Learning Segmentation and Economic Development*

Santiago Franco
Boston University

Jose M. Quintero
University of Chicago

August 2026

Abstract

How does learning from others shape economic development? We answer this question with a quantitative model of multiple labor markets and modern and traditional sectors, capturing key features of dual economies in developing countries. Workers sort across sectors and labor markets and accumulate human capital by learning from peers. Learning opportunities are sector-specific and depend on peer composition within the labor market. The model makes one main prediction: more skilled workers sort into the modern sector and, as a consequence, workers in this sector experience faster wage growth because they are exposed to higher-skilled peers. We test this prediction using longitudinal worker data from Chile and find empirical support for it. Finally, we estimate the model and use it to quantify the aggregate consequences of improving learning opportunities for traditional-sector workers. We find that improving learning opportunities generates meaningful welfare gains, but raises income per capita substantially only when it also increases the returns to human capital, encouraging workers to enter the modern sector.

*Franco: safranco@bu.edu. Quintero: jmquintero925@uchicago.edu. This is a substantially revised version of "(In)formal Growth: Knowledge Dynamics with Learning Segmentation." We thank Alvaro Contreras and Mohammad Jarrahi for outstanding research assistance. We are grateful to Santiago Caicedo, Levi Crews, Masao Fukui, Chang-Tai Hsieh, Greg Kaplan, David Lagakos, Aleksei Oskolkov, Pascual Restrepo, Esteban Rossi-Hansberg, Rob Shimer, Fernando Alvarez, Rodrigo Adão, and Roman David Zarate for insightful discussions. We thank the Subsecretaria de Prevision Social of Chile for access to the data.

1 Introduction

Human interactions are a crucial source of learning. People learn from others in the workplace, in the places they live, or from individuals in similar occupations who perform similar tasks. Lower-income economies tend to be characterized by dual economies, in which workers are employed across modern and traditional sectors that coexist despite large differences in scale. This idea, central to the study of economic development, goes back to Lewis (1954) and is often reflected in agricultural versus non-agricultural productivity gaps, rural versus urban income differences, and the coexistence of many small firms alongside relatively few large firms within the same industries.[1] In turn, such dual economies likely shape who interacts with whom, segmenting the learning opportunities available to workers. In this paper, we study how these segmented learning markets influence workers' skill formation and, through this channel, economic development. How do segmented learning markets affect workers' learning? How does workers' human capital influence their labor market choices? What are the aggregate implications for income and welfare?

We answer these questions using a quantitative model of learning from peers in a dual economy with segmented interactions, combined with worker panel data from Chile. The model features multiple labor markets and includes modern and traditional sectors. Workers differ in skill, choose where to work, and accumulate human capital by learning from peers in their labor market. Learning is segmented by sector within labor markets, so workers' skill accumulation depends on the skill composition of peers in their own sector. The model predicts that higher-skill workers select into the modern sector and, when learning is sufficiently segmented, this sorting generates a dynamic modern-sector wage premium through faster wage growth. We then test these predictions using the Chilean panel, which allows us to track workers across labor markets and between the modern and traditional sectors. Consistent with the model, modern-sector workers earn higher wages and experience faster wage growth. We provide suggestive evidence that this dynamic premium reflects learning from peers: exposure to higher-paid peers predicts higher future wage growth, and modern-sector workers are more exposed to such peers. Finally, we estimate the model and use it to quantify the aggregate income and welfare consequences of segmented learning, as well as the interaction between learning opportunities and workers' incentives to enter the modern sector.

The model is designed to capture key features of dual economies in developing countries. Labor markets differ in productivity and include a modern and a traditional sector. Modern-sector firms operate a scalable constant-returns-to-scale technology but are subject to labor regulations in the form of payroll taxes. Traditional-sector firms avoid these regulations but operate a decreasing-returns-to-scale technology, capturing small-scale production or subsistence self-employment. Workers differ in skill, choose where to work, and accumulate human capital over time. In this sense, our framework relates to recent quantitative dual-economy models with human capital accumulation, such as Ma et al. (2024) and Amaral and Rivera-Padilla (2024). Our main departure is the introduction of a local learning externality: workers learn from peers in their sector but do not internalize how their sectoral choice affects the learning opportunities of others by changing the skill composition of peers within each sector and labor market. This mechanism

[1] Gollin (2014) offers a recent review of the Lewis model and the dual economy concept.

generates different patterns of skill accumulation across sectors and connects our framework to classic work on knowledge spillovers and social learning (Lucas, 2004; Lucas and Moll, 2014), as well as to empirical and quantitative studies of learning from coworkers and urban interactions (Jarosch et al., 2021; Herkenhoff et al., 2024; Crews, 2024; Lhuillier, 2024).

We characterize the steady-state equilibrium of our heterogeneous-worker framework and highlight two key predictions. Within each labor market, higher-skill workers select into the modern sector, where the scalable constant-returns technology makes high productivity particularly valuable despite regulatory costs. In contrast, lower-skill workers find it optimal to operate in the traditional sector, where decreasing returns are less binding at small scale and regulations are avoided. This sorting pattern also shapes workers' learning opportunities. When interactions are sufficiently segmented across sectors, meaning that workers are relatively more exposed to peers in their own sector, modern-sector workers interact with higher-skill peers on average. As a result, they accumulate human capital faster and experience steeper wage growth than traditional-sector workers, generating a dynamic modern-sector wage premium even within the same labor market.

We then examine the model's main prediction using worker panel data from Chile. The country provides a natural setting for our analysis, as its labor market features large formal and informal sectors that map closely onto the model's dual-economy structure. We use the Encuesta de Protección Social (EPS), a nationally representative longitudinal survey that tracks about 17,000 workers over 2002–2016. We classify workers into modern and traditional sectors based on their formality status. Informal workers lack formal labor contracts and are not subject to payroll taxes, consistent with our modeling of the traditional sector. In addition, they face informality costs that rise with output, such as expected fines or avoidance expenses, which generate diminishing returns to production in the traditional sector (Perry et al., 2007; Ulyssea, 2020b). Accordingly, we classify modern-sector workers as those affiliated with pension funds and holding a labor contract. The EPS also provides complete formal and informal employment histories dating back to 1980, allowing us to follow workers across sectors over their careers. With these data in hand, we test the model's main prediction and provide evidence on the learning mechanism that generates it.

First, we show that higher-skilled workers sort into the modern sector. We begin by documenting that workers with higher educational attainment are substantially more likely to work in the modern sector. Then, to capture differences in skill beyond observable characteristics and sector-specific wage differences, we estimate a wage regression with worker fixed effects, controlling for modern-sector status, age, education, occupation, industry, region, firm size, and experience in each sector. We interpret the worker fixed effects as capturing persistent differences in worker skill. The distribution of estimated worker fixed effects among modern-sector workers first-order stochastically dominates that among traditional-sector workers. These results provide evidence consistent with the model's prediction that higher-skilled workers sort into the modern sector.

Second, we provide evidence for the model's dynamic modern-sector wage premium. We analyze individual wage trajectories and examine how wage growth relates to modern-sector participation, controlling for worker characteristics, initial wages, industry, occupation, and firm size. The estimated premium increases over time, reaching approximately 4% after four years. Because our specification includes worker fixed effects, identification comes from workers who change sectoral status, allowing us to compare the wage trajectories of the same worker during periods of modern and traditional employment. The results therefore suggest that transitioning into the modern sector is associated with cumulative wage gains in subsequent years.

Third, we provide suggestive evidence that learning from better-paid peers contributes to the dynamic modern-sector wage premium. Guided by our framework, we use information on workers' region, industry, occupation, and firm size to define the labor markets in which workers interact. Then, in the spirit of Jarosch et al. (2021) and Herkenhoff et al. (2024), we examine whether exposure to higher-paid peers predicts future wage growth. We find that workers exposed to better-paid peers experience significantly higher future wage growth, with a one-standard-deviation increase in peer wages associated with wages that are approximately 4% higher five years later. Moreover, consistent with our sorting results, modern-sector workers are systematically exposed to better-paid peers. Altogether, these findings suggest that peer learning and the sorting of higher-skilled workers into the modern sector contribute to the dynamic modern-sector wage premium.

After confirming the model's main predictions, we turn to its structural estimation using relevant empirical moments from the Chilean data. We recover the parameters governing worker exit, labor-market mobility, and the initial skill distribution using minimum-distance estimators that target the empirical age distribution, the fraction of workers who switch labor markets, and the tail of the entry-wage distribution. We estimate the remaining parameters by simulated method of moments, including those governing the traditional-sector technology, the learning technologies in both sectors, the segmentation of learning across sectors, and the volatility of skills. We target the modern-sector employment share, the variance of modern-sector wages, one-year sector transition probabilities, and five-year within-sector wage-quintile transition probabilities. The estimated model uncovers strong diminishing returns in the traditional sector and substantial learning segmentation. Traditional workers are 2.6 times as likely to learn from a peer in their own sector as from a modern-sector worker who is equally represented in the labor market. The model also reproduces key untargeted moments, including the wage-experience profiles of modern- and traditional-sector workers and the dynamic modern-sector wage premium.

In the final section, we use the estimated model to conduct counterfactual exercises quantifying the role of learning segmentation in shaping income per capita and welfare. We first consider an exercise that reduces the cost of operating in the modern sector by lowering payroll taxes. Effectively, this exercise provides the same learning opportunities to marginal traditional workers who transition into the modern sector. We then contrast this exercise with one in which we directly provide traditional-sector workers with the same learning environment as modern-sector workers by changing their learning technology.

Our first counterfactual exercise reduces the cost of operating in the modern sector by eliminating payroll taxes. This increases the share of modern-sector workers by 12 percentage points and raises income per capita by 1.6% in the long run. The increase in income is fundamentally dynamic. Lower payroll taxes raise the return to human capital and strengthen workers' incentives to transition into the modern sector, where learning opportunities are better and human capital accumulates faster. The welfare gains from improved learning opportunities reach 0.37% in consumption-equivalent terms. These aggregate gains mask substantial heterogeneity. Workers who move into the modern sector gain the most, with average learning-related welfare gains of about 1.2%, while workers who remain in either sector also benefit, though by smaller amounts.

Our second counterfactual directly equalizes learning opportunities across sectors by changing the learning technology faced by traditional-sector workers. This change generates welfare gains for workers who remain in the traditional sector, averaging 0.52% in consumption-equivalent terms. Aggregate welfare gains from learning reach 0.18%, about half of those generated by eliminating payroll taxes. However, despite directly improving learning opportunities for a larger group of workers, aggregate income increases by only 0.02% in the long run.

The contrast between these two exercises highlights an important interaction between learning opportunities and technology adoption. Lowering payroll taxes improves learning opportunities only for workers who transition into the modern sector, but it also raises the value of human capital and encourages workers to make that transition. Better learning and adoption of the modern technology therefore reinforce each other. In contrast, directly improving learning opportunities in the traditional sector reduces workers' incentives to transition into the modern sector, while the value of accumulating human capital increases much less. As a result, more workers remain in the traditional sector, where they operate a smaller-scale technology, despite having access to better learning opportunities. These results provide quantitative evidence of a complementarity between access to better learning opportunities and the incentives to adopt modern technologies.

Related Literature. Our work contributes to four strands of the literature. First, our work is most closely related to the literature on learning from peers. We generalize the ideas in Lucas (2004) by allowing endogenous learning environments in both the modern and traditional sectors, with interactions both within and across sectors, and we provide empirical evidence for this peer-learning mechanism. We are related to the subsequent endogenous growth frameworks of Lucas and Moll (2014), Perla and Tonetti (2014), and Akcigit et al. (2018), but we depart from them by emphasizing how worker sorting across sectors shapes aggregate income in developing economies. We also highlight the idea that workers learn from better peers, as in Jarosch et al. (2021) and Herkenhoff et al. (2024), with the distinction that learning in our framework is not limited to the firm. In this sense, we are closer to Crews (2024) and Lhuillier (2024), who study worker sorting across cities with local learning externalities. Closest to our empirical exercise, Lhuillier (2024) estimates learning from surrounding workers by projecting future wage growth on the wages of nearby individuals in French cities. We implement a similar design across labor markets and sectors in a developing economy. Finally, labor regulations associated with operating in the modern sector generate an endogenous productivity cutoff that separates workers across sectors, akin to the exporting-related sorting mechanism in Perla et al. (2021).

Second, we relate to papers that study how human capital accumulation shapes wage dynamics in developing economies. Building on Lagakos et al. (2018), Ma et al. (2024) show that firm-provided training in developing economies accelerates workers' skill formation, helping explain differences in age-wage profiles between developed and developing economies. We differ from their framework by emphasizing learning from peers as the fundamental driver of differences in human capital accumulation between the modern and traditional sectors. This mechanism can also help rationalize the observed differences in age-wage profiles across developed and developing economies. Along these lines, Engbom (2022) find that countries with higher rates of job-to-job transitions exhibit steeper wage profiles. In contrast, Donovan et al. (2023) document that workers in developing economies transition more frequently across jobs without moving into better-paying ones. Abstracting from labor market frictions, we offer a reconciling explanation for these findings by showing how segmented labor markets in developing countries hinder human capital accumulation, generating flatter wage profiles in the aggregate.

Third, our work connects to the literature on how technological differences across sectors shape aggregate productivity in economic development.[2] Lagakos (2016) shows how retail firms in low-income countries optimally adopt traditional technologies with lower measured productivity when ownership of durable goods is low. Duarte and Restuccia (2019) and Schwartzman (2025) study structural transformation from traditional to modern services and show that the reallocation of demand and workers across these segments is quantitatively important for income differences. We complement these results by showing that worker reallocation across traditional and modern sectors can raise aggregate income by improving the learning opportunities of lower-skill workers.

Closer to our work, Herrendorf and Schoellman (2018) and Amaral and Rivera-Padilla (2024) study how complementarities between workers' human capital investments and firms' technology amplify cross-country income differences. Engbom et al. (2025) shows how low education levels in developing countries give rise to an endogenous dual economy with a prevalent traditional-sector organization of production. We differ from these studies by allowing human capital to accumulate continuously over the life cycle, rather than through a one-time investment, and by emphasizing how peer composition within modern and traditional sectors shapes learning and wage dynamics.

Finally, by linking the modern and traditional sectors to formality status, we contribute to the literature on informal labor markets and their impact on aggregate outcomes. Ulyssea (2020b) provides a comprehensive review of recent studies analyzing firms and workers. Ulyssea (2018) develops a firm dynamics model with intensive and extensive margins of informality, while Dix-Carneiro et al. (2021) evaluates trade gains in economies with large informal labor markets. Both papers use static frameworks that abstract from worker human capital dynamics. In contrast, Meghir et al. (2015) considers a dynamic wage-posting setting but assumes homogeneous and time-invariant worker skills. We depart from these studies by emphasizing wage dynamics arising from ongoing skill accumulation through peer learning. More closely related to our approach are Lopez Garcia (2015) and Bobba et al. (2022), who link informality to schooling investments, and Akcigit et al. (2024) and Lopez-Martin (2019), who study informality's implications for growth. In contrast, we abstract from growth and one-time schooling decisions, and instead emphasize continuous skill accumulation and sectoral mobility over the life cycle, which shape workers' choices between the modern-formal and traditional-informal sectors.[3]

[2] Herrendorf et al. (2014) provides an overview of this literature in the broader context of structural transformation, and Gollin and Rogerson (2025) highlights the recent use of heterogeneous-agent models in this literature.

[3] Ulyssea (2020a) notes that workers' forward-looking formality decisions remain poorly understood.

The rest of the paper is organized as follows. Section 2 outlines our framework and its main predictions. Section 3 describes the data and tests of the model's predictions. Section 4 details our quantification strategy. Section 5 presents our counterfactual exercises, and Section 6 concludes.

2 A Model of Peer Learning in a Dual Economy

In this section, we present a dynamic model of worker sorting across labor markets and between modern and traditional sectors, in which human capital accumulates through learning from peers. We then characterize the stationary equilibrium and derive predictions for worker sorting and wage growth across the two sectors. We test these predictions using Chilean panel data in Section 3.

2.1 Environment

Time is continuous and indexed by $t \geq 0$. The economy is populated by a unit mass of workers who work in one of N labor markets, indexed by $n \in \{1, \ldots, N\}$. Labor markets differ in a productivity term, A_n, which determines the efficiency with which a homogeneous good is produced. Each labor market hosts two sectors, a *modern* sector (M) and a *traditional* sector (T), which differ in the technologies they operate and the regulations they face.

Workers are heterogeneous in their human capital, $z \in \mathbb{R}_+$, which determines the efficiency units of labor they supply. At any instant, a worker is described by the triple (n, S, z), consisting of her labor market n, her sector $S \in \{M, T\}$, and her human capital z. Workers are infinitely lived but exit the economy when a Poisson death shock arrives with intensity $\delta > 0$. Each worker who exits is replaced by a newborn, who draws initial human capital z from an exogenous distribution $\psi(z)$ and is free to choose the labor market in which to begin her career. Workers discount the future at rate $\rho > 0$, derive flow utility $u(c)$ from consumption, and cannot save. Therefore, consumption equals current labor income, inclusive of a proportional government subsidy defined below.

We denote by $N_n^S(z,t)$ the mass of workers in labor market n and sector S with human capital z at time t. The mass of workers with human capital z in market n is $N_n(z,t) = N_n^M(z,t) + N_n^T(z,t)$, and the total population of the market is $L_n(t) = \int_0^\infty N_n(z,t)\,\mathrm{d}z$, with sectoral counterparts $L_n^S(t) = \int_0^\infty N_n^S(z,t)\,\mathrm{d}z$. These cross-sectional distributions are endogenously determined by workers' labor-market and sectoral choices.

2.1.1 Production and Wages

There are two perfectly competitive representative firms in each labor market, one in the modern sector and one in the traditional sector. Both firms produce a homogeneous good that is freely traded across markets and serves as the *numeraire*. The two firms differ in the technologies they operate. A modern-sector firm in market n combines workers in proportion to their human capital, whereas a traditional-sector firm combines workers' human capital through a concave schedule $\phi(\cdot)$:

$$Y_n^M = A_n \int_0^\infty z\,\ell_n^M(z)\,\mathrm{d}z, \qquad Y_n^T = A_n \int_0^\infty \phi(z)\,\ell_n^T(z)\,\mathrm{d}z,$$

where $A_n > 0$ is the productivity of labor market n, and $\ell_n^S(z)$ is the mass of workers with human capital z employed in sector S. Both technologies are linear in labor, so conditional on human capital z, the two sectors differ only in how effectively they translate skill into output. The modern sector is linear in each worker's skill, whereas the traditional sector exhibits diminishing returns to worker skill, capturing small-scale production and subsistence self-employment. Formally, we assume that $\phi(0) = 0$, $\phi(z) \geq 0$, $\phi'(z) > 0$, and $\phi''(z) < 0$ for all z, and that ϕ satisfies the Inada conditions $\lim_{z\to 0} \phi'(z) = \infty$ and $\lim_{z\to\infty} \phi'(z) = 0$.

Labor regulations differ across sectors. The modern-sector firm faces a payroll tax $\tau \geq 0$, while the traditional-sector firm avoids it. Since labor markets are perfectly competitive, firms pay each worker her marginal product of labor, net of regulatory costs, leading to the wage schedules

$$w_n^M(z) = \frac{A_n}{1+\tau} z, \qquad \text{and} \qquad w_n^T(z) = A_n \phi(z), \tag{1}$$

where $w_n^M(z)$ and $w_n^T(z)$ denote the wages of workers with skill z in labor market n in the modern and traditional sectors, respectively. We assume that payroll tax revenues are rebated by a central authority to all workers, regardless of their sector, through a proportional wage subsidy at rate $b(t)$: a worker earning wage w consumes $(1 + b(t))\, w$.

There are two properties of the wage schedules in (1) worth highlighting. First, within a labor market, modern- and traditional-sector wages satisfy a single-crossing property. At low skill levels, traditional-sector wages are higher than modern-sector wages because of both technological differences in production, captured by $\phi(\cdot)$, and the presence of labor regulations in the modern sector. At high skill levels, wages in the modern sector are higher than in the traditional sector because the former increase linearly with worker skill, while the latter increase at a decreasing rate due to the concavity of $\phi(\cdot)$. Hence, absent any dynamics, the wage schedules in (1) imply that, within labor markets, high-skill workers sort into the modern sector and low-skill workers into the traditional sector. Of course, workers care not only about static returns. In Section 2.5, we show how this sorting pattern extends once we account for human capital accumulation.

Second, we assume that firms are perfectly competitive in the sense that they take wages as given. Since production in each sector exhibits constant returns to scale in the number of workers of a given skill level, firms pay each worker her marginal product, leading to the wage schedules in (1). This notion of perfect competition contrasts with Jarosch et al. (2021), who assume that firms can price the learning opportunities they provide to workers. In their framework, wages reflect not only workers' marginal products but also the value of learning at the firm. In contrast, as we explain in Section 2.1.3, learning in our framework occurs at the labor-market level. Hence, we assume that firms cannot price workers' learning opportunities. This assumption implies that wages reflect only technological differences across sectors and labor markets, rather than differences in the learning environments they provide.

2.1.2 Labor Market Mobility

Workers move across labor markets subject to mobility frictions and costs. First, they cannot move instantly across markets. Instead, they remain in their current labor market until they receive a moving opportunity, which arrives according to a Poisson process with intensity λ. Second, when an opportunity arrives, a worker observes a vector of idiosyncratic location tastes $\{\varepsilon_m\}_{m=1}^{N}$, which we assume are independently and identically distributed across workers and markets according to a Type-I extreme value distribution with scale $1/\nu$. The parameter ν governs the dispersion of tastes: as $\nu \to \infty$, tastes vanish and all workers locate in the market offering the highest continuation value, while a smaller ν spreads workers more evenly across markets. After observing the collection of location tastes, a worker in market n may move to any market $m \in \{1, \ldots, N\}$ by paying a bilateral mobility cost ξ_{nm}, with $\xi_{nn} = 0$.

