio increases with development (Inklaar et al., 2019). We follow the literature (Krusell et al., 2000; Hsieh and Klenow, 2007) and introduce a capital-specific technological change parameter, χ, denoting that one unit of the wage sector good can be transformed into χ units of capital. We will let this parameter χ vary across countries to capture the increases in the capital-to-output ratio that occur with development. For analytical tractability, we assume workers make

[17] The wages w^Y and w^O for young and old ages are net of the training costs paid by the workers.

sectoral choices in the beginning of the first period. We allow for workers' reshuffling between sectors in the quantitative analysis.

Consumption Good Production. The consumption good is a composite of goods from the two different sectors: the self-employment (or traditional) sector good C_T and the wage (or modern) sector good C_M:

$$C = \left(\gamma C_T^{\sigma} + (1-\gamma) C_M^{\sigma}\right)^{\frac{1}{\sigma}}.$$

Since the wage sector good is the numeraire, the price of the consumption good is $P = \frac{1}{1-\gamma}\left(\frac{C}{C_M}\right)^{\sigma-1}$.

Self-Employment Sector. Production in the self-employment sector is characterized by a constant-returns-to-scale function:

$$Y_T = A_T N_T,$$

where A_T and N_T denote productivity and labor in this sector respectively. We assume training is not provided to workers in this sector, following the empirical evidence on flat wage-experience profiles for self-employed people (Lazear and Moore, 1984). All the goods produced by the self-employment sector are used for consumption: $Y_T = C_T$. The price of the self-employment-sector good is $P_T = \frac{\gamma}{1-\gamma}\left(\frac{C_T}{C_M}\right)^{\sigma-1}$.

Wage Sector. This sector is characterized by frictional labor markets. There is a unit measure of firms with heterogeneous productivity $z \sim G(z)$. Firms produce a homogeneous good, which is used for consumption and paying training and vacancy costs, and can also be transformed into physical capital. Firms' production function is Cobb-Douglas as in the development-accounting literature (e.g., Caselli, 2005). Once workers and firms are matched, worker i's production in firm j is

$$y_{ji} = A_M z_j h_i^{1-\mu} k_{ji}^{\mu},$$

where A_M denotes the productivity in this sector, z_j is the firm-specific productivity, h_i is worker i's efficiency units of labor (human capital), and k_{ji} is the amount of capital rented by firm j to equip worker i. The elasticity of output to capital is given by $0 < \mu < 1$. Since capital is rented at a constant rate R, the firm chooses the optimal capital level k_{ji}

to maximize net revenue from worker i's production given her level of human capital,

$$\max_{k_{ji}\geq 0}\ r_{ji} = A_M z_j h_i^{1-\mu} k_{ji}^{\mu} - Rk_{ji}.$$

By solving this problem, we can denote $\tilde{r}(z) = (1-\mu)A_M z\left(\frac{\mu A_M z}{R}\right)^{\frac{\mu}{1-\mu}}$ as net revenue per efficiency unit in a firm with productivity z, which facilitates the characterization of training decisions below.[18] $\tilde{r}(z)$ decreases with capital rent R, suggesting that cheaper physical capital induces higher labor productivity. By aggregating output across all workers within each firm and across all firms, we obtain total output in the wage sector:

$$Y_M = A_M \int_j z_j \int_{i\epsilon j} h_i^{1-\mu} k_{ji}^{\mu}\ di\ dj.$$

Job Search and Matching. Firms post vacancies $v(z)$ at the start of each period, with a contract stipulating the wage rate per efficiency unit $w(z)$ and working period, which we assume to be two periods for young workers and one period for old workers. The vacancy cost is defined by $c_v \frac{v^{1+\gamma_v}}{1+\gamma_v}$, and similar to Acemoglu and Hawkins (2014), we require vacancy costs to be strictly convex (i.e., $\gamma_v > 0$), ensuring that firms with different productivity levels coexist. The total number of vacancies is then $V = \int v(z)dG(z)$. The wage distribution of offers is $F(w) = \int_{w(z)<w} v(z)dG(z)/V$.

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 full-time job jointly with all newly born workers. Moreover, a portion η of the remaining old workers search on the job. Therefore, the number 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 wage sector. For analytical tractability, we consider the matching function as $M(\tilde{U}, V) = \min\{\tilde{U}, V\}$, and assume c_v is small enough such that $V > \tilde{U}$, which ensures full employment. As usual, market tightness is $\theta = \frac{V}{\tilde{U}}$.

Contract Enforceability and Workers' Optimal Separation Policy. If old workers who search on the job get an outside offer, they can exert an effort to break their current work contract.[19] Specifically, these workers choose a probability p of breaking

[18]In particular, if k_{ji} yields an internal solution, we obtain $k^*_{ji} = \left(\frac{\mu A_M z_j}{R}\right)^{\frac{1}{1-\mu}} h_i$, and with optimal k^*_{ji}, the net revenue from worker i's production is $r_{ji} = (1-\mu)A_M z_j h_i \left(\frac{\mu A_M z_j}{R}\right)^{\frac{\mu}{1-\mu}}$.

[19]Consistent with the previous literature, firms cannot break work contracts in our setting, since they always benefit from hiring and willingly pay for training costs (Acemoglu and Pischke, 1999).

their current contract, and incur the costs $c_p^{\gamma_p} \frac{p^{1+\gamma_p}}{1+\gamma_p}$ per efficiency unit. The costs represent a friction in job-to-job transitions, with a lower c_p representing lower costs of leaving the firm.[20] We assume $\gamma_p > 0$, such that the marginal cost of breaking contracts increases with probability p. We use the contract-breaking cost as a modeling tool that can be quantitatively adjusted to match different job-to-job transition rates across countries.[21]

In a firm with productivity z, a worker faced with an outside offer w' chooses the optimal leaving probability $p \in [0, 1]$ by solving[22]

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

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

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

This result is intuitive. If the new wage offer is lower than the wage at the current firm, workers do not want to switch jobs, and the investment in breaking the contract is 0. On the other hand, if the new wage offer is large enough ($w' > w(z) + c_p^{\gamma_p}$), workers want to switch firms and will therefore break the contract with a probability of 1. If the cost of breaking contracts increases, workers are less willing to switch firms and thus have a lower probability of breaking the contract.

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

[20] With this, we capture that the enforcement of labor and training contracts or noncompete clauses may vary across countries. In Appendix I.5.1, within the context of a cross-country calibration that targets all 203 country-year observations separately, we find that conditional on GDP per capita, the calibrated contract-breaking costs are negatively correlated with labor union power, unemployment benefits, and the generosity of minimum wages across countries. This result provides support to our modeling of contract-breaking costs: in countries where workers have higher negotiation power or protection in the labor market, our calibrated contract-breaking costs tend to be lower.

[21] In Appendix I.2, we instead adjust the on-the-job search intensity (η) to match these job-to-job transitions across countries, and show that the quantitative results are similar.

[22] We solve for workers' optimal choice of leaving probability p while taking the level of training investments as given for two reasons. First, when the new offer arrives in the beginning of the second period, training has already occurred. Second, firms and workers need to internalize workers' probability of leaving the firm when deciding on the optimal level of training. Thus, they must choose training according to the optimal contract-breaking efforts conditional on each new offer.

the next-period's 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. Training is decided upon and paid for jointly by firms and workers. There is a constant cost c_s per unit time of training, reflecting trainers' wages and material costs.[23]

We assume that training raises general human capital, so its benefits accrue even if the worker changes firms.[24] Moreover, we assume that if s_W and s_F are optimal training levels from workers' and firms' perspectives, respectively, training s will be given by $s = \min\{s_W, s_F\}$. This assumption implies that the training level is determined by the party with lower affordability and is thus quite reasonable. For instance, if firms bear all the training costs, workers may desire very high training levels, yet firms would not like to pay for them. The optimal level of training 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*

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

where, in a firm with productivity z, current wage w, new offers of wage w', a wage distribution of 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 respectively

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

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

Proof: See Appendix F.1. □

[23]In principle, training also reduces trainees' production time. Since the analytical properties of the model will not be affected by these training time costs, we omit them here. However, in the quantitative analysis, we include them to match key features of the data.

[24]We focus on general human capital since Table D.1 suggests that our measure of training focuses on building general skills (such as management, customer handling, or technical skills). The literature also shows that the firm-specific components of human capital have been found to be much less important for wage growth than the general component (Altonji and Shakotko, 1987; Lazear, 2009; Kambourov and Manovskii, 2009).

Proposition 1 explains how optimal training is determined when we have different divisions of training costs. As the share of training costs paid by each group increases, the optimal level of training for that group 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.[25] For the calibrated economy, we find that firm decisions determine training investments, as firms always want lower levels of training than workers. Thus, we now focus on understanding firm-level decisions.

Proposition 2 (Labor Market Frictions and Firms' Training) *Given the offer distribution $F(w)$, in a firm with productivity level z, the firm's optimal training level $s_F(z)$*
(1) increases with costs of breaking contracts c_p;
(2) decreases with exogenous separation rate δ; and
(3) increases if capital rent R is cheaper.

Proof: See Appendix F.2. □

The first two results in Proposition 2 indicate that a higher probability of job separation leads to lower training. These results illustrate how higher contract-breaking costs and lower job destruction generate more training investments in our model. The third result suggests that lower capital rent leads to higher training, as cheaper capital induces higher labor productivity and thus larger returns to training.

Propositions 1 and 2 indicate that higher costs of breaking contracts, lower job destruction, and higher physical capital stocks generate more training investments in our model. However, it is also worth noting that since the self-employment sector features no learning, aggregate training patterns are also affected by changes in the share of workers in this sector. For instance, if the exogenous job destruction rate δ increases, the

[25] In this model, firms are willing to invest in general training. This departure from Becker (1964) is due to frictional labor markets, which allow firms to extract partial rents from training (Acemoglu and Pischke, 1999). We differ from the general training literature (e.g., Acemoglu and Pischke, 1999; Engbom, 2021) in that: (1) we assume the cost shares paid by workers and firms are common across firms; and (2) we add a time cost of training when we take the model to the data. We model the economy in this way because when we include training time costs (which are the main costs of training in the data), having constant cost shares of training across firms can help us jointly match training patterns by firm productivity and aggregate training levels. If we were to instead set these cost shares to maximize firm-worker joint surplus, the model would generate an inverse relationship between training investments and firm productivity, which is counterfactual.

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 further reduces aggregate human capital.

Solving Firms' Optimal Choices In each period, a firm chooses wage $w(z)$, vacancies $v(z)$, and young workers' training $s(z)$ to maximize its value. The firm's value function is detailed in Appendix F.3. Note that $s(z)$ is determined according to Proposition 1, whereas $w(z)$ and $v(z)$ are determined according to the FOCs of the firm's value function. In particular, $w(z)$ is determined by a first-order differential equation, combined with the minimum wage $b\bar{w}$, as in Burdett and Mortensen (1998).[26] 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 eventually as the costs of additional vacancies increase.

Solving Workers' Sectoral Choices. If there is a non-zero measure of workers in both sectors, workers must be indifferent in terms of expected utility between going to the self-employment sector and the wage sector in the beginning of the first period. This indifference condition is provided in Appendix F.4.

Equilibrium. In Appendix F.5, we define the model's general equilibrium.

4 Model Estimation

In this section, we extend our two-period analytical model for the quantitative analysis and match our model to the data.

[26]As shown by Hornstein et al. (2011), search and matching models with reasonable unemployment benefits have difficulty in generating the amount of frictional wage dispersion present in the data. Thus, because of our focus on training decisions, we choose to match the frictional wage dispersion by assuming the lowest wage to be $w_{\min} = b\bar{w}$, where $\bar{w}$ denotes the average wage and b is a constant. We assume that the unemployed will take any job offer, which can be rationalized by low, often negative, values of unemployment benefits. This assumption matches empirical findings of the offer acceptance rate being close to one (van den Berg (1990)). Because under these assumptions unemployment benefits do not affect any other equilibrium outcomes, we abstract from unemployment benefits in the model.

4.1 Quantitative Model

We add some features to closely replicate key aspects of the labor market and economic environment.

Workers. We consider that workers live for $J > 2$ periods with discount rate ρ. We assume that human capital from training depreciates at rate d every period, in line with empirical evidence (Blundell et al., 2019), and overall human capital remains above a lower bound, which we assume to be the level of human capital agents are born with (basic cognitive and physical skills).

It is well-documented that some workers flow between self-employment and wage jobs (Donovan et al., 2020). To capture these flows, we assume that right before job search happens in each period, a proportion τ_c of unemployed and self-employed workers have an opportunity to adjust sectors. We follow Artuc et al. (2010) and introduce idiosyncratic preferences $\{x_M, x_T\}$, distributed i.i.d. across sectors and workers according to a Type-I extreme distribution $e^{-e^{-\tau x}}$. Under these assumptions, a potential mover of age a and human capital h compares the value of staying unemployed $W_M^a(h) + x_M$ with the self-employment value $W_T^a(h) + x_T$, where $W_M^a(h)$ and $W_T^a(h)$ are the income flows of staying in each of the two sectors, respectively.

Firms. Training costs are assumed to be proportional to the average wage $c_s\bar{w}$, while we also consider the opportunity cost of training—each unit time of training also causes a δ_s decrease in unit time for production. With this setup, training costs are proportional to income levels (training benefits), and thus the training gap between countries is not directly driven by income levels. Therefore, the wage income for a worker of age a in firm z is given by $w(z)h - \mu_W(c_s\bar{w} + \delta_s\tilde{r}(z))s^a(z,h)$, which depends on the firm-level wage per efficiency unit of labor $w(z)$, the worker's human capital h, the training time provided for this worker in this firm $s^a(z,h)$, the training costs $(c_s\bar{w} + \delta_s\tilde{r}(z))$, and the worker's share of training costs μ_W. Vacancy costs are proportional to income levels $c_v\bar{w}\frac{v^{1+\gamma_v}}{1+\gamma_v}$, which ensures that labor market frictions are not directly driven by income levels. Finally, we assume firms' productivity to be Pareto-distributed, $G(z) = 1 - z^{-\kappa}$, as is often found empirically (Axtell, 2001).

Exogenous Job-to-Job Moves. We assume that the moving probability p has a lower bound $\underline{p} > 0$, capturing that a portion of job-to-job flows are associated with wage losses

(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. We use the widely employed matching function $M(\tilde{U}, V) = c_M \tilde{U}^{\psi} V^{1-\psi}$ for the wage sector. This matching function yields positive unemployment and reasonable elasticities of matches with regard to searchers $\tilde{U}$ and vacancies V.

Conditions for Simulations. The optimal conditions for the quantitative model provide the same intuition as in our analytical model, and are presented in Appendix G. In this appendix, we show that the optimal levels of training depend on firm productivity and workers' age, and also show how firms' wages and vacancies are determined.

4.2 Calibration

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 16 moments describing labor market dynamics and training investments to identify model parameters. Then, we perform a second calibration for representative economies at 10 different income levels to understand how training investments change with development. To this end, we jointly re-calibrate the parameters $\delta, c_p, A_M,\ A_T$, and χ to match self-employment, job destruction rate, job-to-job transition, income levels, and capital-to-output ratios in each representative economy.

4.2.1 Calibrating the Model to the United States

Pre-assigned parameters. We first directly set some parameters following the literature. We calibrate the model to quarterly data. Thus, we set the quarterly discount rate ρ to 0.01. Each individual works for 40 years, and therefore the lifetime length is set to $J = 160$ quarters. We choose the elasticity of output to capital as $\mu = 0.3$, according to Gollin (2002). The ratio of the lowest wage to the average wage is calibrated to $b = 0.6$ following Hornstein et al. (2011), who calculate the mean-min ratio of wages to be around 1.7 from US labor data. We choose the elasticity of matches to searchers in the matching function to be $\psi = 0.7$, as estimated by Shimer (2005). We use $\frac{1}{1-\sigma} = 3$ for the elasticity of substitution between the self-employment and wage sectors in the aggregate production function as in Feng et al. (2018). We set the on-the-job search intensity to be 0.4 following Faberman et al. (2017), who find that the average number of offers per month received by employed workers is around 40% of that for unemployed people in the US. Finally, according to the Penn World Table, the relative price of consumption to capital formation and annual capital depreciation rate were 1.1 and 0.04 respectively between 1994 and 2007 in the United States. Thus, we set the capital-specific technology for the United States to be $\chi = 1.1$ and the quarterly depreciation rate to be $\delta_k = 0.01$.

Table 1: Pre-assigned Parameters

Parameter	Model	Source
ρ - Discount rate	0.01	Annualized interest rate of 0.04
J - Number of periods	160	40 years of work
μ - Elasticity of output to capital	0.3	Gollin (2002)
b - Ratio of lowest wage to average wage	0.6	Hornstein et al. (2011)
ψ - Elasticity of matches to searchers	0.7	Shimer (2005)
$\frac{1}{1-\sigma}$ - Elasticity of substitution	3	Feng et al. (2018)
η - on-the-job search intensity	0.4	Faberman et al. (2017)
χ - Capital-specific technology (US)	1.1	Penn World Table
δ_k - Capital depreciation rate	0.01	Penn World Table
γ_v - Convexity of vacancy costs	1	Dix-Carneiro et al. (2019)
δ_s - Loss in production hours per unit time of training	0.7	EU-LFS 2004 Training Module

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) find a relatively low convexity value of 0.2 for Swiss firms. We use $\gamma_v = 1$ in our baseline calibration. Second, we calibrate the loss in production hours per unit of training time to be $\delta_s = 0.7$, by taking the average from European countries' labor force surveys. We summarize the information on pre-assigned parameters in Table 1.

Parameters to estimate. The remaining parameters to estimate are the constant in the matching function, c_M; the costs per unit time of training 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 self-employment-sector share in the aggregate production function, γ; the convexity in the function of leaving costs, γ_p; the shape parameter of the Pareto productivity distribution, κ; the exogenous separation rate, δ; the lower bound of leaving probability, $\underline{p}$; the share of training costs paid by firms, μ_F; the depreciation rate of human capital, d; the proportion of workers who have a chance to adjust sectors, τ_c; and the parameter in the distribution of sectoral preferences, τ. Finally, since the relative ratio of A_T and A_M has the same effect as the self-employment-sector share γ in the aggregate production, we normalize the US aggregate productivity to be $A_M = A_T$ and choose A_M such that the output per worker is 1.

Targeted moments and fit. To calibrate those remaining parameters, we target the following moments: the average unemployment rate from 1994 to 2007; the ratio of the number of vacancies to the number of unemployed people from 2000 to 2007 from FRED; the share of self-employment in total employment from 1994 to 2007 from the World Bank; the ratio of capital to annual real GDP from 1994 to 2007 from the Penn World Table; the Pareto parameter of firm employment distribution as estimated by Axtell (2001); workers' average wage growth after job-to-job transitions and the share of job-to-job transitions from high to low wage firms as computed by Haltiwanger et al. (2018); the relative wage-job finding rate (unemployed/self-employed) and the transition rate from unemployment to self-employment for the US from Donovan et al. (2020); 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 the Bureau of Labor Statistics (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. (2018b), 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. We rely on data from Donovan et al. (2020), who provide these two probabilities for many countries across a wide range of income levels. Given that these two moments are targeted for determining job separation and job-to-job frictions when we take the model to a cross-country comparison, we directly predict the two probabilities of interest from Donovan et al. (2020) using per-capita GDP for representative economies at 10 different income levels. Although the predicted values for the US are a little higher than the actual US values, we choose to use the predicted values in order to be consistent with our calibration in the second step for representative economies at different income levels.

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 the economies we consider in this paper. Thus, we first take the average hours of formal training per worker from the data.[27] We then impute overall training intensity for every economy, relying on two assumptions according to the Survey of Employer-provided Training (US-SEPT): the average worker spends two hours in informal training for each hour spent on formal training and there are 50% more workers participating in informal training than in formal training.[28] Table H.1 shows that 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.

