



Department of Economics
Working Paper Series

2021/001

Life Cycle Wage Growth
and Internal Migration

Alberto Rivera-Padilla

February 2021

Life Cycle Wage Growth and Internal Migration^{*}

Alberto Rivera-Padilla[†]

February 2021

Abstract

I document new facts on spatial variation in life cycle wage growth within three countries along the development scale: Brazil, Mexico, and the United States. I find that rich states tend to have steeper experience-wage profiles than poor states in each country. Differences in educational attainment and industry mix can account for a large share of the covariance between income per capita and profile steepness in both developing countries, while differences in occupation types are key in every case. Evidence from internal migrants supports the notion of substantial gaps in learning environment across space. Using a general equilibrium model with human capital accumulation and internal migration, I estimate meaningful gains in labor productivity from inducing migration to places with higher lifetime wage growth and find that spatial differences in learning environment account for a considerable portion of the overall gains.

JEL Codes: E24, J61, O11, O18, R23

Keywords: human capital, internal migration, productivity

^{*}I am grateful to Todd Schoellman for valuable comments and suggestions. I also thank participants at the VMACS Junior Conference for insightful comments and discussions, especially Elisa Giannone and Tommaso Porzio. The usual disclaimer applies.

[†]California State University Fullerton, Fullerton, CA 92834-6848. Email: ariverapadilla@fullerton.edu

1 Introduction

A large literature has studied the importance of human capital accumulation in explaining income differences across countries ([Klenow and Rodríguez-Clare, 1997](#); [Bils and Klenow, 2000](#); [Caselli, 2005](#); [Lagakos et al., 2018b](#)). At the same time, economists have studied the potential aggregate gains from increasing labor mobility within countries. Most of this literature focuses on static gains from reallocating workers to places with higher income ([Restuccia et al., 2008](#); [Gollin et al., 2014](#); [Bryan and Morten, 2018](#)).

In this paper I study the interaction of life cycle wage growth and internal migration based on the following insights. First, economic activity is usually distributed across space, so locations within countries are represented by different types of industries and occupations. Second, learning opportunities are potentially different across economic activities and the possibilities to acquire knowledge from others might depend on the activities that predominate in a place. Then, locations are a possible determinant of human capital accumulation. Taking into account spatial differences in learning environment and the implications for aggregate productivity can provide additional motivation to implement policies that encourage labor mobility across space.

The first part of the paper documents facts on spatial differences in life cycle wage growth based on experience-wage profiles estimated at the state level. To do this, I focus on one rich country, the United States, and two developing countries, Brazil and Mexico. These three countries have rich microdata, are fairly large in terms of territory, and have considerable differences in regional development. The main empirical findings can be summarized as follows. First, rich states tend to have steeper profiles than poor states in all three countries, but the gaps are somewhat larger in Brazil and Mexico. Second, differences in the distribution of educational attainment and industry mix account for a considerable share of the covariance between experience-wage growth and income per capita in both developing countries, while they seem to have little importance for such covariance in the United States. In contrast, differences in the distribution of occupations account for a relatively large share of the

covariance between profile steepness and income per capita in every country.

The findings suggest that some locations within each country offer more learning opportunities. In particular, one possible implication is that workers who live in places with flatter profiles accumulate lower human capital because they do not learn as much in their economic activities. If so, a worker who has accumulated much of his experience in a place with seemingly low learning would have a lower return to experience than a worker who has accumulated most of his experience in a place with seemingly high learning, even if they are working in a similar labor market. Indeed, in the three countries considered, I find that recent migrants from low to high wage growth places have significantly flatter profiles than stayers in those same places, while the same is not true for recent migrants in the opposite direction.

The spatial variation in life cycle wage growth implies that mobility barriers might prevent individuals from accumulating more human capital by migrating to places with more learning opportunities, especially in both developing countries. I motivate this possibility by showing that migration patterns across states and regions suggest that such barriers are higher in Brazil and Mexico than in the United States. Then, I build and estimate a general equilibrium model to quantify the productivity gains from eliminating mobility barriers in each country. The model features regions that are potentially different in their endogenous learning environment, as well as workers who differ in their ability to learn in each place and face migrations shocks throughout their lifetime.

I find that removing migration costs for workers born in places with fewer learning opportunities increases labor productivity by 2.7 percent in the United States, 6.8 percent in Mexico, and 8.7 percent in Brazil. These represent meaningful productivity gains from encouraging internal migration to high learning places. Furthermore, using the structure of the model, I estimate that human capital accumulated through experience accounts for over two-thirds of those productivity gains in both developing countries. The majority of these gains are due to comparative advantage based on learning ability; however, differences in learning environment between regions are quantitatively important and account for between one-fourth and a third of the aggregate gains. Thus, the results of the model imply that there

are important dynamic gains from inducing internal migration.

This paper is related to recent literature arguing that human capital might be more important in explaining income differences across countries than it was previously thought (see e.g. [Erosa et al., 2010](#); [Manuelli and Seshadri, 2014](#)). On the empirical side, [Lagakos et al. \(2018b\)](#) show that experience-wage profiles are steeper in rich countries than in poor countries, and their findings suggest that one possible explanation is that workers in poor countries have fewer learning possibilities. I document new facts on spatial variation in life cycle wage growth for countries with different levels of economic development. These facts shed light on the importance of learning opportunities for workers and how the possibility to migrate internally affects those opportunities. More generally, this paper relates to a large literature on life cycle earnings and human capital accumulation (e.g. [Rubinstein and Weiss, 2006](#), [Huggett et al., 2011](#); [Bagger et al., 2014](#); [Engbom, 2019](#)).

In addition, this paper contributes to the literature studying aggregate gains from encouraging internal migration. [Bryan and Morten \(2018\)](#) build a static model of internal migration and find that removing all barriers to labor mobility has significant positive effects on aggregate productivity in Indonesia.¹ I focus on dynamic benefits from inducing internal migration by quantifying the aggregate productivity gains in a model that features regions with different learning possibilities and individual comparative advantage based on learning ability.² Moreover, I use the model to quantify the importance of local learning environment for the overall gains in each country.

The rest of the paper is organized as follows. Section 2 describes the data, and presents evidence on spatial variation in experience-wage profiles and internal migration in the three countries considered. Then, Section 3 presents a general equilibrium model with human capital accumulation and internal migration. Section 4 describes how the model is taken to the data,

¹In an experimental setting, [Bryan et al. \(2014\)](#) show that returns to internal migration can be large based on inducing seasonal migration in Bangladesh. Other papers focusing on labor mobility between urban and rural activities include [Vollrath \(2009\)](#) and [Lagakos et al. \(2019\)](#). Additionally, [De la Roca and Puga \(2016\)](#) and [Martellini \(2019\)](#) focus on dynamic gains from migrating to big cities in developed countries.

²Papers that develop frameworks with migration across multiple regions in dynamic settings include [Kennan and Walker \(2011\)](#) and [Caliendo et al. \(2019\)](#).

as well as the estimation results. Section 5 presents counterfactual experiments using the calibrated model for each country. Finally, Section 6 concludes.

2 Empirical Evidence

The following sections present the data, methodology and main empirical facts of the paper.

2.1 Data

I use microdata from [IPUMS-International](#) for Brazil, Mexico and the United States. Most of the original data sources are national censuses or, in some cases, household surveys. An advantage of these data is that many variables are harmonized across countries which facilitates their comparison. Moreover, census data allows me to slice the data along multiple important dimensions. The main variables that I use are individual earnings, hours worked, age, education years, occupation, industry, state of birth, current state of residence, and state of residence five years before the census or survey. The time period covered for each country is similar: approximately 1990-2010. Most of the variables are available for all this period with few exceptions.

Throughout this paper I define hourly wages as total earnings divided by hours worked in a year. Individual earnings include any source of labor income, including business income. I restrict the sample to full-time male workers with positive earnings and data on hours worked, including those who are self-employed. The decision to include the latter is that omitting them would restrict the sample considerably in several states of Brazil and Mexico. A possible concern is that part of self-employed income might not be labor but capital income; however, this should not matter for the estimated wage growth of self-employed individuals unless one assumes that the share of their income that should be imputed to capital varies significantly over the life cycle. Nevertheless, another concern is that self-employed income is not accurately measured in surveys. [Lagakos et al. \(2018b\)](#) show that despite such concerns including the self-employed does not seem to affect the comparison of wage growth over the life cycle.

I also restrict the sample to individuals who have data on educational attainment and are working in the private sector. This is motivated by the concern that non-wage compensation might be an important component for workers in the public sector. Furthermore, I follow the literature and define potential experience as $\min\{\text{age} - e_i - 6, \text{age} - 16\}$. This means that individuals start working after they finish school or when they turn 16, whichever comes last. I use age 16 as the cutoff because some individuals might start working when they are young, especially in poorer places of Brazil and Mexico. Finally, I focus on workers with 0 to 40 years of potential experience. Appendix A presents additional evidence based on alternative samples such as focusing on wage workers.