We assume that workers can freely switch sectors within a labor market. Accordingly, we denote by $V_n(z,t)$ the value of a worker with skill z in market n at time t. We later formally define $V_n(z,t)$ in equation (7) as the upper envelope of the sector-specific values.

Conditional on receiving a relocation opportunity, a worker chooses the market that maximizes her continuation value net of mobility costs and idiosyncratic tastes. The continuation value of labor market mobility is

$$\mathcal{M}_n[V](z,t) = \lambda \left(\mathbb{E}_\varepsilon \left[\max_m \{V_m(z,t) - \xi_{nm} + \varepsilon_m\} \right] - V_n(z,t) \right),$$

where the expectation is taken over the idiosyncratic location tastes, ε_m. Under the extreme value assumption, this continuation value takes the closed form

$$\mathcal{M}_n[V](z,t) = \lambda \left[\frac{1}{\nu} \log \left(\sum_{m=1}^{N} e^{\nu(V_m(z,t) - \xi_{nm})} \right) - V_n(z,t) \right], \tag{2}$$

and the probability that a worker with skill z who receives an opportunity in market n relocates to market m is given by

$$m_{nm}(z,t) = \frac{e^{\nu(V_m(z,t)-\xi_{nm})}}{\sum_{k=1}^{N} e^{\nu(V_k(z,t)-\xi_{nk})}}.$$

These shares, together with the death-and-birth process, govern the reallocation of workers across markets in the distributional dynamics below.

2.1.3 Human Capital Accumulation

Workers accumulate skill by learning from other workers in their labor market. However, learning is segmented across sectors and depends on the skill distribution of workers in each sector. Formally, human capital for a worker in market n and sector S follows a geometric Brownian motion:

$$\mathrm{d}\log z_n^S(t) = \gamma Z_n^S(t)\mathrm{d}t + \sigma_S \mathrm{d}W(t), \tag{3}$$

where $\gamma > 0$ governs the importance of learning from peers, $Z_n^S(t)$ is a measure of the stock of human capital in sector S and market n, which we call *vibrancy* and define below, $W(t)$ is a standard Wiener process, and $\sigma_S > 0$ is a sector-specific volatility. The skill evolution equation (3) has an endogenous drift component, which depends on $Z_n^S(t)$, and an exogenous volatility component, disciplined by σ_S, which captures sector-specific shocks to skills.[4]

The *vibrancy* of market n and sector S is defined as

$$Z_n^S(t) = \left[\int_0^\infty z^\theta g_n^S(z,t)\mathrm{d}z\right]^{\frac{1}{\theta}}, \tag{4}$$

where $\theta \geq 1$ and $g_n^S(z,t)$ is the effective peer-skill distribution faced by workers in market n and sector S, which we specify below.[5] The parameter θ governs how different parts of the skill distribution combine in fostering human capital accumulation. In general, vibrancy is higher in environments populated by more skilled peers. When $\theta = 1$, vibrancy is the mean of peer human capital, so skills are perfect substitutes in learning. As $\theta \to \infty$, vibrancy is determined by the human capital of the most skilled worker. Thus, θ controls the extent to which learning depends on the upper tail of the peer-skill distribution within sectors and labor markets.

The effective peer-skill distribution $g_n^S(z,t)$ is a mixture distribution that places weight π^S on workers in sector S and weight $1-\pi^S$ on workers in the other sector, which we denote by $-S$. Hence, the effective peer-skill distribution takes the form

$$g_n^S(z,t) = \frac{\pi^S N_n^S(z,t) + (1-\pi^S)N_n^{-S}(z,t)}{\pi^S L_n^S(t) + (1-\pi^S)L_n^{-S}(t)}. \tag{5}$$

This expression is a well-defined density over workers' skill in market n and sector S: $g_n^S(z,t) \geq 0$ for all z, and $\int_0^\infty g_n^S(z,t)\,\mathrm{d}z = 1$.

The intuition behind (5) is as follows. Workers in market n can potentially be exposed to any worker with skill level z in the same labor market. However, exposure is not uniform across sectors. The parameter π^S captures the degree of *learning segmentation* by measuring how biased the interactions of workers in sector S are toward their own sector. When $\pi^S = 1$, workers in sector S learn only from workers in their own sector, so learning is fully segmented. When $\pi^S = 0$, workers in sector S learn only from workers in the other sector. Finally, when $\pi^S = 1/2$, workers are equally exposed to both sectors and learning is fully integrated.

[4]One example of such sector-specific skill shocks could be differences in working conditions across sectors.

[5]Our general theory also admits $0 < \theta < 1$. However, restricting attention to $\theta \geq 1$ simplifies the sufficient conditions for a sorting equilibrium below.

2.2 Worker Values and Labor Supply

We now characterize workers' dynamic labor supply decisions. A worker in market n and sector S receives flow utility $u\big((1+b(t))\,w_n^S(z)\big)$ from consuming her subsidized wage. Over time, she accumulates human capital according to (3), can reallocate across labor markets when a moving opportunity arrives, and exits the economy when a death shock arrives.

To write the worker's problem, it is useful to express the law of motion for human capital in levels. By Itô's lemma, z evolves as

$$\mathrm{d}z_n^S(t) = \mu_n^S(t)\,z\,\mathrm{d}t + \sigma_S\,z\,\mathrm{d}W(t), \qquad \mu_n^S(t) \equiv \gamma\,Z_n^S(t) + \frac{1}{2}\sigma_S^2. \tag{6}$$

With this notation, the value of a worker with skill z in market n satisfies the Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{aligned}(\rho+\delta)\,V_n(z,t) - \partial_t V_n(z,t) = \max_{S\in\{M,T\}} &\Big\{u\big((1+b(t))\,w_n^S(z)\big) + \mu_n^S(t)\,z\,\partial_z V_n(z,t) + \frac{\sigma_S^2}{2}\,z^2\,\partial_{zz}V_n(z,t)\Big\} \\ &+ \mathcal{M}_n[V](z,t).\end{aligned} \tag{7}$$

Equation (7) reveals four forces that shape workers' sectoral and labor-market decisions. The left-hand side captures the effective discounting of the worker's value, adjusted for the death rate and net of changes in the value function over time. On the right-hand side, the first term is the flow value of consumption. The second term captures the continuation value from the expected growth of human capital, which is endogenously determined by peer composition in each sector. The third term captures the effect of skill volatility on the worker's value, which differs across sectors because skill shocks are sector specific. Finally, the last term is the option value of mobility across labor markets, which does not depend on workers' sectoral choices because they can freely switch sectors within a labor market.

How do workers choose which sector to work in? Since there are no within-market reallocation costs, sector choice is embedded directly in the worker's dynamic problem and the labor supply decision $S_n^*(z,t)$ attains the maximum in (7).

The optimal sectoral choice, together with the choice of destination market embedded in the mobility shares $m_{nm}(z,t)$, fully summarizes workers' policies. We characterize this labor supply decision in Section 2.5.

2.3 Skill Distribution Dynamics

We now derive the law of motion for the distribution of workers' human capital across markets and sectors. First, the flux of mass through skill level z in market n and sector S is

$$J_n^S(z,t) = \big(\mu_n^S(t) - \sigma_S^2\big)\,z\,N_n^S(z,t) - \frac{\sigma_S^2}{2}z^2\,\partial_z N_n^S(z,t). \tag{8}$$

Let $\mathbb{1}_n^S(z,t)$ be an indicator equal to one if sector S is optimal for a worker with skill z in market n, as implied by (7), and zero otherwise. Then, the mass of workers with skill level z in labor market n and sector S evolves according to the Kolmogorov Forward Equation (KFE):

$$\begin{aligned}\partial_t N_n^S(z,t) = -\partial_z J_n^S(z,t) + \lambda \left(\sum_{k=1}^{N} m_{kn}(z,t) N_k(z,t) \mathbb{1}_n^S(z,t) - N_n^S(z,t) \right) \\ + \delta \left[\psi(z) \Pi_n(z,t) \mathbb{1}_n^S(z,t) - N_n^S(z,t) \right],\end{aligned} \tag{9}$$

where $\Pi_n(z,t)$ is the probability that a newborn with skill z chooses market n:

$$\Pi_n(z,t) = \frac{e^{\nu V_n(z,t)}}{\sum_{k=1}^{N} e^{\nu V_k(z,t)}}. \tag{10}$$

The KFE in (9) combines three forces that shape the distribution of workers across sectors and labor markets. Anticipating the sorting equilibrium structure characterized in Proposition 1, sectoral decisions take the form of a threshold, $\bar{z}_n(t)$, such that workers with skill above the threshold choose the modern sector, while workers with skill below it choose the traditional sector. With this structure in mind, we now interpret each of the three terms in (9).

The first term, $-\partial_z J_n^S(z,t)$, captures changes in the distribution induced by human capital dynamics. By (6), workers' skill grows in expectation at rate $\mu_n^S(t)$, reflecting learning from peers, and fluctuates with idiosyncratic shocks scaled by σ_S. Away from the threshold $\bar{z}_n(t)$, both forces reshape the skill distribution within a sector. Near the threshold, they can also induce workers to switch sectors. Traditional workers who accumulate enough human capital find it optimal to move into the modern sector. Similarly, positive skill shocks can push traditional workers above the threshold, while negative shocks can move modern workers back into the traditional sector.

The second term captures the net reallocation of workers across labor markets. Workers who receive a moving opportunity choose among labor markets according to the choice probabilities $m_{kn}(z,t)$. Once in the destination market, they enter the sector that delivers the highest value, as indicated by $\mathbb{1}_n^S(z,t)$. Because the modern-sector threshold varies across labor markets, moving across markets can also induce workers to switch sectors. For example, a worker employed in the modern sector in her origin market may choose the traditional sector in the destination market because of differences in learning opportunities.

Finally, the third term captures death turnover. Workers exit the economy at rate δ, while newborns enter with human capital drawn from $\psi(z)$ and choose market n with probability $\Pi_n(z,t)$. After choosing a market, newborns enter the sector that delivers the highest value. Equation (10) reflects that newborns, like incumbent workers who relocate, draw idiosyncratic location tastes and sort into the market that offers them the highest value. This channel changes the skill distribution because newborns draw new initial skill levels. It also reallocates mass across labor markets and sectors, since newborns may choose a different market or sector from the workers they replace.

The system is closed by a boundary condition stating that mass vanishes in the upper tail:

$$\lim_{z\to\infty} N_n^S(z,t) = 0, \qquad S \in \{M, T\}.$$

2.4 Equilibrium Definition

We are interested in a stationary equilibrium of this economy in which the cross-sectional distributions are time invariant, $\partial_t N_n^S(z,t) = 0$, and so are the value functions, $\partial_t V_n(z,t) = 0$. We therefore drop the time index from all equilibrium objects. Although individual workers relocate across markets, switch sectors, and accumulate skill over their lives, aggregate objects and the joint distribution of workers across markets, sectors, and human capital remain unchanged.

A stationary equilibrium consists of wages $w_n^M(z)$ and $w_n^T(z)$, value $V_n(z)$, labor supply decisions $S_n^*(z)$, sectoral masses $N_n^M(z)$ and $N_n^T(z)$, effective peer-skill distributions $g_n^M(z)$ and $g_n^T(z)$, vibrancies Z_n^M and Z_n^T, and subsidy rate b, such that:

(i) Optimal labor demand. Firms maximize profits, so wages satisfy (1).

(ii) Workers' dynamic problem. The value $V_n(z,t)$ and labor supply $S_n^*(z)$ are given by (7).

(iii) Learning consistency. The effective peer-skill distributions $g_n^S(z)$ and the vibrancies Z_n^S are given by (4) and (5).

(iv) Distribution dynamics. The sectoral masses evolve according to the Kolmogorov forward equation (9), together with the boundary condition that mass vanishes in the upper tail.

(v) Government balanced budget. The government runs a balanced-budget period by period,

$$b \sum_{n=1}^{N} \sum_{S\in\{M,T\}} \int_0^\infty w_n^S(z)\, N_n^S(z)\, \mathrm{d}z = \sum_{n=1}^{N} \int_0^\infty \tau\, w_n^M(z)\, N_n^M(z)\, \mathrm{d}z.$$

(vi) Labor market clearing. Labor demand equals labor supply in every market and sector, $\ell_n^S(z) = N_n^S(z)$, and the population adds up to the unit mass of workers,

$$\sum_{n=1}^{N} L_n(t) = 1.$$

The definition of equilibrium highlights the local learning externality. Individual workers take vibrancies as given when making their labor supply decisions. Nevertheless, these vibrancies are endogenously determined by the resulting mass of workers across markets and sectors. In turn, workers' optimal sectoral choices affect the human capital evolution of others.

We solve for the stationary equilibrium numerically using a finite-difference scheme, following Achdou et al. (2022). Appendix B provides a detailed description of the algorithm.

2.5 Equilibrium Characterization

We now characterize the stationary equilibrium in three steps. First, conditional on the existence of values and an invariant distribution, we provide conditions under which workers sort across sectors, with more productive workers selecting into the modern sector. Second, we show that, under additional conditions, such a *sorting equilibrium* exists. Third, we discuss uniqueness. All proofs, together with auxiliary lemmas, are relegated to Appendix A.

From (7), it is useful to define the *flow advantage* of the modern sector,

$$\Delta_n(z) \equiv \underbrace{u\big((1+b)w_n^M(z)\big) - u\big((1+b)w_n^T(z)\big)}_{\text{static gain}} + \underbrace{\left(\mu_n^M - \mu_n^T\right) zV_n'(z) + \frac{\sigma_M^2 - \sigma_T^2}{2} z^2 V_n''(z)}_{\text{dynamic gain}}, \tag{11}$$

so that $S_n^*(z) = M$ exactly when $\Delta_n(z) \geq 0$. The static gain compares the flow utility workers obtain across sectors, which is determined by the potential wages in each sector. The dynamic gain captures differences in the human capital dynamics offered by each sector. First, sectors differ in expected skill growth, pinned down by μ_n^M and μ_n^T. These expected gains are valued through the slope of the value function. Second, sectors expose workers to different skill shocks, determined by σ_M^2 and σ_T^2, which are valued through the curvature of the value function. Worker sorting takes a cutoff form in market n when $\Delta_n(z)$ crosses zero only once, from below.

We first introduce two functional-form assumptions for utility and returns in the traditional sector under which we obtain sharp sorting results. Importantly, these are the same functional form assumptions that we impose in the quantitative analysis in Section 4.

Assumption 1 (Functional forms)**.** Workers have log utility, $u(c) = \log c$, and the traditional technology is $\phi(z) = z^\eta$, with $\eta \in (0, 1)$.

Under Assumption 1, we can characterize the sorting cutoff in the static version of our model. Let

$$\hat{z} \equiv (1+\tau)^{\frac{1}{1-\eta}}. \tag{12}$$

Then, from (11), absent any dynamic gains, workers with $z \geq \hat{z}$ select into the modern sector, while workers with $z < \hat{z}$ select into the traditional sector. Intuitively, this static cutoff is increasing in τ: higher payroll taxes in the modern sector depress wages in that sector, making it less attractive for marginal workers. Similarly, the cutoff is increasing in η: a more scalable traditional-sector technology increases wages in the traditional sector, making it more appealing for marginal workers. The static cutoff is the same across all markets.

How can the static sorting equilibrium be extended to the dynamic setting? The expression in (11) suggests that, as long as the modern sector provides better learning opportunities than the traditional sector in each market, $Z_n^M \geq Z_n^T$, and the volatility of skill shocks does not differ too much across sectors, the static sorting result should extend to the dynamic learning environment. Assumption 2 introduces a regularity condition on the learning segmentation parameters that helps ensure the former condition holds.

Assumption 2 (Consistent segmentation)**.** The segmentation parameters satisfy $\pi^M + \pi^T \geq 1$.

Assumption 2 rules out learning interactions that are biased toward the opposite sector on net. It holds, in particular, whenever workers in each sector are at least as likely to interact with peers in their own sector as with peers in the other sector, $\pi^M \geq 1/2$ and $\pi^T \geq 1/2$, including the case of random interactions, $\pi^M = \pi^T = 1/2$.[6]

Finally, to ensure that vibrancies across markets are finite and human capital does not grow too fast, we impose a regularity condition on the skill distribution of new workers.

Assumption 3 (Turnover dominates learning)**.** Let $m_\psi(\theta)$ be the θ-th moment of the entry distribution $\psi(\cdot)$ and let $\sigma \equiv \max\{\sigma_M, \sigma_T\}$. Then, $\theta^2\sigma^2/2 < \delta$, $m_\psi(\theta') < \infty$ for some $\theta' > \theta$, and

$$m_\psi(\theta) \leq \kappa \frac{\left(\delta - \theta^2 \frac{\sigma^2}{2}\right)^{\theta+1}}{(\theta+1)^{\theta+1}\,\gamma^\theta\,\delta}, \tag{13}$$

where $\kappa \in (0,1]$ is a constant defined in (A5) in Appendix A.

Assumption 3 requires turnover to be sufficiently high relative to the learning process and the upper tail of the entry skill distribution. In our quantitative exercise, the entry distribution is Pareto with unit scale and shape parameter ϑ, so that $m_\psi(\theta) = \vartheta/(\vartheta - \theta)$ provided $\vartheta > \theta$. Assumption 3 then reduces to a lower bound on the Pareto shape parameter, $\vartheta \geq \theta R/(R-1)$, where R denotes the right-hand side of (13). Hence, the newborn skill distribution must be sufficiently thin-tailed for a stationary equilibrium to exist. Importantly, this condition is more likely to hold when turnover δ is high or the importance of learning from peers, γ, is sufficiently low.

Assumption 3 also has a natural mathematical interpretation. It is a drift condition in the sense of Meyn and Tweedie (1993): worker turnover pulls the skill distribution back toward the entry distribution faster than peer learning spreads it out.

Proposition 1 (Existence of a Sorting Equilibrium)**.** Suppose Assumptions 1–3 hold. Then, there exists $\Sigma > 0$ such that, whenever $|\sigma_M - \sigma_T| < \Sigma$, a stationary equilibrium with positive sorting exists: there are finite cutoffs $\bar{z}_n$ such that $S_n^*(z) = T$ for $z < \bar{z}_n$ and $S_n^*(z) = M$ for $z \geq \bar{z}_n$. Moreover, if learning is sufficiently segmented, $\pi^M > 1/2$ and $\pi^T > 1/2$, then $\bar{z}_n < \hat{z}$ for all n.

Proof. See Appendix A.2.1. □

Proposition 1 formalizes the main mechanism of the model. Since workers sort across sectors based on their human capital, the modern sector attracts the most skilled workers. This concentration of skilled workers makes the modern sector a better learning environment, so forward-looking workers with sufficiently high skill find it optimal to choose that sector. Sorting reinforces itself: skilled workers select into the modern sector, their presence improves learning in that sector, and better learning opportunities foster faster human capital accumulation relative to the traditional sector.

[6] Assumption 2 turns out to be the empirically relevant configuration, since our estimates in Section 4, which do not impose this assumption, reveal own-sector bias in both sectors.

Workers with very low levels of human capital, instead, value more the static gains offered by the traditional sector, at the expense of future wage growth.

This logic also explains why the dynamic cutoff lies below the static cutoff when learning is sufficiently segmented. In the static model, workers choose sectors only based on current wages. In the dynamic model, they also value future skill growth. Since the modern sector provides better learning opportunities, some workers enter the modern sector even when their current wage would be higher in the traditional sector. These marginal workers are willing to exchange lower wages today for better learning opportunities that imply higher wages in the future. Therefore, more workers select into the modern sector than static wage comparisons alone would imply. The condition that sectoral skill volatility is not too different ensures that this learning advantage is not overturned by differences in risk.

Having established the existence of a sorting equilibrium, a natural question arises: when is this equilibrium unique? Conditional on the vector of vibrancies across markets and sectors, $Z \equiv \{Z_n^M, Z_n^T\}_{n=1}^N$, all remaining equilibrium objects are unique. Values are uniquely determined by (7), the cutoff policy is unique because the flow advantage is strictly increasing in $\log z$, the invariant distribution is pinned down by (9), and the subsidy rate is pinned down by the government budget constraint. Any multiplicity must therefore arise in the finite-dimensional fixed-point problem $Z = T(Z)$. Our next result shows that when the learning externality is not too strong, this fixed point is unique as well.

Proposition 2 (Uniqueness)**.** Suppose Assumptions 1–3 hold, and let $|\sigma_M - \sigma_T| < \Sigma$ as in Proposition 1. Then, there exists $\bar{\gamma} > 0$ such that, for all $\gamma \in [0, \bar{\gamma})$, the sorting equilibrium of Proposition 1 is the unique stationary equilibrium whose vibrancies satisfy $Z_n^S \leq Z_H^*$ for all n and S, where the threshold Z_H^* is defined in (A1) in Appendix A.

Proof. See Appendix A.2.2. □

Proposition 2 says that when the peer-learning channel is not too strong in determining human capital accumulation, the sorting equilibrium is unique. As is standard in models with externalities, we cannot rule out multiple equilibria when learning from peers is sufficiently strong. What can differ across equilibria is how much human capital the sorting equilibrium can sustain. In one possible equilibrium, workers anticipate strong learning in the modern sector, enter the sector earlier in their careers, and spend more time accumulating skills at a fast rate. As a result, the stock of human capital in the modern sector remains high, validating workers' initial anticipation. In another possible equilibrium, workers anticipate weak learning, enter the modern sector only when current wages justify it, and accumulate skills more slowly. The same market then settles at a permanently lower human capital stock. Across markets, expectations and history may similarly determine which location becomes the high-vibrancy hub when fundamentals are similar. Equilibria with vibrancies above the threshold Z_H^*, if any exist, are precisely such self-fulfilling high-learning configurations, which is why Proposition 2 restricts attention to vibrancies below the threshold.