Calibrated Parameters. We report the calibrated parameters in Table 2. Our parameters are reasonable compared with the literature. Our parameter γ_s can be interpreted as the diminishing returns of human capital investments (in terms of effective hours) in producing new human capital. Its calibrated value $\gamma_s = 0.22$ is in the ballpark of the estimates in the literature. For instance, Imai and Keane (2004) also find this parameter to be 0.22. Moreover, training a young worker for the full quarter (480 working hours) increases her hourly wage by 6%, which lies within the range of empirical studies on US training returns reviewed by Leuven (2004) and Bassanini et al. (2005),[29] and is consistent with the cross-country training returns documented in Figure A.1. Our calibrated quarterly depreciation rate of human capital from training $d = 0.02$ is similar to the annual depreciation rate of 0.06–0.08 of training returns estimated by Blundell et al. (2019) using British labor surveys.

Table 2: Calibrated Parameter Values

c_M	c_s	c_v	γ_s	γ	γ_p	κ	ζ	$\underline{p}$	μ_F	δ	c_p	d	τ_c	τ	A_M
0.59	0.20	1.84	0.22	0.28	8.69	6.64	0.06	0.12	0.95	0.02	2.23	0.02	0.21	0.94	0.26

[27] We multiply shares of workers exposed to formal training by hours spent on 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.

[28] In the United States, 60% of workers receive formal training and 90% receive informal training.

[29] For example, using the National Longitudinal Surveys (NLSY), Veum (1995) finds that increasing one hour of formal training improves hourly wages by 0.01%. Also using NLSY data, Frazis and Loewenstein (2005) find that 60 hours of formal training increases wages by 3% to 5%. Our calibration implies 4% wage growth for 60 hours of training in one quarter.

4.2.2 Cross-Country Calibration

We calibrate the model for representative economies at 10 different income levels aside from that of the United States. To do this, we re-calibrate a few parameters associated with our mechanisms of interest, namely self-employment, job turnover (captured jointly by job-to-job transitions and job destruction rates), and physical capital endowments. The focus on these mechanisms is motivated by our empirical evidence, the literature, and particularly the fact that the size and nature of these channels radically change with development.[30]

We keep most of the baseline parameters at the US levels and re-calibrate δ, c_p, A_M, A_T, and χ to match income levels, self-employment, the share of workers who stay in the same firm from quarter to quarter, the share of workers who stay employed from quarter to quarter, and the capital-to-output ratios in each representative economy.[31] We show how the model fits the targeted moments in each of our representative economies in Figure H.1, and find that our model matches the targeted data moments very well. In Figure H.2 we show how the values of the re-calibrated parameters change across the representative economies. For details on these calibration results, see Appendix H.2.

4.2.3 Non-targeted Moments and Model Validation

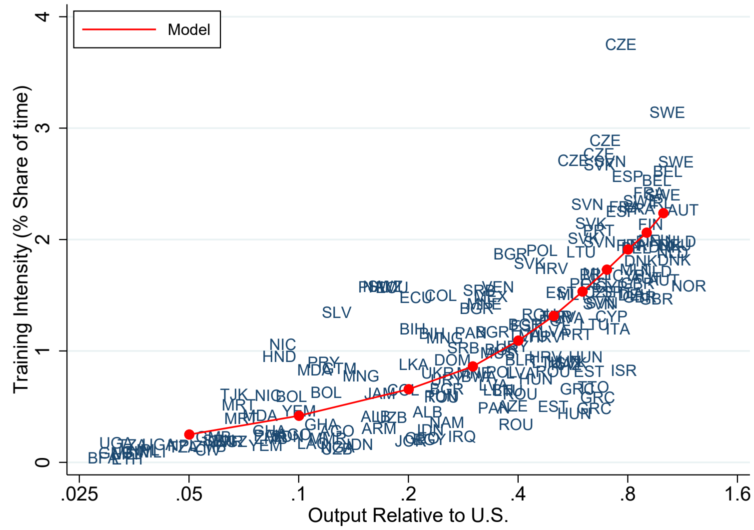
We first turn our attention to the main non-targeted moments we want to analyze: training levels in our representative economies and the elasticity of training with respect to income. We plot the training intensity from the data and model as a function of GDP per capita in Figure 6. The model matches the levels and elasticity of training data with respect to per-capita GDP well.

Now, in Table 3, we compare several non-targeted moments in the model to the data across countries. First, we compare the relationship between different measures of the labor share and income from Gollin (2002). The first measure (adjustment 1) assumes that the labor share in the self-employment sector is equal to 1, while the second measure (adjustment 2) assumes that the labor share in the self-employment sector is identical to its counterpart in the wage sector. Our model captures that the labor share in the wage sector increases with income, whereas the large share of self-employment may induce a

[30] Gollin (2002, 2008) find developing countries exhibit higher shares of self-employment, while Donovan et al. (2020) find that job turnover rates are higher in these economies. The Penn World Table indicates that the capital-to-output ratio increases with development (Inklaar et al., 2019).

[31] According to the Penn World Table, the slope of log capital-to-output ratios on log GDP per capita is 0.27. We use this slope to compute the capital-to-output ratio relative to the US level for each representative economy.

Figure 6: Training in the Model vs Data



Notes: This graph shows the cross-country training intensity (measured in the share of time that an average worker spends in training) as a function of output relative to the US. The red line shows the outcome of the model and in blue we plot each country-year observation from the data.

high labor share in poor countries if the labor share in self-employment is 1. Second, we check our modeling of the capital-specific technology by comparing how the change in the relative price of capital to consumption with respect to income differs between our model and the data. In the model, the price of capital formation relative to consumption is mainly driven by the inverse of capital-specific technology $1/\chi$.[32] We find that both the model and the data predict a decline in the relative price of capital to consumption with respect to income levels, yielding quantitatively similar elasticities.

Table 3: Non-targeted Moments in the Model vs Data

Non-targeted Moments Across Countries	Data	Model
1. Slope of labor shares on log GDPPC (adj 1)	-0.02	-0.03
2. Slope of labor shares on log GDPPC (adj 2)	0.02	0.03
3. Slope of log relative price of capital formation to consumption on log GDPPC	-0.15	-0.23

Notes: The measures of the labor share and income to calculate moments 1 and 2 come from Gollin (2002). The first measure (adjustment 1) assumes the self-employment sector's 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. We use the Penn World Table to compute the slope of log relative price of capital formation to consumption on log GDP per capita.

[32] The price of capital formation relative to consumption is $1/(\chi P)$ in the model, where P is the price of the consumption good. We find that $1/\chi$ strongly declines with income and mainly drives the negative relationship between the price of capital formation relative to consumption and income.

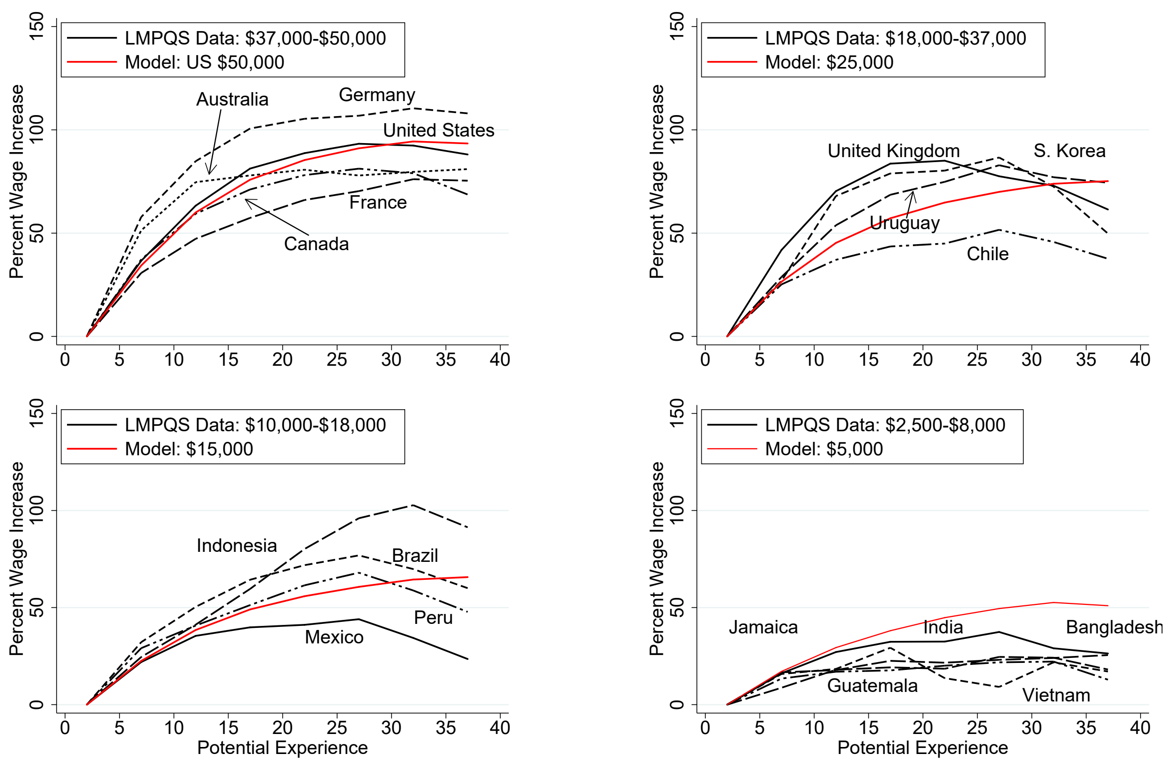
5 Wage Growth, Training, and Income Differences

In the following section, we aim to answer three main questions: (1) What portion of wage growth differences across countries can on-the-job training account for? (Section 5.1), (2) Why do developed economies invest more in training? (Section 5.2) ; and (3) What portion of income differences across countries can on-the-job training account for? (Section 5.3). We also consider the robustness of these results to several model extensions in Section 5.4.

5.1 Cross-Country Wage Growth Differences

We first analyze how much our model, and specifically training, contribute to explaining the differences in workers' lifecycle wage growth between developed and developing economies. Figure 7 plots the experience-wage profiles of the 18 economies studied in Lagakos et al. (2018b) (LMPQS henceforth), which span all income levels, and the corresponding model predictions. Each panel shows the profiles from countries within a

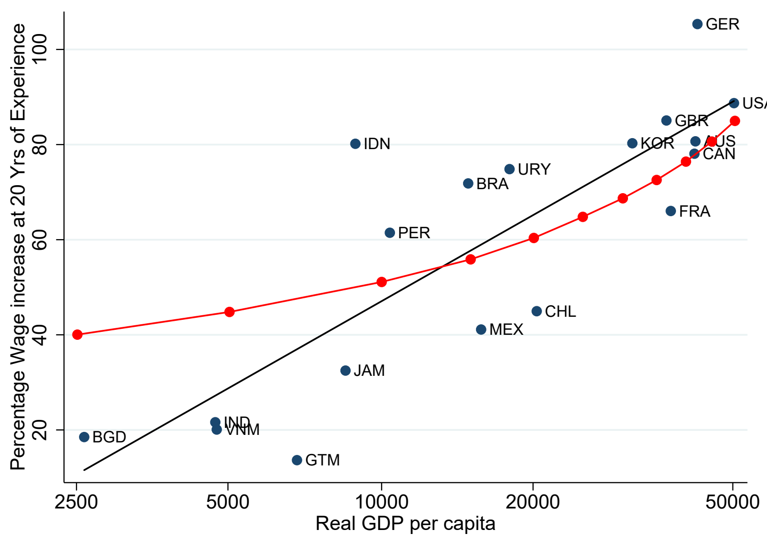
Figure 7: Cross-country Experience-Wage Profiles: LMPQS vs Model



Notes: This figure replicates Figure 2 from Lagakos et al. (2018b) (LMPQS) and adds the wage-experience profiles from our model in red. Along the y-axis we plot the percent increase in wages at each potential experience bin, and along the x-axis we plot potential experience in years.

particular income range and the model's profile for an economy within that same range. Our model matches the wage growth profiles well at all income levels except for those at the bottom of the world income distribution. The calibrated economy at $5,000 of GDP per capita has a steeper experience-wage profile than its counterparts in the data. This suggests that other factors that we do not model may play an important role in explaining the low wage growth in these low-income economies.

Figure 8: Cross-country Experience-Wage Profiles: LMPQS vs Model



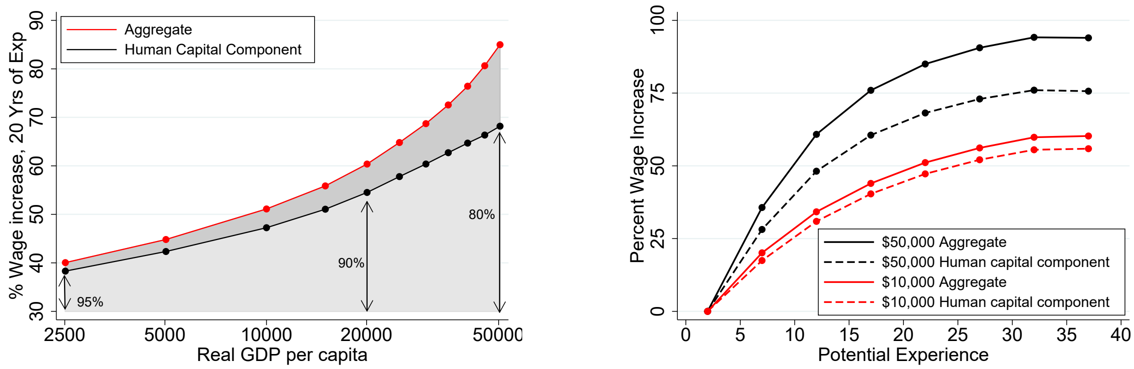
Notes: This figure replicates Figure 3 from Lagakos et al. (2018b) (LMPQS) and adds the returns to experience from our model in red. The slope in the LMPQS data is 26%, while the slope of the model-predicted returns on log per-capita GDP is 15%.

We now turn our attention to quantifying what portion of the cross-country difference in returns to experience our model can account for. To this end, in Figure 8 we plot the cross-country returns to 20 years of experience found by LMPQS and the corresponding model predictions as a function of per-capita GDP. As before, the model matches the wage growth for middle- and high-income countries very well, and overestimates the wage growth for workers in the poorest economies. We then regress these returns on log per-capita GDP, and find a slope of 0.26 in LMPQS and a slope of 0.15 in our model. This implies that our model captures 58% of the cross-country differences in returns to experience.[33]

We then decompose the wage growth predicted by our model across all income levels into human capital (or training) and job turnover components in order to quantify their relative importance. In Panel (a) of Figure 9, we present the model-predicted returns to

[33]Nonetheless, the model captures all of the difference for the economies above $10,000.

Figure 9: Cross-country Experience-Wage Profiles: Composition



(a) Decomposition and Income

(b) Profile Decomposition

Notes: This figure shows the decomposition of model-predicted wage growth into human capital (or training) and job turnover components. Panel (a) plots the returns to 20 years of experience as a function of per-capita GDP, and Panel(b) plots wage-experience profiles. Wage growth stemming from human capital is calculated using the average increase in human capital for workers at each level of potential experience. The residual wage growth stems from job turnover.

20 years of experience as a function of per-capita GDP in the aggregate model and with just the human capital component. We find that the contribution of human capital to wage growth is large for every economy, though it decreases with income. This stems from the high level of job destruction prevalent in the poorest economies, which prevents workers from climbing up the job ladder. As income increases, fewer workers are separated from their jobs and become unemployed, which generates larger increases in wages over the lifecycle through job-to-job transitions. In Panel (b) we plot this decomposition for the full model-predicted wage-experience profiles for two economies with GDP per capita levels of $50,000 and $10,000, and find a similar pattern. We find that the human capital component explains 62% of the differences in workers' wage growth between these two economies, while job turnover explains the remaining 38%.[34] Thus, since our model captures 58% of the cross-country differences in returns to experience, firm-provided training accounts for about 36% of cross-country wage growth differences.

[34] We calculate these numbers by (1) subtracting the 20-year wage increase in the economy with GDP per capita level of $10,000 from that of the country with the $50,000 GDP per capita level in both the full and human capital component models and then (2) taking the ratio between these two subtractions to obtain how much of the differences between these two economies' rates of wage growth can be explained by human capital. There is a difference of 34 percentage points in wage growth at 20 years of experience between these economies, while the difference in wage growth just coming from differences in human capital accumulation is 22 percentage points. Thus, 62% (22/34) of the difference stems directly from differences in training investments.

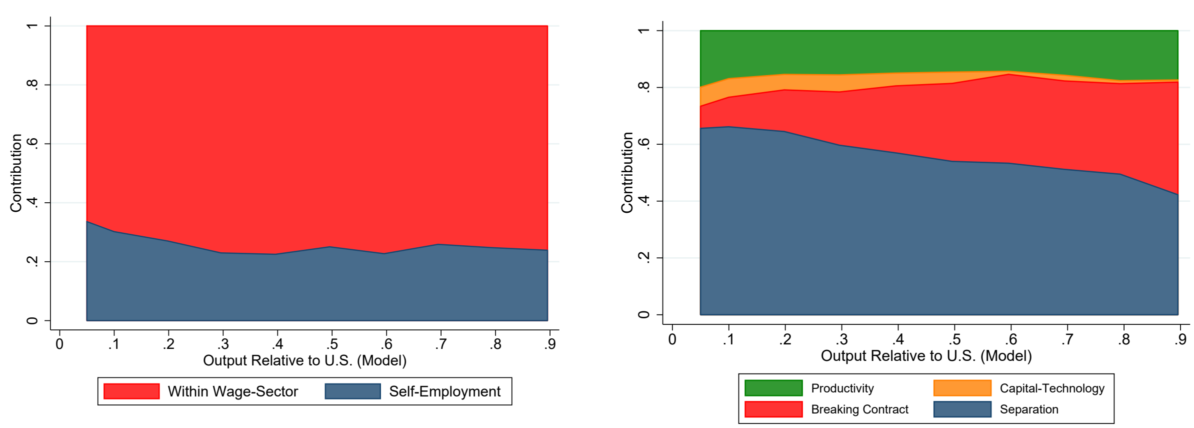
5.2 Training Decomposition

In this section, we aim to understand what drives the lack of training and the role played by each of our channels at different stages of development. We first perform a sectoral accounting analysis that seeks to understand how the sectoral allocation of employment between the self-employment and wage sectors shapes training differences in the aggregate. Denoting S as the training investment in the wage sector and M as the wage-sector employment share, the difference in training between the baseline economy and the United States can be simply decomposed as

$$log(S_{US}M_{US}/S_{base}M_{base}) = log(S_{US}M_{US}/S_{US}M_{base}) + log(S_{US}M_{base}/S_{base}M_{base}).$$

The first term reflects the training increase due to the change in the share of self-employment in total employment, while the second term represents the increase in training in the wage sector, conditional on the sectoral allocation. In Figure 10a we perform this decomposition using the calibrated model. We find that around 35% of cross-country training differences are explained by differences in the share of self-employment in aggregate employment. We also find that the importance of self-employment slightly decreases with income, in line with our finding that the poorest economies have very high self-employment shares and thus few workers exposed to training.

Figure 10: Training Decomposition



(a) Decomposition by Sectoral Component **(b)** Decomposition by Parameter

Notes: Figure (a) shows how (1) changing the self-employment share while keeping training in the wage sector fixed (blue) and (2) changing the wage sector training level while keeping the self-employment share fixed (red) contribute to explaining the difference in training between each economy and the United States. Figure (b) depicts the contribution of each channel to explaining the training gap between the economies at each income level and the United States. The green area represents the contribution from changing A_M and A_T simultaneously, the orange area represents the contribution from changing χ, the red area represents the contribution from changing c_p, and the blue area represents the contribution from changing δ.