2.2 Life Cycle Wage Growth Across Space

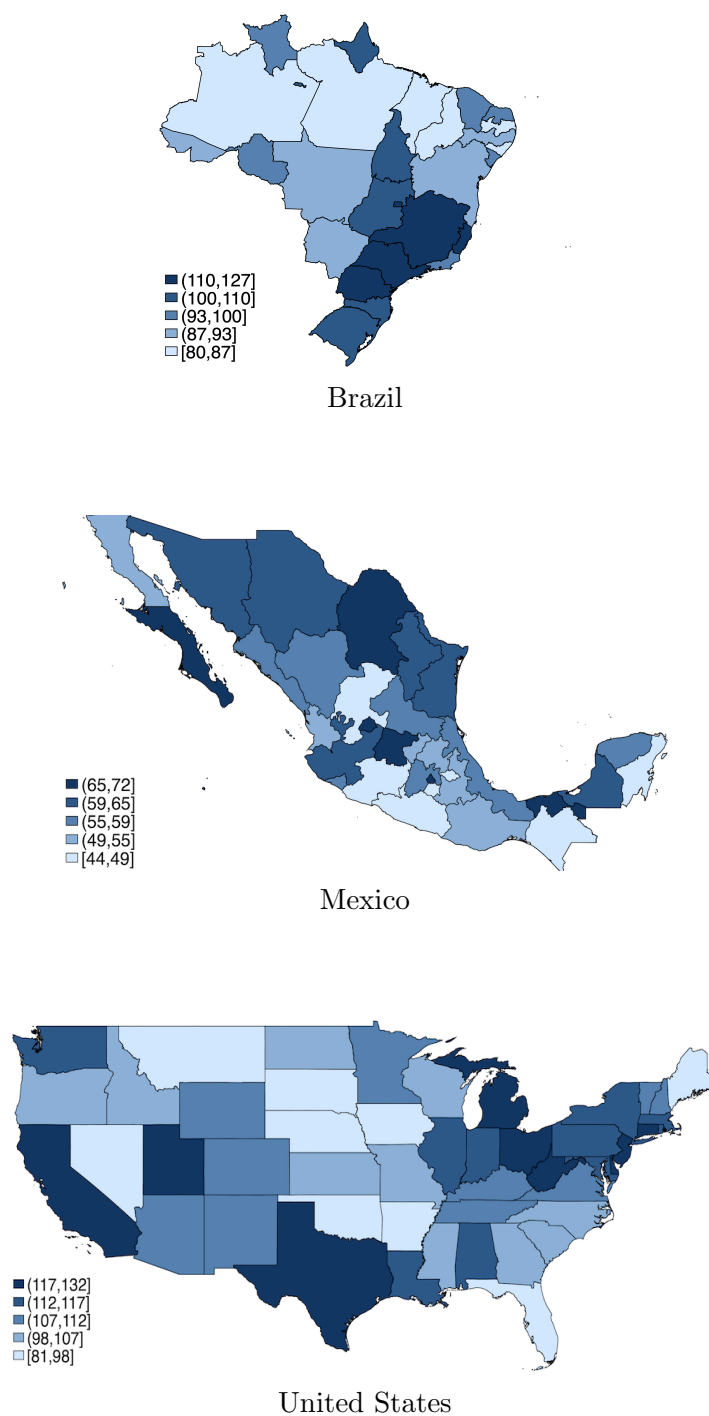
To estimate experience-wage profiles, I use the following specification for each state within a country:

$$\log(w_{ito}) = \alpha + \theta s_{ito} + \sum_{x \in X} \beta_x D_{ito}^x + \mu_t + \psi_o + \epsilon_{ito}, \quad (1)$$

where w_{ito} is the hourly wage of worker i from origin o (state of birth) in year t ; s_{ito} are his years of schooling; μ_t are year dummy variables; ψ_o are origin fixed effects; and D_{ito}^x is a dummy that takes the value of one if a worker is in experience group $x \in X = \{5 - 9, 10 - 14, \dots\}$. The omitted group are workers with less than five years of potential experience. Thus, the coefficient β_x captures the average wage of a worker in experience group x relative to workers with less than five years. This specification is a flexible version of the approach by Mincer (1974) that captures nonlinearities in experience. To ease the notation, I have omitted country and state indices in the specification.

First, to look at spatial variation in life cycle wage growth, Figure 1 presents maps with the height of the estimated profile for every state in each country. To be specific, the maps show the average wage of workers with 25 to 29 years of potential experience relative to workers with less than five years of experience. These results indicate that there is substantial dispersion in profile steepness across states within each country.

Figure 1: Life Cycle Wage Growth Across Space



Notes: These maps display the average wage gain of workers with 25 to 29 of experience relative to workers with 0 to 4 years of experience in percentage terms for each state within a country. Experience-wage profiles are estimated for every state controlling for education, state of birth, and time effects in every case. Alaska and Hawaii are omitted from the U.S. figure. I follow the geographical division of IPUMS-International for every country.

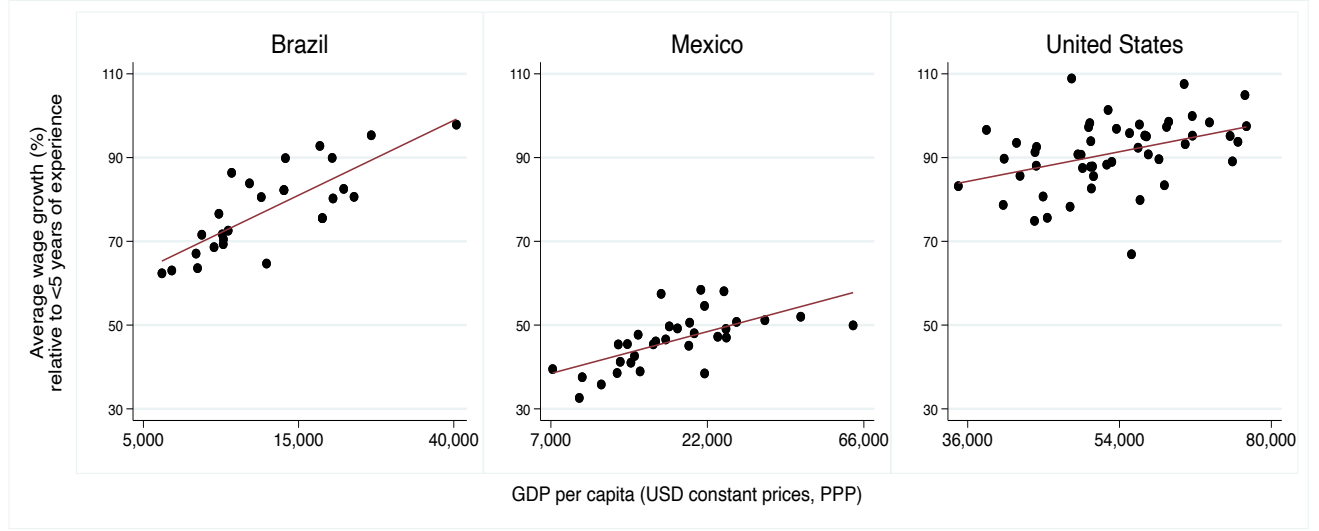
Now, to compare how big are the differences in experience-wage profiles across states in each country, I compute the average profile height for each state relative to workers with less than five years of potential experience, and rank states according to their average height. The ratio between states in the 90th and 10th percentile of the profile height distribution is equal to 1.46 in Brazil, 1.42 in Mexico, and 1.24 for the United States. Therefore, while there is significant spatial variation in profile height in all of these countries, the gaps in profile steepness are larger in both developing countries. The following paragraphs document how the spatial variation in life cycle wage growth is related to differences in income, educational attainment, and types of economic activities across states.

Figure 2 presents the relationship between income per capita and life cycle wage growth among states for each country. Income refers to Gross Domestic Product (GDP) per capita obtained from OECD data and wage growth represents the average height of the experience-wage profile estimated from equation 1. These results show that there is a positive correlation between income and profile steepness in every country. The linear correlation coefficient (using log income) is 0.79 for Brazil, 0.64 for Mexico, and 0.40 for the United States. Additionally, the elasticity of average wage growth with respect to income per capita implied by a linear fit of the data is equal to 0.22 in Brazil, 0.20 in Mexico and 0.19 in the United States. This evidence implies that experience-wage profiles tend to be steeper in rich states than in poor states within each of these countries and the relationship is somewhat stronger in both developing countries, especially in Brazil.³

Next, I use data on recent internal migrants to present additional evidence on spatial differences in life cycle growth. The fundamental idea of this exercise is similar to the insight of [Lagakos et al. \(2018a\)](#). That is, estimate returns to experience for workers who have recently migrated to places with seemingly higher learning from places with seemingly lower learning, and compare them with the returns to experience of native workers or stayers in the new place of residence. This comparison is informative about differences in human capital accumulation

³The evidence on profile height and income per capita does not include the District of Columbia due to being a considerable outlier: GDP per capita is over three times higher than the national average and the experience-wage profile is relatively flat.

Figure 2: Life Cycle Wage Growth and Income Across Space



Notes: Income per capita was obtained from OECD statistics for 2015 and wage growth refers to the average height of the experience-wage profile. These profiles are estimated at the state level in each country controlling for education, state of birth, and time effects in every case. I follow the geographical division of IPUMS-International for every country.

across space since recent migrants in high wage growth places acquired most of their working experience in low growth locations, especially those who move later in life, while native workers acquired their experience in high growth places. If the place of origin did not matter for returns to experience or knowledge acquired through work was not portable across places, then recent migrants in places with steeper profiles should have similar returns to native workers.

Moreover, if recent migrants do have flatter experience-wage profiles in high wage growth places and this reflects differences in learning opportunities across locations, then migrants from high to low growth places should have relatively steep profiles in such locations, or at least the difference in wage growth with respect to native workers should be smaller than the difference between migrants and native workers in high growth places. This is because recent migrants in low growth places coming from locations with higher wage growth would have accumulated some of their human capital in places with larger scope for learning.