In our framework, three forces work against this multiplicity. First, idiosyncratic tastes disperse workers across markets. Second, workers continuously exit the labor market, bringing in new workers from the common entry distribution. Third, productivity differences break symmetry across markets. In our quantitative implementation, we test for multiplicity by initializing our algorithm from different starting distributions. Reassuringly, the predictions we derive next rely only on the sorting structure in Proposition 1, and therefore hold in any sorting equilibrium.

2.6 Model Predictions

The stationary equilibrium delivers two main predictions that organize our empirical analysis. The first is the sorting equilibrium established in Proposition 1. The second is a dynamic modern-sector wage premium. Conditional on the labor market, modern-sector workers experience faster expected wage growth than traditional-sector workers, as formalized in the following lemma.

Lemma 1 (Dynamic modern-sector premium)**.** *Consider a sorting equilibrium as in Proposition 1, and fix a market n. If learning is sufficiently segmented, $\pi^M > 1/2$ and $\pi^T > 1/2$, then there exists $\bar{h} > 0$ such that, for every horizon $h \in (0, \bar{h}]$, the expected difference in log-wage growth between modern- and traditional-sector workers in market n from t to $t+h$ is strictly positive:*

$$\mathbb{E}_t\left[\Delta \log w^M_{n,t+h}\right] - \mathbb{E}_t\left[\Delta \log w^T_{n,t+h}\right] \geq \gamma\left(Z^M_n - \eta\, Z^T_n\right) h > 0.$$

Proof. See Appendix A.2.3. □

Lemma 1 is the main novel testable prediction of our theory. When learning is segmented, workers in both sectors interact more frequently with peers from their own sector. Hence, vibrancies in the stationary sorting equilibrium satisfy $Z^M_n > Z^T_n$. This result reflects the fact that the modern sector attracts more skilled workers, so the modern peer-skill distribution first-order stochastically dominates the traditional one. Consequently, modern-sector workers accumulate human capital faster. Importantly, because the traditional technology is concave, a given increase in human capital raises traditional-sector wages with elasticity $\eta < 1$. Hence, although technological differences across sectors are not necessary for the dynamic premium to arise, they amplify it. The reallocation term reinforces the premium: workers who cross the cutoff from below trade the higher traditional wage for the lower modern wage on the band between $\bar{z}_n$ and $\hat{z}$, which lowers measured traditional wage growth and raises measured modern wage growth.

In the next section, we test Proposition 1 and Lemma 1 in the Chilean data. We first test the sorting prediction using the distribution of estimated worker skill across sectors. We then test the dynamic prediction by estimating, at each horizon h, the average difference in wage growth between modern- and traditional-sector workers.

3 Empirical Analysis

In this section, we test the main model predictions. We begin by describing the data and our definitions of modern and traditional jobs. Our analysis focuses on Chile, where we can exploit rich worker survey data from an important developing economy in which a large fraction of workers still operate in the traditional sector.

3.1 Data

Our main data source is the Chilean *Encuesta de Protección Social* (EPS). The EPS is a comprehensive longitudinal survey designed to gather data on various aspects of social protection. The survey spans five waves conducted in 2002, 2004, 2006, 2009, and 2015.[7] We use two blocks of the survey. First, we use the block containing information on individuals' demographics, such as age and education. Second, we use the block containing individuals' job histories dating back to 1980.[8] The job history includes information on job type, occupation, economic sector, region, weekly hours worked, starting and ending dates, type of contract, pension fund contributions, job category, and the number of workers in the firm. Given our definition of modern and traditional jobs below, we restrict our sample to start in the 2004 wave, the year in which individuals with and without pension affiliation were included, making the survey representative of all workers.

We define modern- and traditional-sector jobs in the data in a way that is consistent with our framework. In the model, modern-sector workers are those employed by firms that comply with labor regulations. Accordingly, we define a modern-sector job as one with a formal labor contract and pension contributions, which maps to the payroll tax in the model. We define a traditional-sector job as any job that does not comply with these labor regulations. In doing so, our definitions of modern and traditional employment coincide with the standard distinction between formal and informal work used in the literature (Ulyssea, 2020b) and typically used by statistical agencies.[9]

The EPS offers four advantages for analyzing employment dynamics among modern and traditional workers. First, the detailed information available for each job allows us to classify workers as modern or traditional based on their contract type and pension fund contributions. Second, the data on disposable monthly income and hours worked allow us to construct hourly wages for jobs starting in 2002.[10] Third, although the survey is conducted only in selected years, the job-history block allows us to reconstruct each worker's employment history since 1980. We aggregate workers' jobs to construct our baseline sample, an unbalanced annual panel with more than 17,000 workers observed between 2002 and 2016.[11] The length and coverage of the panel allow us to track modern and traditional workers over long periods. In particular, we observe 50% of workers in our sample for at least five years and 20% for at least ten years.[12] Finally, although we restrict the analysis to jobs active from 2002 onward because of wage availability, we use workers' job histories since 1980 to measure experience in the modern and traditional sectors separately.[13]

[7] We exclude the 2012 and 2019 waves from our analysis. The 2012 wave is not representative at the national level, and the questions in 2019 changed because of the COVID-19 pandemic.

[8] When an individual enters the survey, they are asked for their job history since 1980. If, in a specific wave, the individual was already part of the survey, they are asked about their job history since the survey's last wave.

[9] Chilean labor law requires workers to be enrolled in the pension system and contribute 10% of their salary.

[10] We construct wages for 2002 and 2003 retroactively using information from workers' job histories.

[11] For workers with multiple jobs in a given year, we retain the job with the highest total hours worked.

[12] In contrast, Meghir et al. (2015) and Samaniego De La Parra and Fernández Bujanda (2020) use panels that follow workers for nine months over roughly one year in Brazil and five quarters in Mexico, respectively.

[13] We compute experience by dividing cumulative hours worked up to a particular year by the number of hours

Table 1: Summary Statistics

	(1) Modern	(2) Traditional	(3) All
Fraction of workers	0.70	0.30	...
Mean log wage	0.81	0.44	0.70
Std. deviation log wage	0.63	0.69	0.67
Mean weekly working hours	46.00	42.19	44.85
Fraction of male workers	0.63	0.59	0.62
Mean experience (years)	15.24	15.38	15.28
Transitions (by initial status)			
Frac. of modern workers next year	0.97	0.08	...
Frac. of traditional workers next year	0.03	0.92	...
Number of observations	86,646	35,114	121,760
Number of workers	...	...	17,260

Notes: Table 1 displays summary statistics for the baseline sample 2002-2016. We include male workers aged 15 to 65 and female workers aged 15 to 60. We exclude workers without remuneration, and public-sector and armed forces employees. Wage is the real hourly wage in 2004 US dollars. The variable experience corresponds to the equivalent in years of a full-time work schedule of 48 hours per week.

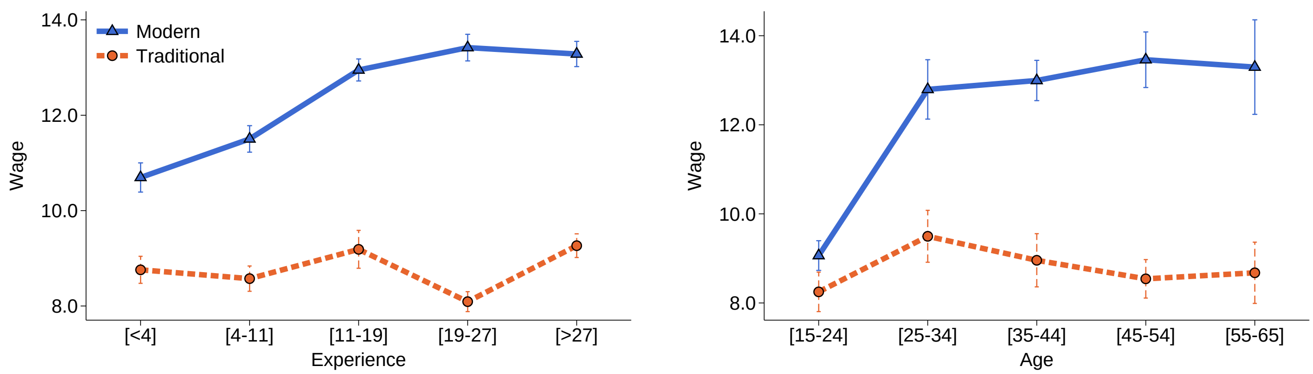
We exclude certain types of workers from our analysis. First, in accordance with Chilean labor regulations, we include male workers aged 15 to 65 and female workers aged 15 to 60, aligning with the country's legal working and retirement ages for men and women, respectively. Second, in our baseline specifications, we omit unpaid workers, as well as public-sector and armed forces employees. Importantly, our data allow us to include self-employed workers and entrepreneurs, who represent a large fraction of traditional-sector workers.[14]

Table 1 displays summary statistics for our baseline sample. Three key points are worth highlighting. First, wages are, on average, higher for modern workers than for traditional workers. Nevertheless, wages among traditional workers are more dispersed than those among modern workers, consistent with previous findings. Second, modern and traditional workers work similar hours per week and have similar years of experience. This similarity supports our modeling choice of abstracting from labor supply decisions and from an explicit allocation of time between working and learning. Third, transition probabilities between the modern and traditional sectors are highly asymmetric. On average, 8% of traditional workers move to the modern sector in the following year, whereas only 3% of modern workers move to the traditional sector. Importantly, these transition probabilities are computed by restricting the sample to active workers, thereby excluding transitions into and out of unemployment. Appendix C provides additional descriptive statistics, including the share of modern workers by age, education, occupation, industry, and firm size.

corresponding to a full-time schedule of 48 hours per week.

[14] We provide robustness exercises in which we restrict the sample to salaried workers only.

Figure 1: Wages Over the Life Cycle for Modern and Traditional Workers



(a) Wage-experience profiles

(b) Wage-age profiles

Notes: Figure 1 displays wage paths over the life cycle for modern and traditional workers. Panel 1(a) shows the average wage for each experience quintile and panel 1(b) for each 10-year age bin. Vertical dashed lines denote 95 percent confidence intervals.

3.2 Testing the Model's Predictions

This section presents empirical evidence on wages for modern and traditional workers that is consistent with our model. We begin by documenting wage-experience profiles, which reveal different wage growth patterns across sectors. We then test the model's main predictions in two steps. First, we present evidence that high-skill workers sort into the modern sector. Second, we test for the presence of a dynamic modern-sector wage premium.

Figure 1 presents wage-experience and wage-age profiles for modern and traditional workers. The left panel displays the wage-experience profiles. We first classify workers as modern or traditional according to our definition, divide them into five experience bins, and compute the average wage within each bin and sector. The right panel displays wage-age profiles, which are constructed analogously using workers' age. Both panels reveal stark differences in life-cycle wage growth between modern and traditional workers.

Figure 1 shows substantial differences in wage levels and wage growth between modern and traditional workers over the life cycle. Focusing first on the wage-experience profiles, modern workers begin their careers with wages around 20% higher than those of traditional workers. This wage gap then widens over the life cycle, reaching more than 40% by the end of workers' careers. The divergence is mainly driven by wage growth among modern workers, while traditional workers' wages remain mostly flat over time. The wage-age profiles display a similar pattern. Although the wage gap early in the life cycle is smaller, at around 12%, it reaches a similar magnitude by the end of workers' careers. Interestingly, the timing of the divergence differs across the two profiles. In the wage-experience profiles, the gap widens gradually over the life cycle, whereas in the wage-age profiles most of the divergence occurs early.

These life-cycle wage patterns for modern and traditional workers align with recent findings on wage growth across countries. Lagakos et al. (2018) document that wages in developed countries increase more over the life cycle than wages in developing economies. Our findings suggest that the modern sector behaves like a "developed-economy labor market", whereas the traditional sector displays patterns typical of a "developing-economy labor market." In Appendix C, we report the same wage profiles restricting the sample to salaried workers. The patterns remain consistent under this alternative specification.

We interpret these initial findings as motivating evidence for the main mechanism of our model. The initial wage differences are consistent with more productive workers sorting into the modern sector. The subsequent widening of the wage gap over the life cycle is also consistent with modern-sector workers experiencing faster human capital accumulation because of better learning opportunities. In the following subsections, we provide more direct tests of both mechanisms.

3.2.1 Worker Sorting

We investigate worker sorting across the modern and traditional sectors in two ways. First, we examine whether workers' observable characteristics that are potentially related to skill are correlated with their sectoral choices. Perhaps not surprisingly, Figure C2 in the Appendix shows that workers with higher educational attainment are more likely to be employed in the modern sector. The share of modern workers among those with no schooling is slightly above 50%. In contrast, more than 80% of workers with two or more years of college education are employed in the modern sector. To the extent that educational attainment is positively correlated with skill, these patterns support the idea that more skilled workers sort into the modern sector.

Second, we study the extent to which workers with higher unobserved skill are more likely to be employed in the modern sector. To do so, we study the modern-traditional wage gap in levels and assess the role of worker sorting in explaining this gap. We estimate the following regression:

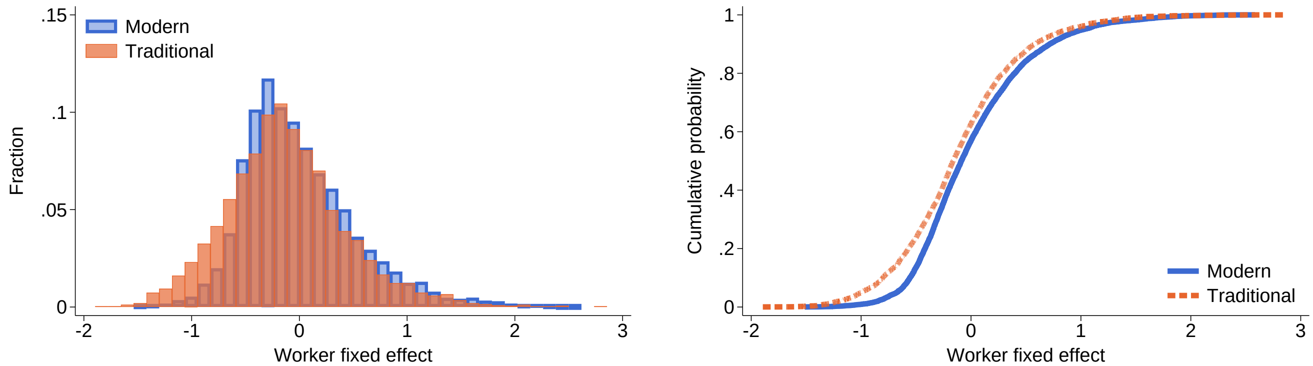
$$\log w_{i,t} = \beta \text{Modern}_{i,t} + \Gamma X_{i,t} + \alpha_t + \alpha_i + \zeta_{i,t}, \tag{14}$$

where $\log w_{i,t}$ is the log wage of worker i at time t, $\text{Modern}_{i,t}$ is an indicator equal to one if worker i is employed in the modern sector in year t and zero if she is employed in the traditional sector, and $X_{i,t}$ is a vector of worker characteristics that includes age, education, gender, occupation, industry, region, firm size, and experience in the modern and traditional sectors. Moreover, α_t and α_i denote time and worker fixed effects, respectively.

The worker fixed effects, α_i, capture time-invariant worker skill that is rewarded in both sectors. Importantly, these fixed effects capture skill differences beyond the observable characteristics included in $X_{i,t}$, such as age, education, and experience in each sector. Our framework predicts that more skilled workers sort into the modern sector. Therefore, through the lens of the model, the distribution of worker fixed effects among modern workers should lie to the right of the corresponding distribution among traditional workers.

Figure 2 presents the distribution of estimated worker fixed effects for modern and traditional workers. We retain one observation per worker and assign each worker to the sector in which she is observed most frequently. The figure provides evidence of worker sorting. Panel 2(a) shows that traditional workers are overrepresented at the bottom of the skill distribution, whereas

Figure 2: Worker Fixed Effects for Modern and Traditional Workers



(a) Histogram

(b) Cumulative distribution

Notes: Figure 2 displays the distribution of the worker fixed effects α_i estimated from equation (14), controlling for age, education, occupation, industry, region, firm size, experience in each sector, and time fixed effects. We keep one observation per worker and assign workers to the sector in which we observe them most often. Panel 2(a) shows the histogram and panel 2(b) the empirical cumulative distribution function of the worker fixed effects for modern and traditional workers.

modern workers are concentrated at the top. Panel 2(b) shows that the cumulative distribution of worker fixed effects for modern workers lies everywhere below that of traditional workers. This implies that the modern-sector skill distribution first-order stochastically dominates the traditional-sector distribution. Moreover, the estimated correlation between workers' fixed effects, α_i, and the modern-sector indicator, $\text{Modern}_{i,t}$, is positive, at approximately 0.10, and statistically significant. Hence, consistent with our model, more skilled workers sort into the modern sector.

Table D1 in Appendix D reports the estimates of β in equation (14). Modern workers earn a wage premium even after conditioning on worker fixed effects and observable characteristics. Through the lens of our model, a wage premium among workers with the same skill level provides evidence of diminishing returns to worker skill in the traditional sector.

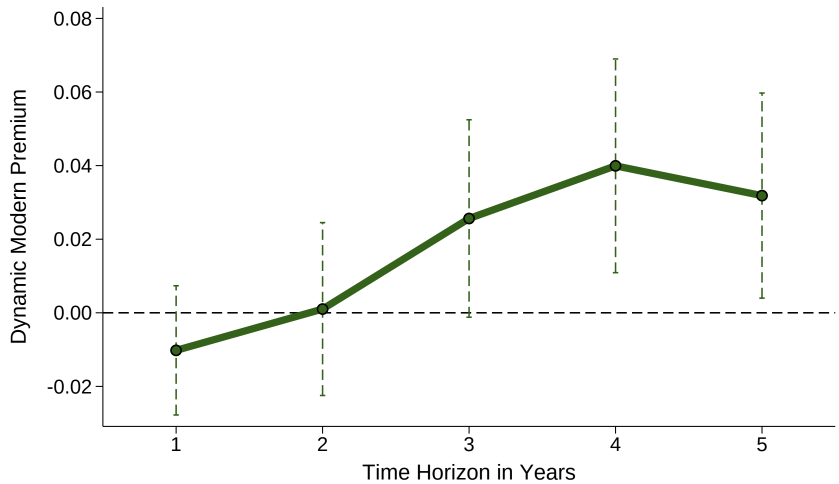
3.2.2 Dynamic Modern-Sector Premium

We now turn to the main novel prediction of our model: the dynamic modern-sector wage premium. We test this prediction in two steps. First, we examine the short- and medium-run effects of working in the modern sector on future wages. Second, we provide evidence that learning from better-paid peers in the modern sector is a plausible mechanism generating these persistent wage differences between modern- and traditional-sector workers.

We begin by studying how employment in the modern sector relates to future wage growth. To do so, we estimate a variant of (14) that relates workers' future wage growth to their sectoral status and a set of observable and unobservable characteristics. Specifically, we estimate the following relationship separately for different time horizons h:

$$\Delta \log w_{i,t+h} = \beta_h \text{Modern}_{i,t} + \Gamma X_{i,t} + \alpha_t + \alpha_i + \zeta_{i,t}, \tag{15}$$

Figure 3: Modern Sector Dynamic Wage Premium



Notes: The figure displays the modern future wage premium based on the estimates of β_h from equation (15) including worker fixed effects. Each dot represents the estimated wage premium at horizon $t+h$: $\exp(\hat{\beta}_h)-1$, for $h = 1, \ldots, 5$. Controls include age, education, gender, occupation, industry, region, experience, wage decile, and time and worker fixed effects. Vertical dotted lines denote 95 percent confidence intervals.

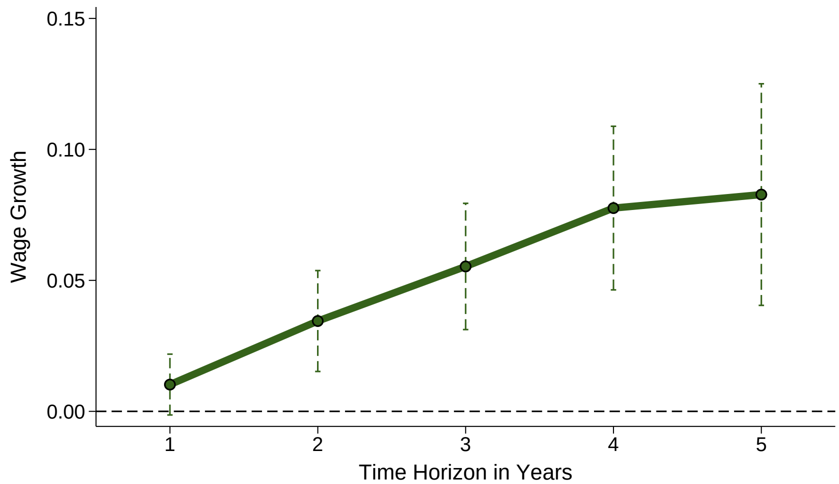
where $\Delta \log w_{i,t+h}$ is the change in worker i's log wage between years t and $t+h$. The vector $X_{i,t}$ includes the same worker characteristics as in (14), together with wage-decile fixed effects. Moreover, α_t and α_i denote time and worker fixed effects, respectively. The parameter of interest is β_h, which captures the average difference in wage growth over h years between modern and traditional workers, conditional on observable characteristics and fixed effects.

Figure 3 presents the estimated future wage premium associated with working in the modern sector across different time horizons. The premium is statistically insignificant at short horizons but increases over time, reaching its peak after four years. At this horizon, modern workers experience wage growth that is approximately 4% higher than that of traditional-sector workers. Crucially, because the specification includes worker fixed effects, identification comes from workers who change sectoral status. This allows us to compare the wage growth trajectories of the same worker during periods of modern and traditional employment and to control for the potential sorting of workers with higher learning ability into the modern sector. Hence, the results suggest that workers who switch from the traditional sector to the modern sector experience cumulative wage gains in subsequent years.