We now explore what portion of the differences in training investments across countries can be explained by differences in labor market frictions, physical capital productivity, and sectoral productivity. To do this, we perform a factor decomposition analysis where we subsequently change the values of the parameters governing these channels. We specifically focus on the five parameters that vary across countries: δ, which shapes job destruction; c_p, which shapes job-to-job transitions; χ, which shapes physical capital intensity; and A_T and A_M, which denote self-employment and wage-sector productivity levels, respectively, and shape income and self-employment shares. For each economy, we simulate the model when changing the value of one of the country-specific parameters (c_p, δ, χ, and jointly A_T and A_M) to match its value in the US economy, and compute the respective change in training. Using this, we then calculate how much of the training gap between each economy and the United States is explained by each channel. We plot the results in Figure 10b.[35]

Most of the difference in training investments across countries is driven by differences in labor market frictions. Differences in the cost of breaking contracts and job destruction jointly explain around 80% of the training differences at all income levels. The higher job separation rates prevalent in low- and medium-income economies that stem from job destruction and job-to-job transitions not only could lead to higher shares of self-employment, but also depress the incentives to invest in training in the wage sector. We also find that the contribution of each of these two components changes with income. In particular, the contribution of job destruction tends to decrease with income, whereas the importance of the cost of breaking contracts increases with income. Thus, for poor economies the most important channel in terms of explaining the lack of training is job destruction, but as income increases, the differences in training stem largely from frictions in job-to-job transitions.

We also find that differences in physical capital productivity and sectoral productivity levels jointly explain the remaining 20% of the training gap. Our results show that the difference in capital-specific technology is an important factor to explain the lack of training in the poorest economies. Its contribution decreases with income, rendering it almost irrelevant to explain differences in training between high-income economies. Finally, we find that differences in sectoral TFP levels play a similar role in determining training differences across all income levels.

There are four main takeaways from the decomposition analyses we perform. First,

[35]Since there may be interactions between the different channels, we normalize the contribution of each factor using the sum of the individual contributions.

sectoral accounting suggests that around one-third of the training gap between poor and rich economies is explained by self-employment shares. Second, labor market frictions are key to explaining training investments. High job separation rates and low contract-breaking costs make job turnover more likely and thus depress the incentives to invest in training in low- and medium-income economies. Third, when we decompose the contribution of these labor market frictions along its two key components, we find that job destruction is the main driver of the lack of training in the poorest economies, while differences in job-to-job transitions are more important in explaining the training differences between more-developed economies. Fourth, the lack of physical capital is important in explaining the lack of training in low-income countries but not in richer countries.

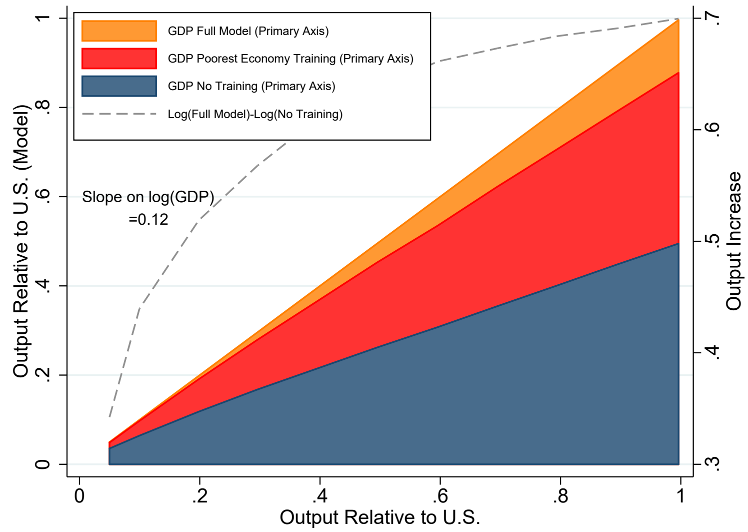
5.3 Explaining Cross-Country Income Differences

We now focus on income differences explained by on-the-job training. 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 along the primary y-axis of Figure 11. In orange we plot the original model, in blue we plot the model with no training, and in red we plot the case where all economies have the same training investments as the poorest economy.[36] Output is the lowest when there is no training. Output increases when we add the poorest economy level of training to the model with no training, and increases even more when we endogenize training, which reflects the fact that training boosts productivity in the aggregate. The heterogeneous increase in output with respect to income shows that adding training improves output more in developed economies than in developing economies.

Using this information, we now quantify the share of income differences across countries that can be explained directly by training in our model. To do this, we plot the difference between the log(per-capita GDP) in the full model and its counterpart in the model with no training along the secondary y-axis. This difference represents the percentage increase in output when we move from the model with no training to the full model. The slope of this percentage increase indicates the share of income differences explained directly by training in our model. Thus, our quantitative model suggests that on-the-job training explains 12% of income differences across countries. The contribution of on-the-job training to cross-country income differences is thus sizeable, given that Lagakos et al. (2018a) show that differences in experience-related human capital explain

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

Figure 11: Income Increase due to Training



Notes: This figure shows the model-predicted per-capita GDP for each model variation along the primary y-axis; and the percentage increase in output from training along the secondary y-axis. Both of these are plotted as functions of GDP per capita. The percentage increase in output from training is calculated as the log change in output from the model with no training (increasing c_s to an extremely large value) to the full model. The slope of 0.12 represents the share of the increase in GDP per capita explained by training in the model. Each observation comes from using the calibrated version of the model for each country.

around 20% of the income differences across countries.

5.4 Model Extensions and Robustness of Results

We consider the robustness of our quantitative results to several model extensions: (1) incorporating learning-by-doing (LBD) for workers; (2) abstracting from contract-breaking costs and adjusting on-the-job search intensities to match different job-to-job transition rates across countries; (3) endogenizing job turnovers by modeling endogenous layoffs related to job tenure (Donovan et al., 2020); (4) allowing for different firm productivity distributions across countries to capture cross-country differences in firm size distributions; and (5) calibrating the model separately to all country-year observations for which we have training data, and targeting the training intensity to discipline country-specific training returns. A summary of these results is presented in Table I.1, which compares the main quantitative results across the baseline and alternate model specifications. Our results are very robust across these different model specifications. The relative importance of firm-provided training fluctuates between 29% and 38% when explaining cross-country wage growth differences, and between 9% and 15% when explaining cross-country income differences across all specifications.[37] For further details about each of the model extensions, along with a discussion of the information in Table I.1, please see Appendix I.

6 Conclusion

Human capital accumulation plays a key role in economic growth. Recent research has highlighted the potential importance of on-the-job human capital accumulation in explaining workers' wage profiles. In this paper, we study one key source of on-the-job human capital accumulation: firm-provided training. We exploit rich enterprise- and worker-level data sources to show that firm-provided training increases with development and that this firm-provided training is the most important source of human capital investments in workers' careers. Then, we build a general equilibrium model with firm heterogeneity and training investments to shed light on the mechanisms giving rise to these facts and their consequences for workers' wage growth.

Our results have several implications for understanding economic growth and conducting policy. First, our data and model suggest that self-employment is key to explaining the lack of on-the-job training in the poorest economies. Thus, our theory suggests that the reallocation of workers away from self-employment triggers human capital gains that add to the productivity gains identified by the literature and stemming from the movement toward higher productivity work. Second, we examine the evolving importance of different channels to explain the training gap at different stages of development, and find that the high level 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 results imply that in order to increase training and productivity, policies that improve the match quality between firms and workers may be desirable in developing economies, whereas policies that improve labor contracts may be more beneficial in richer countries.

Finally, it is worth noting that the importance of on-the-job training could be larger if this type of learning has complementarities with other sources of human capital, such as schooling or co-worker spillovers. A fruitful area for future research would be to study how different sources of human capital accumulation interact with each other and how these interactions matter for countries at different stages of development.

[37] In addition, the results from the exercise that calibrates the model separately to all country-year observations and allows for country-specific training returns suggest that there are no systematic differences in the returns to training across countries, which is consistent with the evidence in Figure A.1 and corroborates our assumption of constant returns to training in the baseline calibration.

References

Acemoglu, D. (1997). Training and Innovation in an Imperfect Labour Market. The Review of Economic Studies, 64(3):445–464.

Acemoglu, D. and Hawkins, W. (2014). Search with Multi-worker Firms. Theoretical Economics, 9.

Acemoglu, D. and Pischke, J.-S. (1998). Why Do Firms Train? The Quarterly Journal of Economics, 113(1):79–119.

Acemoglu, D. and Pischke, J.-S. (1999). Beyond Becker: Training in Imperfect Labour Markets. The Economic Journal, 109:112–142.

Altonji, J. G. and Shakotko, R. A. (1987). Do Wages Rise with Job Seniority? Review of Economic Studies, 54(3):437–459.

Artuc, E., Chaudhuri, S., and McLaren, J. (2010). Trade Shocks and Labor Adjustment: A Structural Empirical Approach. American Economic Review, 100(3).

Axtell, R. L. (2001). Zipf Distribution of U.S. Firm Sizes. Science, 293(5536):1818–1820.

Bagger, J., Fontaine, F., Postel-Vinay, F., and Robin, J.-M. (2014). Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics. American Economic Review, 104(6):1551–96.

Barlevy, G. (2008). Identification of Search Models using Record Statistics. The Review of Economic Studies, 75(1):29–64.

Bassanini, A., Booth, A., Brunello, G., De Paola, M., and Leuven, E. (2005). Workplace Training in Europe. IZA Discussion Paper Series.

Becker, G. (1964). Human Capital. The University of Chicago Press, Chicago.

Bento, P. and Restuccia, D. (2017). Misallocation, Establishment Size, and Productivity. American Economic Journal: Macroeconomics, 9(3):267–303.

Blatter, M., Muehlemann, S., Schenker, S., and Wolter, S. C. (2016). Hiring Costs for Skilled Workers and the Supply of Firm-provided Training. Oxford Economic Papers, 68(1):238–257.

Blundell, R., Costa, M., David, D., and Meghir, A. G. C. (2019). Wages, Experiences and Training of Women over the Lifecycle. NBER Working Paper Series.

Botero, J. C., Djankov, S., Porta, R. L., Lopez-de Silanes, F., and Shleifer, A. (2004). The Regulation of Labor. The Quarterly Journal of Economics, 119(4):1339–1382.

Burdett, K., Carrillo-Tudela, C., and Coles, M. G. (2011). Human Capital Accumulation and Labor Market Equilibrium. International Economic Review, 52(3):657–677.

Burdett, K. and Mortensen, D. T. (1998). Wage Differentials, Employer Size, and Unemployment. International Economic Review, 39(2):257–273.

Caselli, F. (2005). Accounting for Cross-Country Income Differences. In Aghion, P. and Durlauf, S. N., editors, Handbook of Economic Growth, volume 1, chapter 9, pages 679–741. Elsevier.

Ciani, A., Hyland, M. C., Karalashvili, N., Keller, J. L., Ragoussis, A., and Tran, T. T. (2020). Making It Big: Why Developing Countries Need More Large Firms. The World Bank Group.

De la Croix, D., Doepke, M., and Mokyr, J. (2018). Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Preindustrial Economy. The Quarterly Journal of Economics, 133(1):1–70.

Dix-Carneiro, R., Goldberg, P., Meghir, C., and Ulyssea, G. (2019). Trade and Informality in the Presence of Labor Market Frictions and Regulations.

Doepke, M. and Gaetani, R. (2020). Why Didn't the College Premium Rise Everywhere? Employment Protection and On-the-Job Investment in Skills. NBER Working Paper

Series.
Donovan, K., Lu, J., and Schoellman, T. (2020). Labor Market Flows and Development. Working Paper.
Engbom, N. (2021). Labor Market Fluidity and Human Capital Accumulation.
Erosa, A., Koreshkova, T., and Restuccia, D. (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. J., Mueller, A. I., Sahin, A., and Topa, G. (2017). Job Search Behavior among the Employed and Non-Employed. NBER Working Paper Series.
Feng, Y., Lagakos, D., and Rauch, J. (2018). Unemployment and Development. NBER Working Paper Series.
Flinn, C., Gemici, A., and Laufer, S. (2017). Search, Matching and Training. Review of Economic Dynamics, 25:260–297.
Frazis, H. and Loewenstein, M. A. (2005). Reexamining the returns to training: Functional form, magnitude, and interpretation. Journal of Human Resources, 40(2):453–476.
Gollin, D. (2002). Getting Income Shares Right. Journal of Political Economy, 110(2):458–474.
Gollin, D. (2008). Nobody's Business but My Own: Self-employment and Small Enterprise in Economic Development. Journal of Monetary Economics, 55(2):219–233.
Gregory, V. (2021). Firms as Learning Environments: Implications for Earnings Dynamics and Job Search.
Guner, N. and Ruggieri, A. (2022). Misallocation and Inequality.
Hall, R. E. and Jones, C. I. (1999). Why do Some Countries Produce So Much More Output Per Worker than Others? The Quarterly Journal of Economics, 114(1):83–116.
Haltiwanger, J., Hyatt, H., Kahn, L. B., and McEntarfer, E. (2018). Cyclical Job Ladders by Firm Size and Firm Wage. American Economic Journal: Macroeconomics, 10(2):52–85.
Hanushek, E. A. and Woessmann, L. (2012). Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. Journal of Economic Growth, 17(4):267–321.
Hornstein, A., Krusell, P., and Violante, G. L. (2011). Frictional Wage Dispersion in Search Models: A Quantitative Assessment. American Economic Review, 101(7):2873–98.
Hsieh, C.-T. and Klenow, P. J. (2007). Relative Prices and Relative Prosperity. American Economic Review, 97(3).
Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. The Quarterly Journal of Economics, 124(4):1403–1448.
Hsieh, C.-T. and Klenow, P. J. (2010). Development Accounting. American Economic Journal: Macroeconomics, 2(1):207–23.
Imai, S. and Keane, M. (2004). Intertemporal Labor Supply and Human Capital Accumulation. International Economic Review, 45(2):601–641.
Inklaar, R., Woltjer, P., and Albarrán, D. G. (2019). The Composition of Capital and Cross-country Productivity Comparisons. International Productivity Monitor, Centre for the Study of Living Standards, 36:34–52.
Islam, A., Jedwab, R., Romer, P., and Pereira, D. (2019). Returns to Experience and the Misallocation of Labor.
Jones, B. F. (2014). The Human Capital Stock: a Generalized Approach. American Economic Review, 104(11):3752–77.

Kambourov, G. and Manovskii, L. (2009). Occupational Specificity of Human Capital. International Economic Review, 50(1):63–115.

Klenow, P. J. and Rodriguez-Clare, A. (1997). The Neoclassical Revival in Growth Economics: Has it Gone Too Far? NBER macroeconomics annual, 12:73–103.

Krusell, P., Ohanian, L. E., Ríos-Rull, J.-V., and Violante, G. L. (2000). Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. Econometrica, 68(5):1029–1053.

Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2018a). Life-cycle Human Capital Accumulation across Aountries: Lessons from US Immigrants. Journal of Human Capital, 12(2):305–342.

Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2018b). Life Cycle Wage Growth across Countries. Journal of Political Economy, 126(2):797–849.

Lazear, E. P. (2009). Firm-Specific Human Capital: A Skill-Weights Approach. Journal of Political Economy, 117(5):699–726.

Lazear, E. P. and Moore, R. L. (1984). Incentives, Productivity, and Labor Contracts. The Quarterly Journal of Economics, 99(2).

Leuven, E. (2004). A Review of the Wage Returns to Private Sector Training.

Manuelli, R. E. and Seshadri, A. (2014). Human Capital and the Wealth of Nations. American Economic Review, 104(9):2736–762.

Martellini, P., Schoellman, T., and Sockin, J. (2022). The Global Distribution of College Graduate Quality. Working Paper.

Moen, E. and Rosén, A. (2004). Does Poaching Distort Training? The Review of Economic Studies, 71(4):1143–1162.

Poschke, M. (2018). The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology. American Economic Journal: Macroeconomics, 10(3):1–41.

Poschke, M. (2019). Wage Employment, Unemployment and Self-Employment across Countries. Working Paper.

Restuccia, D. and Rogerson, R. (2013). Misallocation and Productivity. Review of Economic Dynamics, 16(1):1–10.

Rubinstein, Y. and Weiss, Y. (2006). Post Schooling Wage Growth: Investment, Search and Learning. In Hanushek, E. and Welch, F., editors, Handbook of the Economics of Education, volume 1, chapter 1, pages 1–67. Elsevier.

Schoellman, T. (2012). Education Quality and Development Accounting. The Review of Economic Studies, 79(1):388–417.

Shimer, R. (2005). The Cyclical Behavior of Equilibrium Unemployment and Vacancies. American Economic Review, 95(1):25–49.

van den Berg, G. J. (1990). Search Behaviour, Transitions to Non-participation and the Duration of Unemployment. The Economic Journal, 100(402):842–865.

Veum, J. R. (1995). Sources of Training and Their Impact on Wages. Industrial and Labor Relations Review, 48(4):812.

Wasmer, E. (2006). General Versus Specific Skills in Labor Markets with Search Frictions and Firing Costs. American Economic Review, 96(3):811–831.

Yamaguchi, S. (2010). Job Search, Bargaining, and Wage Dynamics. Journal of Labor Economics, 28(3):595–631.

Online Appendix

A Literature Review on Effects of OTJ Training

In this section, we perform a review of the labor literature that explores the link between on-the-job training, human capital accumulation, and wages. With this, we provide comprehensive evidence on the importance of firm-provided training in shaping workers' human capital accumulation, and thus the scope of the cross-country differences in on-the-job training investments we document to drive cross-country wage growth and income differences.

We summarize the evidence found by 86 studies in Table A.1. This table is divided into two panels distinguishing between two types of studies: those that focus on worker-level outcomes and thus generally employ worker-level data, and those that focus on firm- or industry-level outcomes and generally employ establishment-level data. Within each of these panels, the table further distinguishes the studies by country. If multiple specifications or outcomes are explored in one study, these are summarized separately within the studies. Note some studies employ linked employer-employee data and are thus able to look at both worker- and firm-level outcomes. The results of these studies are separated into each of the corresponding panels. The same is true for studies that explore the effects of training in several countries separately.

The studies encompass many different types of contexts and time periods, though there are fewer studies set in developing countries than in developed countries. In developing countries, the literature has primarily focused on quantifying the impact of government-sponsored active labor market policies such as vocational training programs on worker outcomes (see Card et al. (2011), Attanasio et al. (2011), Hirshleifer et al. (2016), Alfonsi et al. (2020), and Caicedo et al. (2021), among others). McKenzie (2017) and Card et al. (2018) provide reviews of this literature, and show that in most settings, these training programs increase both earnings and the probability of employment, and that the size of these effects tends to rise over time.

We now turn our attention to Panel A in the table, which refers to studies that focus on worker-level outcomes. Work in this area documents overwhelmingly positive and often significant effects of work-related training on wages and other variables such as promotions, job continuity, productivity and skill content of tasks. We cannot directly compare the magnitudes of these effects across studies since the length and scope of the training may vary. In particular, some studies feature contemporaneous training, which may reduce the time spent on working and thus earnings (see for example Fialho et al. (2019)). Importantly, this panel also suggests that these effects are present in a variety of settings and time periods, suggesting that the positive impacts of on-the-job training on workers' human capital and wages exist at all levels of development. Although the identification strategy varies across studies, it is encouraging that those that exploit random assignment into training find large and positive results (De Grip and Sauermann (2012), Adhvaryu et al. (2018)).

Panel B summarizes studies that focus on firm-, establishment- or industry-level outcomes. The number of studies in this panel is considerably smaller than in Panel A since firm-level variables are less widely available than worker-level variables. Much work in this area has focused on empirically estimating the production function to study the impact of on-the-job training on workers' productivity (measured as value added per worker) in a variety of settings, finding overwhelmingly positive effects. Other work in this area focuses on disentangling the impact of on-the-job training on wages from its impact on productivity. These studies generally find that the productivity gain from firm-sponsored training is substantially higher compared to the wage gain, indicating both that on-the-job training is linked to human capital acquisition, and that firms have an incentive to pay for training investments. For example, Konings and Vanormelingen (2015) use firm-level data from Belgium and find that increasing the share of trained workers by 10 percentage points is associated with 1.7% to 3.2% higher productivity, and 1.0% to 1.7% higher average wage. The findings of this study, together with many others, justify why firms pay for general human capital investments.

Moreover, there is evidence suggesting that the effects of the work-related firm-provided training consistently exceed that of other types of training interventions, and particularly those in which the employer is not actively involved (What Works - Centre for Local Economic Growth (2016); Hansson (2008)). A report reviewing the impact of on-the-job training programs, What Works - Centre for Local Economic Growth (2016), highlights the importance of employers' involvement by showing that training opportunities which are job-related, in-firm, and co-designed by employers tend to be much more impactful for worker wages and employment trajectories than other training opportunities.