To do this comparison, I first classify states with high wage growth as those whose average profile height is in the upper quartile of the profile height distribution across states in each country. Second, workers are defined as recent migrants if they moved from a state classified as low wage growth to a state with high wage growth within the last 5 years, while native

Table 1: Internal Migrants and Life Cycle Wage Growth
Average Height of Experience-Wage Profiles

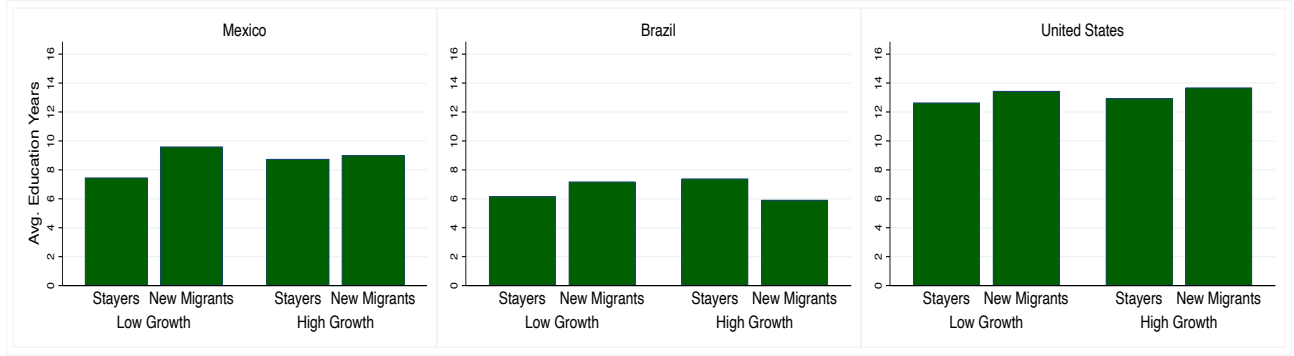
Country	Low Growth States		High Growth States	
	New Migrants	Stayers	New Migrants	Stayers
Brazil	76.1	77.0	54.5	97.9
Mexico	51.2	44.1	44.1	55.2
United States	86.3	85.3	79.3	98.4

Notes: States are classified as high growth places if they are in the upper quartile of the profile height distribution across states in each country. New migrants are workers who moved between low and high growth states within the last five years, and stayers are workers who live in a high (low) growth state and were born in a high (low) growth state. For each group of states and type of worker, I estimate experience-wage profiles and compute the average wage gain of workers with more than five years of potential experience relative to workers with less than 5 years in percentage terms.

workers or stayers are those who were born and are still living in a state classified as high wage growth. To be clear, the idea in this exercise is that a migrant with 30 years of potential experience who migrated within the last five years would have accumulated at least 25 years of experience in his place of origin. Then, I estimate equation 1 for stayers and recent migrants who live in high growth places and compute the average height of the experience-wage profile. This means the comparison is done within types of workers, so I am estimating the average wage of recent migrants with five or more years of experience relative to the average wage of recent migrants with less than five years, and the same is true for stayers. An equivalent exercise is done for workers in low growth places.

Table 1 presents the results of the estimation. These evidence is consistent with the hypothesis that places with flatter profiles offer fewer learning opportunities to workers over the life cycle. Recent migrants from low to high growth places have significantly flatter profiles than stayers in those locations. The difference is especially large for the case of Brazil, where the average height for native workers is 43 percentage points higher, but in both Mexico and the United States the gap in average height is considerable. In contrast, recent migrants from high to low growth places seem to have relatively high returns to experience relative to stayers in those places. In fact, for Mexico and the United States, recent migrants have a steeper profile than native workers in low growth places. In the case of Brazil, recent migrants in

Figure 3: Education years of new migrants and stayers

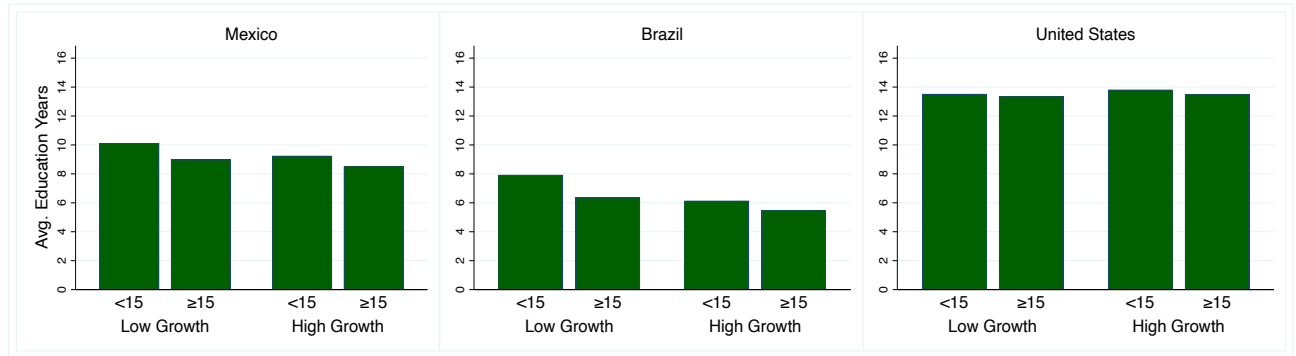


Notes: This figure presents average education years for recent migrants and stayers in each region. States are classified as high growth places if they are in the upper quartile of the profile height distribution across states in each country. New migrants are workers who moved between low and high growth states within the last five years, and stayers are workers who live in a high (low) growth state and were born in a high (low) growth state.

such locations do not have a higher average profile height than stayers but the gap is much smaller than the one observed in the opposite direction. These results also show that recent migrants from high to low wage growth places have steeper profiles than migrants in the opposite direction in every country. Moreover, stayers in high growth places have the steepest experience-wage profile which also supports the idea that those locations offer more scope for lifetime learning.

A possible concern with the previous results is that selection might be playing a role. For instance, recent migrants from high growth places could have a higher learning ability than recent migrants from low growth places. To evaluate this possibility based on the idea that learning ability and schooling are positively correlated, Figure 3 compares average education years of stayers and recent migrants by region in each country. In Mexico and the United States, migrants are positively selected in both regions and differences are not large in most cases. Recent migrants in low growth places of Mexico do have a high education compared to other categories; however, note that in Table 1 those migrants have flatter profiles than workers who stay in high growth places despite the difference in education years. Furthermore, the fact that recent migrants from low growth places are positively selected on education in Mexico and the United States provides support for the idea that location characteristics matter for human capital accumulation. In the case of Brazil, recent migrants in high growth places are negatively

Figure 4: Education years of new migrants by years of potential experience



Notes: This figure presents average education years for recent migrants by years of potential experience (more or less than 15 years). States are classified as high growth places if they are in the upper quartile of the profile height distribution across states in each country. New migrants are workers who moved between low and high growth states within the last five years.

selected on education which could partially explain the large difference in measured returns to experience. Below, I explore further the importance of education for spatial differences in life cycle wage growth.

Another related concern with the results in Table 1 is that selection on ability might depend on experience. For instance, experience-wage profiles of recent migrants from low growth regions might be flatter because workers who move earlier in life are positively selected relative to those who move later in life and the strength of such selection is stronger than for migrants in the opposite direction. To explore this possibility, Figure 4 compares average education years for migrants with different potential experience (more or less than 15 years). In every country and region workers with less experience have more education though the differences are small in most cases. The two exceptions are low growth regions in Brazil and Mexico where the difference in education years is moderately large. Note that if the latter reflects selection on ability related to experience, this could partially explain the difference in profile steepness with respect to stayers in high growth regions but not the difference with respect to recent migrants in those places. Overall, the comparison of education across migrants and stayers does not provide definitive evidence that selection is the main driver of differences in measured returns to experience in Table 1, though it could be playing a role especially in the case of Brazil. The model presented below attempts to separate the individual and local component that determine the return to experience in a location.

I now exploit the data to explore different hypothesis for the spatial covariance between income per capita and life cycle wage growth. It is crucial to look at evidence that helps distinguish the importance of different potential explanations, since this will determine the relevance of policies related to encouraging internal migration. For example, one possible explanation for the findings presented above is that poor states tend to have less educated workers and experience-wage profiles are flatter for this group of workers. The latter has been documented for different countries (see [Lemieux, 2006](#); [Lagakos et al., 2018b](#)). If most of the variation in life cycle wage growth across states is due to differences in the distribution of educational attainment, then the possibility to migrate during the working lifetime is less important, as well as other location characteristics, and policies should focus on closing the gaps in educational attainment across states. Therefore, I carry out multiple accounting exercises to explore the importance of three factors - education, occupation, and industry - for the covariance of income and experience-wage growth across space in each country.

To quantify the importance of differences in the distribution of educational attainment, I do the following exercise. First, I categorize workers by years of schooling: 0-6 (some or all primary school); 7-12 (some or all secondary school); and 13 or more (college and graduate school). Second, for each state and education group in a country, I estimate experience-wage profiles based on equation 1 and compute the average profile height. Next, for each state in a country, I compute a counterfactual profile using the average height of the actual profile by education group and weighting these averages with the distribution of education categories in states that are in the upper quartile of the profile height distribution. In other words, the counterfactual profile uses the actual wage growth of each education group in a state and assumes that the distribution of educational attainment is the same as in the states with steepest profiles. In the extreme case where all the variation in life cycle wage growth across states comes from differences in the distribution of educational attainment, there would not be any spatial dispersion in counterfactual profiles.

To complete the exercise, I run a linear regression of average profile height on log income per capita for the actual and counterfactual profiles, and compare the slope across cases. A

Table 2: Accounting for Spatial Variation in Life Cycle Wage Growth
Covariance with income per capita

Country	Actual	Counterfactual		
		Education	Occupations	Industries
Brazil	17.9	11.2	9.8	11.0
Mexico	9.1	4.5	5.3	3.6
United States	17.7	16.1	11.0	17.2

Notes: This table presents the slope of a regression of profile height on log gdp per capita. This is done at the state level for actual and counterfactual experience-wage profiles. The latter are estimated assuming: (i) that every state has the same distribution of educational attainment; (ii) the same distribution of workers across occupation categories; and (iii) the same distribution of workers across industries. These profiles are estimated for every state controlling for education, state of birth, and time effects in every case.

smaller coefficient in the counterfactual case means that education is quantitatively important since the covariance with income has decreased. The second column in Table 2 shows the actual slope and the third column the counterfactual slope based on educational attainment. These results imply that education accounts for a small share of the covariance between income and life cycle wage growth in the United States since both slopes are similar. In contrast, education can account for a much larger share of the covariance in both developing countries, especially in Mexico where the counterfactual slope is half of the actual slope. Thus, this evidence suggests that differences in the distribution of educational attainment are important for the relationship between regional development and experience-wage growth in developing countries but not in the United States. One explanation for the latter is that U.S. states are more homogeneous in terms of educational attainment, so the variation in life cycle wage growth must be explained by other variables.⁴

Next, I do a similar accounting exercise to measure the importance of differences in the types of occupations that predominate in each state for the covariance between income and

⁴These accounting results may underestimate the importance of education in explaining the variation in experience wage profiles if years of schooling do not capture the full effects of education on human capital accumulation. It could be the case that there are substantial differences in education quality between rich and poor states, which could account for a large portion of the variation in life cycle wage growth. That said, differences in education quality are probably smaller among states within countries than between rich and poor countries.

experience-wage growth. In this case, I define three occupation categories based on the type of skills they require: high cognitive skills, medium cognitive skills, and manual skills. The basic idea is that cognitive occupations have a larger scope for learning than manual occupations, thus, workers in states where cognitive occupations predominate have higher human capital accumulation over the life cycle. In practice, I use the following classification based on the International Standard Classification of Occupations (ISCO). High cognitive occupations include legislators and managers, professionals, and technicians and associate professionals; medium cognitive occupations are clerks, service workers, occupations in shop or market sales, and crafts and trade workers; and manual occupations include agricultural workers, plant or machine operators and assemblers, and elementary occupations.