We do not condition on workers remaining in the modern sector in subsequent years. This is consistent with our model, in which human capital is portable across sectors, and it matches the construction of the dynamic premium in Lemma 1, which classifies workers by their sector at the start of the horizon only. In addition, the estimated effects become statistically insignificant after five years. This is natural in our setting, since workers may switch sectors over time in response to idiosyncratic skill shocks.

Which forces drive the dynamic modern-sector wage premium? Following Jarosch et al. (2021) and Herkenhoff et al. (2024), we examine whether interactions with better-paid peers contribute to

Figure 4: Peer Effects on Wage Growth



Notes: Figure displays the effect of peer wages on workers' future wage growth based on the estimates of β_h from equation (16). Peers are defined as all other workers in the same Region × Year × Industry × Occupation × Firm-Size Bracket. Each dot represents the estimated effect at the time horizon $t+h$. Controls include age, education, gender, occupation, industry, region, experience, wage decile, and time fixed effects. Vertical dotted lines denote 95 percent confidence intervals.

workers' wage growth. In our framework, workers interact with others in the same labor market, and access to more skilled peers accelerates human capital accumulation. Accordingly, we define a labor market in the data as a region × industry × occupation × firm-size bracket cell in a given year. For each worker i, we then compute the average wage of all other workers in the same labor market, denoted by $w_{-i,t}$, as a proxy for peer skill. We estimate the following reduced-form equation:

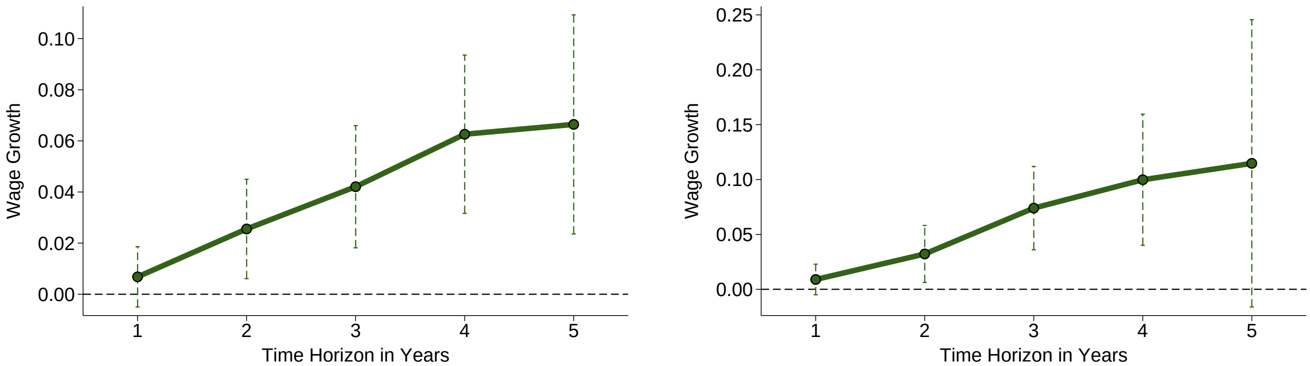
$$\Delta \log w_{i,t+h} = \beta_h \log w_{-i,t} + \Gamma X_{i,t} + \zeta_{i,t}, \tag{16}$$

where β_h captures how exposure to better-paid peers affects a worker's future wage growth, and $X_{i,t}$ contains the same worker-specific controls as in our previous specifications.

Figure 4 presents the estimated effects of peer quality on wage growth across different time horizons. We find strong evidence that workers exposed to better-paid peers experience significantly higher future wage growth. This effect increases with the time horizon, suggesting that peer quality has a cumulative impact on skill accumulation. In terms of magnitude, a one-standard-deviation increase in the average wage of a worker's peers is associated with wages that are approximately 4% higher five years later. These findings align with the literature on learning from coworkers (Jarosch et al., 2021; Herkenhoff et al., 2024) and learning in cities (Crews, 2024; Lhuillier, 2024). They also provide suggestive evidence that the quality of workers' interactions plays an important role in shaping their wage trajectories.

It is natural to consider alternative mechanisms that could drive the dynamic modern-sector premium. Our data allow us to examine whether the peer-learning channel is confounded by two other forces. First, sectors may differ in characteristics other than peer composition that affect wage growth. For instance, consistent with the results of Ma et al. (2024), modern-sector workers may have greater access to on-the-job training, leading to faster human capital accumulation than

Figure 5: Peer Effects on Wage Growth, Robustness



(a) Modern-sector status

(b) Wage-bill growth

Notes: The figure reports estimates of β_h from equation (16) under two robustness specifications. In Panel 5(a) we control for an indicator for modern-sector employment, Modern_{it}. In Panel 5(b) we control for the growth of the labor market's total wage bill between $t-2$ and $t-1$, $t-1$ and t, and t and $t+1$. The remaining controls are the same as in Figure 4. Vertical dotted lines denote 95 percent confidence intervals.

traditional-sector workers. To assess whether our peer-learning estimates are confounded by other sector-specific channels, such as training, we estimate a variant of (16) that directly controls for workers' modern-sector status.[15] Panel 5(a) reports the estimates of β_h from this specification. The estimates remain similar to those in Figure 4, suggesting that the peer-learning channel is not driven solely by other sector characteristics that affect human capital accumulation.

Second, an alternative mechanism is that the estimated effects reflect labor-market-level productivity gains that are passed on to workers rather than learning from peers. Following Jarosch et al. (2021), we estimate a version of (16) that controls for growth in the labor market's total wage bill between $t-2$ and $t-1$, $t-1$ and t, and t and $t+1$.[16] Figure 5(b) displays the estimates, which remain positive and increase with the horizon, suggesting that the results are not driven by common shocks to workers' labor markets. Although the standard errors increase at longer horizons, the point estimates are even larger in magnitude.

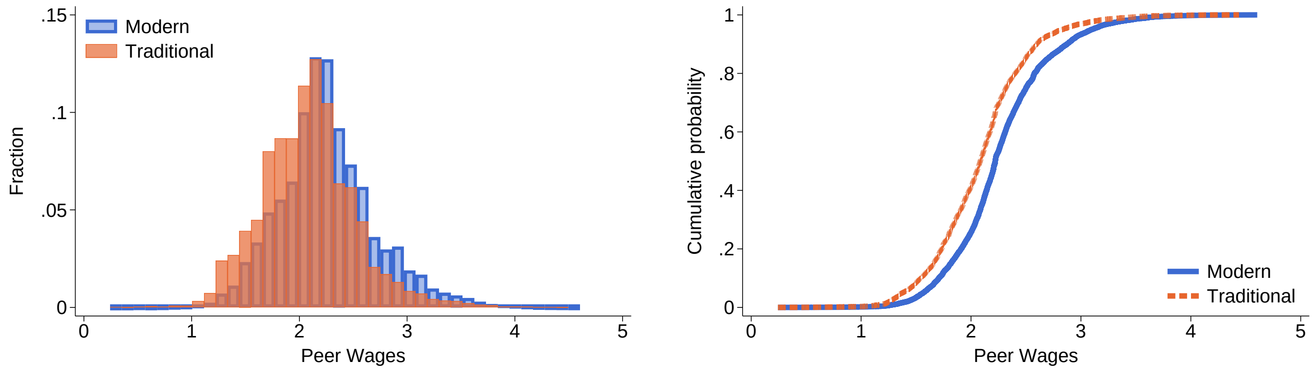
Having established that peers matter for wage growth, we now examine whether modern-sector workers systematically have access to better peers than traditional-sector workers. Figure 6 compares the distribution of peer wages, $w_{-i,t}$, for workers in the two sectors.

The figure reveals that modern workers indeed have better-paid peers. On average, the peers of modern workers earn wages approximately 20% higher than the peers of traditional workers. In addition, the distribution of peer wages for modern workers first-order stochastically dominates that of traditional workers. Because peer wages average over all workers in a labor market, the two distributions naturally overlap: modern and traditional workers who share the same labor market face similar peer groups. The dominance of the modern-sector distribution therefore reflects the sorting of modern workers into labor markets with better-paid peers. This finding provides the

[15] Perhaps surprisingly, participation in formal training programs appears rare in our data. Only 2% of job spells report participation in an employment or training program.

[16] Recall that our data do not allow us to identify the establishments or firms in which workers are employed.

Figure 6: Distribution of Peer Wages for Modern and Traditional Workers



(a) Histogram

(b) Cumulative distribution

Notes: Figure 6 displays the distribution of log peer wages, $\log w_{-i,t}$, for modern and traditional workers. Panel 6(a) shows the histogram and panel 6(b) the empirical cumulative distribution function. Peers are defined as workers in the same Region × Year × Industry × Occupation × Firm-Size Bracket.

missing link in our narrative: modern workers experience higher wage growth partly because they interact with better-paid peers, and these interactions contribute to faster skill accumulation.

We note that this evidence captures differences in peer exposure across labor markets rather than learning segmentation within them. It is rare to observe workers' interactions directly within a labor market, and our data are no exception. Our model addresses this limitation by allowing the frequency of interactions between modern and traditional workers within a labor market to differ. We estimate this learning segmentation in our quantitative analysis and study its implications.

Finally, we provide suggestive evidence supporting our modeling choice for the learning technology. Equation (3) assumes that workers can learn from peers across the entire skill distribution within a labor market, from the least to the most skilled. One way to assess this assumption is to compare the effects of better-paid peers for workers at the top and bottom of the wage distribution. If workers learned only from more skilled peers, we would expect stronger peer effects among low-wage workers, who have a larger pool of more skilled workers from whom to learn. Conversely, if workers learned only from less skilled peers, we would expect stronger effects among high-wage workers. To explore this asymmetry, we estimate the following specification:

$$\Delta \log w_{i,t+h} = \beta_h^+ \, \text{Above}_{i,t} \times \log w_{-i,t} + \beta_h^- \, \text{Below}_{i,t} \times \log w_{-i,t} + \Gamma X_{i,t} + \zeta_{i,t}, \tag{17}$$

where $\text{Above}_{i,t}$ and $\text{Below}_{i,t}$ are indicator variables equal to one when worker i is above or below the mean wage of her peers, respectively.[17]

Table D2 in Appendix D reports the estimates of β_h^+ and β_h^- across different time horizons. Both groups display positive and sustained responses to peer quality at every horizon. While the point estimates are larger for workers above the mean, both sets of coefficients are economically and statistically significant, and they converge at the longest horizons. We interpret this as suggestive

[17] We also include the non-interacted indicator $\text{Above}_{i,t}$ in the vector of controls $X_{i,t}$.

evidence that workers learn from the full pool of peers in their labor market rather than only from a subset of them.

3.2.3 Taking Stock and Alternative Mechanisms

In this section, we provided evidence for the model's main predictions and the mechanism connecting them. Modern-sector workers earn higher wages throughout their careers, and a substantial share of this wage gap reflects the sorting of higher-skilled workers into the modern sector. We also documented a dynamic modern-sector wage premium, with modern workers experiencing faster future wage growth than traditional workers. Consistent with the model, exposure to better-paid peers predicts higher subsequent wage growth, and modern-sector workers are systematically exposed to such peers. These sorting patterns create stronger learning opportunities in the modern sector, contributing to faster human capital accumulation and steeper wage trajectories, while wages in the traditional sector remain relatively flat over the life cycle.

Peer learning, however, may not be the only force shaping human capital accumulation across sectors. Although our peer-effect estimates remain robust after controlling for modern-sector status, more detailed information on firm-provided training, such as that used by Ma et al. (2024), would be necessary to quantify the contribution of on-the-job training to the wage trajectories we document. Traditional-sector workers may also perform tasks that provide weaker incentives to accumulate skills than those typically performed by modern-sector workers. We do not expect this channel to be quantitatively dominant, since all our specifications include industry and occupation fixed effects. Moreover, as shown in Table 1, modern and traditional workers work similar hours. While we abstract from workers' allocation of time between working and learning, standard human capital accumulation models such as Ben-Porath (1967) view hours worked as informative about this tradeoff. Through the lens of these models, the similarity in working hours suggests that workers in the two sectors allocate similar shares of their time to these activities.

Search and matching frictions provide another potential explanation for the different wage trajectories observed in the data. An argument similar to that of Lagakos et al. (2018) regarding differences in life-cycle wage profiles across developed and developing countries applies in our setting. Traditional-sector labor markets may be subject to different search and matching frictions from those in the modern sector, generating flatter wage growth over the medium and long run. The findings of Donovan et al. (2023) suggest that traditional-sector workers may not experience less labor market mobility, but may instead be exposed to riskier jobs and a more precarious job ladder. Under this interpretation, they change jobs more frequently without obtaining the wage gains typically associated with moving up the job ladder. Distinguishing between human capital and search-based explanations is difficult even with detailed panel data such as ours. We therefore view search and matching frictions as a valid and potentially important alternative mechanism. Nevertheless, the relationship we document between peer exposure and subsequent wage growth provides separate support for the peer-learning channel. Quantifying the relative contributions of these mechanisms remains an important avenue for future research.

Taken together, our findings support worker sorting and peer learning as important forces shaping wage trajectories across sectors, while leaving room for complementary mechanisms. In the next section, we quantify the model and assess the aggregate importance of peer learning and learning segmentation for income and welfare.

4 Model Quantification

This section describes the quantification of the model. We first summarize the parameters introduced in the theory and explain the strategies used to estimate them. We then discuss the variation in the data that identifies the parameters and characterize the decentralized equilibrium under the estimated parameter values.

4.1 Setup

Because time is continuous in the model, we interpret each interval $[t, t+1)$ as one year. We average wages and transition probabilities across the modern and traditional sectors over the 2002–2016 period and interpret these averages as representing the economy's steady state. We also adopt the functional-form assumptions for utility and traditional-sector technology stated in Assumption 1 and assume that the entry skill distribution is Pareto with scale parameter normalized to one and shape parameter ϑ.

We define labor markets in a manner consistent with the empirical exercises. The EPS reports information on 14 regions, 10 industries, 10 occupations, and firm size. To make the estimation tractable, we abstract from firm size and define a labor market as a Region $\times$ Industry $\times$ Occupation cell.[18] We then group regions, industries, and occupations into three categories each. For regions, we distinguish among the north, Santiago, and the south. For industries and occupations, we construct high, medium, and low wage groups. Our quantitative exercises therefore include $N = 27$ labor markets. Table D3 in the Appendix describes the composition of each group.[19]

4.2 Estimation

Under our functional-form assumptions, the model contains 13 parameters:

$$\left\{\rho, \tau, \nu, \vartheta, \lambda, \delta, \eta, \theta, \gamma, \pi^M, \pi^T, \sigma_M, \sigma_T\right\},$$

in addition to labor-market productivity, $\{A_n\}_{n=1}^N$, and bilateral migration costs, $\{\xi_{nm}\}_{n,m=1}^N$. We divide these parameters into three groups according to the strategy used to estimate each of them.

External Calibration (3 parameters): ρ, τ, ν

We set workers' discount rate, ρ, to 0.05, a standard value in the literature (Akcigit et al., 2021). We set the payroll tax in the modern sector to $\tau = 0.10$, consistent with Chilean statutory regulations.

[18] Appendix C.3 shows that the main results from our analysis are robust to omitting firm size as a control.

[19] This grouping is necessary because of computational constraints. Without it, the model would contain 1400 labor markets, each requiring us to track a separate skill distribution.

Finally, we set the inverse dispersion parameter governing workers' idiosyncratic preferences across labor markets to $\nu = 1/1.5$, following Traiberman (2019).

Given this value of ν, we estimate bilateral migration costs, $\{\xi_{nm}\}_{n,m=1}^{N}$, nonparametrically using the bilateral migration flows across labor markets observed in the EPS. To do so, we impose symmetric migration costs, $\xi_{nm} = \xi_{mn}$. Appendix B.4 describes the recovery procedure.

Minimum Distance Estimated Parameters (3 parameters): δ, λ, ϑ

We estimate the death rate, δ, from the age distribution implied by the model. Letting $H(a)$ denote the age CDF, the model implies the log-linear relationship

$$\log\left(1 - H(a)\right) = -\delta a.$$

We therefore estimate δ by regressing $\log\left(1 - H^e(a)\right)$ on age a without a constant, where $H^e(a)$ denotes the empirical age CDF.

For the migration rate, λ, we use the worker panel and set it equal to the overall fraction of workers who change labor markets in a given year. In the model, migration opportunities arrive according to a Poisson process with intensity λ, so that over a one-year interval, a fraction $1 - e^{-\lambda} \approx \lambda$ of workers receive at least one opportunity to move. Using the EPS, we identify movers and non-movers and compute λ as the average fraction of workers who move across labor markets between consecutive years.[20]

We estimate the Pareto shape parameter ϑ from the cross-section of entry wages. Because modern-sector wages are proportional to skill, $w_n^M(z) = A_n z/(1+\tau)$, the upper tail of newborn wages within each labor market inherits the Pareto tail index ϑ, up to the market-specific scale factor $A_n/(1+\tau)$. We define modern-sector newborns as formal workers aged 25 or younger with fewer than two years of experience and use each worker's first observed wage. Within each labor market, we rank workers by wage and retain the top quartile. We then estimate the rank-size regression

$$\log\left(\text{rank}_{i,n} - \tfrac{1}{2}\right) = c_n - \vartheta \log w_{i,n} + \zeta_{i,n},$$

where the market fixed effect c_n absorbs differences in the local wage scale, while the coefficient on log wages identifies ϑ. The $-\frac{1}{2}$ rank correction reduces the small-sample bias of the OLS tail-exponent estimator (Gabaix and Ibragimov, 2011).

Our estimation yields $\delta = 0.04$, $\lambda = 0.101$, and $\vartheta = 2.66$. The estimated value of δ implies that approximately 4% of workers exit employment each year and that active workers have an expected remaining work life of about 25 years. This value exceeds Chile's overall mortality rate, which is natural because δ captures not only death but also transitions from employment into unemployment or out of the labor force. Our estimate of λ lies toward the lower end of the range reported by Crews (2024) for the United States. Finally, our estimate of ϑ implies a degree

[20]Because workers who receive a relocation opportunity may optimally remain in their current market, this estimate provides a lower bound on the arrival intensity of relocation opportunities.

of dispersion in entry skills broadly consistent with Hsieh et al. (2019), who calibrate a Fréchet distribution of talent with a shape parameter of 2 to match wage dispersion in the United States.

Simulated Method of Moments (7 parameters): $\eta, \theta, \gamma, \pi^M, \pi^T, \sigma_M, \sigma_T$

We estimate the remaining seven parameters using the Simulated Method of Moments (SMM), jointly with an inversion procedure that recovers the market-specific productivity shifters $\{A_n\}_{n=1}^N$.

The productivity inversion proceeds as follows. Given the previously estimated parameters and a candidate vector for the remaining parameters, we recover the market-specific productivity shifters by exactly matching average wages in each labor market. Formally, the wage equation (1) implies

$$A_n = \frac{\overline{w}_n}{(1+\tau)^{-1} s_n^M \mathbb{E}_n[z|M] + s_n^T \mathbb{E}_n[z^\eta|T]}, \tag{18}$$

where $\overline{w}_n$ is the average wage in market n, s_n^M and s_n^T are the shares of modern and traditional workers in market n, respectively, and $\mathbb{E}_n[z|M]$ and $\mathbb{E}_n[z^\eta|T]$ are their corresponding conditional average marginal products. We use the observed average wage in each labor market to construct the numerator on the right-hand side of (18). Constructing the denominator, in contrast, requires solving the model for a candidate vector of productivities and parameter values to obtain the equilibrium objects entering the expression. We therefore iterate on the vector $\{A_n\}_{n=1}^N$ until convergence, conditional on the parameter vector. This inversion is nested within the SMM estimation routine described below.

Turning to the SMM estimation, we estimate the seven structural parameters $\eta, \theta, \gamma, \pi^M, \pi^T, \sigma_M, \sigma_T$ using ten moments. Four are aggregate cross-sectional moments reported in Table 1: the modern-sector employment share, the one-year transition probabilities across sectors, and the variance of log wages in the modern sector. The remaining six moments are the diagonal elements of the five-year wage-quintile transition matrices within each sector. To construct these moments, we define quintiles of the sector-specific wage distributions and restrict the sample to workers observed in the same sector at both endpoints. We exclude the bottom and top diagonal elements in each sector because the smooth Brownian specification has difficulty matching the high persistence observed at the extremes of the wage distribution.

We deliberately construct the sectoral and wage-quintile transition moments over different horizons. The one-year sector transition probabilities primarily capture diffusion across the sectoral cutoff and are therefore informative about σ_S. In contrast, five-year persistence within sector-specific wage quintiles reflects the cumulative effect of drift over workers' careers and is therefore informative about the parameters governing expected skill growth.

The identification follows naturally from these differences across moments. Persistence within sector-specific wage quintiles helps identify the sectoral drifts, μ^S, and therefore the parameters that determine learning from peers: the learning intensity, γ, the power mean governing vibrancy, θ, and the learning segmentation parameters, π^S. Persistence in the upper wage quintiles is particularly informative about μ^M, while persistence in the lower quintiles is more informative about μ^T. Conditional on the drifts identified by these moments, the one-year sector transition probabilities pin down σ_M and σ_T through the frequency with which skill shocks move workers across the sectoral cutoff. Perla et al. (2021) use a related strategy to estimate the parameters governing firm productivity when it follows a geometric Brownian motion, as in our setting. Finally, the modern-sector employment share and the variance of log wages help identify the technology curvature, η, and the modern-sector volatility, σ_M.