A.1 Returns to Training

We now examine whether these training returns vary systematically across countries. To do this, we plot the returns to training predicted by these studies against the GDP per capita in the year and country where the study was conducted. In order to increase the number of studies we can compare, the studies we focus on examine the effects of the incidence of on-the-job training on wages.[38] Further, in order to contemplate information from overlapping countries and time periods, we construct a 5-year average per country of the training returns found by different studies. We plot the results of this in Figure A.1, and find that are no systematic differences in the returns to training across countries.

[38]Nevertheless, it is worth noting that the definition of "incidence of training" may vary across studies.

Table A.1: Studies examining the effects of firm-provided training on wages and human capital accumulation

Study	Period	Dataset	Sample	Training Definition	Outcome, Estimation Method	Effect Sizes
Panel A: Studies that focus on worker-level outcomes						
United States						
Duncan and Hoffman (1979)	1975	PSID	White men Black men White women Black women	Years of OJT	Log Wage, OLS	0.0538*** 0.0592*** 0.0853*** 0.0664***
Lillard and Tan (1986)	1983 1969–1980	CPS NLS Young Men	Men Young Men	Incidence of OJT (informal) Incidence of OJT (company) Number of OJT Spells	Log Wage, OLS	0.224*** 0.044*** 0.119***
Gronau (1988)	1977–1979	PSID	Women Men	Incidence of OJT	Log Wage, SEM	0.7910*** 0.6643***
Mincer (1988)	1968–1982	PSID	Men	One Year of OJT	Yearly Wage Growth, OLS	0.044***
Brown (1989)	1976–1984	PSID	Heads of Household	Years of OJT	Log Wage, FE	0.21***
Lynch (1992)	1980–1983	NLSY79	Young Non-College Graduates	Weeks of OJT	Log Wage,, OLS+SC Log Wage, FE	0.002 -0.0002
Bishop (1994)	1982–1987	EOPP-NCRVE, NFIBS	New Hires	Incidence of OJT	Log Starting Wage, OLS Log Current Wage, OLS Subjective Starting Prod., OLS Subjective Current Prod., OLS	0.019 -0.013 0.095*** -0.003
Bartel (1995)	1986–1990	Personnel records for Manufacturing Co.	Professional employees	Incidence of OJT Days of OJT	Log Wage, IV+FE	0.106*** 0.016***
Hill (1995)	1967–1984	NLS Mature Women	Middle-Aged Women	Incidence of OJT (1967-1977) Incidence of OJT (1977-1984)	Log Wage, FD	0.02 0.06*
Veum (1995)	1986–1990	NLSY79	Young	Hours of OJT Incidence of OJT	Log Wage, OLS Log Wage, FE Log Wage, OLS Log Wage, FE	0.004 0.008 0.0728* 0.0897**
Krueger and Rouse (1998)[1]	1991–1995	Records for Manufacturing Co. Records for Service Co.	Production Workers All Workers	Incidence of OJT	Log Wage, FE	0.004*** -0.002
Loewenstein and Spletzer (1999)	1988-1991	NLSY79	Young	Number of OJT Spells	Log Wage, OLS	0.044**
Parent (1999)	1979–1991	NLSY79	Young	Years of OJT	Log Wage, OLS Log Wage, HT	0.1692*** 0.1216***
Marcotte (2000)	1966-1981 1979-1994	NLS NLSY79	White Young Men	Incidence of OJT	Log Wage, OLS	0.14*** 0.105***
Frazis and Loewenstein (2005)[2]	1979-2000	NLSY79	Young	Hours of OJT	Log Wage, FE Log Wage, FE (cube root spec.)	0.003 0.036
Hamil-Luker (2005)	1977–1987 1988–1998	NLS Young Women NLSY79	Young Women Young Women	Incidence of OJT	Log Wage, RE	0.05*** 0.07***
Blanchflower and Lynch (2007)	1979–1988	NLSY79	Young Non-College Graduates	Incidence of OJT	Log Wage, OLS	0.08
Hannagan et al. (2010)[3]	1996–2008	CAWP, Emily's List Almanac American Pol.	Female Pro-Choice Democrat Candidates in Contested Elections for US House (Q1 of PSM)	Endorsement/Training Emily's List	Victory, PSM	8.839***
Canada						
Parent (2003)	1991–1995	FSLS	Young Men Young Women	Incidence of Firm-Sponsored OJT	Log Wage, FE	0.1034*** 0.0168
Yoshida and Smith (2005)	1999–2000	WES	Native-born Men	Number of OJT Courses	Wage Growth OLS	0.0009*
Germany						
Lechner (1999)	1990-1994	GSOEP	East Germany	Incidence of OJT (6 months after)	Full-time Employed, PSM Monthly earnings, PSM	0.30* 867
Pischke (2001)	1986–1989	GSOEP	Prime-age	Years of OJT	Log Wage, FE Log Wage, FE+Indiv. Growth FE	0.026 0.038
Kuckulenz and Zwick (2003)	1998–1999	BIBB/IAB	Full-time West Germany	Incidence of OJT	Log Wage, OLS+SC	0.15***
Albert et al. (2010)	1995–2001	ECHP	Prime-Age	Incidence of OJT	Log Wage, FE	-0.030
Görlitz (2011)†	2007–2010	WeLL	Manufacturing or Service Workers	One course of OJT Two courses of OJT Three courses of OJT	Log Wage, Tobit+Firm FE+SC	0.005 0.022 -0.018
Mazza (2015)	1993–2008	GSOEP	Prime-age	Incidence of OJT (in 1990–1993) Incidence of OJT (in 1997–2000) Incidence of OJT (in 2001–2004) Incidence of OJT (in 2005–2008)	Log Wage, FE	0.036*** 0.016** 0.004 -0.006
Tamm (2018)†	2007–2010	WeLL	Manufacturing or Service Workers	Incidence of OJT	Non-routine tasks, Worker FD + Firm FE Non-routine tasks, Worker FD + Firm FE	0.0086** 0.0005
France						
Goux and Maurin (2000)†	1988–1993	FQP	Prime-age	Incidence of Employer-provided OJT Employer-provided OJT	Log Wage, OLS+SC Log Wage, OLS	-0.057 0.066***
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	0.000
Albert et al. (2010)	1995–2001	ECHP	Prime-Age	Incidence of OJT	Log Wage, FE	-0.046

Study	Period	Dataset	Sample	Training Definition	Outcome, Estimation Method	Effect Sizes
United Kingdom						
Booth (1991)	1987	BSAS	Full-time Employed Men Full-time Employed Women	Incidence of Formal OJT	Log Wage, OLS	0.106*** 0.166***
Booth (1993)	1980–1987	BHPS	College-Graduate Men College-Graduate Women	Days of OJT during first year of job	Log Wage, FE	-0.0004 0.002***
Blundell et al. (1996)	1981–1991	NCDS	Young Men Young Women	Incidence of Employer-Provided OJT	Log Wage, Quasi-Difference	0.041* 0.003
Arulampalam and Booth (2001)	1981–1991	NCDS	Young Men	Incidence of OJT	Wage Growth, OLS+SC	0.342**
Booth et al. (2003)	1991–1996	BHPS	Full-time Employed Men	Incidence of OJT	Log Wage, OLS Log Wage, FE	0.033** 0.010
Vignoles et al. (2004)	1991–2000	NCDS	Middle-Aged Men	Incidence of OJT	Wage Growth, OLS Wage Growth, IV	0.048*** 0.050
Booth and Bryan (2005)	1998–2000	BHPS	Full-time Employed Prime-Age	Incidence of OJT employer-sponsored)	Log Wage, FE	0.024***
Addison and Belfield (2007)†	1998–2004	WERS	Firms w/ 10+ workers	Incidence of firm-provided OJT	Log wage (in 1998), OLS Log wage (in 2004), OLS	0.0312*** -0.0023
Blanchflower and Lynch (2007)	1965–1981	NCDS	Young Non-College Graduates	Incidence of OJT	Log Wage, OLS	0.024***
Gielen (2007)	1998–2003	BHPS	Prime-age Men	Incidence of OJT	Log Wage, FE Separation from curr. firm, FE	0.035** -0.196**
Metcalfe and Sloane (2007)†	2004	WERS	Firms w/ 5+ workers	Incidence of Training	Log Wage, RE	0.065***
Albert et al. (2010)	1995–2001	ECHP	Prime-Age	Incidence of OJT	Log Wage, FE	0.001
Almeida-Santos et al. (2010)	1998–2005	BHPS	Prime-Age White-Collar Prime-Age Blue-Collar Prime-Age White-Collar Prime-Age Blue-Collar	Incidence of Firm-Sponsored OJT Days of Firm-Sponsored OJT	Log Wage, FE	0.0074*** 0.0010 0.0343*** 0.0173
Melero (2010)	1991–2002	BHPS	Men Women Men Women	Incidence of OJT	Promotion w/ curr. firm, Logit FE Quits to better job, Logit FE	1.037 1.198** 1.074 1.250**
Italy						
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	0.038***
Albert et al. (2010)	1995–2001	ECHP	Prime-Age	Incidence of OJT	Log Wage, FE	0.095
Spain						
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	0.017
Albert et al. (2010)	1995–2001	ECHP	Prime-Age	Incidence of OJT	Log Wage, FE	-0.020
Netherlands						
Groot (1995)	1952–1983	Brabant Survey	Middle Aged	Hours of OJT	Log Wage, SEM	-0.014
Leuven and Oosterbeek (2004)	1999	EPIO	Prime-age	Incidence of OJT	Log Wage, OLS Log Wage, IV	0.030 -0.063
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	-0.030
Leuven and Oosterbeek (2008)	2001	EPIO	Prime-age	Incidence of OJT	Log Wage, OLS Log Wage, OLS+SC	0.106*** 0.009
De Grip and Sauermann (2012)	2008–2009	Personnel records of Call Center	Call Agents	Incidence of OJT	Log Prod.[4], OLS+Random Assign.	0.0882***
Fouarge et al. (2013)	1994–2006	LSP	Prime-age Low Educated	Incidence of OJT	Log Wage, FE	0.026***
Belgium						
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	0.026*
Greece						
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	0.060*
Sweden						
Regnér (2002)[5]	1968–1991	LNU	Prime-age	Incidence of OJT Training (1mo–1yr duration) Incidence of OJT Training (1yr or more)	Log Wage, OLS	0.056* 0.146*
Evertsson (2004)	1994–1998	SSLC	Prime-age Men	Incidence of Employer -provided OJT (General)	Log Wage, OLS	0.08***
Portugal						
Budria and Telhado Pereira (2004)	1998-2000	PLFS	Men Women	Incidence of OJT	Log Wage, OLS+SC	0.2183*** 0.3649***
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	0.105***
Albert et al. (2010)	1995–2001	ECHP	Prime-Age	Incidence of OJT	Log Wage, FE	-0.030
Austria						
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	0.004
Switzerland						
Gerfin (2004)	1998-2000	SLFS	Full-time Employed Men	Incidence of Employer-Sponsored OJT Incidence of OJT	Log Wage, PSM+DiD	0.018* 0.023**
Denmark						
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	0.0260***
Finland						
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	0.038**
Norway						
Schøne (2004)	1989–1993	NSOE	Private-Sector Employees	Incidence of employer-provided OJT (in 1989)	Log Wage, OLS Log Wage, FE	0.044*** 0.006
Ireland						
Bassanini et al. (2007)	1995-2001	ECHP	Prime-age	Incidence of OJT	Log Wage, FE	0.005
McGuinness et al. (2014)	2003	NCPP SEAEW	Employees	Incidence of employer-provided OJT	Log Net Wage, OLS Log Net Wage, PSM	0.036 0.038*

Study	Period	Dataset	Sample	Training Definition	Outcome, Estimation Method	Effect Sizes
Russia						
Travkin and Sharunina (2016)	2003-2011	RLMS-HSE	Prime-age	Incidence of Employer-Sponsored OJT	Log Wage, OLS Log Wage, Double DiD	0.162*** 0.083***
Japan						
Kawaguchi (2006)	1994–1998	JPSC	Middle-aged Women	Incidence of Employer-provided OJT	Log Wage, FD	0.046**
Thailand						
Almeida and Faria (2014)†	2004	WBES	Manufacturing Workers	Incidence of Employer-provided OJT	Log Wage, PSM	0.045**
Malaysia						
Almeida and Faria (2014)†	2002	WBES	Manufacturing and Service Workers	Incidence of Employer-provided OJT	Log Wage, PSM	0.077***
New Zealand						
Gibson (2003)	1996	ETS	Prime-age White Prime-age Non-White	Incidence of Employer-provided OJT	Log Wage, MLE+SC	0.083*** 0.150***
India						
Adhvaryu et al. (2018)	2013–2015	Personnel records of Garment Company	Female Workers	Incidence of OJT	Log Efficiency[6], OLS+Random Assign. Log Wage, OLS+Random Assign.	0.108** 0.00492*
Peru						
Arriagada (1990)	1985	PLSS	Prime-age Men	Incidence of OJT	Log Wage, OLS	0.132*
Tanzania						
Beyer de (1990)†	1980	Enterprise Survey	Skilled Manual Workers	Incidence of OJT	Log Wage, OLS	0.155*
Biesebroeck (2007)†	1991–1993	WBRED	Manufacturing Workers	Incidence of OJT	Log Wage, Firm FE	-0.007
Kahyarara and Teal (2008)†	1997–2000	MES	Manufacturing Workers	Incidence of OJT (current) Incidence of OJT (past)	Log Wage, OLS Log Wage, FE Log Wage, OLS Log Wage, FE	0.218** 0.003 0.059 -0.011
Kenya						
Beyer de (1990)†	1980	Enterprise Survey	Skilled Manual Workers	Incidence of OJT	Log Wage, OLS	0.039
Biesebroeck (2007)†	1991–1993	WBRED	Manufacturing Workers	Incidence of OJT	Log Wage, Firm FE	0.099**
Rosholm et al. (2007)†	1995	RPED	Manufacturing Workers	Incidence of OJT (formal) Incidence of OJT (informal)	Log Wage, OLS Log Wage, PSM Log Wage, OLS	0.367*** 0.542*** -0.009
Zambia						
Rosholm et al. (2007)†	1995	RPED	Manufacturing Workers	Incidence of OJT (formal) Incidence of OJT (informal)	Log Wage, OLS Log Wage, PSM Log Wage, OLS Log Wage, PSM	0.114 0.076 0.224*** 0.279*
Zimbabwe						
Biesebroeck (2007)†	1991–1993	WBRED	Manufacturing Workers	Incidence of OJT	Log Wage, Firm FE	0.175***
Cross–Country (Pooled Data from Various Countries)						
Salas-Velasco (2009)	1999	CHEERS	Recent European College Graduates	Incidence of OJT	Log Wage, OLS+SC	0.5236
Fialho et al. (2019)	2012–2015	PIAAC	OECD	Incidence of OJT (Formal) Incidence of OJT (Non-formal)	Log Wage, OLS+SC	-0.045*** 0.112***

Panel B: Studies that focus on firm-level outcomes

Study	Period	Dataset	Sample	Training Definition	Outcome, Estimation Method	Effect Sizes
United States						
Holzer et al. (1993)	1986–1990	MJOB	Manufacturing firms w/ -500 workers	Prop. workers trained	Δ Log Scrap Rate, OLS	-0.068**
Bartel (1994)	1983–1986	Columbia Survey	Compustat II	Prop. workers trained	Growth in net sales per worker, OLS	0.39***
Loewenstein and Spletzer (1999)	1982	EOPP	New hires	Weeks of OJT	Log average wage, OLS Log average wage, FE	0.028** 0.038***
Black and Lynch (2001)	1994	EQW-NES	Private firms	Log no. workers in training	Log sales per worker, OLS Log sales per worker, FE+SC Log sales per worker, GMM	-0.002 0.001 0.004
Bassi et al. (2002)	1996–1998	ASTD	Publicly-traded firms	Prop. workers trained	Median wage growth, Q4-Q1	5p.p
Frazis and Loewenstein (2005)[2]	1982	EOPP	New hires	Incidence of OJT (formal)	Δ log average wage, OLS Δ log average wage, OLS (quadratic spec.)	0.014 0.031
Sepúlveda (2009)[7]	1988–1998	NLSY, Industry data(BLS)	2-digit manufacturing	Prop. workers trained	Δ production, OLS+SC Δ production, OLS Δ production, FE	0.676* 0.095*** 0.067*
Canada						
Dostie (2013)†	1999–2006	WES	Representative Firm Sample	Prop. workers trained (classroom) Prop. workers trained (on-the-job)	Log VA per worker, FE Log VA per worker, FE+GMM Log VA per worker, FE Log VA per worker, FE+GMM	0.048*** 0.072 0.017 0.022
Germany						
Zwick (2006)	1997–2001	IAB	Firms with workers covered by SS	Prop. workers trained	Log VA per worker, OLS Log VA per worker, FE+SC	0.145* 0.761**
France						
Ballot et al. (2002)[8]	1981–1993	HRA, Financials, SE	Large manufacturing firms	Stock of Training[9]	Log VA per worker, OLS Log VA per worker, GMM Log average wage, OLS Log average wage, GMM	0.116 0.260 0.124 0.154

Study	Period	Dataset	Sample	Training Definition	Outcome, Estimation Method	Effect Sizes
United Kingdom						
Schonewille (2001)	1988–1996	UK LFS, OECD ISB	1-digit industries	Hours of OTJ Training	Gross VA, OLS	-0.01
Dearden et al. (2006)[10]	1983–1996	UK LFS, ACP	3-digit industries	Prop. workers trained	Log VA per worker, RE	0.70***
					Log VA per worker, FE	0.696***
					Log VA per worker, GMM	0.602***
					Log average wage, RE	0.344***
					Log average wage, FE	0.365**
					Log average wage, GMM	0.351***
Addison and Belfield (2007)†	1998–2004	WERS	Firms w/ 10+ workers	Prop. trained workers	Subjective Productivity (in 1998), OLS	0.0169***
					Subjective Productivity (in 2004), OLS	0.0162**
Metcalfe and Sloane (2007)†	2004	WERS	Firms w/ 5+ workers	Prop. workers trained	Subjective Prod., Oprobit	0.29*
Italy						
Conti (2005)[11]	1996–1999	ILFS, AIDA	Industry-by-region	Prop. workers trained	Log VA per worker, FE	0.349**
					Log VA per worker, FD+FE	0.376**
					Log average wage, FE	0.215
					Log average wage, FD+FE	0.287*
Colombo and Stanca (2014)	2002-2004	Excelsior, AIDA	Firms w/ 50+ workers	Prop. workers trained	Log VA per worker, OLS	0.08***
					Log VA per worker, FE	0.045**
					Log VA per worker, GMM	0.072***
Belgium						
Konings and Vanormelingen (2015)	1997-2006	Belfirst	All Belgian Firms	Prop. workers trained	Log VA per worker, OLS	0.46***
					Log VA per worker, ACF	0.243***
Sweden						
Ballot et al. (2002)[7]	1981–1993	IUI	Large manufacturing firms	Stock of Training[9]	Log VA per worker, OLS	0.067
					Log VA per worker, GMM	0.058
					Log average wage, OLS	0.092
					Log average wage, GMM	0.063
Portugal						
Almeida and Carneiro (2009)	1995–1999	Census of Large Firms	Firms w/ 100+ workers	Stock of Training[39]	Log VA per worker, FD	0.0013***
					Log VA per worker, FD+GMM	0.0006*
Finland						
Maliranta and Asplund (2007)†	1999–2001	FLEED, FSS, CVT2	Private firms	Days of training per worker	Log VA per worker, FE	-0.001
Jones et al. (2012)	2000–2004	OP Group records	Co-op Banks	Training Expenditures	Log VA per worker, FE	0.0701***
					Log VA per worker, GMM	0.0535
Russia						
Tan et al. (2007)	2005	Russia LME	Large and Medium Firms	Prop. workers trained (formal)	Log VA, OLS	0.225***
					Log average wage, OLS	0.016***
Cross–Country (Pooled Data from Various Countries)						
Gonzalez-Velosa et al. (2016)	2006–2010	WBES	Latin America	Prop. workers trained	Solow Residual, FE+SC	-0.001
				Incidence of OJT (Non-formal)	Solow Residual, FE+SC	0.112***

Notes: *Estimation Method Acronyms:* OLS denotes ordinary least squares. In Panel A, FE denotes individual fixed effects, while in Panel B it denotes firm fixed effects. RE denotes random effects model. SC denotes various forms of selection correction (Heckit, Hurdle, etc). SEM is simultaneous equation method. FD is first difference. HT corresponds to the Hausmann-Taylor method. GMM is generalized method of moments. PSM is propensity score matching. IV is instrumental variable approach. Q1, Q2, Q3 and Q4 denote quartiles one through four respectively. DiD denotes difference in differences. ACF denotes the estimation procedure proposed by Ackerberg et al. (2006). *Effects Considered:* We focus on studies examining the effects of job-related on-the-job training. The estimates mostly refer to the effects of completed training, though in some studies this may not be the case. In studies where the distinction between training with current and previous employers is made, we consider training completed with the current employer generally, though evidence shows similar effects for both types of training (Loewenstein and Spletzer (1997), Parent (1999)).