Then, using these groups of occupations, I estimate counterfactual experience-wage profiles assuming that every state has the same distribution of workers across occupations as states that are in the upper quartile of profile height distribution. Similar to the case of education, if all the variation in life cycle wage growth across states was due to the fact that high cognitive occupations predominate in states with steeper profiles, then there should not be any variation in the counterfactual profiles in which the distribution of occupations is the same for every state.

The fourth column in Table 2 shows that occupations can account for a large share of the covariance between income and life cycle wage growth across states in every country. That is, the counterfactual slope is considerably smaller than the actual slope in every case. These results imply that differences in the distribution of occupations across states are quantitatively important for the spatial variation in experience-wage profiles. Additionally, the last column in Table 2 presents a similar accounting exercise using industries instead of occupations. For this, I consider four types of industries: agriculture, manufacturing, and services (low-skilled and high-skilled).⁵ According to these results, differences in industry mix can also account for a substantial amount of the covariance between income and life cycle wage growth in both

⁵I use the industry classification of IPUMS-International. Manufacturing includes mining and extraction, construction, and electricity, gas, water and waste management. Low-skilled services include wholesale and retail trade, hotels and restaurants, transport, storage and communication, and private household services.

Table 3: Accounting for Spatial Variation in Life Cycle Wage Growth
Covariance with income per capita: workers with less than 12 years of education

Country	Actual	Counterfactual	
		Occupations	Industries
Brazil	17.5	12.9	10.6
Mexico	9.8	6.0	3.0

Notes: This table presents the slope of a regression of profile height on log gdp per capita. This is done at the state level for actual and counterfactual experience-wage profiles. The latter are estimated assuming: (i) that every state has the same distribution of educational attainment; (ii) the same distribution of workers across occupation categories; and (iii) the same distribution of workers across industries. These profiles are estimated for every state controlling for education, state of birth, and time effects in every case.

developing countries, but not in the United States where the counterfactual slope does not change much.

Now, the previous results imply that both education and the local mix of economic activity matter for the covariance of income and life cycle wage growth across states in both developing countries. To highlight the quantitative importance of local economic activity in Mexico and Brazil, I do a similar accounting exercise to the ones described above but restricting the sample to workers with less than 12 years of education. This group of workers has the flattest experience-wage profile in every country and accounts for approximately 70% of the workforce in Brazil and Mexico in the data. The results in Table 3 imply that the local mix of economic activity is important for the spatial covariance between income per capita and life cycle wage growth, independently from education. In the case where differences in educational attainment are driving most of the importance of occupation and industry, there should not be a large change in the covariance between income and experience-wage growth based on counterfactual profiles from workers with similar education but this is far from being the case.

Lastly, recent work has shown that big cities are a important factor in terms of spatial variation in life-cycle wage growth (De la Roca and Puga, 2016; Martellini, 2019). In this case, the correlation between urban population share and average height of experience-wage profile among states is equal to 0.22 in Mexico, 0.43 in the United states, and 0.65 in Brazil.

Thus, the data in this paper are consistent with the notion that urbanization has a positive relationship with human capital accumulation over the life cycle.⁶

To summarize, the evidence presented in this section suggests that there are significant differences in the scope for lifetime learning across space within countries. This implies that migration possibilities from places with low learning to places with high learning might be key for human capital accumulation, especially for young workers, and this is particularly important in developing countries where barriers to internal migration might be larger. To explore this idea, the next section documents differences in migration flows across the three countries considered. Lastly, it is worth mentioning that none of the results in this section are considered causal relationships since, for example, unobserved individual characteristics might explain some these patterns; however, they are suggestive that there is substantial variation in learning environment across space.

2.3 Internal Migration

This section presents migration facts for the three countries considered. There are two simple measures of internal migration that can be constructed with the available data and compared across countries. One could be considered a long run interstate migration rate and is the share of workers who are no longer living in their state of birth. The second is a short run interstate migration rate and is the share of workers who have moved across states within the last five years. I calculate these two statistics for each country based on the same sample of workers that was used in the previous section.⁷

The estimates presented in Table 4 show that 37 percent of U.S. workers have moved from their state of birth compared to 26 and 20 percent of workers in Mexico and Brazil, respectively. Furthermore, the share of workers who have moved to a different state within the last five years is more than two times larger in the United States (11 percent) than in

⁶The relatively low correlation in Mexico is particularly influenced by one state with moderate urbanization and a steep profile (state of Tabasco). If this state is omitted, the correlation is equal to 0.39.

⁷For the United States, data on interstate migration within the last five years is only available for the years 1990 and 2000. Patterns of internal migration in the United States have been studied in detail by [Molloy et al. \(2011\)](#) and [Kaplan and Schulhofer-Wohl \(2017\)](#).

Table 4: Migration Rates (percentage of population)

Country	Interstate		Interregional	
	Long Run	Short Run	Long Run	Short Run
Brazil	19.9	4.0	13.9	2.8
Mexico	25.8	5.7	14.3	3.1
United States	36.6	10.8	25.8	7.3

Notes: Long run migration refers to workers who are no longer living in their state of birth and short run migration to workers who moved across states within the last five years. For regional rates, states are grouped using the classification by the BEA (U.S., 8 regions), IBGE (Brazil, 5 regions) and CONABIO (Mexico, 8 regions). These results are based on the whole sample of male workers in the private sector with 0 to 40 years of experience. I follow the geographical division of IPUMS-International for every country.

both developing countries (6 percent in Mexico and 4 percent in Brazil). These facts support the notion that the United States is a benchmark of high labor mobility across space, while workers in developing countries face higher barriers to migration. This also means that workers in developing countries who are born in a state with less scope for lifetime learning might face higher migration barriers than U.S. workers in a similar situation.

I now present an additional measure of internal migration. It aims to address the possibility that most of the relatively high migration in the United States is due to its geopolitical division, so workers move between states that are close to each other in a narrow geographical region. To evaluate this, I calculate long and short run migration rates across regions in each country based on regional divisions established by a government or public agency.⁸ The third and fourth columns of Table 4 present the interregional migration rates, which show that taking into account broader geographical units does not change the qualitative findings. That is, the share of U.S. workers that have moved from their region of birth is more than 10 percentage points higher than in Mexico and Brazil, and the share of U.S. workers who have moved between regions in the last five years is more than two times larger than in both developing countries.

To sum up, in each of the three analyzed countries there is substantial variation in the

⁸I use the eight regions used by the Bureau of Economic Analysis (BEA) in the U.S.; the five regions used by the Brazilian Institute of Geography and Statistics (IBGE); and the eight economic regions used by the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad(CONABIO) in Mexico.

steepness of experience-wage profiles across space, and in every country richer states tend to have steeper profiles. To account for the latter, variation in the distribution of educational attainment and industry mix seems to matter in Brazil and Mexico, but not so much in the United States. In contrast, variation in the distribution of workers across occupations accounts for a relatively large share of the covariance between income and experience-wage growth in every case. These findings support the idea that labor markets in some states offer a larger scope for lifetime learning than others. The latter implies that migration possibilities might be key to accumulate human capital through working experience and barriers to leave places with fewer learning opportunities can play an important role in explaining productivity differences across countries.

The next section presents a general equilibrium model with internal migration and human capital accumulation. The main goal of the model is to quantify the aggregate productivity gains from reducing barriers to labor mobility in each country, particularly between places with different wage growth over the life cycle, and measure the importance of local learning environment for migration gains. The model considers multiple determinants of labor mobility, differences in human capital accumulation across places, and general equilibrium effects of increasing internal migration. The model also provides a structural interpretation to the apparent differences in returns to experience across space, taking into account different sources of individual heterogeneity.

3 Model

The setting in the model is as follows. There are two regions in the economy indexed by $j \in \{L, H\}$. These regions are potentially different in the learning environment they offer to workers over the life cycle. Workers are born in one region and have the option to move at some point of their lifetime. Therefore, from the worker point of view, regions represent an origin o and a destination d . Time is discrete, a unit mass of workers is born every period in the economy, and I focus on the steady state.

3.1 Workers

Workers are indexed by i , have linear preferences over a final consumption good and discount the future with factor β . Moreover, workers differ in their learning ability in each location a_{ijs} , which also depends on a given level of individual education s . The idiosyncratic learning ability can be interpreted as comparative advantage to perform and learn in the economic activities that predominate in each location. Additionally, each worker is endowed with one unit of time that he supplies inelastically in the local labor market and workers are not allowed to borrow or lend. Then, each worker lives for T periods and maximizes expected lifetime utility:

$$\max_{\{m_{i,t}\}_{t=1}^T} E \left[\sum_{t=1}^T \beta^{t-1} (\mu_j w_{ij,t} - \tau_{i,t} \mathcal{I}_{[m_{i,t}=d]}) \right] \quad (2)$$

subject to

$$w_{ij,t} = \omega_j h_{i,t}, \quad \forall t \text{ and } j \in \{o, d\}, \quad (3)$$

$$m_{i,t} \in \begin{cases} \{d, o\}, & \text{if } m_{i,x} = o, \forall x < t, \\ \{\emptyset\}, & \text{otherwise,} \end{cases}$$

$$h_{i,t+1} = h_{i,t} + a_{ijs} \mathcal{L}_j h_{i,t}^{-\alpha}, \quad \forall t \geq 1 \text{ and } j \in \{o, d\}, \quad (4)$$

where $w_{ij,t}$ are labor earnings in life period t , which are a function of the wage rate per efficiency unit ω_j and worker's human capital $h_{i,t}$. Note that labor income is used by every worker to buy the final consumption good so that $c_{ij,t} \equiv w_{ij,t}$. Furthermore, μ_j are exogenous amenities in region j which are broadly defined and govern compensating wage differentials across places; $m_{i,t}$ is the migration decision by worker i in period t (workers start in their origin); $\tau_{i,t}$ are migration costs in utility units and $\mathcal{I}_P$ is an indicator function of the set P that takes the value of one when the worker decides to migrate from his place of origin. Migration costs are given by $\tau_{i,t} = \bar{\tau} - \psi b_{i,t}$, so they include a common element that is time-invariant plus an idiosyncratic shock experienced by each worker.