Overall, we use 10 moments to estimate seven parameters. Let $\Theta = (\eta, \theta, \gamma, \pi^M, \pi^T, \sigma_M, \sigma_T)$ denote the vector of parameters to be estimated. We implement the SMM by minimizing the sum of squared symmetric percentage differences between the model-implied moments, $\Psi^m(\Theta)$, and their empirical counterparts, Ψ^d:

$$\min_{\Theta} \sum_{i=1}^{10} \left(\frac{\Psi_i^m(\Theta) - \Psi_i^d}{0.5\left(\Psi_i^m(\Theta) + \Psi_i^d\right)} \right)^2 .$$

We search over the parameter space using the TikTak global optimization algorithm of Arnoud et al. (2019).[21] At each iteration of the optimization routine, we invert the equilibrium mapping in (18) to recover $\{A_n\}_{n=1}^N$ nonparametrically.

Table 2 reports the estimates. Panel A presents the estimated SMM parameters. We find strong diminishing returns in the traditional sector, with η well below one. We also estimate the power-mean parameter governing sectoral vibrancy to be 1.77. This value is somewhat lower than the estimate reported by Crews (2024) for the United States, who focus on interactions within cities and incorporate dynamic agglomeration forces. The difference likely reflects both our narrower definition of labor markets and the absence of agglomeration forces in our framework. Our estimates also imply a sharply asymmetric degree of learning segmentation across the two sectors. For modern workers, π^M is close to 1/2, so their learning interactions are essentially unbiased. Compared with a random-interactions benchmark, a modern worker is only about 5 percent more likely to learn from another modern worker than from a traditional worker. Traditional workers, by contrast, exhibit a strong own-sector bias, with π^T well above 1/2. They are roughly two and a half times more likely to learn from a worker in their own sector than from one in the modern sector. Finally, we estimate skill volatility to be similar across the two sectors, although somewhat higher for modern workers.[22]

Panel B compares the moments in the data with their model counterparts. Overall, the model fits the 10 moments well using only seven parameters. Among the aggregate moments, it closely matches the one-year sector transition probabilities and the variance of log wages in the modern sector, although it falls slightly short of reproducing the modern-sector employment share. Despite its parsimonious specification for skill dynamics in both sectors, the model also performs well in matching the sector-specific transition probabilities across wage quintiles.

[21] We use a Sobol sequence of 5,000 starting points and a simplex method for local optimization, initialized from the 50 best candidates.

[22] Although we do not impose the restrictions in Assumptions 2 and 3 during estimation, the resulting estimates satisfy the parameter restrictions required by the sufficient conditions for the sorting equilibrium.

Table 2: SMM estimation results

Parameter	Description	Estimate
Panel A: Parameter Estimates		
η	Decreasing returns, traditional	0.520
θ	Power-mean, vibrancy	1.770
γ	Peer-learning intensity	0.001
π^M	Learning segmentation, modern	0.512
π^T	Learning segmentation, traditional	0.719
σ_M	Volatility of log skill, modern	0.136
σ_T	Volatility of log skill, traditional	0.096

Moment	Data	Model
Panel B: Targeted Moments		
Modern-sector employment share	0.70	0.65
Variance of log wages, modern sector	0.39	0.42
Modern-to-modern transition	0.97	0.95
Traditional-to-traditional transition	0.92	0.91
Modern $q_2 \to q_2$ transition	0.30	0.33
Modern $q_3 \to q_3$ transition	0.37	0.36
Modern $q_4 \to q_4$ transition	0.48	0.50
Traditional $q_2 \to q_2$ transition	0.37	0.44
Traditional $q_3 \to q_3$ transition	0.33	0.35
Traditional $q_4 \to q_4$ transition	0.39	0.36

Notes: Panel A reports the SMM parameter estimates obtained by minimizing the weighted distance between the model and data moments. Panel B displays the empirical moments computed from the EPS and their model counterparts at the estimated parameters.

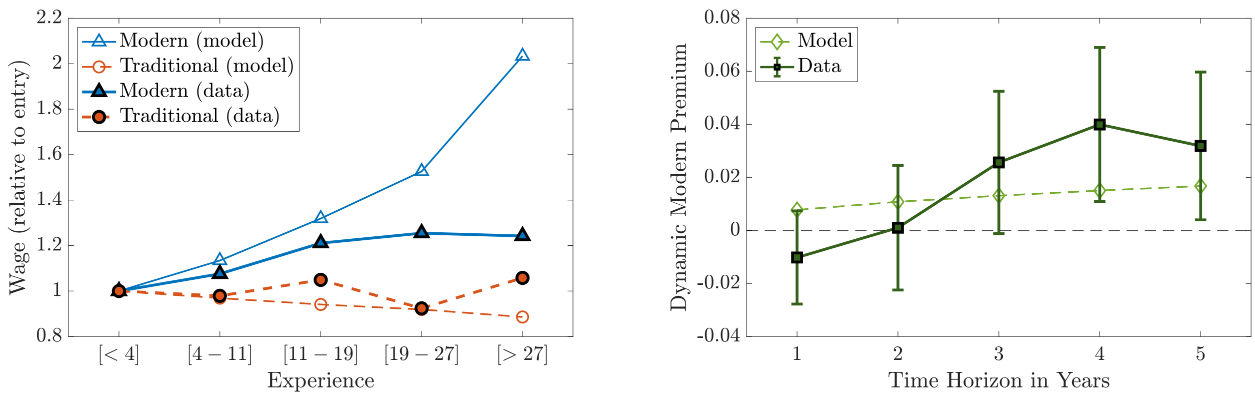
4.3 Over-identification Exercises and Stationary Equilibrium

This subsection presents a set of over-identification and model-fit exercises. We use nontargeted moments to assess the identification of the estimated parameters and then discuss key qualitative properties of the stationary equilibrium under those estimates.

Figure 7 presents our over-identification exercises, comparing two important nontargeted moments in the model and the data. Figure 7(a) first reports wage-experience profiles for modern and traditional workers. We normalize average wages in each experience bin by the wage level in the first bin, so the figure shows how wages evolve over the life cycle for each group in the model and the data. Figure C11 presents the corresponding comparison in levels, without normalizing entry wages. In that figure, the data series coincide with those reported in Figure 1(a).

The model performs well overall. Although we do not target any moments related to life-cycle wage growth, it matches the wage profiles of both modern and traditional workers reasonably well, both qualitatively and quantitatively. For modern workers, the model closely tracks wage growth through 19 years of experience. Beyond that point, however, it does not reproduce the plateau observed in the data and therefore overpredicts wage growth in the last two experience bins. For traditional workers, the model tracks wage growth closely across all experience bins. As Figure

Figure 7: Life Cycle Wage Growth and Dynamic Modern Premium: Model vs Data



(a) Life-cycle wage growth **(b)** Dynamic modern premium

Notes: Figure 7 compares nontargeted moments in the model and in the data. Panel 7(a) displays average wages by experience for modern and traditional workers, with each series normalized by its own entry level. Model averages are computed over the same experience-quintile bins as in Figure 1, weighting ages by surviving sector mass. Panel 7(b) displays the dynamic modern premium: the model series is the raw sectoral gap in expected wage growth, $\mathbb{E}[\Delta \log w_{t+h} \mid M] - \mathbb{E}[\Delta \log w_{t+h} \mid T]$, and the data series reports the estimates of β_h from equation (15). Vertical bars denote 95 percent confidence intervals.

C11 shows, this fit also holds in levels. Even without normalizing entry wages, the model closely replicates both the level and growth of traditional-sector wages. For modern workers, by contrast, the model predicts wages above those observed in the data throughout the life cycle.

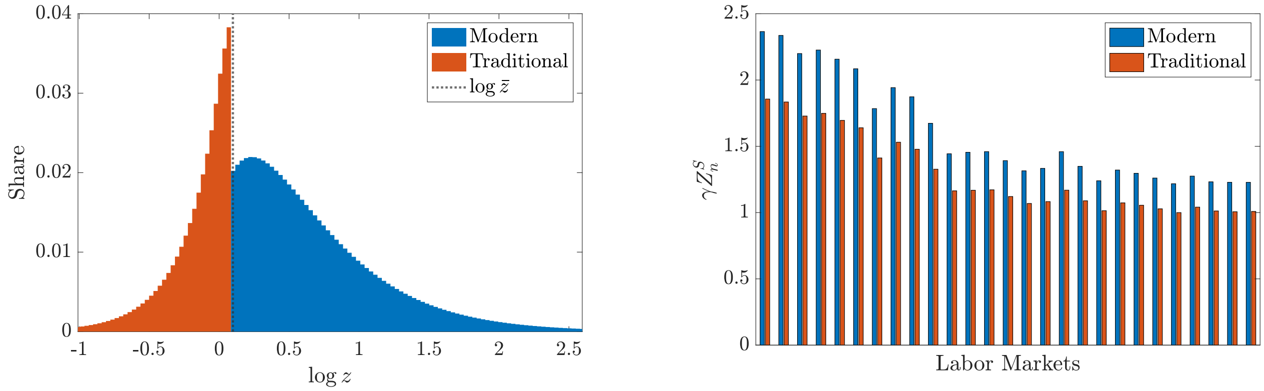
Figure 7(b) turns to the dynamic modern-sector wage premium. The model series reports the difference in expected wage growth between modern and traditional workers over horizons of one to five years, which corresponds to the model counterpart of the estimated coefficients β_h in equation (15). As before, none of these moments are targeted in the estimation. The model reproduces the two main features of the data. Qualitatively, the premium is positive and increases with the time horizon. Quantitatively, the model-implied premium lies within the 95 percent confidence intervals of the empirical estimates at every horizon. The main discrepancy appears at short horizons. In the data, the premium is close to zero during the first two years and emerges only afterward, whereas in the model modern-sector workers begin to pull ahead immediately.

Taken together, we view these results as evidence that the model can replicate the wage dynamics observed in the data, both qualitatively and quantitatively. We now turn to the stationary equilibrium under the estimated parameters, focusing on how workers sort across sectors and on the learning environments offered by different labor markets.

Figure 8 illustrates key properties of the stationary equilibrium. Figure 8(a) first shows the skill distribution of workers in the median market in terms of productivity, A_n.[23] The y-axis denotes the share of workers at each skill level within the market, while the vertical dotted line marks the log of the modern-sector cutoff, $\bar{z}_n$. Consistent with Proposition 1, the estimated stationary equilibrium features sorting across sectors, with higher-skill workers employed primarily in the

[23] Appendix Figure C12 displays the skill distributions for all markets.

Figure 8: Decentralized Equilibrium



(a) Skill distribution, median market **(b)** Learning intensity

Notes: Figure 8 illustrates the stationary equilibrium at the estimated parameters. Panel 8(a) displays the stationary distribution of workers over log skill by sector in the median market by productivity A_n, expressed as a share of the market's workers. The dotted vertical line denotes the modern-sector cutoff $\log \bar{z}$. Panel 8(b) displays the learning drifts γZ_n^S by market and sector, normalized by the smallest drift across markets and sectors. Markets are sorted by productivity, with the most productive market on the left.

modern sector and lower-skill workers in the traditional sector.

The figure also shows a jump in the distribution immediately to the right of the modern-sector cutoff. This jump reflects the higher volatility of idiosyncratic skill shocks in the modern sector relative to the traditional sector. Modern workers just above the cutoff experience larger skill fluctuations, which may push them farther up the skill distribution or below the cutoff, inducing them to switch into the traditional sector. By contrast, traditional workers just below the cutoff experience less volatile shocks and therefore move across skill levels less frequently. As a result of this asymmetry in skill volatility, workers pool just below the modern-sector cutoff. This pattern is not unique to the median market and appears across all labor markets, as shown in Figure C12.

Second, Figure 8(b) displays estimated average skill growth across the economy. The figure reports the drifts γZ_n^M and γZ_n^T in (3) across markets and sectors. We order markets from most to least productive and normalize all drifts by the smallest one.

The figure reveals substantial heterogeneity in average skill growth across markets and sectors. Modern-sector workers in every market accumulate skills faster than traditional-sector workers in any market, and within markets their human capital growth rates are about 20% higher. Average human capital accumulation is not monotonic in market productivity because markets also differ in bilateral migration costs. Markets with higher migration costs attract fewer workers, since forward-looking individuals value the option to relocate in response to future labor-market preference shocks, which can weaken local learning opportunities. As a result, workers in more productive markets with higher migration costs may accumulate less human capital than workers in less productive markets with lower migration costs.

5 Counterfactual Analysis

With the estimated model, we conduct two counterfactual exercises to explore the role of peer learning in economic development in the Chilean context. First, we consider a policy exercise that studies the aggregate effects of changing the costs of operating in the modern sector. Specifically, we consider counterfactual economies in which we vary the payroll tax τ. Changes in τ affect the skill composition of workers across sectors and, in turn, the learning opportunities available to workers in each sector. Second, we study alternative specifications of the learning technology. In particular, we consider counterfactual economies with different configurations of the learning segmentation parameters, π^M and π^T, and trace out how the degree of learning segmentation faced by modern and traditional workers affects aggregate income and welfare.

Do we expect learning segmentation to always be a barrier to economic development? The answer, perhaps surprisingly, is no. Workers impose a learning externality on others. When they choose labor markets and sectors, they do not internalize that these decisions affect other workers' learning opportunities. Eliminating learning segmentation altogether can actually reduce the learning opportunities of the most skilled workers, because they are now exposed to less skilled peers, creating a crowding-out effect. At the same time, eliminating segmentation increases the learning opportunities of traditional-sector workers, who become more exposed to higher-skilled counterparts. Whether, in the aggregate, removing learning segmentation raises or lowers income and welfare is therefore a quantitative question. Importantly, this crowding-out effect also appears in semi-endogenous growth models such as Buera and Oberfield (2016), who show that in a world with free trade, idea diffusion can reduce growth in some countries.

5.1 Reducing the Cost of Operating in the Modern Sector

In our first counterfactual exercise, we study how changes in modern-sector regulation affect peer composition and aggregate outcomes. We implement this exercise by varying the payroll tax around its baseline value of $\tau = 0.10$, considering four counterfactual economies with payroll taxes 5 and 10 percentage points above and below the baseline, corresponding to $\tau = 0.15$, $\tau = 0.20$, $\tau = 0.05$, and $\tau = 0$. We focus on the effects of these changes on four aggregate outcomes relative to the baseline economy: the share of modern workers, income per capita, the labor productivity gap between the modern and traditional sectors, and the consumption-equivalent welfare change coming from workers' learning. This last object captures the change in the continuation value from learning, corresponding to the skill-drift term $\mu_n^S z \partial_z V_n$ in (7). This is precisely the margin through which workers value the peer composition of their sector.

We focus on the learning channel when comparing welfare across counterfactual economies. One could instead consider the full welfare change, incorporating all the components of workers' value. However, changes in payroll taxes induce redistribution through the government lump-sum transfer, so aggregate welfare comparisons depend on how gains and losses are weighted across workers with different skill levels. Focusing on the learning component isolates the welfare effects generated by changes in peer composition, which is the mechanism at the core of our paper. Table D4 in the Appendix provides a full welfare decomposition for each value of τ, weighting workers equally.

Table 3: Aggregate Effects of Changing the Payroll Tax, τ

(1)	(2)	Change from baseline (%)		
		(3)	(4)	(5)
τ	Share Modern	Income	Sector Gap	CE Learning
0.20	0.50	-2.95	5.19	-0.36
0.15	0.57	-1.32	2.02	-0.18
0.10	0.65	...	...	...
0.05	0.71	0.97	-0.71	0.20
0.00	0.77	1.58	-0.23	0.37

Notes: Table 3 reports the effects of changing the payroll tax, τ, relative to the baseline stationary equilibrium with $\tau = 0.10$. Column (2) shows the modern-sector employment share, column (3) the change in income per capita, column (4) the change in the labor productivity gap between the modern and traditional sectors, and column (5) the consumption-equivalent welfare change coming from the learning channel. All changes are reported in percentage terms.

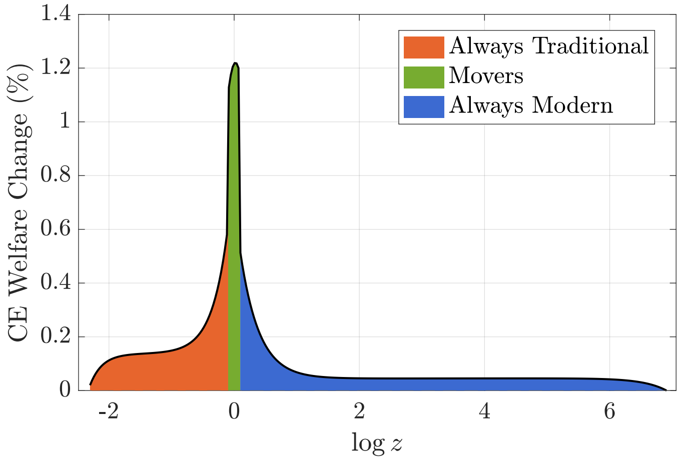
Table 3 reports the aggregate effects of changing the cost of operating in the modern sector. The table compares steady states across different values of τ. The first two columns report the payroll tax and the share of modern workers in the economy, respectively. Lower payroll taxes increase the modern-sector employment share. However, because of the curvature of the traditional-sector production technology, the least productive workers still find it optimal to remain in the traditional sector even when payroll taxes are reduced to zero.

Column 3 of Table 3 shows the change in income per capita, which declines monotonically with the payroll tax. Lowering the cost of operating in the modern sector increases aggregate income by 1.58% when payroll taxes are eliminated. Conversely, raising the payroll tax to 0.20 reduces income per capita by 2.95%. These changes are driven by two forces. First, workers reallocate immediately across sectors with different production technologies, generating a static effect on aggregate income. Second, this reallocation changes workers' learning opportunities and therefore the evolution of human capital over time, leading to a different long-run skill distribution and an additional dynamic effect on income. Figure 10 below decomposes these two margins.

Productivity gaps between the modern and traditional sectors respond systematically to changes in payroll taxes, as shown in Column 4. The key mechanism is worker selection. When payroll taxes increase, marginal workers leave the modern sector for the traditional sector. Because these workers are less productive than the average modern worker but more productive than the average traditional worker, average productivity rises in both sectors, with a larger increase in the modern sector, so the productivity gap widens. When payroll taxes fall, the opposite occurs. Marginal workers enter the modern sector, lowering average productivity in both sectors, with a larger decline in the modern sector and therefore a narrower productivity gap.

Finally, the last column of Table 3 reports the welfare effects generated through peer learning. We find that workers gain on average when the cost of working in the modern sector falls. In consumption-equivalent terms, the welfare gains from improved learning opportunities reach 0.37%

Figure 9: Heterogeneous Welfare Gains from Peer Learning with No Payroll Taxes



Group	Share	CE Learning
Always Traditional	0.20	0.35
Movers	0.16	1.19
Always Modern	0.65	0.17

Notes: Figure 9 decomposes the welfare gains from peer learning in the counterfactual economy with $\tau = 0$. In the left panel, the x-axis denotes log skill and the y-axis denotes consumption-equivalent welfare gains from the learning channel. The shaded areas divide workers into three groups: always-traditional workers (orange), movers (green), and always-modern workers (blue). The right panel reports the baseline population share of each group and the corresponding mass-weighted average consumption-equivalent welfare gain from learning.

when payroll taxes are completely eliminated, while increasing the payroll tax to 0.20 generates learning-related welfare losses of 0.36%. These effects arise because lower payroll taxes induce marginal workers to enter the modern sector, where they face better learning opportunities.

However, the aggregate welfare effect of completely removing payroll taxes in the modern sector masks substantial heterogeneity across workers. The policy changes sectoral composition and creates three groups that are affected differently. First, movers are marginal workers who switch from the traditional to the modern sector after the policy change. Second, the policy also affects workers who remain in their original sector, the "always traditional" and "always modern" workers. These groups may gain or lose through the peer-learning channel depending on how peer composition changes and on how much workers value future human capital accumulation.[24] Therefore, the aggregate welfare effect reported in the last column averages across workers who are affected differently by the policy. We decompose these heterogeneous effects below.

Figure 9 decomposes the welfare gains from learning when payroll taxes are completely eliminated across workers with different skill levels. The x-axis denotes workers' log skill, while the y-axis denotes the consumption-equivalent welfare gains from the learning channel. The shaded areas map the three groups described above into the skill distribution: always-traditional workers lie below both average modern-sector cutoffs, movers lie between the baseline and counterfactual cutoffs, and always-modern workers lie above both. The table next to the figure reports the baseline population share of each group and its average welfare gain from learning.

The policy has heterogeneous effects on workers depending on their skill. First, the largest welfare gains are concentrated among movers. By switching into the modern sector, these workers gain access to better learning opportunities, generating welfare gains of around 1.2% on average. Movers

[24]This valuation is captured by $V'(z)$, which measures the value of an additional unit of human capital.

also benefit from a higher valuation of human capital. Eliminating payroll taxes raises modern-sector wages, increasing the value of accumulating skills as future increases in z translate into higher future earnings. Both improved learning opportunities and a higher valuation of human capital therefore contribute to the large welfare gains experienced by this group.

Second, workers who remain in the modern sector also experience welfare gains, although these are the smallest in the economy, averaging 0.17%. Similar to movers, these gains are driven primarily by a higher valuation of human capital. Learning opportunities in the modern sector, by contrast, are largely unchanged. The estimated value of θ places relatively more weight on higher-skilled workers in the peer distribution, so the influx of less-skilled workers has little quantitative effect on modern-sector vibrancy. Thus, although one might initially expect the entry of lower-skilled workers to worsen learning opportunities for incumbent modern workers, this effect is quantitatively negligible under the estimated parameters.

Third, workers who remain in the traditional sector experience welfare gains of around 0.35% on average. Unlike modern-sector incumbents, these gains reflect both channels in roughly equal measure. On the one hand, learning opportunities in the traditional sector improve. Because traditional workers partly learn from modern-sector peers, the expansion of the modern sector raises traditional-sector vibrancy and increases skill growth in the average market. On the other hand, traditional workers also place a higher value on accumulating human capital. Eliminating the payroll tax raises modern-sector wages and lowers the skill threshold for entering that sector, making a future switch more valuable. This increases the marginal value of skill even for workers who remain below the cutoff. Thus, low-skilled workers benefit from the policy despite never entering the modern sector themselves, both because the larger modern sector improves their learning environment and because the human capital they accumulate becomes more valuable.