† This study uses linked employer-employee data.

[1] Although the training programs considered here were partially sponsored by the government, we still include them as part of firm-provided training since it targeted workers that were already employees at the firm, the content was tailored to the specific needs and requests of the firm, and some of the costs were assumed by the firm.

[2] This study reports the marginal effects of training at the median. No standard errors are provided.

[3] Although the training/endorsement program here is not provided by the employer directly, we still consider it because Emily's List is a PAC that is very involved with the needs of the Democratic party.

[4] Productivity is measured in this study as the inverse of the average time taken to handle a customer call.

[5] The comparison group here is individuals with training of less than 1 month in duration.

[6] Productivity is measured here as the pieces produced divided by the target quantity of pieces per unit time.

[7] This study uses industry-level data in conjunction with the NLSY.

[8] This study reports the marginal effects of training at the median. No standard errors are provided.

[9] The stock of training is measured via an equation linking the stock of human capital in the firm with the number of employees, the hours of training per employee and the depreciation rate of human capital.

[10] This study uses firm-level data in conjunction with a labor force survey: the UK LFS.

[11] This study uses firm-level data in conjunction with a labor force survey: the ILFS.

Figure A.1: Returns to Training and Development

Notes: This graph summarizes the returns to training found by studies summarized in Panel A Table A.1. The estimates correspond to papers that use worker-level data, and examine the impact of the incidence of on-the-job training on log wages. We exclude outlier estimates larger than 0.25, and exclude data prior to 1980. In order to contemplate information from overlapping countries and time periods, we construct a 5-year average per country of the training returns found by different studies. The definition of "incidence of training" may vary across studies, and as such these results should be interpreted with caution.

Appendix A References

Ackerberg, D., Caves, K., and Frazer, G. (2006). Structural Identification of Production Functions.

Addison, J. T. and Belfield, C. R. (2007). Unions, Training and Firm Performance. Journal of Labor Market Research, 40(4):361–381.

Adhvaryu, A., Kala, N., and Nyshadham, A. (2018). The skills to Pay the Bills: Returns to On-the-Job Soft Skills Training. NBER Woking Paper Series.

Albert, C., García-Serrano, C., and Hernanz, V. (2010). On-the-job training in Europe: Determinants and wage returns. International Labour Review, 149(3):315–341.

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., and Vitali, A. (2020). Tackling Youth Unemployment: Evidence From a Labor Market Experiment in Uganda. Econometrica, 88(6):2369–2414.

Almeida, R. and Carneiro, P. (2009). The return to firm investments in human capital. Labour Economics, 16(1):97–106.

Almeida, R. K. and Faria, M. (2014). The wage returns to on-the-job training: Evidence from matched employer-employee data. IZA Journal of Labor and Development, 3(19).

Almeida-Santos, F., Chzhen, Y., and Mumford, K. (2010). Employee training and wage dispersion: White-and blue-collar workers in Britain. Research in Labor Economics, 30(4821):35–60.

Arriagada, A. M. (1990). Labor market outcomes of non-formal training for male and female workers in Peru. Economics of Education Review, 9(4):331–342.

Arulampalam, W. and Booth, A. L. (2001). Learning and earning: Do multiple training events pay? A decade of evidence from a cohort of young British men. Economica, 68(271):379–400.

Attanasio, O., Kugler, A., and Meghir, C. (2011). Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial. American Economic Journal: Applied Economics, 3(3):188–220.

Ballot, G., Fakhfakh, F., and Taymaz, E. (2002). Who benefits from training and R&D: The firm or the workers? A study on panels of French and Swedish firms. British Journal of Industrial Relations, 44(June):1–30.

Bartel, A. P. (1994). Productivity Gains from the Implementation of Employee Training Programs. Industrial Relations: A Journal of Economy and Society, 33(4):411–425.

Bartel, A. P. (1995). Training, Wage Growth, and Job Performance: Evidence from a Company Database. Journal of Labor Economics, 13(3):401–425.

Bassanini, A., Booth, A., Brunello, G., De Paola, M., and Leuven, E. (2007). Workplace Training in Europe. In Brunello, G., Garibaldi, P., and Wasmer, E., editors, Education and Training in Europe. Oxford Scholarship Online.

Bassi, L. J., Ludwig, J., Mcmurrer, D. P., and Buren, M. V. (2002). Profiting from Learning: Firm-Level Effects of Training Investments and Market Implications. Singapore Management Review, 24(3):61–77.

Beyer de, J. (1990). The incidence and impact on earnings of formal training provided by enterprises in Kenya and Tanzania. Economics of Education Review, 9(4):321–330.

Biesebroeck, J. V. (2007). Wage and Productivity Premiums in Sub-Saharan Africa. NBER Working Paper Series.

Bishop, J. H. (1994). The Impact of Previous Training on Producitivity and Wages. In Lynch, L. M., editor, Training and the Private Sector, pages 161–200. University of Chicago Press, Chicago.

Black, S. E. and Lynch, L. M. (2001). How to compete: The impact of workplace practices and information technology on productivity. Review of Economics and Statistics, 83(3):434–445.

Blanchflower, D. G. and Lynch, L. M. (2007). Training at Work : A Comparison of U.S. and British Youths. In Lynch, L. M., editor, Training and the Private Sector: International Comparisons, pages 233–260. University of Chicago Press, Chicago.

Blundell, R., Dearden, L., and Meghir, C. (1996). The Determinants and Effects of Work-Related Training in Britain. Technical report, The Institute for Fiscal Studies.

Booth, A. L. (1991). Job-Related Formal Training: Who Receives It and What Is It Worth? Oxford Bulletin of Economics and Statistics, 53(3):281–294.

Booth, A. L. (1993). Private sector training and graduate earnings. Review of Economics & Statistics, 75(1):164–170.

Booth, A. L. and Bryan, M. L. (2005). Testing some predictions of human capital theory: New training evidence from Britain. Review of Economics and Statistics, 87(2):391–394.

Booth, A. L., Francesconi, M., and Zoega, G. (2003). Unions, work-related training, and wages: Evidence for British men. Industrial and Labor Relations Review, 57(1):68–91.

Brown, J. N. (1989). Why Do Wages Increase with Tenure? On-the-Job Training and Life-Cycle Wage Growth Observed within Firms. The American Economic Review, 79(5):971–991.

Budria, S. and Telhado Pereira, P. (2004). On the Returns to Training in Portugal. IZA Discussion Paper Series.

Caicedo, S., Espinosa, M., and Seibold, A. (2021). Unwilling to Train? Firm Responses to the Colombian Apprenticeship Regulation. Econometrica, Forthcomin.

Card, D., Ibarrarán, P., Regalia, F., Rosas-Shady, D., and Soares, Y. (2011). The labor market impacts of youth training in the dominican republic. Journal of Labor Economics, 29(2):267–300.

Card, D., Kluve, J., and Weber, A. (2018). What works? A meta analysis of recent active labor market program evaluations. Journal of the European Economic Association, 16(3):894–931.

Colombo, E. and Stanca, L. (2014). The impact of training on productivity: Evidence from a panel of italian firms. International Journal of Manpower, 35(8):1140–1158.

Conti, G. (2005). Training, productivity and wages in Italy. Labour Economics, 12(4):557–576.

De Grip, A. and Sauermann, J. (2012). The Effects of Training on Own and Co-worker Productivity: Evidence from a Field Experiment. Economic Journal, 122(560):376–399.

Dearden, L., Reed, H., and Van Reenen, J. (2006). The Impact of Training on Productivity and Wages: Evidence from British Panel Data. Oxford Bulletin of Economics and Statistics, 68(4):397–421.

Dostie, B. (2013). Estimating the returns to firm-sponsored on-the-job and classroom training. Journal of Human Capital, 7(2):161–189.

Duncan, G. J. and Hoffman, S. (1979). On-The-Job Training and Earnings Differences by Race and Sex. The Review of Economics and Statistics, 61(4):594.

Evertsson, M. (2004). Formal On-the-Job Training: A Gender-Typed Experience and Wage-Related Advantage? European Sociological Review, 20(1):79–94.

Fialho, P., Quintini, G., and Vandeweyer, M. (2019). Returns to different forms of job related training: Factoring in informal learning. OECD Social, Employment and Migration Working Papers.

Fouarge, D., Schils, T., and de Grip, A. (2013). Why do low-educated workers invest less in further training? Applied Economics, 45(18):2587–2601.

Frazis, H. and Loewenstein, M. A. (2005). Reexamining the returns to training: Functional form, magnitude, and interpretation. Journal of Human Resources, 40(2):453–476.

Gerfin, M. (2004). Work-Related Training and Wages: An Empirical Analysis for Male Workers in Switzerland. IZA Working Paper Series.

Gibson, J. (2003). Do Lower Expected Wage Benefits Explain Ethnic Gaps in Job-related Training? Evidence from New Zealand. Motu Working Paper Series.

Gielen, A. C. (2007). Performance Pay, Training and Labor Mobility. IZA Discussion Paper Series.

Gonzalez-Velosa, C., Rosas, D., and Flores, R. (2016). Firm innovation and productivity in Latin America and the Caribbean: The engine of economic development. In Grazzi, M. and Pietrobelli, C., editors, Firm Innovation and Productivity in Latin America and the Caribbean: The Engine of Economic Development, pages 137–166. Inter-American Development Bank, Washington D.C.

Görlitz, K. (2011). Continuous training and wages: An empirical analysis using a comparison-group approach. Economics of Education Review, 30(4):691–701.

Goux, D. and Maurin, E. (2000). Returns to firm-provided training: Evidence from French worker-firm matched data. Labour Economics, 7(1):1–19.

Gronau, R. (1988). Sex-Related Wage Differentials and Women's Interrupted Labor Careers-the Chicken or the Egg. Journal of Labor Economics, 6(3):277–301.

Groot, W. (1995). The wage effects of investments in enterprise-related training. Empirical Economics, 20(1):133–147.

Hamil-Luker, J. (2005). Women's wages: Cohort differences in returns to education and training over time. Social Science Quarterly, 86(SPEC. ISS.):1261–1278.

Hannagan, R. J., Pimlott, J. P., and Littvay, L. (2010). Does an EMILY's list endorsement predict electoral success, or Does EMILY pick the winners? PS - Political Science and Politics, 43(3):503–508.

Hansson, B. (2008). Job-Related Training and Benefits for Individuals: A Review of Evidence and Explanations. OECD Education Working Papers.

Heckman, J. J., Lalonde, R. J., and Smith, J. A. (1999). The economics and econometrics of active labor market programs. Handbook of Labor Economics, 3(1):1865–2097.

Hill, E. T. (1995). Labor Market Effects of Women's Post-School-Age Training. ILR Review, 49(1):138–149.

Hirshleifer, S., McKenzie, D., Almeida, R., and Ridao-Cano, C. (2016). The Impact of Vocational Training for the Unemployed: Experimental Evidence from Turkey. Economic Journal, 126(597):2115–2146.

Holzer, H. J., Block, R. N., Cheatham, M., and Knott, J. H. (1993). Are Training Subsidies for Firms Effective? The Michigan Experience. Industrial and Labor Relations Review, 46(4):625.

Jones, D. C., Kalmi, P., and Kauhanen, A. (2012). The effects of general and firm-specific training on wages and performance: Evidence from banking. Oxford Economic Papers, 64(1):151–175.

Kahyarara, G. and Teal, F. (2008). The Returns to Vocational Training and Academic Education: Evidence from Tanzania. World Development, 36(11):2223–2242.

Kawaguchi, D. (2006). The incidence and effect of job training among Japanese women. Industrial Relations, 45(3):469–477.

Kluve, J. (2010). The effectiveness of European active labor market programs. Labour Economics, 17(6):904–918.

Konings, J. and Vanormelingen, S. (2015). The Impact of Training on Productivity and Wages: Firm-Level Evidence. The Review of Economics and Statistics, 97(May):485–497.

Krueger, A. and Rouse, C. (1998). The effect of workplace education on earnings, turnover, and job performance. Journal of Labor Economics, 16(1):61–94.

Kuckulenz, A. and Zwick, T. (2003). The impact of training on earnings-differences between participant group and training forms. ZEW Discussion Papers.

Lechner, M. (1999). The Effects of Enterprise-Related Training in East Germany on Individual Employment and Earnings. Annales d'Économie et de Statistique, (55/56):97–128.

Leuven, E. and Oosterbeek, H. (2004). Evaluating the effect of tax deductions on training. Journal of Labor Economics, 22(2):461–488.

Leuven, E. and Oosterbeek, H. (2008). An Alternative Approach to Estimate the Wage Returns to Private-Sector Training. Journal of Applied Econometrics, 23(4):423–434.

Lillard, L. A. and Tan, H. W. (1986). Private Sector Training Who Gets It and What Are Its Effects.

Technical report.
Loewenstein, M. A. and Spletzer, J. R. (1997). Delayed formal on-the-job training. Industrial and Labor Relations Review, 51(1):82–99.
Loewenstein, M. A. and Spletzer, J. R. (1999). General and specific training evidence and implications. Journal of Human Resources, 34(4):731–733.
Lynch, L. M. (1992). Private-Sector Training and the Earnings of Young Workers. The American Economic Review, 82(1):299–312.
Maliranta, M. and Asplund, R. (2007). Training and Hiring Strategies to Improve Firm Performance. ETLA Discussion Papers.
Marcotte, D. E. (2000). Continuing education, job training, and the growth of earnings inequality. Industrial and Labor Relations Review, 53(4):602–623.
Mazza, E. (2015). Wage Effects of Workplace Training in Germany. Technical report.
McGuinness, S., O'Connel, P. J., and Kelly, E. (2014). The impact of training programme type and duration on the employment chances of the unemployed in Ireland. Economic and Social Review, 45(3):425–450.
McKenzie, D. (2017). How effective are active labor market policies in developing countries? A critical review of recent evidence. World Bank Research Observer, 32(2):127–154.
Melero, E. (2010). Training and promotion: Allocation of skills or incentives? Industrial Relations, 49(4):640–667.
Metcalfe, R. and Sloane, P. J. (2007). Human Capital Spillovers and Economic Performance in the Workplace in 2004: Some British Evidence. IZA Discussion Paper Series.
Mincer, J. (1988). Job training, wage growth and labor turnover. Technical report.
Parent, D. (1999). Wages and mobility: The impact of employer-provided training. Journal of Labor Economics, 17(2):298–317.
Parent, D. (2003). Employer-supported training in Canada and its impact on mobility and wages. Empirical Economics, 28(3):431–459.
Pischke, J.-S. (2001). Continuous Training in Germany. Journal of Population Economics, 14:523–548.
Regnér, H. (2002). The effects of on-the-job training on wages in Sweden. International Journal of Manpower, 23(4).
Rosholm, M., Nielsen, H. S., and Dabalen, A. (2007). Evaluation of training in African enterprises. Journal of Development Economics, 84(1):310–329.
Salas-Velasco, M. (2009). Beyond lectures and tutorials: Formal on-the-job training received by young European university graduates. Research in Economics, 63(3):200–211.
Schøne, P. (2004). Firm-financed training: Firm-specific or general skills? Empirical Economics, 29(4):885–900.
Schonewille, M. (2001). Does training generally work? International Journal of Manpower, 22(1/2):158–173.
Sepúlveda, F. (2009). Training and productivity: Evidence for US manufacturing industries. Oxford Economic Papers, 62(3):504–528.
Tamm, M. (2018). Training and changes in job Tasks. Economics of Education Review, 67(11787):137–147.
Tan, H., Savchenko, Y., Gimpelson, V., Kapeliushnikov, R., and Lukiyanova, A. (2007). Skills Shortages and Training in Russian Enterprises. IZA Discussion Paper Series.
Travkin, P. and Sharunina, A. (2016). The returns to training in Russia: A difference-in-differences analysis. International Journal of Training and Development, 20(4):262–279.
Veum, J. R. (1995). Sources of Training and Their Impact on Wages. Industrial and Labor Relations Review, 48(4):812.
Vignoles, A., Galindo-Rueda, F., and Feinstein, L. (2004). The labour market impact of adult education and training: A cohort analysis. Scottish Journal of Political Economy, 51(2):266–280.
What Works - Centre for Local Economic Growth (2016). Evidence Review 1: Employment Training. Technical report, London School of Economics.
Yoshida, Y. and Smith, M. R. (2005). Training and the earnings of immigrant males: Evidence from the Canadian workplace and employee survey. Social Science Quarterly, 86(Special Issue: Income, Poverty and Opportunity):1218–1241.
Zwick, T. (2006). The impact of training intensity on establishment productivity. Industrial Relations, 45(1):26–46.

B Data Sources

For the main analysis, we rely on enterprise- and worker-level surveys in developed and developing economies. For developing countries, we use the World Bank Enterprise Survey (WB-ES). For developed countries, on the other hand, we rely on data from the European Union. Specifically, we use the European Union Labor Force Survey (EU-LFS), the Adult Education Survey (EU-AES), and the Continuing Vocational Training (EU-CVT) enterprise survey.

We also rely on worker qualification data from Germany for further empirical validation. The data were collected by the BIBB (Bundesinstitut fr Berufsbildung, Bonn), a federal agency devoted to vocational education. The data include measures of on-the-job skill acquisition, formal education, and occupational skill requirements. The data collection strategy was designed to cover a representative sample of 20,000 to 35,000 members of the German labor force. The survey is repeated every 6 years for a different set of subjects, yielding a repeated cross-sectional structure spanning from 1979 to 2018. We consider three waves of the data (1986, 1992, and 1999) since the questions regarding skill acquisition remain stable across these waves.[40]

We also rely on a secondary US data source for calibration and model validation. To calibrate the model to the benchmark US economy, 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. Finally, we rely on data from Donovan et al. (2020) for measures of job destruction and job-to-job transitions.

C Details on Training Definition

We first carefully define training and its characteristics to ensure consistency across different data sources. We separate the sources of workers' skill acquisition into four categories that allow for data comparability and also for meaningful economic interpretations through the lens of the model. We present a summary of these learning sources in Table C.1. The categories rank from the most structured and planned type of learning (schooling) to the least structured (informal learning). For expositional purposes, and because 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). We present detailed information on each source of learning below.

[40]Using the data, we build three measures of workers' skill acquisition: training, self-learning, and learning-by-doing. The definition of training encompasses the two types of training introduced before: formal training, which captures workers who learned how to do their job via company or external training; and informal training, which captures workers who learned how to do their job via instruction from colleagues or superiors. The definition of self-learning captures workers who learned how to do their job via self-directed learning outside the workplace and thus broadly matches the definition of informal learning. Learning-by-doing captures workers who learned how to do their job through the work itself. The questions on skill acquisition also consider school (middle and high school, vocational school and university), but we leave this out to focus on on-the-job human capital accumulation.