According to equation (4), human capital is accumulated over time through learning by

doing.⁹ That is, workers increase their level of human capital with working experience and their rate of learning depends on individual ability a_{ijs} and endogenous local environment given by $\mathcal{L}_j = \gamma_j H_j^{\gamma_j}$, where H_j is total human capital in region j , which a worker takes as given, and γ_j is parameter that governs the strength of local human capital spillovers in life cycle wage growth. This parameter captures differences in occupation or industry mix, as well as other factors that might increase learning possibilities in a particular region such as the existence of big cities. Parameter α governs the decline of learning opportunities over the life cycle. Therefore, this specification is based on the idea that workers who are close to the end of their working life have less potential or room for learning. Moreover, the production of human capital allows for the possibility that workers with higher learning ability have steeper experience-wage profiles, so selection can be an important determinant of spatial differences in life cycle wage growth. Initial human capital h_1 is normalized to one for every individual.

In this model, a worker would like to migrate as soon as possible if the potential destination gives him a higher stream of lifetime utility, that is, if he has a comparative advantage in his potential destination. The individual shock $b_{i,t}$ captures the fact that workers face random shocks that make migration more or less costly over time. This shock is observed by the worker at the beginning of every period before making a migration decision. Then, the problem of each worker can be summarized as follows. A worker is born in one of the two regions and can choose to move at any point in time if his potential destination offers higher lifetime utility net of migration costs. Then, if a worker chooses to migrate, the decision is irreversible and they incur migration costs in the period they move.¹⁰ Lastly, after workers make their migration choice, production and consumption take place.

⁹For other examples of human capital models based on learning by doing see [Imai and Keane \(2004\)](#), [Gemici and Wiswall \(2014\)](#), [Fan et al. \(2015\)](#), and [Blandin \(2018\)](#).

¹⁰In this model, the simplification of allowing just one move is not restrictive given that workers who move to a place with higher lifetime income do not have incentives to leave. This would be different in a model with location-specific shocks over the life cycle. [Kennan and Walker \(2011\)](#) document that the average numbers of moves in a lifetime are lower than two in the United States, and the majority of second moves are returns to home (thus, less likely to be based on comparative advantage or job opportunities). [Bryan and Morten \(2018\)](#) provide similar evidence for Indonesia.

3.2 Production

The final consumption good in the economy is produced according to

$$Y = \left[\sum_{j \in \{L, H\}} (A_j Z_j)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}, \quad (5)$$

where Z_j is the demand for efficiency units from region j ; A_j is a productivity parameter in each region j ; and $\rho \in [0, \infty)$ is the elasticity of substitution across goods produced in different regions. The final good producer must pay the wage per efficiency unit ω_j to each worker and maximizes profits every period solving

$$\max_{Z_j, j \in \{L, H\}} Y - \sum_j \omega_j Z_j, \quad (6)$$

where the final consumption good has been used as the numeraire. Note that this problem is equivalent to regions producing differentiated intermediate goods using a linear technology, $A_j Z_j$, which are then supplied to the final producer in the economy.

3.3 Equilibrium Analysis

In equilibrium, the final good market has to clear in each period as well as the local labor market. The latter means that regional demand for efficiency units has to be equal to total supply, that is,

$$\begin{aligned} \omega_j &= (A_j)^{\frac{\rho-1}{\rho}} \left(\frac{Y}{Z_j} \right)^{1/\rho}, \\ Z_j &= \int_{i \in \Omega_j} h_i dF_i \equiv H_j, \quad j \in \{L, H\}, \end{aligned} \quad (7)$$

where Ω_j is the set of individuals who choose to live in region j and F_i is a cumulative probability function of individual characteristics. The parametrization of the latter is described in the next section. For $\rho > 0$, local wages decrease with labor demand which tends to limit

the concentration of population in one place.

As stated above, workers in the model migrate if they have a comparative advantage in their potential destination or if a random shock leads them to leave their place of origin. Then, because comparative advantage is based on learning ability, one feature of the equilibrium is that movers are positively selected on human capital. This means that a potential gain from relaxing migration costs in the model economy is that workers who are “misallocated” can move to places where they can learn more based on their individual talents for economic activities in such locations. This is a dynamic version of the static comparative advantage that is common in the literature. The existence of idiosyncratic shocks weakens the strength of selection by introducing randomness in the workers who move, but it is still the case that conditional on having a comparative advantage in the potential destination ($a_{ids} > a_{ios}$), workers with higher learning ability are more likely to move. This is important given the concern that spatial differences in experience-wage profiles are driven by unobserved learning ability.

Moreover, to see how individual ability and $\mathcal{L}_j$ determine the measured return to experience in each region, assume that α is equal to -1 so that $h_{i,t+1} = h_{i,t}(1 + a_{ijs}\mathcal{L}_j)$. This assumption would imply that the rate of learning is constant over the life cycle, but it is helpful to highlight the interaction between selection on ability and the true return to experience. That said, taking log approximations, in such case we can define the change in log earnings from $t - 1$ to t for a given worker in region j as $\Delta \log(w_{ij,t}) = a_{ijs}\mathcal{L}_j$. Then, the average one-period wage gain for workers in a particular region is given by

$$\overline{\Delta \log(w_{ij,t})} = \mathcal{L}_j E[a_{ios} | i \in \Omega_j], \quad j \in \{L, H\} \quad \text{and} \quad \forall t. \quad (8)$$

Therefore, the measured return to experience in one region can be high if individuals who live there have a high learning ability or if the true return to experience captured by $\mathcal{L}_j$ is high. Furthermore, based on the same assumptions, the difference in average wage across regions for

workers with t years of experience who have *not* migrated is

$$\begin{aligned} \overline{\log(w_{iH,t})} - \overline{\log(w_{iL,t})} = & \log(\omega_H) - \log(\omega_L) + \\ & (t-1) \left[\mathcal{L}_H E[a_{iHs}] - \mathcal{L}_L E[a_{iLs}] \right], \quad i : m_{i,x} = o, \forall x \leq t. \end{aligned} \tag{9}$$

This expression emphasizes that the potential gains from inducing migration to high learning places include a static gain and a human capital accumulation gain, and the latter is a function of a comparative advantage effect from learning ability and the local learning environment. If dynamic gains due to differences in $\mathcal{L}_j$ are large, then there is more potential for policy-induced migration gains. In the quantitative section, I decompose the productivity gains coming from each of these factors when workers are induced to migrate from the low learning region.

4 Estimation

This section describes how I take the model to the data. First, I present the parametrization of the model. Then, I calibrate it using a simulated method of moments. I assume that the number of life periods in the model T is equal to eight and each period represents five years of potential experience. Furthermore, region H in the model represents states with high wage growth in the data, similar to the empirical results presented in Table 1. That is, for each country, I classify states with high wage growth as those with an average height of the experience-wage profile in the upper quartile of the profile height distribution across states, while the rest of the states are classified as low wage growth. There are two parameters that I set outside of the model: the discount factor β is set equal to 0.81 based on a five year period and the elasticity of substitution across regions ρ is set equal to 6. This is a moderately lower value than the one used for models with disaggregated locations (see [Allen and Arkolakis, 2014](#); [Bryan and Morten, 2018](#)).

Additionally, I set initial population shares in the model equal to the shares of workers born in each region in the data. The latter results in the following initial population share for region H in each country: 0.47 in the U.S., 0.42 in Brazil, and 0.36 in Mexico. Thus, according to

the classification based on profile heights, states with low wage growth concentrate a relatively large amount of population in Mexico in comparison to the other two countries.

4.1 Parametrization

This section describes the parametrization of the model. First, location-specific individual learning ability is defined as $a_{ijs} = \phi_s \epsilon_{ij}$ where ϕ_s is a scale factor for education category $s \in \{led, hed\}$ and ϵ_{ij} represents an unobserved component of learning ability that is location specific. Each worker belongs to either of two schooling categories: (i) no high school completed or less than 12 schooling years (*led*), or (ii) at least high school completed (*hed*). The initial shares of workers in these categories is set by region of origin based on the data sample in each country. Thus, the calibration of the model takes into account differences in the distribution of educational attainment across regions. Furthermore, I assume ϵ_{ij} is independent and identically distributed across individuals and regions, and follow the human capital literature by assuming that it is distributed according to a log-normal distribution so that $\log(\epsilon) \sim N(0, \sigma)$, where σ governs the variation in learning ability. In addition, the idiosyncratic component of migration costs $b_{i,t}$ is independently and identically distributed across individuals and over time, and I assume that $b \sim \text{Logistic}(0, 1)$. This type of distribution is commonly used in frameworks of location choice. Then, since migration costs are given by $\tau_{i,t} = \bar{\tau} - \psi b_{i,t}$, parameters $\bar{\tau}$ and ψ are used to target migration moments as explained below.

4.2 Method of Moments

This section presents the calibration of the model using a simulated method of moments. To be specific, I use model-simulated data to compute the value of relevant moments and match them with their actual data counterparts. I normalize productivity A_H and compensating differential μ_H to one, so that the model matches relative wages and migration flows across regions. I also normalize the ability scale factor for workers with less than 12 years of education ϕ_{led} to one and match the difference in experience-wage profiles between education categories.

Then, there are nine parameters that need to be estimated: $A_L, \tau, \mu_L, \psi, \sigma, \gamma_j, \alpha, \phi_{hed}$, for

$j \in \{L, H\}$. To do this, I target the following set of moments: the ratio of average wage in region L (low growth) to average wage in region H (high growth); difference in variance of log hourly wages between workers with 25-30 years of potential experience and those with less than five years; average height of the experience-wage profile in each region; average wage of a worker with 10-15 years of potential experience relative to a worker with less than five years of experience; height difference of the experience-wage profile between education categories; the short-run migration rate in the economy; difference in migration rate between young and older workers; and the population share in region L .