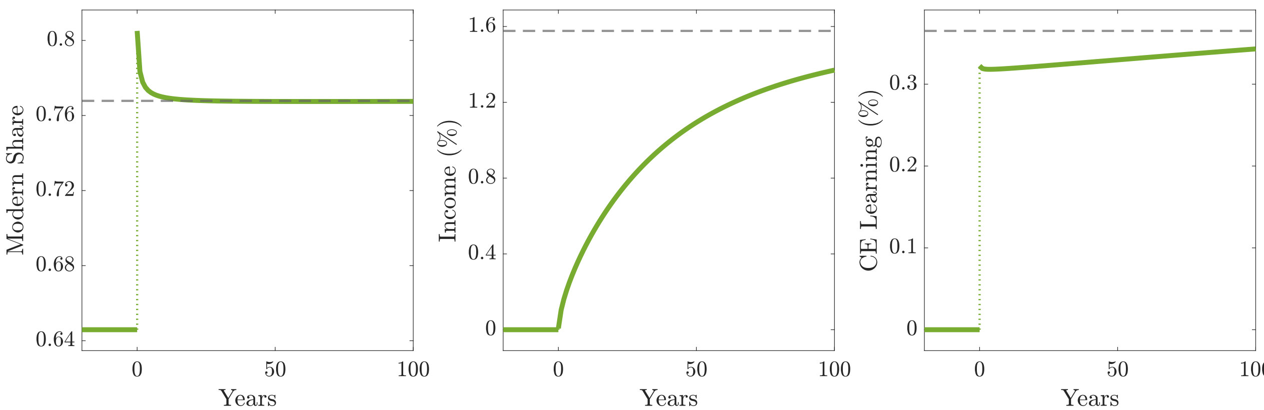
Finally, the effects of raising payroll taxes are roughly symmetric. Increasing τ to 0.20 generates mirror-image welfare losses, concentrated among workers who switch from the modern to the traditional sector. Figure C13 in Appendix C shows this pattern.

5.1.1 Transitional Dynamics

The previous analysis focused on comparisons across steady states. We now extend the analysis to study transition dynamics for the case in which payroll taxes are completely eliminated. Computing these transitions is challenging because workers interact through the skill distributions across sectors and labor markets. As a result, at every instant along the transition path, we need to solve for a fixed point in several infinitesimal equilibrium objects. To address this problem, we build on Bilal and Goyal (2026), who extend the Sequence Space Jacobian (SSJ) method of Auclert et al. (2021) to continuous time. We defer the details of the implementation to Appendix B.6.

Figure 10 reports the transition paths for the share of modern workers, income per capita, and the learning-channel welfare gains. The modern-sector share responds almost immediately to the reform. Because workers can freely switch sectors, removing the payroll tax makes the modern sector attractive right away to a large group of workers who were previously close to the cutoff. As a result, the modern-sector share jumps from 0.65 to 0.80 on impact, temporarily overshooting its

Figure 10: Transitional Dynamics After Eliminating Payroll Taxes



Notes: Figure 10 displays the transition paths following an unanticipated permanent elimination of the payroll tax at $t = 0$. The panels report the share of modern workers, income per capita as a percent deviation from the baseline stationary equilibrium, and the consumption-equivalent welfare change from the learning channel. Dashed horizontal lines indicate the new stationary equilibrium with $\tau = 0$. Transition paths are computed over a 300-year horizon, while the figure displays the first 100 years.

long-run level of 0.77, with 90% of the adjustment completed within six years. The overshooting reflects the shape of the baseline skill distribution. Slower learning in the traditional sector leaves many workers concentrated just below the modern-sector cutoff, and these workers switch as soon as the tax is removed. Over time, this concentration fades as workers accumulate more human capital in the modern sector and new low-skill workers enter the economy.

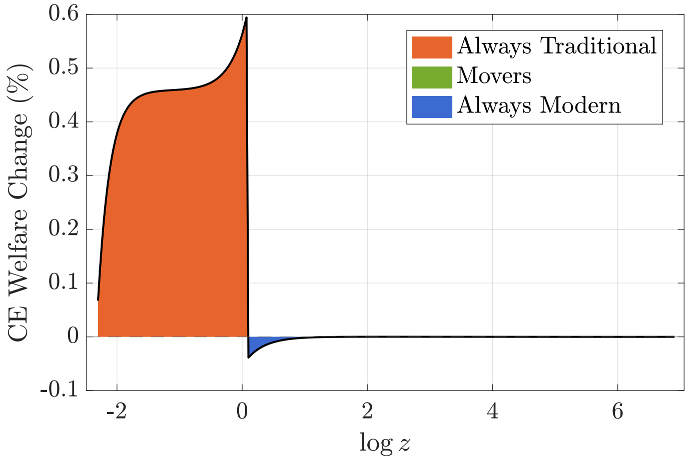
Income per capita, by contrast, responds much more slowly. Even though many workers switch sectors immediately after the reform, this reallocation has almost no effect on income on impact. Most of the long-run increase in income comes instead from workers gradually accumulating more human capital as their learning opportunities improve. Income closes half of the gap to the new steady state after 26 years and 90% of the gap after roughly 120 years. This slow transition reflects the feedback between peer learning and the skill distribution. Better peers raise skill growth, but peer quality itself improves only gradually as workers accumulate human capital, making the adjustment much more persistent than demographic turnover alone would imply.

Finally, the learning-channel welfare gains respond immediately to the reform. They jump to 0.32% on impact, about 88% of the long-run effect. Unlike income, welfare moves right away because workers value the better learning opportunities they expect to receive in the future. The gains then rise gradually to their long-run level of 0.37% as the modern sector expands and peer composition improves across labor markets.

5.2 Changing the Learning Technology for Traditional Workers

In our second counterfactual exercise, we directly equalize the learning opportunities of modern and traditional workers by changing the learning technology. Formally, we change the learning segmentation parameter for traditional workers such that $\pi^T = 1 - \pi^M$. Under this restriction, modern and traditional workers face the same vibrancy in every labor market and therefore the same learning opportunities. At the estimated parameters, this counterfactual is close to a random-

Figure 11: Heterogeneous Welfare Gains from Peer Learning with Equal Learning Technology



Group	Share	CE Learning
Always Traditional	0.35	0.52
Movers	0.02	-0.04
Always Modern	0.63	-0.01

Notes: Figure 11 decomposes the welfare gains from peer learning in the counterfactual economy with $\pi^T = 1 - \pi^M$. In the left panel, the x-axis denotes log skill and the y-axis denotes consumption-equivalent welfare gains from the learning channel. The shaded areas divide workers into three groups: always-traditional workers (orange), movers (green), and always-modern workers (blue). The right panel reports the baseline population shares and the mass-weighted average consumption-equivalent welfare gain from learning.

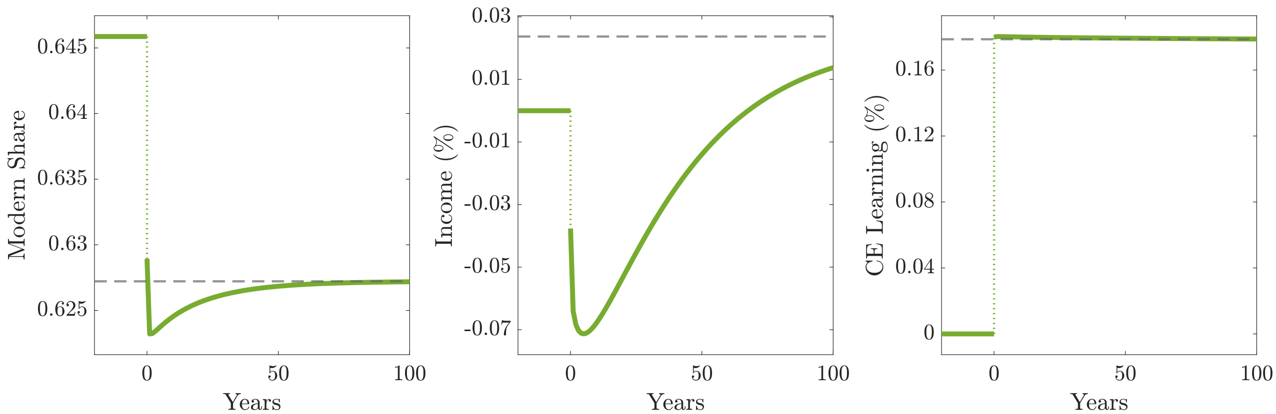
interactions benchmark, with learning parameters $\pi^M = 0.51$ and $\pi^T = 0.49$. Thus, traditional-sector workers move from a learning technology strongly biased toward interactions with other traditional workers to one in which they learn almost equally from workers in both sectors.

We first describe the heterogeneous welfare effects across workers. Figure 11 shows that equalizing learning opportunities across sectors generates welfare gains for workers who always remain in the traditional sector. Workers in this group gain 0.52%, in consumption-equivalent terms, as their learning opportunities improve significantly. In contrast, workers who remain in the modern sector are largely unaffected by the new environment, with only those close to the cutoff experiencing small welfare losses. As Figure 12 below shows, the share of modern workers remains nearly constant, declining by only 2 percentage points. Therefore, only a very small fraction of workers move from the modern to the traditional sector, and the associated welfare losses are negligible.

Turning to the aggregate effects, Figure 12 illustrates the transitional dynamics and highlights the new steady-state values. As noted above, the share of modern workers decreases by 2 percentage points when learning opportunities are equalized across the two groups. The figure shows, however, that the share of modern workers falls more sharply on impact and then gradually recovers before converging to its new long-run value. The initial decline reflects the fact that the thresholds, $\bar{z}_n$, jump on impact, as workers' forward-looking values adjust immediately when the learning technology changes. Subsequently, some traditional workers accumulate human capital more rapidly because they face better learning opportunities and transition into the modern sector, explaining the gradual recovery.

The responses of aggregate income and aggregate welfare gains from learning are small under the new learning environment. Income declines slightly on impact, reflecting the immediate movement of marginal workers from the modern to the traditional sector. Because these workers operate

Figure 12: Transitional Dynamics After Equalizing Learning Technologies



Notes: Figure 12 displays the transition paths following an unanticipated permanent equalization of the learning technologies, $\pi^T = 1 - \pi^M$, at $t = 0$. The panels report the share of modern workers, income per capita as a percent deviation from the baseline stationary equilibrium, and the consumption-equivalent welfare change from the learning channel. Dashed horizontal lines indicate the new stationary equilibrium with equal learning technologies. Transition paths are computed over a 300-year horizon, while the figure displays the first 100 years.

an inferior technology given their skill level, income per capita falls. Subsequently, workers in the traditional sector accumulate human capital at a faster pace, leading income to recover and eventually increase by 0.02% in the new steady state. In contrast, welfare gains from learning from peers jump on impact by 0.18% and remain at this new long-run level thereafter.

5.3 Discussion

Why do the two counterfactual exercises imply such different responses? At first glance, one might expect the counterfactual that equalizes learning technologies across sectors to generate larger effects. Equalizing learning opportunities directly improves the learning environment of all workers initially in the traditional sector. In contrast, removing τ improves learning opportunities primarily for workers who switch from the traditional to the modern sector. Hence, a much larger group of workers is directly exposed to better learning opportunities in the former exercise.

There is, however, an important difference between the two exercises. When learning opportunities improve by lowering the cost of operating in the modern sector, human capital also becomes more valuable in the economy. Removing τ increases the share of modern workers by 12 percentage points and raises income per capita by 1.6%. As more workers operate the superior modern technology, wages increase, and so does the value of accumulating human capital. Workers therefore have stronger incentives to accumulate skills and transition into the modern sector. Hence, the increase in income is fundamentally dynamic. Lowering τ raises the return to human capital, which strengthens incentives to accumulate skills and gradually increases aggregate income.

In contrast, equalizing learning opportunities by changing the learning technology makes working in the modern sector less attractive. Once workers can access the same learning opportunities in both sectors, traditional workers have weaker incentives to transition into the modern sector and instead remain traditional, where they operate a smaller-scale technology. The welfare gains from improved learning opportunities are nevertheless sizable, increasing by 0.18%, compared with 0.37% when τ is eliminated. Thus, equalizing learning technologies generates approximately half of the welfare-through-learning gains from eliminating τ, despite directly improving learning opportunities for a larger group of workers. The reason is that traditional workers facing the improved learning technology value additional human capital less than workers in the economy without payroll taxes, where the returns to accumulating human capital are higher. Taken together, these results highlight an important complementarity between access to better learning opportunities and the returns to accumulating human capital.

6 Conclusion

This paper has studied how segmented learning markets shape human capital accumulation and economic development. We developed a quantitative model in which workers sort across labor markets and between a modern and a traditional sector, accumulating skills by learning from peers within their sector. The model predicts that higher-skill workers select into the modern sector and that modern-sector workers experience faster wage growth. We document this dynamic premium in Chilean panel data and provide evidence that it reflects learning from better-paid peers. Our counterfactuals show that improving learning opportunities for traditional-sector workers can generate meaningful welfare gains, but that the effects on aggregate income depend crucially on the returns to accumulating human capital. Lowering the cost of operating in the modern sector improves workers' learning environment while also raising the value of human capital. This encourages workers to adopt the modern technology, generating long-run income gains. In contrast, directly improving learning opportunities in the traditional sector produces welfare gains but little aggregate income growth because it weakens workers' incentives to transition into the modern sector. A central message of the paper is therefore that learning opportunities and the returns to human capital are complementary forces in economic development. Policies that improve learning are most effective when workers also have incentives to use those skills with better technologies.

Our framework, inevitably, abstracts from several potentially important mechanisms. Perhaps the most important is labor-market frictions. Understanding how search, matching, and job mobility interact with human capital accumulation in developing economies would provide a natural extension of our analysis. A second promising avenue is to give firms a more active role. Just as workers face different incentives to accumulate skills depending on their learning environment, firms in developing countries may face different incentives to improve their technologies depending on the other firms operating in their markets. Finally, our framework can be used to study optimal policy in dual economies with peer learning. Throughout the paper, we emphasize that workers do not internalize how their sectoral choices affect the learning opportunities of others. This local externality creates scope for policy interventions that can improve worker welfare.

References

Achdou, Y., Han, J., Lasry, J.-M., Lions, P.-L., and Moll, B. (2022). Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach. *The Review of Economic Studies*, 89(1):45–86.

Akcigit, U., Akgunduz, Y. E., Alp, H., Quintero, J. M., and Cilasun, S. M. (2024). Good Firms, Bad Policies: The Dynamics of Informality and Industry Policy in Shaping Economic Growth.

Akcigit, U., Alp, H., and Peters, M. (2021). Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries. *American Economic Review*, 111(1):231–275.

Akcigit, U., Caicedo, S., Miguelez, E., Stantcheva, S., and Sterzi, V. (2018). Dancing with the Stars: Innovation Through Interactions.

Amaral, P. and Rivera-Padilla, A. (2024). Cross-country income dispersion, human capital, and technology adoption.

Arnoud, A., Guvenen, F., and Kleineberg, T. (2019). Benchmarking global optimizers. Working Paper 26340, National Bureau of Economic Research.

Auclert, A., Bardóczy, B., Rognlie, M., and Straub, L. (2021). Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models. *Econometrica*, 89(5):2375–2408.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*, 75(4, Part 1):352–365. Publisher: The University of Chicago Press.

Bilal, A. and Goyal, S. (2026). Some Pleasant Sequence-Space Arithmetic in Continuous Time. Working Paper.

Bobba, M., Flabbi, L., and Levy, S. (2022). Labor Market Search, Informality, and Schooling Investments. *International Economic Review*, 63(1):211–259.

Buera, F. J. and Oberfield, E. (2016). The Global Diffusion of Ideas.

Crews, L. (2024). A Dynamic Spatial Knowledge Economy.

Dix-Carneiro, R., Goldberg, P. K., Meghir, C., and Ulyssea, G. (2021). Trade and Informality in the Presence of Labor Market Frictions and Regulations.

Donovan, K., Lu, W. J., and Schoellman, T. (2023). Labor Market Dynamics and Development*. *The Quarterly Journal of Economics*, 138(4):2287–2325.

Duarte, M. and Restuccia, D. (2019). Relative prices and sectoral productivity. *Journal of the European Economic Association*, 18(3):1400–1443.

Engbom, N. (2022). Labor Market Fluidity and Human Capital Accumulation.

Engbom, N., Malmberg, H., Porzio, T., Rossi, F., and Schoellman, T. (2025). Economic development according to chandler. Technical report, National Bureau of Economic Research.

Gabaix, X. and Ibragimov, R. (2011). Rank — 1/2: A simple way to improve the ols estimation of tail exponents. *Journal of Business and Economic Statistics*, 29(1):24–39.

Gollin, D. (2014). The lewis model: A 60-year retrospective. *Journal of Economic Perspectives*, 28(3):71–88.

Gollin, D. and Rogerson, R. (2025). Structural change and macro development: beyond the one-sector growth model. *Oxford Review of Economic Policy*, 41(2):291–312.

Herkenhoff, K., Lise, J., Menzio, G., and Phillips, G. M. (2024). Production and Learning in Teams. *Econometrica*, 92(2):467–504.

Herrendorf, B., Rogerson, R., and Ákos Valentinyi (2014). Chapter 6 - growth and structural transformation. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 2 of *Handbook of Economic Growth*, pages 855–941. Elsevier.

Herrendorf, B. and Schoellman, T. (2018). Wages, human capital, and barriers to structural transformation. *American Economic Journal: Macroeconomics*, 10(2):1–23.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The allocation of talent and U.S. economic growth. *Econometrica*, 87(5):1439–1474.

Jarosch, G., Oberfield, E., and Rossi-Hansberg, E. (2021). Learning From Coworkers. *Econometrica*, 89(2):647–676.

Lagakos, D. (2016). Explaining cross-country productivity differences in retail trade. *Journal of Political Economy*, 124(2):579–620.

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

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, 22(2):139–191.

Lhuillier, H. (2024). Should I Stay or Should I Grow?

Lopez Garcia, I. (2015). Human Capital and Labor Informality in Chile: A Life-Cycle Approach.

Lopez-Martin, B. (2019). Informal Sector Misallocation. *Macroeconomic Dynamics*, 23(8):3065–3098. Publisher: Cambridge University Press.

Lucas, R. E. and Moll, B. (2014). Knowledge Growth and the Allocation of Time. *Journal of Political Economy*, 122(1):1–51. Publisher: The University of Chicago Press.

Lucas, Jr., R. E. (2004). Life earnings and rural-urban migration. *Journal of Political Economy*, 112(S1):S29–S59.

Ma, X., Nakab, A., and Vidart, D. (2024). Human Capital Investment and Development: The Role of On-the-Job Training. *Journal of Political Economy Macroeconomics*, 2(1):107–148. Publisher: The University of Chicago Press.

Meghir, C., Narita, R., and Robin, J.-M. (2015). Wages and Informality in Developing Countries. *American Economic Review*, 105(4):1509–1546.

Meyn, S. P. and Tweedie, R. L. (1993). Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3):518–548.

Perla, J. and Tonetti, C. (2014). Equilibrium Imitation and Growth. *Journal of Political Economy*, 122(1):52–76.

Perla, J., Tonetti, C., and Waugh, M. E. (2021). Equilibrium Technology Diffusion, Trade, and Growth. *American Economic Review*, 111(1):73–128.

Perry, G. E., Arias, O., Fajnzylber, P., Maloney, W. F., Mason, A., and Saavedra-Chanduvi, J. (2007). *Informality: Exit and Exclusion.* The World Bank.

Samaniego De La Parra, B. and Fernández Bujanda, L. (2020). Increasing the Cost of Informal Workers: Evidence from Mexico.

Schwartzman, M. (2025). From street markets to shopping malls: The transformation of the service sector. Technical report.

Traiberman, S. (2019). Occupations and import competition: Evidence from denmark. *American Economic Review*, 109(12):4260–4301.

Ulyssea, G. (2018). Firms, Informality, and Development: Theory and Evidence from Brazil. *American Economic Review*, 108(8):2015–2047.

Ulyssea, G. (2020a). Formal and Informal Firm Dynamics.

Ulyssea, G. (2020b). Informality: Causes and Consequences for Development. *Annual Review of Economics*, 12(1):525–546.

Appendix

A Theory Appendix

Appendix A.1 states and proves the auxiliary results behind the equilibrium characterization of Section 2.5. Appendix A.2 proves Propositions 1 and 2 and Lemma 1.

A.1 Auxiliary Results and Derivations

Throughout, we work with log skill, $x = \log z$, write $v_n(x) = V_n(e^x)$ for the envelope value in these units, and let

$$Z_H^* \equiv \frac{\delta - \theta^2\sigma^2/2}{(\theta+1)\,\gamma}, \qquad \sigma \equiv \max\{\sigma_M, \sigma_T\}, \tag{A1}$$

which is positive under Assumption 3. The product γZ_H^* does not depend on γ, a fact we use repeatedly. The first lemma establishes the properties of the value function needed for the sorting argument. We state it for the worker problem induced by an arbitrary vibrancy vector because the existence proof applies the result before constructing an equilibrium.

Lemma 2 (Shape of the value function)**.** *Fix vibrancies $\{Z_n^S\}$ with finite induced values, and suppose Assumption 1 holds. Then, for every n: (a) v_n is strictly increasing, with slopes bounded by $\eta/(\rho+\delta) \le v_n' \le 1/(\rho+\delta)$; and (b) v_n is convex.*

Proof of Lemma 2. (a) The increments of $\log z$ in (3) do not depend on a worker's own skill. Fix $x_1 > x_0$ and consider any feasible plan for a worker starting at (n, S, x_0). By a feasible plan, we mean a sectoral and relocation choice process adapted to the exogenous randomness: the Wiener path, the arrival times of relocation opportunities, and the taste draws. Such a plan is feasible from any initial skill, and the value function is the supremum of discounted expected payoffs over this class of plans.