Table C.1: Human Capital Sources and Examples

How Structured		**Firm Sponsored**	Non-Firm Sponsored
	1. Schooling	MBA paid by firm	MBA self-financed
	2. Formal Training	Firm-organized presentation	Pre-employment training (license/certification)
	3. Informal Training	Guided o-t-j Training Job Rotation	-
	4. Informal Learning	-	Self-learning (e.g. Reading Journals)

Notes: Our definition of Schooling reflects "Formal Education and Training," according to ISCED 2011, while both Formal Training and Informal Training are categories within "Non-Formal Education and Training" from ISCED 2011. The definitions of Formal and Informal Training follow the definitions in the WB-ES and EU-CVT. The different sources of human capital are ordered along two key features: (1) how structured the learning is and (2) the financing source for the educational investment (firm vs. worker sponsored).

Schooling: According to ISCED 2011, formal education and training is defined as "education that is institutionalized, intentional and planned through public organizations and recognized private bodies and in their totality constitutes 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: According to ISCED 2011, 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 EU-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 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 currently engaging in an apprenticeship or training contract should not be considered as taking part in CVT. Random learning and initial vocational training are explicitly excluded and measured separately. These courses are typically separated from the active workplace (for example, they take place in a classroom or at a training institution), show a high degree of organization by a trainer, and the content is designed for a group of learners (e.g., a curriculum exists).

As defined by the EU-CVT survey, “Other forms of CVT” that we refer to as *informal training* are geared toward 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 are typically tailored 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: “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.”

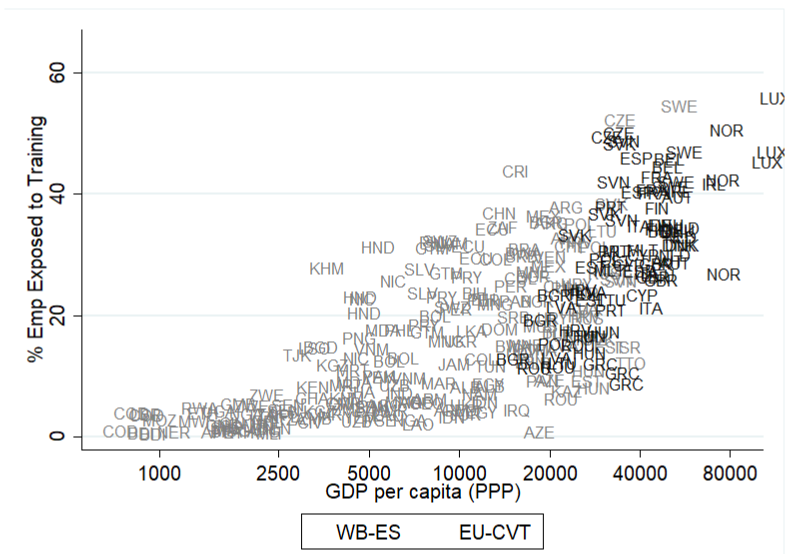
5. Participation in conferences, workshops, trade fairs, and lectures: “Participation (instruction received) in conferences, workshops, trade fairs and lectures are considered as training actions only when they are planned in advance and if the primary intention of the person employed for participating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

Initial vocational training is defined as a formal education program (or a component thereof) where working time alternates between periods of education and training at the workplace and in educational institutions or training centers. This program consists of learning activities for workers who are new at their jobs.

Informal learning: Informal learning 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. Some examples include: learning by using printed material; learning by using computers; learning through media (television, radio or videos); learning through guided tours as museums; and learning by visiting learning centers as libraries."

D Additional Empirical Results

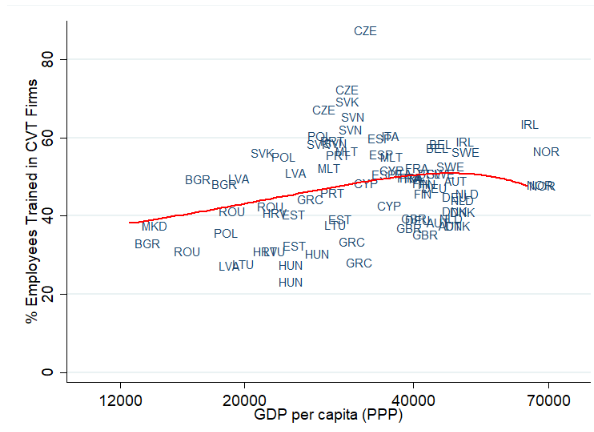
Figure D.1: Share of Formally Trained Employment (Full Sample)



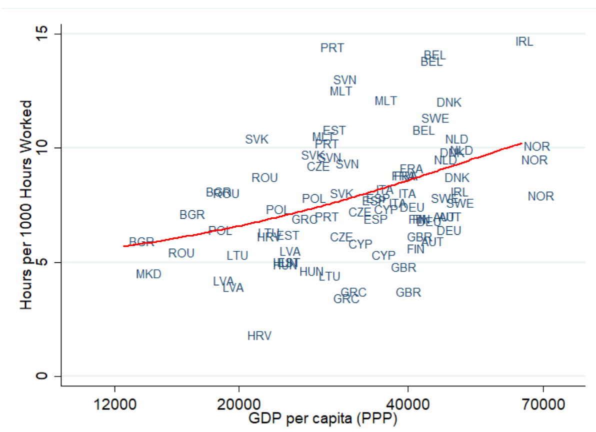
Notes: The share of formally trained employment follows from adjusting the share of workers who receive training from firms by the share of self-employment. Data on the share of employees trained within firms come from the WB-ES for all developing economies and from the EU-CVT for European economies. 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 waves with data from 2005–2017 for which we have firm weights, and we plot all countries with no restrictions.

Figure D.2: Margins of On-the-Job Training within the Wage Sector

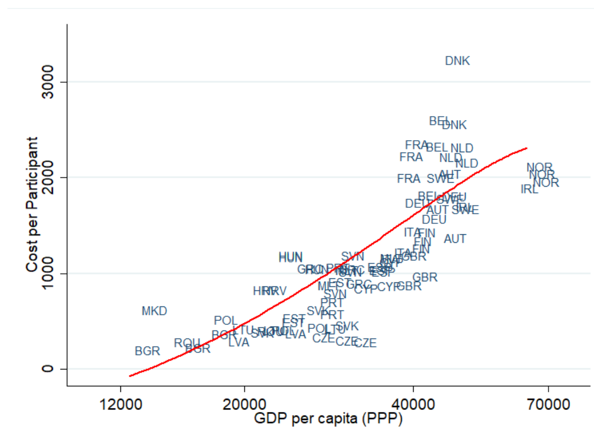
(a) % Firms Offering Training



(b) % of Participants within Firm



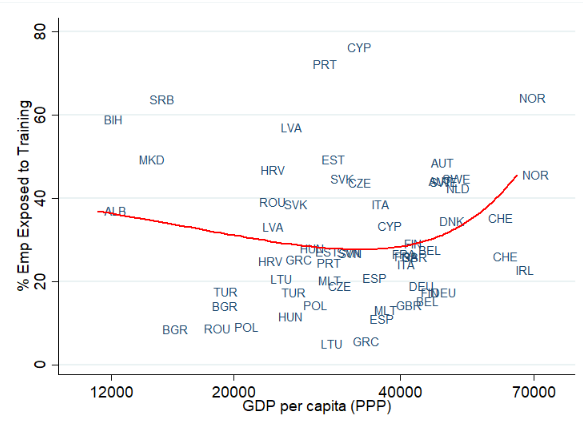
(c) Hours per 1000 Hrs Worked



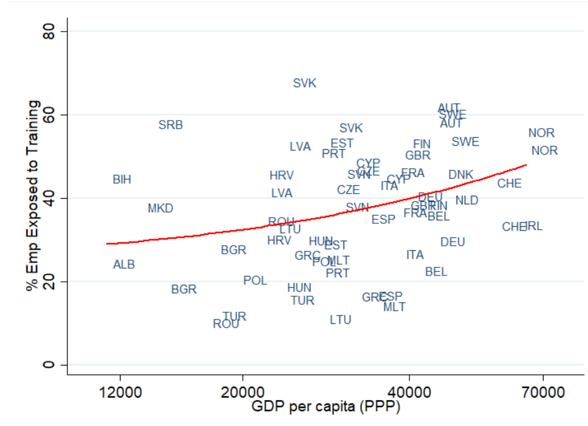
(d) Cost per Participant

Notes: This figure shows all margins of training. Panel (a) shows the share of firms offering training, which was 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. Panel (c) shows the training hours per 1000 hours worked by all employees (those who did and who did not participate in the training) in the firms offering training. Panel (d) shows the total cost of training per participant, which includes both direct and indirect training costs (wages of trainers and wages lost by not working during training). Data come from the EU-CVT.

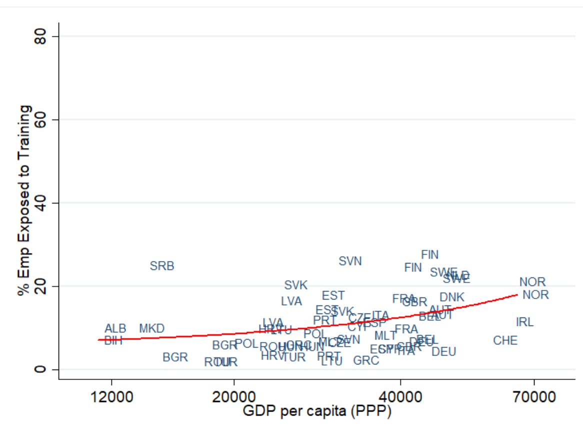
Figure D.3: Informal Learning



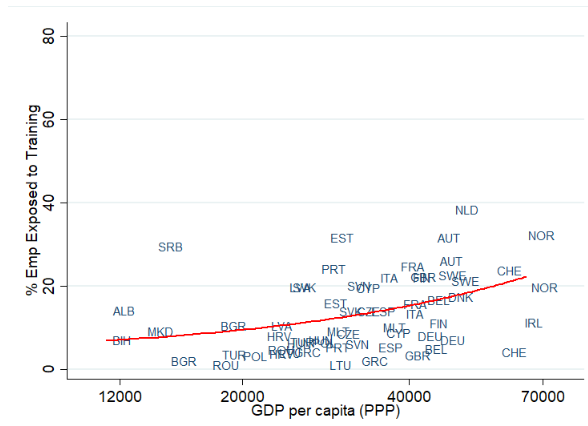
(a) Through Family and Friends



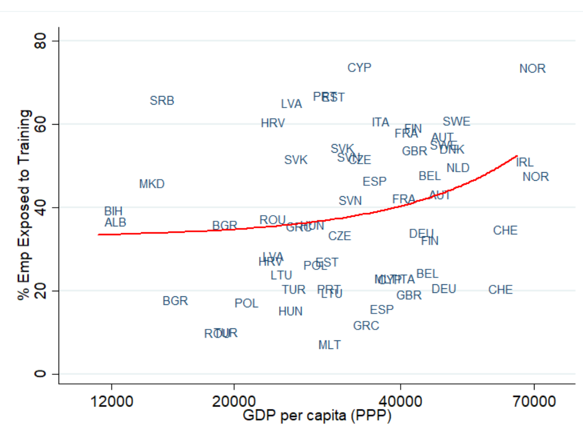
(b) Using Printed Material



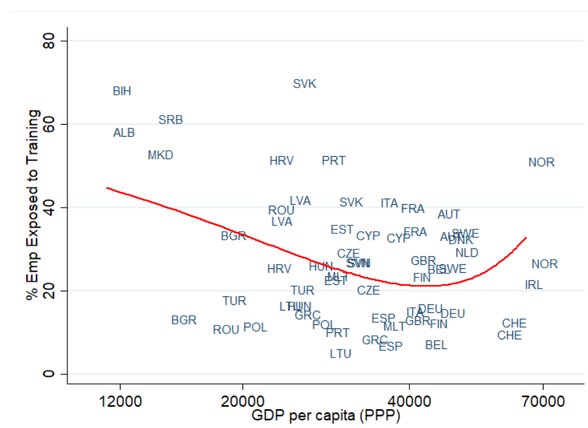
(c) Attending Learning Centers



(d) Tours on Relevant Sites



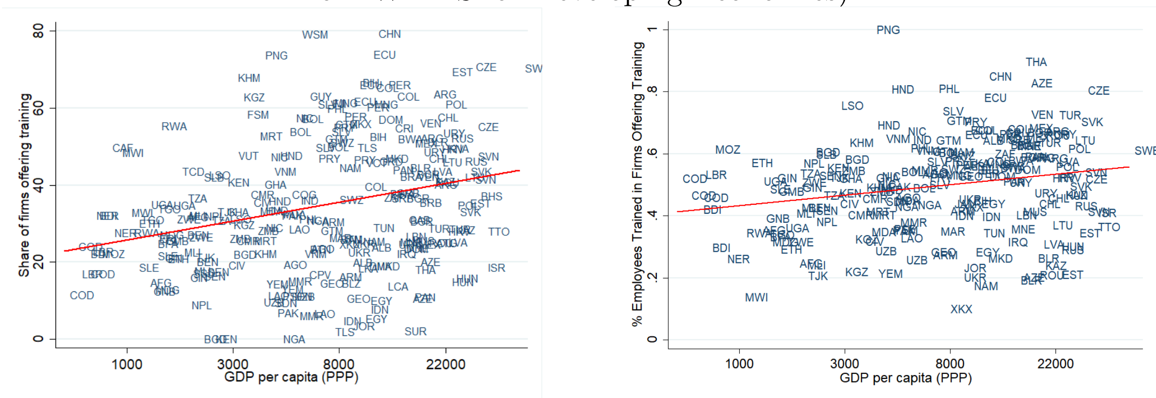
(e) Using Computer



(f) Using Media

Notes: These figures show participation rates in different types of informal learning: learning through family and friends (Panel (a)), using printed material (Panel (b)), attending learning centers (Panel (c)), tours of learning sites (Panel (d)), using computers (Panel (e)), and using media (Panel (f)). Data come from the EU-AES.

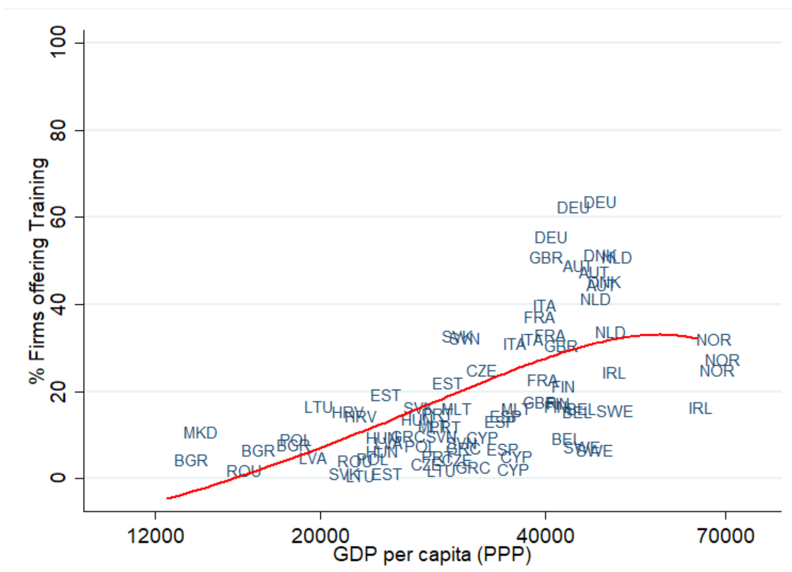
Figure D.4: Margins of On-the-Job Training within the Wage Sector (Using Data from WB-ES for Developing Economies)



(a) % Firms Offering Training **(b)** % Participants within Firm

Notes: Panel (a) shows the share of firms offering training in the WB-ES, and Panel (b) shows the share of participants per firm in the manufacturing and service sectors weighted by the WB-ES-provided weights. For the WB-ES, we use the standardized waves with data from 2005–2017, and we plot all countries with no restrictions.

Figure D.5: Share of Firms Offering IVT



Notes: This figure shows the share of firms offering initial vocational training (IVT). IVT includes coaching workers on job-specific skills for a new job or teaching workers general knowledge about the firm as they enter a new job. Data come from the EU-CVT.

Table D.1: Training Purpose by country (EU-CVT)

	Europe	Germany	France	United Kingdom	Italy	Spain
General IT	19.7	29.7	15.3	21.0	11.6	22.0
Professional IT	13.5	11.0	15.1	10.3	17.4	10.1
Management	27.2	24.1	24.7	46.1	19.6	17.8
Team working	25.9	25.0	15.7	45.3	23.4	20.0
Customer handling	31.4	38.4	22.0	47.2	23.4	23.2
Problem solving	21.7	27.1	13.5	32.1	22.7	12.5
Office administration	20.0	24.1	25.6	20.1	16.2	18.0
Foreign language	11.5	10.6	13.3	3.2	11.3	18.2
Technical	65.7	64.6	70.9	79.9	57.6	55.6
Oral or written communication	8.6	6.2	7.2	19.9	4.4	3.5
Numeracy and/or literacy	4.2	2.1	2.3	16.3	2.2	1.8
Other skills and competences	14.7	22.9	16.7	0.5	23.3	19.8

Notes: This table shows the share of enterprises providing CVT courses by type of skill targeted and country. A particular course may cover more than one category. The first column shows calculations for all European countries present in the survey. In the remaining columns, we present the results for the top 5 populated countries in Europe. We use the weighting factors provided by the survey.

E Training Decomposition

We exploit the rich EU-AES data containing information on training participation, occupation, industry, education level of workers, and firm size in order to account for how much of the cross-country training differences documented above stems from differences in observables, particularly differences in the share of workers in "high training" bins. To this end, we decompose the trend in on-the-job training using shift-share accounting analysis. For example, we split occupations into bins according to the mean training levels in the EU-AES, as shown in Figure E.1, and then assume that all economies have the richest economies' share of workers in each occupation category.[41] By comparing the slope of the original measure with the one that assumes the same occupation structure for all economies, we can calculate the share of the increase in training driven by the larger share of workers in high-training occupations.

We show the main results in Table E.1. The results suggest that richer countries have larger shares of workers in industries and occupations that require more training. Moreover, richer economies also have more highly educated workers and more employment in larger firms, which taken together partially explain the positive correlation between training and development. The most important individual factor is the occupation het-

[41] We split occupations and industries into three bins each (high, medium and low training) according to Panels (a) and (b). The occupation bins we consider encompass Professionals, Technicians, and Clerical Support (high training); Managers, Service and Sales, and Craft Workers (medium training); and Machine Operators, Elementary Occupations, and Skilled Agriculture (low training). The industry bins we consider encompass from Health and Social Work to Real State (high training), from Other Services to Manufacturing (medium training), and from Wholesale and Retail to Households as Employers (low training). We split education and firm size into three and four bins each, respectively, following the definitions in Panels (c) and (d).

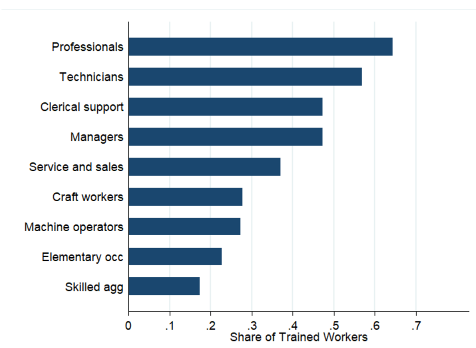
erogeneity, accounting for 11% of the slope of training with respect to per-capita GDP. All these factors jointly explain 21% of the increase in training, which implies these observables drive only a small fraction of our results.

Table E.1: Accounting for On-the-Job Training

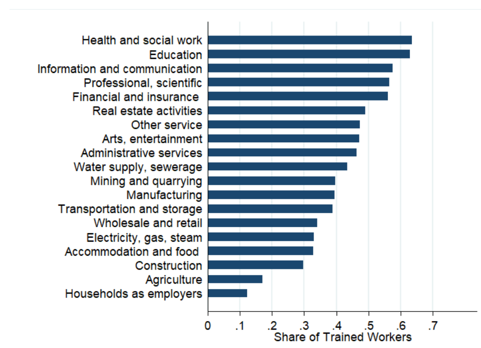
	Industry	Occupation	Education	Firm Size	All
Share Accounted For (%)	8	11	5	4	21

Notes: This table reports the share of the slope of training with respect to per-capita GDP that is accounted for by the industry, occupation, education, or firm size structure. For each of these categories, we split the sample into bins according to the mean training levels discussed in Appendix E. Then, we calculate the share of workers in each bin for the top 10% richest economies in the sample and use those shares as weights for all the economies. Finally, we regress each measure on per-capita GDP and calculate the "share accounted for" as 1-$\hat{\beta}/\beta$, where $\hat{\beta}$ is the coefficient of the regression using the richest economies' structure and β is the coefficient of the regression with each particular economy's structure.