The internal calibration involves nine parameters and moments, and the identification can be explained as follows. First, productivity parameter A_L , compensating wage differential μ_L , and common migration costs $\bar{\tau}$ govern the average wage gap across regions as well as the flow of migrants between each origin and destination. That is, a low value of A_L tends to decrease wages in region L relative to region H , so then μ_L must adjust to keep the right amount of people across regions. On the other hand, a higher value of $\bar{\tau}$ decreases the flow of migrants in both directions given the distribution of idiosyncratic shocks. To be more specific, I use $\bar{\tau}$ to target the share of workers in the economy who migrated across regions within the last 5 years in the data (one model period). Additionally, a low value of ψ means that idiosyncratic shocks are less important and most migration happens early in life based on comparative advantage. Thus, the value of ψ governs how steep is the life cycle profile of migration rates.

To obtain the value of moments related to experience-wage profiles, I estimate a similar specification by region in both the model and data,¹¹

$$\log(w_i) = \alpha + \sum_{x \in X} \beta_x D_i^x + \psi_o + \epsilon_i, \quad (10)$$

where w_i is the hourly wage of worker i , D_i^x is a dummy that takes the value of one if a

¹¹The specification in the data includes year and state of birth dummy variables. Also, in the data, I control for education years to be consistent with the profiles estimated in the empirical section; otherwise, the shape of the profiles can change considerably because older workers with low education tend to have relatively low wages, which pushes down the average wage of workers with more than 30 years of potential experience. That is because I assume individuals start working when they turn 16 or when they finish school, whichever comes last, so relatively old workers with low education show up as workers with high potential experience.

Table 5: Calibration by Method of Moments

Parameter	Value			Description	Moment
	US	Mexico	Brazil		
A_L	0.95	0.84	0.75	Labor productivity	Avg. wage of region L relative to region H
μ_L	1.23	1.48	1.25	Compensating differentials	Population share in region L
$\bar{\tau}$	2.84	4.43	6.51	Migration costs	Short run migration rate in the economy
ψ	1.62	1.76	2.02	Variance of idiosyncratic shocks	Diff. in migration rate btw peak and early life
σ	1.18	1.87	1.57	Variance of learning ability	Diff. in wage variance btw peak and early life
ϕ_{hed}	1.30	1.80	1.27	Learning ability with high education	Diff. in wage growth across education groups
α	1.92	1.29	1.03	Human capital production curvature	Wage growth in beginning of working life
γ_L	0.151	0.043	0.108	Learning environment spillovers	Wage growth during working life
γ_H	0.190	0.059	0.147		

Notes: This table presents the calibrated values of each parameter based on a simulated method of moments done for each country independently. The short run migration rate refers to worker who have moved within the last five years. Region H in the model represents the states with the steepest experience-wage profiles. See text for details.

worker is in experience group $x \in X = \{5 - 9, 10 - 14, \dots\}$, and ψ_o is dummy for region of origin. The same specification is estimated for each region and education category. Thus, the calibration is equivalent to an indirect inference approach of matching values and statistics of β_x in the model and the data. Parameter γ_j governs the learning rate through experience in each region by determining the strength of human capital spillovers in life cycle wage growth and, therefore, is related to the profile steepness in each place. Furthermore, parameter α governs the curvature of human capital production, so that high values tend to put most of the wage growth in the beginning of the life cycle and flatten the profile later in life. Parameter ϕ_{hed} governs the mean learning ability of workers with high level of schooling, and thus the difference in profile height between education groups. Lastly, distribution parameter σ governs the variation of hourly wages over the life cycle because a higher variation of learning ability increases the variance of wages for workers with more experience.

The calibrated value of each parameter is presented in Table 5. These results show that compensating differentials must be somewhat higher in region L to match the right flow of migrants across regions. This could represent, for example, the fact that richer regions tend

Table 6: Method of Moments: Model Fit

Moment	Brazil		Mexico		US	
	Model	Data	Model	Data	Model	Data
Avg. wage of region L relative to region H	0.70	0.70	0.76	0.76	0.89	0.89
Diff. in wage variance btw peak and early life	0.33	0.33	0.26	0.26	0.15	0.14
Short run migration rate in economy (%)	2.13	2.14	3.30	3.29	5.30	5.29
Diff. in migration rate btw peak and early life	1.27	1.30	1.20	1.21	3.87	3.86
Population share in region L	0.46	0.46	0.35	0.35	0.43	0.43
Wage growth during working life in region L	1.78	1.78	1.46	1.46	1.89	1.89
Wage growth during working life in region H	1.96	1.96	1.54	1.54	2.01	2.01
Wage growth in beginning of working life	1.61	1.61	1.36	1.37	1.72	1.73
Diff. in profile height btw high and low education	0.14	0.14	0.19	0.19	0.13	0.13

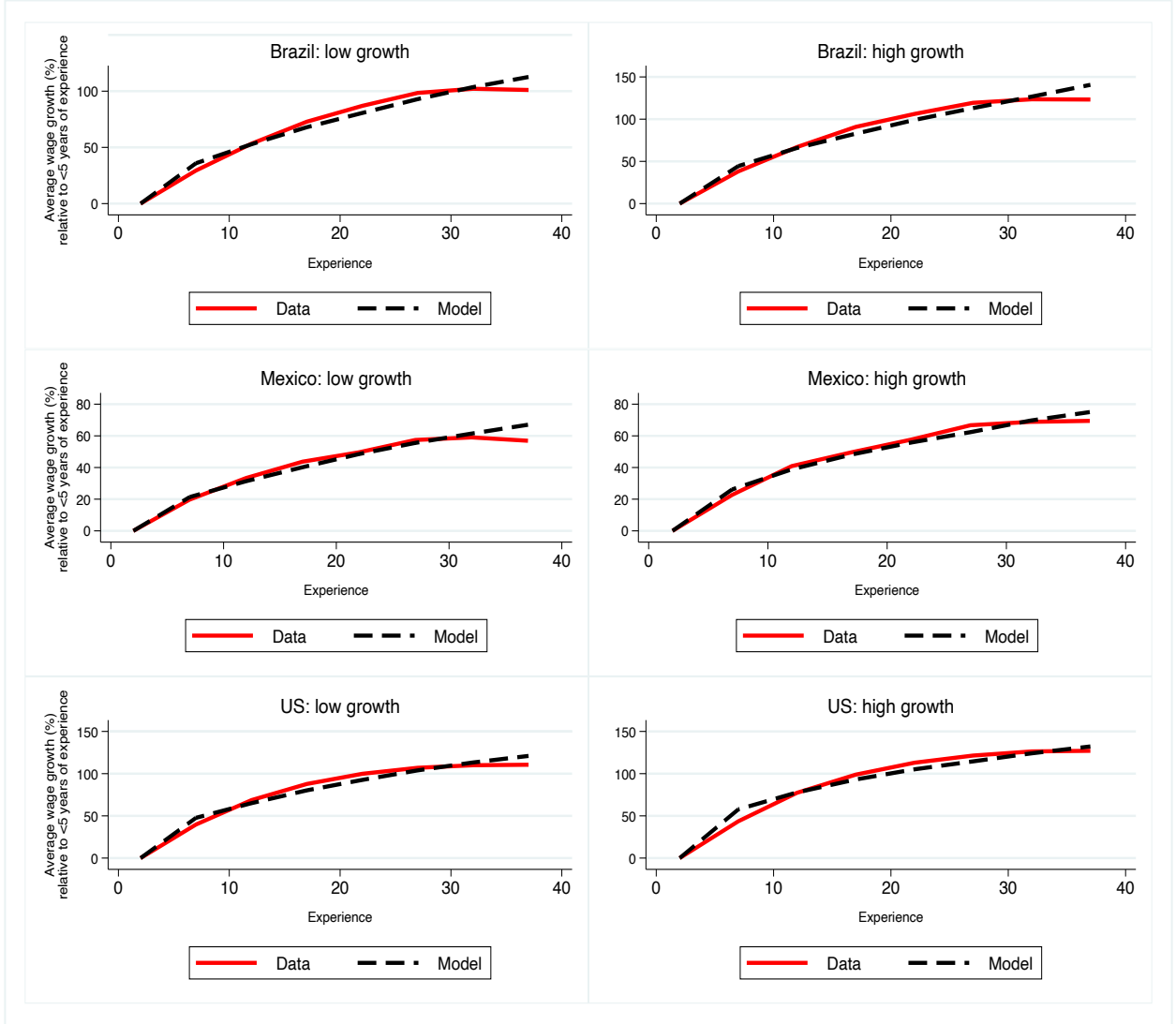
Notes: This table presents the results of a simulated method of moments done for each country independently. Region H in the model represents the states with the steepest experience-wage profiles. The short run migration rate refers to worker who have moved within the last five years. Peak of the working life refers to workers with 25-30 years of potential experience. The difference in experience-wage profile height between education groups is based on average profile height.

to have higher costs of living. Furthermore, the results show that migration costs $\bar{\tau}$ are higher in both developing countries than in the United States. In fact, as a share of average lifetime income, trade costs are equal to 16 percent in the United States, compared to 32 percent in Mexico and 38 percent in Brazil.¹² Note that the ranking of migration costs implied by the calibration is consistent with the ranking of internal migration shares presented in Section 2.3; that is, the United States has the lowest migration costs and highest migration rates, whereas Brazil has the highest migration costs and lowest migration rates.

The calibration results also imply that there are considerable regional differences in the strength of learning spillovers in each country (the rate of human capital accumulation is governed by γ_j). Thus, once variation in individual learning ability is taken into account to match experience-wage profiles in the model and data, it is the case that returns to experience captured by the endogenous learning rate are higher in regions with steeper profiles. Furthermore, the calibration results imply that the learning premium for workers with high

¹²The magnitudes of the permanent migration costs are similar to the ones estimated by Bryan and Morten (2018). Heise and Porzio (2019) find smaller migration costs in a model with home-bias preference and labor market frictions within regions.