Consider now a worker starting at x_1 who follows the same plan under the same realizations of the exogenous randomness. Since skill increments are independent of a worker's own skill, the two skill paths satisfy

$$x_t(x_1) = x_t(x_0) + (x_1 - x_0),$$

and the two workers visit the same sequence of markets and sectors.

Under Assumption 1, flow utility is affine in current log skill, with slope η in the traditional sector and slope 1 in the modern sector:

$$\log\left((1+b)\,w_n^T\right) = \log(1+b) + \log A_n + \eta\, x$$

and

$$\log\left((1+b)\,w_n^M\right) = \log(1+b) + \log A_n - \log(1+\tau) + x.$$

Mobility costs and taste draws enter additively, while the death time is independent of skill. Therefore, along any given plan, the payoff gain from starting at x_1 rather than x_0 is bounded below by $\eta\,(x_1 - x_0)/(\rho+\delta)$ and above by $(x_1 - x_0)/(\rho+\delta)$.

Applying the lower bound to plans that approximate the optimum at x_0, and the upper bound to plans that approximate the optimum at x_1, gives

$$\frac{\eta\,(x_1 - x_0)}{\rho + \delta} \;\le\; v_n(x_1) - v_n(x_0) \;\le\; \frac{x_1 - x_0}{\rho + \delta}.$$

Since $x_1 > x_0$ and $\eta > 0$, the lower bound implies that v_n is strictly increasing. The two inequalities also deliver the stated slope bounds.

(b) Consider again any fixed feasible plan. The skill path can be written as $x_t = x_0 + I_t$, where I_t is independent of x_0. Flow utility is therefore affine in x_0 along every realization, and each fixed plan delivers a discounted expected payoff that is affine in x_0. Since v_n is the supremum of these affine functions, it is convex. $\square$

The second lemma connects sorting to workers' learning environments. We say that the allocation in market n is *sorted at* $\bar{z}_n$ if N_n^T places no mass above $\bar{z}_n$ and N_n^M places no mass below $\bar{z}_n$.

Lemma 3 (Sorting ranks learning environments)**.** *Suppose the allocation in market n is sorted at some $\bar{z}_n$ and Assumption 2 holds. Then g_n^M first-order stochastically dominates g_n^T, and consequently $Z_n^M \ge Z_n^T$ for every θ. Both inequalities are strict if both sectors are populated and $\pi^M + \pi^T > 1$.*

Proof of Lemma 3. If either sector is empty, the two peer-skill distributions in (5) coincide and the ranking is trivial, so suppose both sectors are populated. Define $\hat{g}_n^S \equiv N_n^S / L_n^S$ as the within-sector skill densities.

By (5), the effective peer-skill distributions are mixtures of the same two within-sector distributions:

$$g_n^S = \omega^S \hat{g}_n^M + (1 - \omega^S)\,\hat{g}_n^T,$$

where

$$\omega^M = \frac{\pi^M L_n^M}{\pi^M L_n^M + (1 - \pi^M) L_n^T}, \qquad \omega^T = \frac{(1 - \pi^T) L_n^M}{(1 - \pi^T) L_n^M + \pi^T L_n^T}.$$

Direct computation shows that $\omega^M \ge \omega^T$ if and only if $\pi^M \pi^T \ge (1 - \pi^M)(1 - \pi^T)$, or equivalently $\pi^M + \pi^T \ge 1$. Under Assumption 2, this condition holds independently of the relative size of the two sectors.

Next, sorting orders the two components of these mixtures. Since the allocation is sorted at $\bar{z}_n$, the support of $\hat{g}_n^T$ lies below that of $\hat{g}_n^M$. Hence, $\hat{g}_n^M$ first-order stochastically dominates $\hat{g}_n^T$. Because g_n^M places weakly more weight on the dominating distribution than g_n^T, it follows that g_n^M first-order stochastically dominates g_n^T.

It remains to translate this ranking into vibrancy. For $\theta > 0$, z^θ is increasing, so first-order stochastic dominance implies

$$\int z^\theta g_n^M \, \mathrm{d}z \ge \int z^\theta g_n^T \, \mathrm{d}z.$$

Taking the power $1/\theta > 0$ preserves the inequality, giving $Z_n^M \geq Z_n^T$. For $\theta < 0$, z^θ is decreasing, so the inequality between the integrals is reversed. Taking the negative power $1/\theta$ reverses it once more, again giving $Z_n^M \geq Z_n^T$. The case $\theta \to 0$ follows by continuity.

Finally, if both sectors are populated and $\pi^M + \pi^T > 1$, then $\omega^M > \omega^T$. Moreover, sorting implies that the two within-sector skill distributions differ. The stochastic dominance of g_n^M over g_n^T is therefore strict, as is the resulting ranking $Z_n^M > Z_n^T$ whenever the relevant moments are finite, which holds everywhere the lemma is applied below. □

The third lemma delivers the sorting result at the core of Proposition 1. Given learning environments that favor the modern sector and common skill volatilities, optimal labor supply takes a cutoff form, with the cutoff bounded between two primitive thresholds.

Lemma 4 (Labor supply is a cutoff)**.** *Fix vibrancies with $Z_n^T \leq Z_n^M \leq Z_H^*$ in market n, suppose $\sigma_M = \sigma_T$, and let Assumption 1 hold. Then labor supply in market n follows a cutoff policy: there exists $\bar{z}_n$ such that $S_n^*(z) = T$ for $z < \bar{z}_n$ and $S_n^*(z) = M$ for $z > \bar{z}_n$, with*

$$\underline{z} \leq \bar{z}_n \leq \hat{z}, \qquad \underline{z} \equiv \exp\left(\frac{\log(1+\tau) - \gamma Z_H^*/(\rho+\delta)}{1-\eta}\right). \tag{A2}$$

Moreover, $\bar{z}_n < \hat{z}$ whenever $Z_n^M > Z_n^T$.

Proof of Lemma 4. With $\sigma_M = \sigma_T$, the Itô corrections in (6) cancel, so $\mu_n^M - \mu_n^T = \gamma\,(Z_n^M - Z_n^T)$. The curvature term in (11) therefore drops out. Under Assumption 1, the subsidy and $\log A_n$ also cancel from the static gain. Writing the flow advantage in log skill gives

$$\Delta_n(x) = (1-\eta)\,x - \log(1+\tau) + \gamma\left(Z_n^M - Z_n^T\right) v_n'(x). \tag{A3}$$

We first show that $\Delta_n(x)$ crosses zero exactly once. The first term in (A3) is strictly increasing and unbounded in both directions. The dynamic term is nonnegative and nondecreasing because $v_n' \geq 0$ by Lemma 2(a) and v_n' is nondecreasing by Lemma 2(b). Hence, Δ_n is strictly increasing with slope at least $1-\eta$ and $\lim_{x\to\infty} \Delta_n(x) = \infty$.

In the other direction, the slope bound in Lemma 2(a), together with $Z_n^M - Z_n^T \leq Z_H^*$, implies that the dynamic term is bounded above by $\gamma Z_H^*/(\rho+\delta)$. Therefore, $\lim_{x\to-\infty} \Delta_n(x) = -\infty$. It follows that Δ_n crosses zero exactly once, so labor supply follows a cutoff policy.

We next establish the bounds on the cutoff. Using the same upper bound on the dynamic term,

$$\Delta_n(x) \leq (1-\eta)x - \log(1+\tau) + \gamma Z_H^*/(\rho+\delta) < 0$$

for $x < \log \underline{z}$. Hence, $\bar{z}_n \geq \underline{z}$.

For the upper bound, consider the static crossing point $\hat{x} = \log(1+\tau)/(1-\eta)$. At this point,

$$\Delta_n(\hat{x}) = \gamma(Z_n^M - Z_n^T)\,v_n'(\hat{x}) \geq 0.$$

Since Δ_n is strictly increasing and crosses zero only once, this implies $\bar{z}_n \leq \hat{z}$. Finally, Lemma 2(a) gives $v_n' \geq \eta/(\rho+\delta) > 0$, so $\Delta_n(\hat{x}) > 0$ whenever $Z_n^M > Z_n^T$. In this case, the upper bound is strict and $\bar{z}_n < \hat{z}$. □

The fourth lemma establishes a uniform lower bound on stationary market populations. This result makes the constant κ in Assumption 3 primitive.

Lemma 5 (Uniform population floor)**.** *Let Assumption 1 hold, and fix any vibrancy vector satisfying $Z_n^S \leq Z_H^*$ for all n and S. Then the stationary populations induced by the associated policies satisfy $L_n \geq \underline{L}$ for every n, where*

$$\underline{L} \equiv \frac{\delta}{\delta+\lambda}\frac{e^{-\nu\bar{\Delta}}}{N}, \qquad \bar{\Delta} \equiv \frac{1}{\rho+\delta}\left[\max_{n,m}|\log A_n - \log A_m| + \frac{\gamma Z_H^*}{\rho+\delta} + \frac{\lambda}{\nu}\log N\right]. \tag{A4}$$

Proof of Lemma 5. We first bound value differences across markets. Consider a point at which $v_n - v_m$ is maximized. At this point, the first derivative of the difference is zero and the second derivative is nonpositive. Comparing the two Bellman equations, three terms can generate differences in values across markets.

First, flow utility differences at a common skill and sector are bounded by $\max_{n,m}|\log A_n - \log A_m|$. Second, differences in learning drifts are bounded by γZ_H^*. Using the slope bound from Lemma 2(a), their contribution to the value difference is therefore bounded by $\gamma Z_H^*/(\rho+\delta)$. Finally, consider the mobility terms. Under the Type-I extreme value distribution, the continuation value of a relocation opportunity takes the log-sum-exp form in (2), which lies between the value of the best destination and that value plus $(\log N)/\nu$, the option value of the taste draws. The mobility terms therefore differ across origins by at most $(\lambda/\nu)\log N$, since the option to move dampens rather than amplifies value differences. Discounting these terms at $\rho+\delta$ gives

$$|v_n - v_m| \leq \bar{\Delta}$$

at every skill level.

This bound also places a lower bound on the probability that newborn workers enter any given market. Because location tastes are Gumbel, entry probabilities take the logit form in (10) and place strictly positive probability on every market. Since entry involves no mobility cost, the entry probabilities satisfy

$$\Pi_n(z) \geq \frac{e^{-\nu\bar{\Delta}}}{N}$$

for every z and n. Hence, market n receives a stationary entry inflow at rate at least $\delta\, e^{-\nu\bar{\Delta}}/N$.

Finally, population leaves market n at a rate no greater than $\lambda+\delta$, since relocation opportunities arrive at rate λ and death occurs at rate δ. Stationarity therefore implies

$$L_n \geq \frac{\delta}{\delta+\lambda}\frac{e^{-\nu\bar{\Delta}}}{N} = \underline{L},$$

which is the bound defined in (A4). $\square$

The final lemma converts the entry-moment bound in Assumption 3 into an a priori upper bound on the vibrancies that any sorted allocation can generate.

Lemma 6 (Vibrancy cap)**.** *Define*

$$\kappa \equiv \frac{\underline{L}}{c_\pi}, \qquad c_\pi \equiv \max_{S\in\{M,T\}} \frac{\max\{\pi^S,\, 1-\pi^S\}}{\min\{\pi^S,\, 1-\pi^S\}}, \tag{A5}$$

with $\underline{L}$ given by (A4).[25] *Under Assumption 3, every sorted stationary allocation whose learning drifts are bounded by γZ_H^* generates vibrancies bounded above by Z_H^*: $Z_n^S \le Z_H^*$ for all n and S.*

Proof of Lemma 6. Fix a sorted stationary allocation whose learning drifts are bounded by γZ_H^*. The stationary law of the worker process exists and is unique by Step 2 of the proof of Proposition 1. Importantly, that step relies only on the bound on learning drifts and does not use the result of this lemma.

We first bound the effective peer-skill distributions. By (5), the numerator of each distribution satisfies

$$\pi^S N_n^S(z) + (1-\pi^S)\, N_n^{-S}(z) \le \max\{\pi^S, 1-\pi^S\}\, N_n(z),$$

while its denominator satisfies

$$\pi^S L_n^S + (1-\pi^S)\, L_n^{-S} \ge \min\{\pi^S, 1-\pi^S\}\, L_n.$$

Using the population floor from Lemma 5, it follows that

$$\left(Z_n^S\right)^\theta = \int_0^\infty z^\theta\, g_n^S(z)\,\mathrm{d}z \;\le\; \frac{c_\pi}{\underline{L}} \int_0^\infty z^\theta \sum_{k=1}^{N} N_k(z)\,\mathrm{d}z.$$

We next bound the aggregate skill moment on the right-hand side. The drift condition established in Step 2 of the proof of Proposition 1, condition (CD2) of Meyn and Tweedie (1993) for the test function $V(n,z) = z^\theta$, together with their Theorem 4.3(ii), gives

$$\int_0^\infty z^\theta \sum_{k=1}^{N} N_k(z)\,\mathrm{d}z \;\le\; \frac{d}{c} \;=\; \frac{(\theta+1)\,\delta\, m_\psi(\theta)}{\delta - \theta^2\sigma^2/2}, \tag{A6}$$

where c and d are defined in that step.

Combining these two bounds gives

$$\left(Z_n^S\right)^\theta \le \frac{c_\pi}{\underline{L}}\, \frac{(\theta+1)\,\delta\, m_\psi(\theta)}{\delta - \theta^2\sigma^2/2}.$$

[25] The constant c_π is finite for interior segmentation, $\pi^S \in (0,1)$.

Therefore, $Z_n^S \leq Z_H^*$ for all n and S whenever

$$\frac{c_\pi}{\underline{L}} \frac{(\theta+1)\,\delta\, m_\psi(\theta)}{\delta - \theta^2\sigma^2/2} \;\leq\; (Z_H^*)^\theta \,.$$

This condition rearranges exactly to (13) with $\kappa = \underline{L}/c_\pi$. Assumption 3 therefore delivers the desired vibrancy cap. $\square$

Remark 1 (Matching conditions)**.** Under the cutoff policy, the general KFE (9) reduces to a two-region system. For $z < \bar{z}_n$, $N_n^T(z) = N_n(z)$, while for $z > \bar{z}_n$, $N_n^M(z) = N_n(z)$. At the cutoff, the stationary distribution satisfies continuity of mass and flux:

$$N_n^T(\bar{z}_n^-) = N_n^M(\bar{z}_n^+),$$

and

$$J_n^T(\bar{z}_n^-) = J_n^M(\bar{z}_n^+).$$

Because the learning drift jumps at $\bar{z}_n$, the stationary density generically exhibits a kink at the cutoff. These are the matching conditions we impose in the numerical solution of Appendix B.

A.2 Proofs Main Text

A.2.1 Proof of Proposition 1

Proof. We first establish existence for $\sigma_M = \sigma_T = \sigma$ in Steps 1–4, and then extend the construction to small volatility gaps in Step 5. Let $\mathcal{K} \equiv \{(Z_n^M, Z_n^T)_{n=1}^N : 0 \leq Z_n^T \leq Z_n^M \leq Z_H^*\}$, a compact convex subset of $\mathbb{R}^{2N}$, and fix $Z \in \mathcal{K}$.

Step 1: worker problem. Given the learning drifts implied by Z, the envelope equation (7) defines a discounted monotone system. The sector max and the mobility term $\frac{\lambda}{\nu}\log\sum_m e^{\nu(V_m - \xi_{nm})}$ are monotone and 1-Lipschitz in the sup norm. Hence, the Bellman operator is a contraction with modulus $\lambda/(\rho+\delta+\lambda)$ in a supremum norm weighted by $1+|x|$, under which flow payoffs are bounded with logarithmic utility. This delivers a unique solution with at most linear growth in x, jointly continuous in Z. Achdou et al. (2022) study analogous coupled systems.

Given this solution, Lemmas 2 and 4 apply. Labor supply therefore follows a cutoff policy, with thresholds $\bar{z}_n(Z) \in [\underline{z}, \hat{z}\,]$. Each threshold is unique because the flow advantage in (A3) has slope at least $1-\eta > 0$.

The thresholds are also continuous in Z. The values $v_n(\cdot\,; Z)$ are convex, so locally uniform convergence of the value functions implies convergence of their derivatives v_n' at every point of differentiability. The flow advantage Δ_n is therefore jointly continuous in its arguments, and its unique zero varies continuously with Z.

Step 2: stationary distribution. Given the policies, the worker process (n_t, z_t) is non-explosive. Log skill follows a Brownian motion with drift in $[0, \gamma Z_H^*]$ and volatility at most σ, while relocation and death occur at the finite rates λ and δ. Upon death, rebirth skill is drawn from ψ. Hence, sample paths remain in compact sets on bounded time intervals almost surely.

Death and rebirth also ensure existence and uniqueness of the stationary law. Death arrives at rate δ independently of the state and resets the worker to the entry law $\psi \otimes \Pi$, which does not depend on the pre-death state. The transition kernels therefore satisfy the uniform minorization

$$P^t(x, \cdot) \geq \int_0^t \delta e^{-\delta s} \big((\psi \otimes \Pi)\, P^{t-s} \big)(\cdot)\, \mathrm{d}s,$$

whose right-hand side is independent of x and has total mass $1 - e^{-\delta t}$. The whole state space is therefore a small set in the sense of Meyn and Tweedie (1993). Coupling two copies of the process at their first common death time then gives existence and uniqueness of the stationary law, together with uniform exponential convergence to it in total variation at rate $e^{-\delta t}$.

The drift condition is needed only to control moments of the stationary distribution. Consider the test function $V(n, z) = z^\theta$, which lies in the domain of the extended generator $\mathcal{A}$. The diffusion part contributes at most $\big(\theta \gamma Z_n^S + \theta^2 \sigma^2 / 2\big)\, z^\theta$, mobility leaves V unchanged, and death and rebirth contribute $\delta(m_\psi(\theta) - z^\theta)$. It follows that

$$\mathcal{A}V \;\leq\; -c\, V + d, \qquad c \equiv \delta - \theta \gamma Z_H^* - \frac{\theta^2 \sigma^2}{2} = \frac{\delta - \theta^2 \sigma^2 / 2}{\theta + 1} > 0, \qquad d \equiv \delta\, m_\psi(\theta), \tag{A7}$$

which is condition (CD2) of Meyn and Tweedie (1993) with $f = V + 1$. Their Theorem 4.3(ii) then bounds the stationary expectation of V by d/c, delivering (A6).

We also need a slightly higher moment below. Assumption 3 guarantees a finite entry moment at an exponent above θ. Since c remains positive at $\theta + \varepsilon$ by continuity, the same argument gives a uniform bound on the stationary $(\theta + \varepsilon)$-moment over $\mathcal{K}$ for some $\varepsilon > 0$.

Finally, the stationary distribution inherits the sorting structure from Step 1 because sector choice is a deterministic function of the worker's state under the cutoff policy. Its aggregate density therefore solves the stationary KFE (9), together with the matching conditions in Remark 1.

Step 3: fixed point. Let $T(Z)$ collect the vibrancies implied by the stationary distribution through (4) and (5). By Lemma 3, T preserves the ranking $T_n^M(Z) \geq T_n^T(Z)$. Moreover, Lemma 6 implies that these vibrancies remain bounded above by Z_H^*. Hence, T maps $\mathcal{K}$ into itself.

We next establish continuity. The stationary law varies continuously with Z through the policies and learning drifts, by its renewal representation and dominated convergence. The mixture weights in (5) are also continuous because their denominators are bounded below by $\min\{\pi^S, 1 - \pi^S\}\, \underline{L}$ from Lemma 5. Finally, the uniform $(\theta + \varepsilon)$-moment bound established in Step 2 implies uniform integrability at the exponent θ. The resulting θ-moments, and therefore T, are continuous in Z.

Since $\mathcal{K}$ is compact and convex and T is a continuous self-map of $\mathcal{K}$, Brouwer's fixed point theorem yields a fixed point $Z^* = T(Z^*)$.

Step 4: subsidy. At the fixed point Z^*, compute the subsidy rate from the government budget:

$$b^* = \frac{\sum_{n=1}^{N} \int_0^\infty \tau\, w_n^M(z)\, N_n^M(z)\, \mathrm{d}z}{\sum_{n=1}^{N} \sum_{S\in\{M,T\}} \int_0^\infty w_n^S(z)\, N_n^S(z)\, \mathrm{d}z}.$$

The integrals are finite because first moments are finite for $\theta \geq 1$. The numerator is strictly positive because entry cohorts place positive mass above $\bar{z}_n \leq \hat{z}$.

Under logarithmic utility, the proportional rebate raises all values by the common amount $\log(1+b^*)/(\rho+\delta)$ and therefore does not affect workers' decisions. In particular, the sector comparison, the log-sum-exp mobility terms, and the logit choice probabilities are all invariant to a common shift in values.

The allocation constructed in Steps 1–3, together with b^*, therefore satisfies conditions *(i)* to *(vi)* of the equilibrium definition.

Step 5: small volatility gaps. We first extend the cutoff argument to sufficiently small differences in sectoral volatilities. A volatility gap enters the worker's problem through the curvature term $\frac{1}{2}(\sigma_M^2 - \sigma_T^2)(v_n'' - v_n')$ in the log form of (11) and through the Itô corrections in the learning drifts. Both terms vanish as $\sigma_M \to \sigma_T$.

On each side of a candidate cutoff, the HJB equation is uniformly elliptic with smooth coefficients. The equation itself can therefore be used to express v_n'' in terms of v_n, v_n', and the flow payoff. This provides a uniform bound on v_n'' over the compact region $[\log \underline{z}, \hat{x}\,]$ where the flow advantage can vanish, uniformly over $\mathcal{K}$. It follows that

$$\partial_x \Delta_n \geq (1-\eta) - O(|\sigma_M^2 - \sigma_T^2|)$$

on this region. Hence, for sufficiently small volatility gaps, the flow advantage remains strictly increasing and retains a unique zero. The cutoff structure therefore survives. Steps 2–4 then go through unchanged because Lemma 6 is stated using $\sigma = \max\{\sigma_M, \sigma_T\}$.