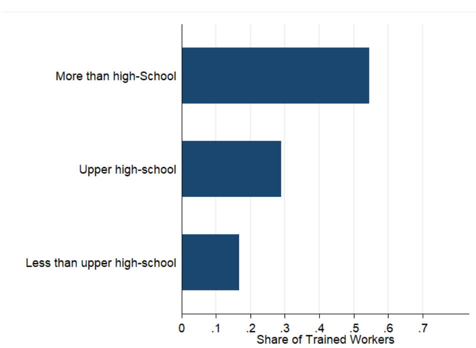
Figure E.1: Training Levels by Industry, Occupation, Education, and Firm Size



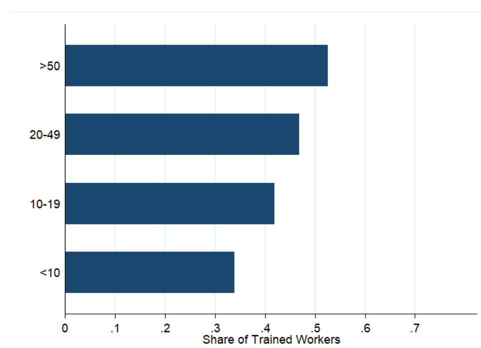
(a) By Occupation



(b) By Industry



(c) By Education Level



(d) By Firm Size

Notes: These figures show employees' participation rates in training by occupation, industry, education, and firm size categories. Data come from the EU-AES.

F Analytical Model: Additional Results

F.1 Proof of Proposition 1

We derive the optimal training levels chosen by the worker and the firm by considering their respective optimization problems.

A young worker in firm z will choose her optimal training level to maximize the present value of income:

$$\max_s \underbrace{w(z)}_{\text{current wage}} - \underbrace{\mu_W c_s}_{\text{worker's per-unit training costs}} \times \underbrace{s}_{\text{training level}} + \frac{1}{1+\rho}\Bigg\{\underbrace{\delta \int w' dF(w')}_{\text{U back to a firm}} \times \underbrace{(1+\zeta s^{\gamma_s})}_{\text{next-period human capital}}$$

$$+ (1-\delta)\Bigg[\underbrace{\left(1 - \eta \int p(w')dF(w')\right) w(z)}_{\text{if stay in current firm}} + \underbrace{\eta \int p(w')w'dF(w')}_{\text{if move to new firm}} - \underbrace{\eta \int c_p^{\gamma_p} \frac{p(w')^{1+\gamma_p}}{1+\gamma_p} dF(w')}_{\text{cost of breaking contract}}\Bigg] \times \underbrace{(1+\zeta s^{\gamma_s})}_{\text{next-period human capital}}\Bigg\}. \tag{F.1}$$

If we define $MR_W(z)$ as in Proposition 1, the first-order condition with respect to the level of training s yields the worker's desired level of training:

$$s_W(z) = \left(\frac{\zeta \gamma_s MR_W(z)}{(1+\rho)\mu_W c_s}\right)^{\frac{1}{1-\gamma_s}}. \tag{F.2}$$

The firm will choose the optimal training level for this worker to maximize its net profits:

$$\max_s \underbrace{\tilde{r}(z) - w(z)}_{\text{current profits}} - \underbrace{\mu_F c_s}_{\text{firm's per-unit training costs}} \times \underbrace{s}_{\text{training level}} +$$

$$\frac{1}{1+\rho} \times \underbrace{(1-\delta)\left(1 - \eta \int p(w')dF(w')\right)(\tilde{r}(z) - w(z))}_{\text{future profits, from workers who stay}} \times \underbrace{(1+\zeta s^{\gamma_s})}_{\text{next-period human capital}}. \tag{F.3}$$

If we define $MR_F(z)$ as in Proposition 1, the first-order condition with respect to training s yields the firm's desired level of training:

$$s_F(z) = \left(\frac{\zeta \gamma_s MR_F(z)}{(1+\rho)\mu_F c_s}\right)^{\frac{1}{1-\gamma_s}}. \tag{F.4}$$

F.2 Proof of Proposition 2

We examine how changes to the cost of breaking contracts c_p, the separation rate δ, and the capital rent R affect the firm's optimal training level $s_F(z)$ derived in Proposition 2.

For changes in the cost of breaking contracts c_p, first notice that for each outside offer w', the leaving probability $p(w(z)), w')$ decreases with c_p. Thus, given the offer distribution $F(w)$, the probability of a worker leaving after on-the-job search, $\eta \int p(w')dF(w')$, also decreases with c_p. This implies:

$$\frac{\partial s_F(z)}{\partial c_p} = -\frac{1}{1-\gamma_s}\left(\frac{\zeta\gamma_s}{(1+\rho)\mu_F c_s}\right)^{\frac{1}{1-\gamma_s}} MR_F(z)^{\frac{\gamma_s}{1-\gamma_s}}(1-\delta)(\tilde{r}(z)-w(z))\frac{\partial \eta \int p(w')dF(w')}{\partial c_p} > 0. \tag{F.5}$$

For changes in the separation rate δ, we notice:

$$\frac{\partial s_F(z)}{\partial \delta} = -\frac{1}{1-\gamma_s}\left(\frac{\zeta\gamma_s}{(1+\rho)\mu_F c_s}\right)^{\frac{1}{1-\gamma_s}} MR_F(z)^{\frac{\gamma_s}{1-\gamma_s}}\left(1-\eta\int p(w')dF(w')\right)(\tilde{r}(z)-w(z)) < 0. \tag{F.6}$$

Finally, for changes in capital rent R, we notice:

$$\frac{\partial s_F(z)}{\partial R} = \frac{1}{1-\gamma_s}\left(\frac{\zeta\gamma_s}{(1+\rho)\mu_F c_s}\right)^{\frac{1}{1-\gamma_s}} MR_F(z)^{\frac{\gamma_s}{1-\gamma_s}}(1-\delta)\left(1-\eta\int p(w')dF(w')\right)\frac{\partial \tilde{r}(z)}{\partial R} < 0. \tag{F.7}$$

which follows because $\frac{\partial \tilde{r}(z)}{\partial R} = -\mu A_M z(\mu A_M z)^{\frac{\mu}{1-\mu}} R^{-\frac{1}{1-\mu}} < 0$.

F.3 Firm's Problem

A firm's value function can be written as

$$J(z, l^O_{-1}, w_{-1}, F_{-1}) = \max_{\{w,v,s\}} \underbrace{l^O_{-1}(1-\delta)\left(1-\eta\int p(w_{-1}, w')dF(w')\right)(\tilde{r}(z)-w_{-1})}_{\text{profits from remaining workers}} - \underbrace{\frac{c_v v^{1+\gamma_v}}{1+\gamma_v}}_{\text{vacancy costs}}$$

$$+ \underbrace{\frac{v}{\theta}\frac{\tilde{r}(z)-w-\mu_F c_s s}{1+\eta(1-\delta)+\delta}}_{\text{profits from hiring young workers}} + \underbrace{\frac{v}{\theta}\frac{\eta(1-\delta)\int p(w', w)dF_{-1}(w')\bar{l}(w)+\delta\bar{l}}{1+\eta(1-\delta)+\delta}(\tilde{r}(z)-w)}_{\text{profits from hiring old workers}} + \frac{J(z, l^O, w, F)}{1+\rho}$$

$$s.t. \quad l^O = \frac{v}{\theta}\frac{1}{1+\eta(1-\delta)+\delta}(1+\zeta s^{\gamma_s}), \quad F = \Gamma(F_{-1}), \quad w \geq b\bar{w},$$

where we use the subscript -1 to denote variables that are determined in the last period. l^O_{-1} is the total supply of efficiency units by old workers before exogenous separations, and $F_{-1}(w)$ is the wage distribution of job offers during the last period. The first term on the right-hand side represents the net profits generated by all the workers who remain in the firm from the previous period. The second term represents the total vacancy costs. The third term represents the profits from hiring young workers net of training costs. The fourth term represents the profits from poaching old workers who are willing to move to the current firm.[42]

[42]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 $\bar{l} = 1 + \int \zeta s_{-1}(z)^{\gamma_s}dF_{-1}(w_{-1}(z))$.

F.4 Solving Workers' Sectoral Choices

If there is a non-zero measure of workers in both sectors, workers must be indifferent in terms of expected utility between going to the self-employment sector and the wage sector in the beginning of the first period:

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

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.

F.5 Equilibrium

We now define the model's general equilibrium in the steady state.

Definition 1 *The general equilibrium for this economy is given by*
(1) workers' decisions over consumption $\{c_i^Y, c_i^O\}$, savings k_i^Y, and sector to work in;
(2) workers' decisions over optimal contract-breaking probability $\{p(w, w')\}$;
(3) firms' decisions over physical capital, wages, and vacancy posting $\{k(z), w(z), v(z)\}$;
(4) the joint decision of human capital accumulation $\{s_F(z), s_W(z)\}$;
(5) aggregate prices $\{P_T, P\}$; and
(6) perceived law of motion for firms' wage distribution $\Gamma(F_{-1}(w))$,
such that:
(i) given prices, wage distribution, and human capital accumulation, (1) solves the households' consumption, saving, and sectoral choices problems, and (2) solves the workers' contract-breaking problem;
(ii) given prices, workers' leaving rates and wage distribution, human capital accumulation, market tightness, and the perceived law of motion, (3) solves the firm's problem.
(iii) given prices, wage distribution, and workers' leaving rates, (4) solves the optimal training problem for firms and workers;
(iv) perceptions are correct;
(v) workers' total savings equal the total amount of capital demanded by firms; and
(vi) the wage-sector output equals the sum of consumption of the wage-sector good, capital investments, costs of breaking contracts, and vacancy and training costs, and the self-employment-sector output equals consumption of the self-employment-sector good.

G Quantitative Model: Conditions for Simulations

Wage Workers' Value. With linear utility, workers' utility is determined by the discounted income flows that are earned with current human capital and potential future human capital accumulation. For a worker of age a and human capital h in a firm with productivity z, we denote $W_M^a(h,z)$ as the worker's value. We index the value of unemployed workers by $W_M^a(h)$, and following Bagger et al. (2014), we assume that unemployment is equivalent to employment in the least productive firm: $W_M^a(h) = W_M^a(h, z_{min})$. This assumption resolves the complication of allowing for heterogeneous reservation wages for workers of different human capital levels and ages. With $\theta = \frac{V}{\tilde{U}}$ denoted as market tightness, 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.

First, note that in the last period of employees' lifetime ($a = J$), employees have no incentive to accumulate human capital. Thus, we can obtain

$$W_M^J(h,z) = w(z)h.$$

For younger employees ($a < J$), we can obtain their values by backward induction:

$$W_M^a(h,z) = \underbrace{w(z)h - \mu_W(c_s\bar{w} + \delta_s\tilde{r}(z))s^a(h,z)}_{\text{wage income net of training costs}} + \underbrace{\frac{\delta}{1+\rho}W_M^{a+1}(h')}_{\text{value if being separated exogenously in the next period}}$$

$$+ \underbrace{\frac{1-\delta}{1+\rho}\left[W_M^{a+1}(h',z) + \eta\theta q(\theta)\int p^{a+1}(h',z,z')\left(W_M^{a+1}(h',z') - W_M^{a+1}(h',z)\right) - c_p^{\gamma_p}h'\frac{(p^{a+1}(h',z,z'))^{1+\gamma_p}}{1+\gamma_p}dF(w(z'))\right]}_{\text{value if staying or transitioning from job to job in the next period}}.$$

$h' = e_M(h) = \bar{h} + (1-d)(h-\bar{h}) + \zeta(s^a(h,z))^{\gamma_s}$ is the next-period's human capital for employees, where d captures depreciation of human capital from training above workers' innate human capital ($\bar{h} = 1$), and $s^a(h,z)$ is the optimal training level as described below. To simplify our notation, we use $e_M(h)$ to represent employees' human capital evolution. $p^{a+1}(h',z,z')$ is the probability of leaving firm z conditional on an offer from a firm with productivity z', which is obtained by maximizing the benefits from leaving the firm:

$$\max_{p\in[\underline{p},1]} p \times (W_M^{a+1}(h',z') - W_M^{a+1}(h',z)) - c_p^{\gamma_p}h'\frac{p^{1+\gamma_p}}{1+\gamma_p}.$$

Self-employed Workers' Value. We denote $W_T^a(h)$ as the value of income flows for a self-employed worker of age a and human capital h. Even though the self-employment income does not rely on human capital, the dependence of $W_T^a(h)$ on human capital reflects that self-employed workers may switch to the wage sector. Note that in the last period of self-employed workers' lifetimes ($a = T$), we have:

$$W_T^J(h) = A_T P_T.$$

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

$$W_T^a(h) = \underbrace{A_T P_T}_{\text{self-employment income}} + \underbrace{\frac{1}{1+\rho}\left[\tau_c \Lambda_M^{a+1}(h') W_M^{a+1}(h') + \left(1 - \tau_c \Lambda_M^{a+1}(h')\right) W_T^{a+1}(h')\right]}_{\text{value if transitioning to wage sector or staying self-employed in the next period}},$$

where $h' = e_U(h) = \bar{h} + (1-d)(h - \bar{h})$ reflects no human capital investment for unemployed and self-employed workers. Given assumptions about sectoral preferences in Section 4.1, the probability of choosing the wage sector for a potential switcher is:

$$\Lambda_M^{a+1}(h') = \frac{e^{\tau W_M^{a+1}(h')}}{e^{\tau W_M^{a+1}(h')} + e^{\tau W_T^{a+1}(h')}},$$

which is similar to a Logit probability.

Firms' Value. Denote $F^a(h, z)$ as the firm value of a worker of age a and human capital level h after being hired. We have:

$$F^a(h, z) = \underbrace{\left(\tilde{r}(z) - w(z)\right) h - \mu_F \left(\delta_s \tilde{r}(z) + c_s \bar{w}\right) s^a(z, h)}_{\text{revenue net of wage and training costs}}$$
$$+ \underbrace{\frac{1-\delta}{1+\rho}\left(1 - \eta\theta q(\theta) \int p^{a+1}(h', z, z') dF(w(z'))\right) F^{a+1}(h', z)}_{\text{value if the worker stays in the firm in the next period}},$$

where $h' = e_M(h)$ is defined above.

Training. Firms' optimal training is determined by

$$\mu_F \left(\delta_s \tilde{r}(z) + c_s \bar{w}\right) = \zeta \gamma_s s_F^a(h, z)^{\gamma_s - 1} \frac{\partial F^a(h, z)}{\partial h'},$$

where $\frac{\partial F^a(h,z)}{\partial h'}$ is firms' return for an extra efficiency unit of human capital in the next period. And workers' optimal training is determined by

$$\mu_W \left(\delta_s \tilde{r}(z) + c_s \bar{w}\right) = \zeta \gamma_s s_W^a(h, z)^{\gamma_s - 1} \frac{\partial W_M^a(h, z)}{\partial h'},$$

where $\frac{\partial W_M^a(h,z)}{\partial h'}$ is workers' return for an extra efficiency unit of human capital in the next period. The optimal training is $s^a(h, z) = \min(s_F^a(h, z), s_W^a(h, z))$. In comparison with our analytical model, the optimal training level now depends on the present value of all future returns, adjusted for the depreciation rate of training as well as workers' separation rates (for firms).

Employment Distribution. Denote $N_M^a(h, z)$ as the measure of workers of age a and human capital h in all firms with productivity z right before job search happens.

Similarly, we denote $U_M^a(h)$ as the measure of unemployed workers right before job search happens, and after sectoral adjustments occur. Then, the size of searchers in the wage sector is the sum of the unemployed and on-the-job searchers,

$$\tilde{U} = \sum_{a=1}^{T} \left[\int U_M^a(h) dh + \eta \int \int N_M^a(h,z) dh dz \right].$$

For the entering cohort endowed with human capital $\bar{h}$, the amount of unemployed searchers is $U_M^1(\bar{h}) = \Lambda_M^1(\bar{h})$ and $U_M^1(h) = 0 \ \forall \ h > \bar{h}$, with existing employment $N_M^1(h,z) = 0 \ \forall \ h$ (thus no on-the-job searchers as the entering cohort starts by looking for jobs). Denote $N_T^a(h)$ as the number of self-employed workers, and the self-employed population for the entering cohort is $N_T^1(\bar{h}) = 1 - \Lambda_M^1(\bar{h})$ and $N_T^1(h) = 0 \ \forall \ h > \bar{h}$. The following equations characterize the evolution of these measures, accounting for human capital formation, job search, and exogenous and endogenous job separations,

$$N_M^{a+1}(h', z) = \underbrace{(1-\delta) \int_{h'=e_M(h)} \left[1 - \eta\theta q(\theta) \int p^a(h,z,z') dF(w(z')) \right] N_M^a(h,z) dh}_{\text{workers that stay in the last-period job search and are not exogenously separated this period}}$$

$$+ \underbrace{(1-\delta)\theta q(\theta) f(w(z)) w'(z) \left[\int_{h'=e_M(h)} U^a(h) dh + \eta \int_{h'=e_M(h)} \int N_M^a(h,y) p^a(h,y,z) dy dh \right]}_{\text{last-period hires that are not exogenously separated this period}};$$

$$U_M^{a+1}(h') = \underbrace{\left[1 - \tau_c(1 - \Lambda_M^{a+1}(h'))\right]}_{\text{share of U staying in wage sector}} \left(\underbrace{\int \frac{\delta}{1-\delta} N_M^{a+1}(h',z) dz}_{\text{exog separations this period}} + \underbrace{(1-\theta q(\theta)) \int_{h'=e_U(h)} U_M^a(h) dh}_{\text{last-period unemployed searchers still w/o jobs}} \right)$$

$$+ \underbrace{\tau_c \Lambda_M^{a+1}(h') \int_{h'=e_U(h)} N_T^a(h) dh}_{\text{people from self-employment to wage sector}} ;$$

$$N_T^{a+1}(h') = \underbrace{\tau_c(1 - \Lambda_M^{a+1}(h'))}_{\text{share of U to self-employment}} \left(\underbrace{\int \frac{\delta}{1-\delta} N_M^{a+1}(h',z) dz}_{\text{exog separations this period}} + \underbrace{(1-\theta q(\theta)) \int_{h'=e_U(h)} U_M^a(h) dh}_{\text{last-period unemployed searchers still w/o jobs}} \right)$$

$$+ \underbrace{\left[1 - \tau_c \Lambda_M^{a+1}(h')\right] \int_{h'=e_U(h)} N_T^a(h) dh}_{\text{people staying self-employed}}.$$

Vacancies and Wage Determination. We can then obtain the optimal condition for vacancies:

$$\underbrace{c_v \bar{w} v(z)^{\gamma_v}}_{\text{costs of posting a vacancy}} = \underbrace{\sum_{a=1}^{T} \frac{q(\theta)}{\tilde{U}} \left[\eta \int \int p^a(h,y,z) N_M^a(h,y) F^a(h,z) dy dh + \int U_M^a(h) F^a(h,z) dh \right]}_{\text{benefits from hiring on-the-job or unemployed searchers by posting a vacancy}}.$$

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

$$\sum_{a=1}^{T} \frac{q(\theta)}{\tilde{U}} \left[\eta \int \int p^a(h,y,z) N_M^a(h,y) \frac{\partial F^a(h,z)}{\partial w(z)} dy dh + \eta \int \int \frac{\partial p^a(h,y,z)}{\partial w(z)} N_M^a(h,y) F^a(h,z) dy dh \right.$$
$$\left. + \int U_M^a(h) \frac{\partial F^a(h,z)}{\partial w(z)} dh \right] = 0.$$

This differential equation can be evaluated numerically. Combined with the lowest wage $b\bar{w}$, we can iterate the wage structure $w(z)$ until convergence. When the model abstracts from human capital, the wage differential equation can be analytically written in a similar way as in Burdett and Mortensen (1998).