Figure 5: Model Fit: Experience-Wage Profiles by Region

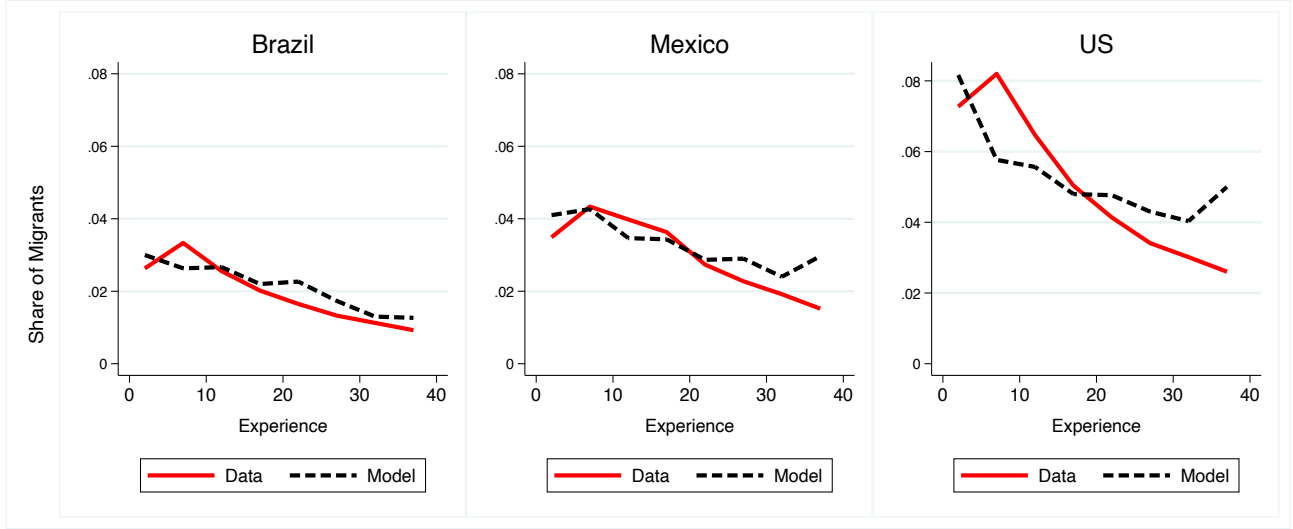


Notes: This figure compares experience-wage profiles in the actual data and model-simulated data by country and type of region. High growth region refers to states that are in the upper quartile of the profile height distribution within each country, while low growth region includes the remaining states.

education ϕ_{hed} is substantial among the countries considered and must be particularly large in Mexico to match the difference in profile height between education groups.

Table 6 presents the fit of the model based on the simulated method of moments. The model does very well in matching the targeted moments. It is worth highlighting that Brazil is the country with the largest wage gap between regions, the most variation in hourly wages over the life cycle, and the lowest migration rate. Moreover, the height difference of experience-wage profiles between education groups is especially large in Mexico, so the composition of

Figure 6: Model Fit: Short Run Migration Shares by Experience



Notes: This figure compares short run migration shares across regions by amount of experience in the actual data and model-simulated data.

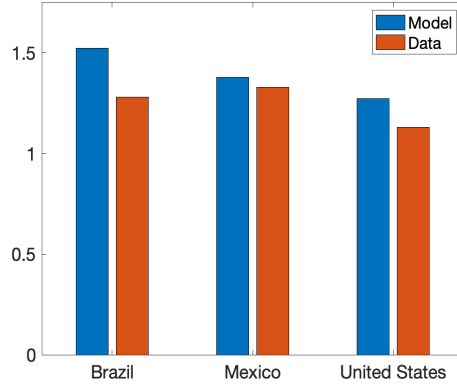
educational attainment is potentially more important in this country for the regions considered.

Furthermore, Figure 5 shows that the model does remarkably well in fitting the experience-wage profile by type of region in each country. Note that only three profile moments were targeted for each country in the calibration and the model is able to match the different profiles to a substantial extent. Lastly, Figure 6 compares the share of workers who migrated within the last five years by amount of potential experience. The model matches well the fact that migration declines with experience. That is because workers would like to move as soon as possible to their potential destination if that place gives them a higher lifetime income, while older workers might not find it optimal to incur migration costs even if they receive a positive shock in their potential destination.

4.3 Model Validation

Next, I validate the model by comparing non-targeted facts in the model and data. First, Figure 7 presents the average wage of migrants (those who left their place of origin at some point in time) relative to stayers in the model and data. The model is consistent with the fact that migrants have higher average wages in every country and the magnitude of relative

Figure 7: Model Validation:
Average wage of migrants relative to stayers



Notes: This figure compares the average wage of migrants and stayers in the data and model for each country.

Table 7: Model validation:
Difference in returns to experience between recent migrants and stayers

Country	Low Growth Region		High Growth Region	
	Model	Data	Model	Data
Brazil	-8.90	-0.93	-23.70	-43.36
Mexico	8.31	7.06	-25.56	-11.14
United States	6.53	1.01	-29.43	-19.12

Notes: This table presents the difference in average height of experience-wage profile between recent migrants and stayers in each region. See text for details.

wages is fairly close to the data counterparts. The model can replicate this because migrants are positively selected in terms of learning ability.

Next, I compare the average height of experience-wage profiles of recent migrants and stayers in each region. This was an important piece of evidence presented in Section 2.2 to argue that migrants bring with them some of the knowledge they acquired in their place of origin. Thus, to evaluate the model in this dimension, I calculate the difference in average profile height between type of workers within each region. Table 7 shows that the model matches quite well the differences in the data. That is, recent migrants in the high growth region have a considerable flatter profile than stayers in that place, whereas recent migrants in

the low growth region have a moderately steeper profile than stayers in that region. Moreover, the model matches the fact that recent migrants in low growth places of Brazil have a flatter profile than stayers in that place, but the difference is much smaller than in the other region.

To summarize, the calibrated model matches relevant moments of the data related to life cycle wage growth across space and internal migration in each country, and is consistent with other important non-targeted moments of migrations patterns and experience-wage profiles between migrants and stayers.

5 Quantitative Experiments

This section presents counterfactual experiments using the calibrated model. The goal is to quantify the gains in productivity, measured as total output per worker, from increasing labor mobility in a setting that allows for different human capital accumulation across space and measure the importance of local learning environment for such gains. First, I focus on quantifying the gains of inducing higher migration to region H , which is the place with higher income and more learning possibilities in every country. In the model, this first experiment is done by setting the common and permanent component of migration costs $\bar{\tau}$ equal to zero for individuals born in region L , keeping barriers constant in the opposite direction. Reducing migration costs in a particular group of states could be done through targeted policies such as migration subsidies, transport infrastructure projects, or training programs.

The results presented in the second column of Table 8 show that labor productivity increases by 2.7 percent in the United States, 6.8 percent in Mexico, and 8.7 percent in Brazil. Thus, there are moderate but important gains in labor productivity from encouraging migration to the richer region with more scope for lifetime learning. It is worth highlighting that productivity gains are larger in both developing countries where migrations costs are higher, but even in the United States there are gains from increasing labor mobility to high learning places. These results are informative about the importance of migration costs for internal mobility and labor productivity in a dynamic setting; however, from a policy perspective it is

Table 8: Inducing migration to high learning region
Percentage change in labor productivity

Country	No migration costs	Matching US migration rate
Brazil	8.67	6.98
Mexico	6.78	4.01
US	2.74	-

Notes: This table presents the results of counterfactual experiments relative to the baseline economy. In this case, the common component of migration costs $\bar{\tau}$ is set equal to zero only for workers who were born in region L , while keeping them constant in the opposite direction.

not realistic to think about zero migration costs.

Therefore, in a second experiment, I take the United States as a high mobility benchmark and reduce $\bar{\tau}$ in low learning regions of Mexico and Brazil to match the share of workers who have left their place of origin within the last five years (one period in the model) in the United States. In other words, I generate the same short run migration rate of the U.S. economy in Brazil and Mexico by reducing migration costs for workers born in low learning places. Then, this experiment represents a more realistic goal for targeted migration policies in both developing countries. In this case I also keep migration costs unchanged for individuals born in region H . The third column of Table 8 shows that labor productivity increases by 7.0 and 4.0 percent in Brazil and Mexico, respectively. These results are more conservative than the previous experiment, but represent meaningful productivity gains taking into account that selection plays an important role in these results; that is, new migrants in region H tend to have a lower learning ability in that place than workers who migrated before the policy change.

Now, one crucial question that arises from the previous results is how much of the labor productivity gain is due to human capital accumulation and how much is coming from static gains. To decompose these two contributions, I calculate an alternative value of total output in the counterfactual economy that takes the counterfactual location for each individual but fixes the human capital profiles to their value in the baseline economy. Then, I aggregate these alternative labor supplies using the production function in equation (5) and calculate output

Table 9: Inducing migration to high learning region
Decomposition of labor productivity gain (percentage)

Country	All human capital	Location learning environment
Brazil	68.1	32.4
Mexico	69.2	26.6
US	84.1	34.4

Notes: This table presents the decomposition of productivity gain from increasing migration to the high learning region in each country. The latter is done by setting the common component of migration costs $\bar{\tau}$ equal to zero only for workers who were born in region L , while keeping them constant in the opposite direction. The decomposition in the second column is based on fixing the human capital profile of each worker to its baseline value. The decomposition in the third column is based on fixing the local learning rate $\mathcal{L}_j$ to its baseline value, letting individual ability change.

per worker in the counterfactual economy. The latter represents the gain in labor productivity coming from reallocation of workers to the region where labor earnings are higher in a given moment of time (static gain), and the difference between this gain and the actual gain observed in the counterfactual economy represents the contribution of human capital.

The results presented in Table 9 imply that human capital accounts for a large share of the labor productivity gains in every country. To be specific, according to the decomposition described above, human capital accounts for 68 and 69 percent of the gains in Brazil and Mexico, respectively, and 84 percent of the gains in the United States. The static productivity gap across regions (A_j) is smaller in the latter case, so dynamic gains are relatively larger in that country. That said, a large share of the human capital gains is due to comparative advantage based on learning ability. To separate this effect from the local learning environment, I allow ability to change based on location choices but not the regional learning rate $\mathcal{L}_j$, then calculate the new value of output to compare it with the total gain in productivity. The last column in Table 9 shows that local learning environment accounts for around 30 percent of the aggregate productivity gains across all countries, and between a fourth and a third of the aggregate gains in both developing countries.

These results suggest that human capital might be an important contributor to migration gains in settings where learning opportunities differ substantially across space. Moreover,

Table 10: Increasing migration in the economy
Percentage change in labor productivity

Country	No migration costs	Removing all barriers
Brazil	8.65	10.95
Mexico	6.66	9.34
US	3.42	4.12

Notes: This table present the results of counterfactual experiments relative to the baseline economy. In column 2, the common component of migration costs $\bar{\tau}$ is set equal to zero for every worker in the economy. In column 3, migration costs are set to zero and compensating differentials are eliminated (i.e. $\mu_H = \mu_L = 1$).

they imply that even if we consider regions with large cross-sectional income differences, as both developing countries in this case, human capital can account for a large fraction of the productivity gains from inducing migration to rich, high learning places.

The experiments described in previous paragraphs were based on introducing changes to induce more migration to the high learning region in each country, keeping barriers constant for migration in the opposite direction. Next, I quantify the productivity gains of reducing migration costs in the whole economy. The second column in Table 10 presents the results of setting $\bar{\tau}$ equal to zero for workers born in every region. In this case, labor productivity increase by 3.4 percent in the United States, 6.7 percent in Mexico, and 8.7 percent in Brazil. Note that in this experiment migration also increases from region H to region L , which means that workers are moving to places with lower productivity and less scope for lifetime learning. However, the gains in human capital accumulation from higher agglomeration based on comparative advantage compensate for the latter effects, so the aggregate gains are similar to the previous case.

The previous results show that compensating differentials act as a barrier to labor mobility in the economy since they keep or attract workers to places with lower productivity. Thus, I now consider an experiment where I remove all barriers in the economy by eliminating migration costs and compensating differentials. In terms of policy, amenities differences could be addressed with regulations regarding factors such as costs of living or pollution in richer

regions (see [Bryan and Morten, 2018](#)). According to the results presented in column 3 of Table 10, removing all barriers leads to a 11.0 and 9.3 percent productivity boost in Brazil and Mexico, respectively. Thus, if the goal of encouraging migration is to increase labor productivity in developing countries, an important implication of these results is that policies could focus on increasing labor mobility to places that offer more learning opportunities through working experience.

6 Conclusion

This paper documents new facts on spatial variation in life cycle wage growth for three countries with different levels of economic development: Brazil, Mexico, and the United States. First, rich states tend to have steeper experience-wage profiles than poor states in every country, but the gaps are larger in both developing countries. Second, differences in the distribution of educational attainment and industry mix can account for a considerable share of the covariance between income per capita and experience-wage growth in Brazil and Mexico, but they seem to be less relevant in the United States. In contrast, spatial differences in the distribution of occupations can account for a relatively large share of the covariance in every case. This suggests that workers face different learning opportunities depending on the economic activities that predominate in their location. Indeed, in each of the countries considered, recent migrants from low to high wage growth places have flatter experience-wage profiles than stayers in the destination place.

In order to quantify the potential productivity gains of increasing labor mobility across space, I build a general equilibrium model with human capital accumulation and internal migration. I estimate meaningful gains in labor productivity from removing migration barriers to high learning places in Brazil and Mexico. Moreover, using the structure of the model, I find that spatial differences in learning environment are quantitatively important and account for approximately 30 percent of the aggregate gains.

References

- Allen, Treb, and Costas Arkolakis.** 2014. “Trade and the Topography of the Spatial Economy.” *The Quarterly Journal of Economics*, 129(3): 1085–1140.
- Bagger, Jesper, François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin.** 2014. “Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics.” *American Economic Review*, 104(6): 1551–96.
- Bils, Mark, and Peter J. Klenow.** 2000. “Does Schooling Cause Growth?” *American Economic Review*, 90(5): 1160–1183.
- Blandin, Adam.** 2018. “Learning by Doing and Ben-Porath: Life-cycle Predictions and Policy Implications.” *Journal of Economic Dynamics and Control*, 90(C): 220–235.
- Bryan, Gharad, and Melanie Morten.** 2018. “The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia.” *Journal of Political Economy*, Forthcoming.
- Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak.** 2014. “Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh.” *Econometrica*, 82.
- Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro.** 2019. “Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock.” *Econometrica*, 87(3): 741–835.
- Caselli, Francesco.** 2005. “Accounting for Cross-Country Income Differences.” 1: 679–741.
- De la Roca, Jorge, and Diego Puga.** 2016. “Learning by Working in Big Cities.” *The Review of Economic Studies*, 84(1): 106–142.
- Engbom, Niklas.** 2019. “Worker Flows and Wage Growth over the Life-Cycle: A Cross-Country Analysis.” Working Paper.

- Erosa, Andrés, Tatyana Koreshkova, and Diego Restuccia.** 2010. “How Important Is Human Capital? A Quantitative Theory Assessment of World Income Inequality.” *Review of Economic Studies*, 77(4): 1421–1449.
- Fan, Xiadong, Ananth Seshadri, and Christopher Taber.** 2015. “Estimation of Life-Cycle Model with Human Capital, Labor Supply and Retirement.” Working Paper.
- Gemici, Ahu, and Matthew Wiswall.** 2014. “Evolution of Gender Differences in Post-Secondary Human Capital Investments: College Majors.” *International Economic Review*, 55(1): 23–56.
- Gollin, Douglas, David Lagakos, and Michael E. Waugh.** 2014. “The Agricultural Productivity Gap.” *The Quarterly Journal of Economics*, 129: 939–993.
- Heise, Sebastian, and Tommaso Porzio.** 2019. “Spatial Wage Gaps in Frictional Labor Markets.” Working Paper.
- Huggett, Mark, Gustavo Ventura, and Amir Yaron.** 2011. “Sources of Lifetime Inequality.” *American Economic Review*, 101(7): 2923–2954.
- Imai, Susumu, and Michael P. Keane.** 2004. “Intertemporal Labor Supply and Human Capital Accumulation.” *International Economic Review*, 45(2): 601–641.
- IPUMS.** 2019. Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.2. Minneapolis, MN: IPUMS. <https://doi.org/10.18128/D020.V7.2>.
- Kaplan, Greg, and Sam Schulhofer-Wohl.** 2017. “Understanding The Long-Run Decline In Interstate Migration.” *International Economic Review*, 58: 57–94.
- Kennan, John, and James R. Walker.** 2011. “The Effect of Expected Income on Individual Migration Decisions.” *Econometrica*, 79(1): 211–251.

- Klenow, Peter, and Andrés Rodríguez-Clare.** 1997. “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?” In *NBER Macroeconomics Annual 1997, Volume 12. NBER Chapters*, 73–114. National Bureau of Economic Research, Inc.
- Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman.** 2018*a*. “Life-Cycle Human Capital Accumulation across Countries: Lessons from US Immigrants.” *Journal of Human Capital*, 12(2): 305–342.
- Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman.** 2018*b*. “Life Cycle Wage Growth across Countries.” *Journal of Political Economy*, 126(2): 797–849.
- Lagakos, David, Mushfiq Mobarak, and Michael E. Waugh.** 2019. “The Welfare Effects of Encouraging Rural-Urban Migration.” Working Paper.
- Lemieux, Thomas.** 2006. “The “Mincer Equation” Thirty Years After Schooling, Experience, and Earnings.” , ed. S. Grossbard-Schechtman *Jacob Mincer: A Pioneer of Modern Labor Economics*, 127–145.
- Manuelli, Rodolfo E., and Ananth Seshadri.** 2014. “Human Capital and the Wealth of Nations.” *American Economic Review*, 104(9): 2736–62.
- Martellini, Paolo.** 2019. “The City-Size Wage Premium: Origins and Aggregate Implications.” Working Paper.
- Mincer, Jacob.** 1974. “Schooling, Experience, and Earnings.” *New York: Columbia University Press*.
- Molloy, Raven, Christopher L. Smith, and Abigail Wozniak.** 2011. “Internal Migration in the United States.” *Journal of Economic Perspectives*, 25(3): 173–196.
- Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu.** 2008. “Agriculture and aggregate productivity: A quantitative cross-country analysis.” *Journal of Monetary Economics*, 55: 234–250.

- Rubinstein, Yona, and Yoram Weiss.** 2006. “Post Schooling Wage Growth: Investment, Search and Learning.” , ed. Erik Hanushek and F. Welch Vol. 1 of *Handbook of the Economics of Education*, Chapter 1, 1–67. Elsevier.
- Vollrath, Dietrich.** 2009. “How Important are Dual Economy Effects for Aggregate Productivity?” *Journal of Development Economics*, 88: 325–334.

A Profiles and income per capita: alternative samples

In the main text experience-wage profiles were estimated based on a restricted sample: full-time male workers in the private sector. In this section I compare the relationship between experience-wage profile steepness and income per capita using alternative samples. The comparison is based on regressing average profile height on log income per capita across states. First, I consider a larger sample that includes both male and female workers regardless of the time worked and economic sector. This sample includes workers that could be considered relatively less attached to the labor market. The third column in Table 11 shows that there is a positive covariance between income and profile steepness when a larger sample of workers is considered, though the value is smaller in every country. Additionally, I restrict the baseline sample to wage or salaried workers due to measurement concerns related to self-employment. The last column in Table 11 shows that there is also a positive covariance between income per capita and profile steepness of wage workers across states in every country. This covariance is somewhat smaller in both developing countries compared to the baseline case, though it is moderately larger for the United States. Overall, these results show that the positive relationship between regional development and life cycle wage growth holds for alternative samples.

Table 11: Income per capita and life cycle wage growth across states

Country	Baseline	Max sample	Wage workers
Brazil	17.9***	12.7***	11.5***
Mexico	9.1***	4.3**	4.9***
United States	17.7***	14.2***	19.3***

Notes: This table shows the results of regressing average experience-wage profile height on log GDP per capita across states. Profile height is relative to workers with less than five years of potential experience. Max sample includes male and female workers regardless of time worked and economic sector. Wage workers restricts the baseline sample to salaried or wage workers. Income per capita was obtained from OECD statistics for 2015 and wage growth refers to the average height of the experience-wage profile. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.