H Quantitative Model: Additional Results

H.1 Baseline Calibration: Targeted Moments

Table H.1: Targeted Moments in the Model vs Data

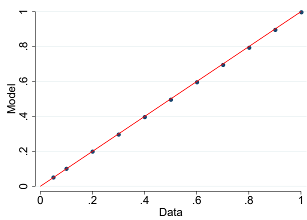
Moments	Data	Model
1. Moments: labor market		
1.1 Unemployment rate (%)	6.5	5.4
1.2 Ratio of #Vacancies to #Unemployed	0.55	0.52
1.3 Self-employment sector employment share (%)	6.0	4.0
1.4 Pareto parameter of firm size distribution	1.06	1.13
1.5 % workers remaining in same firm after one quarter	0.95	0.95
1.6 % workers remaining employed after one quarter	0.97	0.97
1.7 Workers' avg wage growth after job-to-job transition	0.13	0.13
1.8 % job-to-job transition from high to low wage firms	0.22	0.22
1.9 Relative wage-job finding rate (unemployed/self-employed)	9.0	9.0
1.10 Transition rate from unemployment to self-employment	0.04	0.04
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 in firms with 50–99 employees	1.2	1.1
2.3 Ratio of training costs to wage costs of training	0.24	0.26
2.4 Percent wage increase of 20 years' experience (%)	88	85
2.5 Percent wage increase of 40 years' experience (%)	89	92
3. Other moments		
3.2 Output per worker (normalization)	1	1

H.2 Cross-Country Calibration Results

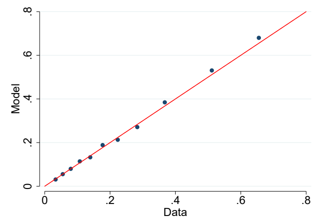
Figure H.1 shows how the model fits the targeted moments in each of our representative economies. The x-axis captures the empirical estimates of each moment while the y-axis captures the model estimates of each moment. We also plot the 45-degree line to aid the comparison between the model and data.

Figure H.2 shows how the values of the re-calibrated parameters change across the representative economies. In Panel (a) we plot the productivity level in the wage sector relative to the productivity level in the self-employment sector. As expected, although both productivity levels increase with GDP per capita, productivity in the wage sector grows faster than its counterpart in the self-employment sector. In Panels (b) and (c) we plot the parameters shaping the labor market dynamics. The job destruction rate δ decreases with income, while the cost of breaking contracts c_p increases with income. Finally, in Panel (d) we plot the physical-capital-specific productivity, which increases with income and thus generates larger capital-output ratios in rich economies.

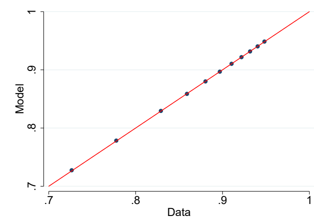
Figure H.1: Cross-country Targeted Moments



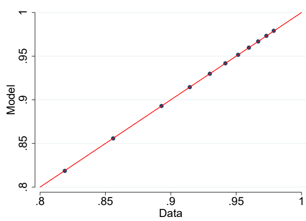
(a) GDP per capita



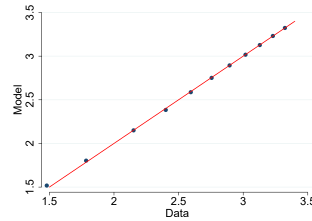
(b) Self-employment Share



(c) % Emp in Same Firm



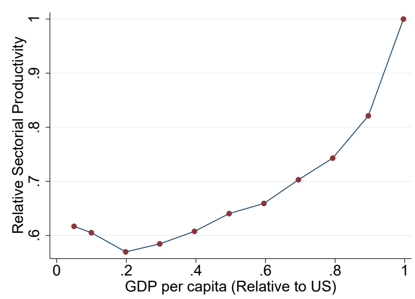
(d) % Emp still Employed



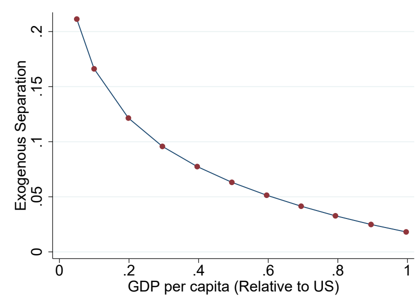
(e) Capital/Output Ratio

Notes: 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 \$2,500, \$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).

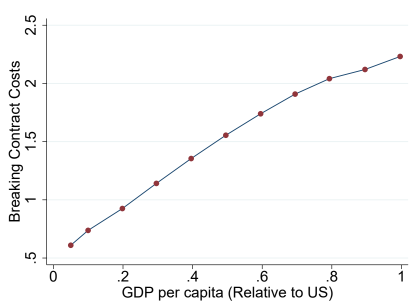
Figure H.2: Cross-country Calibrated Parameters



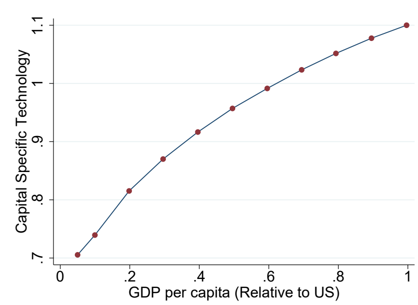
(a) Relative Prod A_M/A_T



(b) Exogenous Separation



(c) Break Contract Costs



(d) Capital Productivity

Notes: This figure shows how the values of the re-calibrated parameters change across our representative economies. We consider 10 representative economies at income levels of \$2,500, \$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).

I Quantitative Model: Extensions

In this section, we present extensions to our quantitative model and discuss how the results change in each case. In particular, in each subsection below we consider the robustness of our quantitative results to: (1) incorporating learning-by-doing (LBD) for workers; (2) abstracting from contract-breaking costs; (3) allowing for endogenous layoffs; (4) allowing for different firm productivity distributions across countries; and (5) calibrating the model separately to all country-year observations for which we have training data, and targeting the training intensity to discipline country-specific training returns.

Table I.1: Wage Growth and Income Accounting in Baseline and Extensions

	Slope of % wage increase at 20 yrs-exp on log income	**% wage increase at 20 yrs-exp in $50,000 Economy**	**% wage increase at 20 yrs-exp in $10,000 Economy**	**Share of world income differences captured**
	Panel A: Data and Model Moments			
1. Data	26%	89%	47%	100%
2. Benchmark	15%	85%	51%	100%
3. Model with LBD	16%	87%	50%	100%
4. Model with no contract-breaking costs	16%	87%	48%	100%
5. Model with screening of workers	15%	80%	46%	100%
6. Model with different productivity dist	16%	85%	50%	100%
7. Model with country-specific calibration	12%	90%	57%	100%
	Panel B: Decomposing Contribution of Model Channels			
<u>Benchmark</u>:				
Training	62%	80%	92%	12%
Job turnover	38%	20%	8%	-
<u>Model with LBD</u>:				
LBD	16%	23%	27%	-
Training	47%	57%	67%	9%
Job turnover	37%	20%	6%	-
<u>Model with no contract-breaking costs</u>:				
Training	59%	78%	93%	14%
Job turnover	41%	22%	7%	-
<u>Model with endogenous layoffs</u>:				
Training	60%	78%	91%	12%
Job turnover	40%	22%	9%	-
<u>Model with different productivity dist</u>:				
Training	61%	80%	93%	14%
Job turnover	39%	20%	7%	-
<u>Model with country-specific calibration</u>:				
Training	64%	83%	94%	15%
Job turnover	36%	17%	6%	-

Notes: "LBD" is short for "learning-by-doing." Panel A reports the main moments we focus on in this paper. Panel B explains the contribution of the model channels to each main moment predicted by the model.

Table I.1 summarizes and compares the main quantitative results across the baseline and alternate model specifications. Panel A compares the scope of each model specification to explain the main data moments we consider.[43] Panel B, on the other hand, presents: (1) the relative importance of firm-provided training and job turnover to explain the cross-country wage growth differences predicted by the model; and (2) the overall share of world income differences explained by firm-provided training.

I.1 Incorporating Learning-by-Doing

In this subsection, we discuss how our results change when we incorporate learning-by-doing (LBD) into our model. With this we allow for an additional mechanism of human capital acquisition for workers, with potential implications for lifecycle wage growth and cross-country income differences.

In our main model specification we abstracted from LBD for two main reasons. First, the lack of empirical evidence on cross-country differences in LBD makes it difficult to evaluate the scope of LBD to drive cross-country differences in human capital and wage growth. Second, as shown in Fact 2 of Section 2.4, firm-provided training is the main source of adult education. In particular, our data on workers' sources of skill acquisition in Germany indicate that only about 32% of workers report LBD as their primary source of on-the-job skill acquisition, while this value is about 62% in the case of training.

To study how our results change in the presence of LBD, we extend the quantitative model to allow for exogenous increases in human capital due to LBD, in addition to the endogenous increases in human capital due to firm-worker training decisions presented in the benchmark model. In particular, the evolution of employed workers' human capital becomes $h' = e_M(h) = \bar{h} + (1-d)(h-\bar{h}) + \zeta(s^a(h,z))^{\gamma_s} + \zeta_{LBD}$.

Compared with the human capital evolution in the baseline quantitative model, there is an additional exogenous human capital increment denoted by ζ_{LBD}, which is identical for all employed workers in the economy and does not require any investments. We calibrate ζ_{LBD} separately for each representative economy so that the lifetime human capital gain from LBD accounts for 32% of the overall lifetime human capital increment for workers in that economy, matching the data from Germany. We re-calibrate all other model parameters for the US and in the cross-country calibration to match the targeted data moments in Table H.1 and Figure H.1.

Column 1 of Panel A of Table I.1 suggests that both the benchmark and LBD models explain around 60% of the cross-country wage growth differences found in the data.[44] The difference between the models lies in the relative importance of firm-provided training and LBD in explaining wage growth and income differences across countries, presented in Panel B. In the benchmark model, firm-provided training explains 62% of the cross-country wage growth differences predicted by the model. This percentage drops to 47%

[43]All the models target per capita GDP, and thus perfectly match overall income differences across countries by construction.

[44]In addition, both models are able to closely match the percent increase in wages after 20 years of experience in economies with incomes above $10,000.

in the model with LBD.[45] This difference is fully accounted for by LBD, which explains 16% of cross-country wage growth differences. Taken jointly, these results suggest that firm-provided training accounts for about 36% of cross-country wage growth differences in the benchmark model, and 29% in the LBD model.

Furthermore, Column 4 in Panel B suggests that firm-provided training explains 9% of income differences across countries in the model with LBD, compared to 12% in the benchmark model. This evidence suggests that even when we include LBD, firm-provided training is still the main factor explaining wage growth and income differences across countries.

I.2 Abstracting from Contract-Breaking Costs

In our baseline model, we relied on cross-country differences in contract-breaking costs in order to match the differences in job-to-job transitions that exist across countries. These differences in contract-breaking costs also captured heterogeneity in contract quality that exists around the world and shapes the incentives for firms to provide training investments. As a robustness check, we now instead set contract-breaking costs to be zero ($c_p = 0$), and adjust on-the-job search intensities (η) to match the different probabilities of job-to-job transitions across countries.[46] We re-calibrate all model parameters for the US and in the cross-country calibration to match the targeted data moments in Table H.1 and Figure H.1, and find that workers in poorer countries have higher on-the-job search intensities consistent with the higher job-to-job mobility in these settings.

Table I.1 suggests that the quantitative results of this model extension are very similar to our baseline results. Interestingly, in this model version we find that training has a slightly larger impact on cross-country income differences than in the baseline model. This stems from the fact that the high on-the-job search intensities of employed workers in poor countries increase the total amount of job searchers in these settings relative to when contract-breaking costs had to be paid. This lowers the probability that unemployed workers meet employers and obtain job offers, which in turn increases the unemployment rate and results in lower returns to on-the-job training in poor countries.

I.3 Allowing for Endogenous Layoffs

In our baseline model, we assumed that job destruction was exogenous. As a robustness check, we now embed endogenous layoffs into our model, following Donovan et al. (2020) who document that workers with lower job tenures are more likely to be separated from jobs, particularly in developing settings. We assume that when a job match is formed between firm z and a worker of human capital h, there is a random draw of the match-specific productivity $\epsilon \sim \mathbf{P}(\epsilon)$. Thus, the job's productivity is given by $\epsilon z h$.

[45] Similarly, LBD lowers the contribution of firm-provided training to within-country wage growth from 80% to 57% in the economy with \$50,000 GDP per capita, and from 92% to 67% in the economy with \$10,000 GDP per capita.

[46] In our baseline model, on-the-job search intensity is calibrated using US data and assumed to be identical across countries.

For simplicity, we assume $\mathbf{P}(\epsilon = 1) = \frac{1}{2}$ and $\mathbf{P}(\epsilon = 0) = \frac{1}{2}$.[47] Upon the job match being formed, the worker and the firm both observe the true value of ϵ with probability λ. If the match-specific productivity remains unknown, it can be observed in each period through production with probability ξ. If a match is revealed to be low-productivity ($\epsilon = 0$) at any point, then it is endogenously destroyed by both parties.

For each representative economy, we choose λ and ξ to match employment exit rates in two job tenure bins: 0–6 months and 6–12 months, since these are informative about the information asymmetries prevalent upon hiring, and the speed of information revelation afterward. As expected, we find that information asymmetries and endogenous layoffs are more severe in poorer countries, given that employment exit rates decline faster with job tenure in these settings (Donovan et al., 2020). We re-calibrate all other model parameters for the US and in the cross-country calibration to match the targeted data moments in Table H.1 and Figure H.1.

Table I.1 suggests that the quantitative results of this model extension are very similar to our baseline results. However, in this model version the role of training in accounting for lifetime wage profiles becomes slightly less important relative to the baseline results since endogenous layoffs disproportionately increase the unemployment rates among the young, who tend to have lower job tenure but also larger rates of training when employed.

I.4 Allowing for Different Firm Productivity Distributions across Countries

In our baseline model, we assumed the firm productivity distribution to be identical across countries in order to focus our attention on the role of self-employment shares, job turnover rates, and physical capital endowments in driving training differences across countries.

As a robustness check, we now allow firm productivity distributions to vary across countries in order to match the relative absence of large firms in developing countries shown by several papers in the development literature (see Ciani et al. (2020) for a review). To do this, we adjust the shape parameter of the firm productivity distribution (κ) in our representative economies to match the slope of the standard deviation of firm-level employment with respect to income, which is 0.18 as found by Poschke (2018). In line with the findings in the literature, we find that poorer countries have a higher κ, indicating that poorer countries have fewer productive firms. We re-calibrate all other model parameters for the US and in the cross-country calibration to match the targeted data moments in Table H.1 and Figure H.1.

Table I.1 suggests that the quantitative results of this model extension are very similar to our baseline results. Nevertheless, this extension increases the scope of training to explain cross-country income differences since less productive firms have lower returns from providing training, leading to larger cross-country training gaps.

[47] This simplification is motivated by Faberman et al. (2017) who show that less than half of contacts between employers and prospective employees translate into offers in the US.

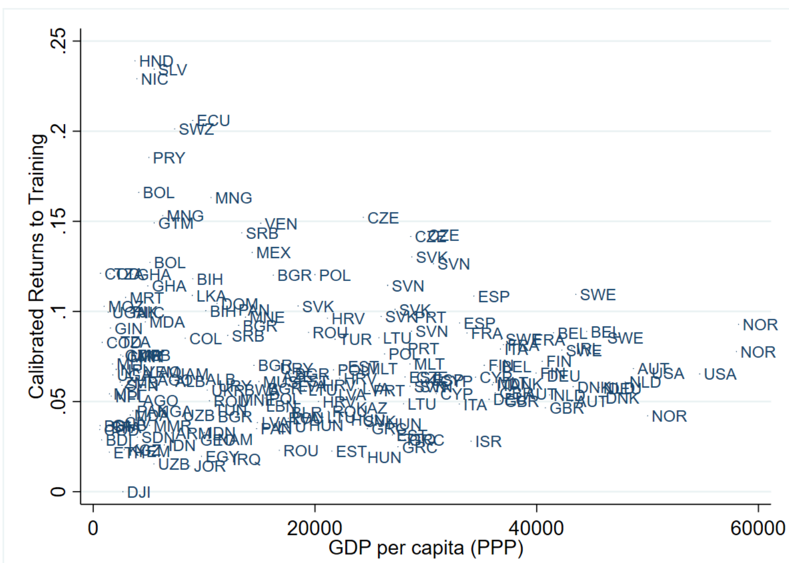
I.5 Country-Specific Calibration

In our baseline model, we introduced cross-country heterogeneity by calibrating our model to match representative economies at 10 different income levels.

As a robustness check, we now calibrate the model separately to all 203 country-year observations for which we have data on training. To do this, we target the same set of moments as in the baseline cross-country calibration for each of these country-years, and in addition, we target the training intensity to discipline country-specific training returns. Specifically, we keep the calibrated parameters from the US baseline calibration, but re-calibrate δ, c_p, A_M, A_T, χ, and ζ in each country-year to match 6 different moments: (1) real GDP per capita; (2) self-employment share; (3) the share of employed people remaining in the same firm after a quarter; (4) the share of employed people remaining employed after a quarter; (5) capital-output ratio; and (6) training intensity.

Table I.1 suggests that the quantitative results from the cross-country calibration are similar to our baseline results. Figure I.1 further plots the calibrated returns to training (ζ) across all country-years used in the cross-country calibration. This figure suggests that there are no systematic differences in the returns to training across countries,[48] which is consistent with evidence in Figure A.1 and corroborates our assumption of constant returns to training in the baseline calibration.

Figure I.1: Calibrated Returns to Training in Country-Specific Calibration



Notes: This graph plots the calibrated returns to a full quarter of training (ζ) for each country-year used in the cross-country calibration.

[48]We find that the slope of calibrated returns to training on log GDP per capita is -0.003, with a standard deviation of 0.003 (p-value=0.32).

I.5.1 Contract-Breaking Costs and Workers' Labor Market Power

Within this country-specific calibration exercise, we can also explore the relationship between our calibrated contract-breaking costs and measures of workers' labor market power across countries. In Table I.2, we regress the calibrated country-level contract-breaking costs on different measures of workers' labor market power, using data on each country's labor market institutions from Botero et al. (2004). We find that conditional on GDP per capita, the calibrated contract-breaking costs are negatively correlated with labor union power, unemployment benefits, and the generosity of minimum wages. This result provides support to our modeling of contract-breaking costs: in countries where workers have higher negotiation power or protection in the labor market, our calibrated contract-breaking costs tend to be lower.

Table I.2: Calibrated Contract-breaking Costs and Labor Market Institutions

Dependent variable	Log(Calibrated Contract-breaking Costs)			
	(1)	(2)	(3)	(4)
Log(GDP per capita)	0.405***	0.456***	0.387***	0.437***
	(0.021)	(0.028)	(0.024)	(0.031)
Labor union power	-0.031			0.008
	(0.069)			(0.067)
Unemployment benefits		-0.211***		-0.188**
		(0.075)		(0.075)
Minimum wage			-0.091**	-0.068*
			(0.037)	(0.037)
Obs	137	137	137	137
R-squared	0.841	0.854	0.848	0.858

Note: The dependent variable is the calibrated contract-breaking cost (c_p) for each country-year used in the cross-country calibration. The independent variables are drawn from Botero et al. (2004) who measure the labor market institutions for a set of countries. Labor union power is a variable normalized between 0 and 1, with a larger value representing more power for labor union. Unemployment benefits is a variable normalized between 0 and 1, with a larger value representing more generous unemployment benefits for workers. Minimum wage is a variable normalized between 0 and 1, with a larger value representing more generous minimum wages. The detailed definitions of these independent variables are available in Botero et al. (2004). Robust standard errors are in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.