It remains to establish the strict upper bound on the cutoff. Suppose $\pi^M > 1/2$ and $\pi^T > 1/2$. Under common volatilities, the cutoffs constructed above are interior. Since $\sigma_S > 0$, the stationary law places positive mass on both sides of every interior cutoff, so both sectors are populated. The strict version of Lemma 3 then implies $Z_n^M > Z_n^T$ in every market. Hence,

$$\Delta_n(\hat{x}) > 0$$

at $\sigma_M = \sigma_T$, which implies $\bar{z}_n < \hat{z}$.

This strict inequality continues to hold for sufficiently small volatility gaps. Suppose otherwise. Then there would exist a sequence of volatility gaps converging to zero and associated equilibria for which $\bar{z}_n \geq \hat{z}$ in at least one market. The corresponding vibrancies lie in the compact set $\mathcal{K}$. Since the estimates in Steps 1–3 are uniform and continuous jointly in (Z, σ_M, σ_T), we can take a convergent subsequence whose limit is a common-volatility equilibrium satisfying $\min_n \Delta_n(\hat{x}) \leq 0$.

This contradicts the strict inequality established above. Because there are finitely many markets, the same compactness argument yields a single threshold $\Sigma > 0$ such that $\bar{z}_n < \hat{z}$ for every market whenever the volatility gap is below Σ. $\square$

A.2.2 Proof of Proposition 2

Proof. We first consider $\sigma_M = \sigma_T$. Volatility gaps below Σ are handled using the same perturbation and compactness arguments as in Step 5 of the proof of Proposition 1. It is convenient to work with learning drifts rather than vibrancies. Let $\mu = \gamma Z$ and define the fixed box

$$\mathcal{D} \equiv \{\mu : 0 \leq \mu_n^T \leq \mu_n^M \leq \gamma Z_H^*\},$$

whose upper bound γZ_H^* does not depend on γ. The corresponding map on learning drifts is

$$\widetilde{T}(\mu) \equiv \gamma\, T(\mu/\gamma).$$

Step 1: reduction to the drift fixed point. Consider any stationary equilibrium with $Z_n^S \leq Z_H^*$. Its learning drifts lie in $[0, \gamma Z_H^*]$, so the moment bound (A6) applies. Combining this bound with the mixture bound in the proof of Lemma 6 gives

$$Z_n^S \leq \bar{Z} \equiv \left[\frac{c_\pi}{\underline{L}} \, \frac{(\theta + 1)\, \delta\, m_\psi(\theta)}{\delta - \theta^2 \sigma^2/2} \right]^{1/\theta},$$

where $\bar{Z}$ does not depend on γ.

We next show that, for sufficiently small γ, every capped equilibrium follows a cutoff policy. In such an equilibrium, the flow advantage (A3) has slope at least $(1 - \eta) - 2\gamma \bar{Z} \sup |v_n''|$. As in Step 5 of the proof of Proposition 1, the second derivative is uniformly bounded on a compact interval around $\hat{x}$ containing every zero of the flow advantage in a capped equilibrium, regardless of the sign of $Z_n^M - Z_n^T$. For sufficiently small γ, the slope is therefore strictly positive, so the flow advantage has a unique zero and labor supply follows a cutoff policy.

Every capped equilibrium is thus sorted. Lemma 3 then ranks its learning environments, so its drift vector lies in $\mathcal{D}$. Uniqueness among capped equilibria therefore reduces to uniqueness of the fixed point of $\widetilde{T}$ on $\mathcal{D}$.

Step 2: contraction. We now show that $\widetilde{T}$ is a contraction on $\mathcal{D}$ for sufficiently small γ. Drifts affect $\widetilde{T}$ through two channels: their effect on workers' values and policies, and their effect on the stationary distribution.

Consider first the value channel. A drift perturbation of size $\|\mu - \mu'\|$ changes values by $O(\|\mu - \mu'\|)$ in the weighted norm because the max and mobility operators are 1-Lipschitz. Interior gradient estimates imply that the derivatives v_n' change by the same order, uniformly over $\mathcal{D}$ since drifts and volatilities are bounded. Applying the implicit function theorem to (A3), the resulting cutoff movements are therefore of order $O(\|\mu - \mu'\|)/(1 - \eta)$.

Consider next the distribution channel. The renewal representation of the stationary law bounds the effect of a drift perturbation on the θ-moments by a constant times $\|\mu - \mu'\|$. This constant depends only on δ, θ, σ, and $m_\psi(\theta)$ because the age integral converges at the γ-free rate c in (A7). Cutoff movements also relabel a band of workers, with an effect on sectoral moments proportional to the cutoff movement. This follows because the stationary density is bounded near the cutoff by the renewal representation and the mixture denominators are bounded below by the market floor in Lemma 5.

Combining these two channels, the induced vibrancies change by at most a constant times $\|\mu - \mu'\|$. The output of $\widetilde{T}$ carries an additional factor γ through (B1), while all constants in the preceding bounds depend only on the γ-free quantities γZ_H^*, $\bar{Z}$, and $m_\psi(\theta)$. Hence, for some constant C independent of γ,

$$\|\widetilde{T}(\mu) - \widetilde{T}(\mu')\| \leq C\gamma \|\mu - \mu'\|$$

for all $\mu, \mu' \in \mathcal{D}$. For $\gamma < \bar{\gamma} \equiv 1/C$, the map is a contraction and therefore has a unique fixed point.

At $\gamma = 0$, the result is immediate. Learning is inactive, every cutoff equals $\hat{z}$, and $\widetilde{T} \equiv 0$. □

A.2.3 Proof of Lemma 1

Proof. Consider first workers who remain in (n, S) over $[t, t+h]$. Integrating (3) and using the log-linear wage schedules in (1) under Assumption 1, expected log-wage growth is $\gamma Z_n^M h$ in the modern sector and $\eta \gamma Z_n^T h$ in the traditional sector. The Brownian increment vanishes in expectation, while the wage elasticity with respect to skill is one in the modern sector and η in the traditional sector. Hence, the difference in expected wage growth coming from learning is

$$\gamma(Z_n^M - \eta Z_n^T)h.$$

This term is strictly positive. Since $\pi^M, \pi^T > 1/2$, we have $\pi^M + \pi^T > 1$. Moreover, the cutoffs are interior and the nondegenerate diffusion places stationary mass on both sides of the cutoff, so both sectors are populated. The strict version of Lemma 3 therefore gives $Z_n^M > Z_n^T$. It follows that

$$\gamma(Z_n^M - \eta Z_n^T) = \gamma(Z_n^M - Z_n^T) + (1 - \eta)\gamma Z_n^T > 0.$$

We next account for workers who switch sectors within the horizon. Workers who cross the cutoff change the wage schedule on which their wage is measured. On the interval $(\bar{z}_n, \hat{z})$, the modern-sector wage is below the traditional-sector wage. Thus, a traditional worker who switches into the modern sector records a wage-level loss relative to remaining traditional, while a modern worker who switches into the traditional sector records the corresponding gain. Both movements increase the difference in wage growth between workers initially in the modern and traditional sectors.

Over a short horizon, the mass of workers crossing the cutoff is of order $\sqrt{h}$ and is concentrated near the cutoff. Because $\bar{z}_n < \hat{z}$, the wage difference around the cutoff is strictly positive. The resulting reallocation term is therefore nonnegative and of order $\sqrt{h}$. Migration and death occur at rates

λ and δ, so their effects on average wage growth are of order h and are dominated at sufficiently short horizons. Death is neutralized by conditioning on workers observed at both dates.

Collecting these terms, there exists $\bar{h} > 0$ such that, for every $h \in (0, \bar{h}]$,

$$\mathbb{E}_t\left[\Delta \log w^M_{n,t+h}\right] - \mathbb{E}_t\left[\Delta \log w^T_{n,t+h}\right] \geq \gamma\left(Z^M_n - \eta\, Z^T_n\right) h > 0.$$

$\square$

B Computational Appendix

This appendix describes the algorithm we use to compute the stationary equilibrium, based on finite-difference methods in the spirit of Achdou et al. (2022). Throughout, we work with log skill, $x \equiv \log z$. By (3), log skill follows a Brownian motion with constant drift within each market and sector,

$$\bar{\mu}_n^S = \gamma\, Z_n^S, \qquad S \in \{M, T\}, \tag{B1}$$

where the vibrancies Z_n^S are given by (4), evaluated at the stationary peer-skill distributions (5). The bar distinguishes the drift of *log* skill from the level drift μ_n^S in (6), which includes the Itô correction $\sigma_S^2/2$. Since $\gamma \geq 0$ and $Z_n^S \geq 0$, these drifts are nonnegative, a sign restriction we exploit in the discretization below. We discretize x on a uniform grid $\{x_j\}_{j=1}^J$ with spacing Δx and approximate the continuous-time model by a discrete-time model with period length Δt. Our implementation uses $J = 300$ nodes spanning $z \in [0.1,\, 1000]$ and $\Delta t = 0.003$ years.

The algorithm has three nested components. Given drifts and a subsidy rate, Appendix B.1 solves the workers' problem (7) using Howard's policy iteration, delivering values and sector cutoffs. Given values, cutoffs, and drifts, Appendix B.2 solves the stationary KFE (9) as a single sparse linear system. Appendix B.3 then closes the model by expressing the stationary equilibrium as a fixed point in the $2N$ vibrancies and solving it by damped iteration over the map generated by the two inner solvers. Finally, Appendix B.6 describes how we compute transitional dynamics.

B.1 Solving the HJB Equation

Fix drifts (B1) and a subsidy rate b. In logs, the stationary envelope HJB equation (7) becomes

$$(\rho + \delta)\, V_n(x) = \max_{S \in \{M,T\}} \Big\{ u_n^S(x) + \bar{\mu}_n^S\, \partial_x V_n(x) + \frac{\sigma_S^2}{2}\, \partial_{xx} V_n(x) \Big\} + \mathcal{M}_n[V](x), \tag{B2}$$

where $u_n^S(x) = \log\big((1+b)\, w_n^S(e^x)\big)$ is flow utility under the wage schedules (1),[26] and $\mathcal{M}_n[V](x)$ is the mobility option value (2), with time arguments omitted.

We approximate (B2) by the Bellman equation of a discrete-time Markov chain on the grid. Over a period of length Δt, a worker in sector S of market n discounts the future by the survival-adjusted factor $\beta \equiv 1 - \Delta t\,(\rho + \delta)$ and transitions to neighboring nodes with probabilities

$$p_+^{n,S} = \frac{\Delta t}{\beta} \left(\frac{\bar{\mu}_n^S}{\Delta x} + \frac{\sigma_S^2}{2\Delta x^2} \right), \qquad p_-^S = \frac{\Delta t}{\beta}\, \frac{\sigma_S^2}{2\Delta x^2}, \qquad p_0^{n,S} = 1 - p_+^{n,S} - p_-^S, \tag{B3}$$

which we collect in a row-stochastic tridiagonal matrix P_n^S. At the two boundary nodes, the outward move is suppressed, so the grid boundaries are reflecting. We upwind the drift forward because $\bar{\mu}_n^S \geq 0$. All three probabilities are nonnegative provided $\Delta t\,\big(\rho+\delta+\bar{\mu}_n^S/\Delta x+\sigma_S^2/\Delta x^2\big) \leq 1$,

[26] In the implementation, the logarithm is guarded by a linear extension below a small consumption floor (10^{-4}); the floor never binds on the equilibrium wage schedules.

which is satisfied by our time step. The resulting discrete Bellman equation is

$$V_n(x_j) = \max_{S\in\{M,T\}} \Big\{ \Delta t\, u_n^S(x_j) + \beta \left[P_n^S V_n\right]_j \Big\} + \Delta t\, \mathcal{M}_n[V](x_j), \tag{B4}$$

which converges to (B2) as the grid spacing and time step shrink.

We solve (B4) by Howard's policy iteration on the $N \times J$ state space. Given a candidate envelope V^k, the policy step assigns to each node the sector with the larger Bellman return,

$$S_n^k(x_j) = \operatorname*{argmax}_{S\in\{M,T\}} \Big\{ \Delta t\, u_n^S(x_j) + \beta \left[P_n^S V_n^k\right]_j \Big\},$$

where the mobility term drops out because it is common to both sectors.

Given this policy, the evaluation step solves the linear system implied by (B4),

$$(1 + \Delta t\, \lambda)\, V_n^{k+1}(x_j) - \beta \left[P_n^{S^k} V_n^{k+1}\right]_j = \Delta t\, u_n^{S^k}(x_j) + \frac{\Delta t\, \lambda}{\nu} \log \Bigg(\sum_{m=1}^{N} e^{\nu\left(V_m^k(x_j) - \xi_{nm}\right)} \Bigg), \tag{B5}$$

which evaluates the log-sum-exp inside $\mathcal{M}_n[V]$ at the lagged iterate while treating its own-value term, $-\lambda V_n$, implicitly. Each evaluation step requires a single sparse solve over the NJ stacked states. We alternate between the policy and evaluation steps until values converge in the sup norm, using a tolerance of 10^{-5}.

Relative to direct iteration on (B4), Howard's policy iteration replaces many small value updates with exact solves under candidate policies and converges in a small number of iterations. From a cold start, we initialize values at the annuitized modern flow value, $V^0 = u^M/(\rho + \delta)$. Inside the equilibrium loop of Appendix B.3, we instead warm-start from the values obtained in the previous iteration.

Given the converged envelope, sector-specific values follow from one application of each sector's Bellman operator,

$$V_n^S = \Delta t\, u_n^S + \beta\, P_n^S V_n + \Delta t\, \mathcal{M}_n[V], \qquad S \in \{M, T\},$$

with $V_n = \max\{V_n^M, V_n^T\}$ node by node. The labor supply policy is determined by the sign change of $\Delta_n \equiv V_n^M - V_n^T$. Letting j denote the first node at which the policy switches from the traditional to the modern sector, linear interpolation gives

$$\bar{x}_n = x_j - \frac{\Delta_n(x_j)}{\Delta_n(x_{j+1}) - \Delta_n(x_j)}\, \Delta x. \tag{B6}$$

If one sector is selected everywhere, we assign the cutoff to the corresponding end of the grid.

B.2 Solving the Kolmogorov Forward Equation

Given values, cutoffs, and drifts, the stationary distribution solves (9) with $\partial_t N_n^S = 0$. Under the cutoff policy, sector choice is a deterministic function of the worker's state. It therefore suffices to solve for the aggregate densities $N_n(x)$ and recover the sectoral densities as

$$N_n^M(x) = N_n(x)\,\mathbb{1}\{x > \bar{x}_n\}, \qquad N_n^T(x) = N_n(x)\,\mathbb{1}\{x \le \bar{x}_n\}.$$

We discretize the Brownian part of (9) in conservative form using a finite-volume scheme. Each grid node represents a cell that exchanges mass with its neighbors through common faces, with each face flux added to one cell and subtracted from the other. Mass conservation therefore holds cell by cell, including at the cutoff face where the drift and diffusion coefficients jump. This is the discrete counterpart of the continuity of mass and flux at $\bar{z}_n$ imposed in Appendix A.

For this reason, we do not construct the forward operator by transposing the HJB matrices of Appendix B.1. The transposed scheme assigns the cutoff face the coefficients of a single node, misrepresenting the flux precisely where the coefficients jump. The conservative scheme instead imposes flux continuity by construction.

Concretely, drop the market index and let cell j carry the coefficients $(\bar{\mu}_j, \sigma_j^2)$ of its sector, modern if and only if $x_j > \bar{x}_n$. Conditional on receiving neither a mobility nor a death shock over the period, an event with probability $1 - \Delta t\,(\lambda + \delta)$, mass crosses the face between cells j and $j+1$ upward and downward with probabilities

$$q_j^+ = s\left(\frac{\bar{\mu}_j + \bar{\mu}_{j+1}}{2\,\Delta x} + \frac{\sigma_j^2}{2\Delta x^2}\right), \qquad q_j^- = s\,\frac{\sigma_{j+1}^2}{2\Delta x^2}, \qquad s \equiv \frac{\Delta t}{1 - \Delta t\,(\lambda + \delta)}, \tag{B7}$$

so the drift is transported using the face-averaged $\bar{\mu}$ and forward-only upwinding, while each cell diffuses across a face at the rate implied by its own σ^2. We impose zero flux at the ends of the grid. Collecting these probabilities yields a column-stochastic tridiagonal matrix Q_n for each market. Every column sums to one, so any mass leaving a cell is exactly transferred to an adjacent cell.

The stationary distribution then satisfies the discrete-time balance

$$\boldsymbol{N} = \big(1 - \Delta t\,(\lambda + \delta)\big)\,Q\,\boldsymbol{N} + \Delta t\,\lambda\,\boldsymbol{M}(V)\,\boldsymbol{N} + \Delta t\,\delta\,\boldsymbol{\psi}, \tag{B8}$$

where $\boldsymbol{N} \in \mathbb{R}^{NJ}$ stacks the market densities and $Q = \mathrm{diag}(Q_1, \ldots, Q_N)$. The mobility operator reallocates mass across markets at fixed skill according to the logit shares of Section 2,

$$[\boldsymbol{M}(V)\boldsymbol{N}]_n(x_j) = \sum_{k=1}^{N} m_{kn}(x_j)\,N_k(x_j),$$

while the entry vector has components

$$[\boldsymbol{\psi}]_n(x_j) = \psi(x_j)\,\Pi_n(x_j),$$

with Π_n denoting the newborn location probabilities in (10).

Equation (B8) is a sparse linear system in $\boldsymbol{N}$, which we solve with a single factorization; no time iteration is required. Moreover, Q and $\boldsymbol{M}(V)$ are column-stochastic, Π_n sums to one across markets, and ψ is normalized to integrate to one on the grid. Summing (B8) over all states implies that the stationary distribution integrates to one exactly, so no renormalization is required.

B.3 Solving for the Stationary Equilibrium

The outer loop combines the two solvers into a stationary equilibrium as defined in Section 2.4. The key observation, anticipated in the discussion of uniqueness preceding Proposition 2, is that conditional on the vector of vibrancies $Z = \{Z_n^M, Z_n^T\}_{n=1}^N$, all remaining equilibrium objects are pinned down sequentially. Wages follow from optimal labor demand, condition (i), in closed form from (1). Values and the cutoff policy follow from the workers' dynamic problem, condition (ii), computed as in Appendix B.1. The stationary distribution follows from the distribution dynamics, condition (iv), computed as in Appendix B.2. Finally, the subsidy rate follows from budget balance, condition (v):

$$b = \tau \, \frac{\sum_{n=1}^N \int w_n^M(z)\, N_n^M(z)\, \mathrm{d}z}{\sum_{n=1}^N \sum_{S\in\{M,T\}} \int w_n^S(z)\, N_n^S(z)\, \mathrm{d}z}. \tag{B9}$$

Labor market clearing, condition (vi), holds by construction: firms absorb all workers at the posted wages, while the unit-mass condition is preserved exactly by the KFE scheme.

The remaining equilibrium condition is learning consistency, condition (iii). The vibrancies implied by the computed sectoral distributions through (4) and (5) must coincide with those workers take as given when solving their problem. A stationary equilibrium is therefore a fixed point $Z = T(Z)$ of the finite-dimensional map $T : \mathbb{R}_+^{2N} \to \mathbb{R}_+^{2N}$ that, given vibrancies, computes the remaining equilibrium objects and returns the implied vibrancies. We solve this fixed point by damped successive approximation on T, or equivalently, by (B1), by iterating on the drifts $\bar{\mu}_n^S = \gamma Z_n^S$:

Step 0. Initialize. Set $N_n^0(x) = \psi(x)/N$, so newborn skills are spread evenly across markets, with a common sector cutoff at the grid midpoint; alternatively, warm-start from a previously computed equilibrium. Compute Z^0 from (4) and (5), set $\bar{\mu}^0 = \gamma Z^0$, and compute b^0 from (B9).

Step 1. Worker problem. Given $(\bar{\mu}^r, b^r)$, solve (B4) to obtain values V^r and cutoffs $\bar{x}^r$.

Step 2. Distribution. Solve (B8) at $(\bar{\mu}^r, \bar{x}^r, V^r)$ to obtain $\widetilde{N}$. Relax toward the previous iterate, $N^{r+1} = \omega_N \widetilde{N} + (1-\omega_N)\, N^r$, and split the resulting distribution into sectors at $\bar{x}^r$.

Step 3. Learning consistency. Compute $\widetilde{Z}$ from (4) and (5) at N^{r+1} and obtain the implied drifts $\gamma\widetilde{Z}$.

Step 4. Convergence. If $\max_{n,S} \left|\gamma \widetilde{Z}_n^S - \bar{\mu}_n^{S,r}\right| / \bar{\mu}_n^{S,r}$ is below tolerance, stop. Otherwise, set $\bar{\mu}^{r+1} = \omega\, \gamma \widetilde{Z} + (1-\omega)\, \bar{\mu}^r$, update b^{r+1} from (B9), and return to Step 1, warm-starting the value solver at V^r.

The two relaxations ($\omega = \omega_N = 0.2$; convergence tolerance 10^{-2}) temper the feedback loop at the core of the model. A larger modern sector raises Z_n^M, which increases $\bar{\mu}_n^M$ and, in turn, attracts more workers into the modern sector. Undamped iteration can overshoot through this feedback and spuriously absorb the traditional sector. At convergence, conditions (i)–(vi) hold simultaneously up to the stated tolerance. Finally, as discussed after Proposition 2, we probe for multiple equilibria by re-running the algorithm from different initial distributions.

B.4 Recovering Migration Costs

We recover the bilateral migration costs $\{\xi_{nm}\}_{n,m=1}^N$ from migration flows observed in the EPS. We first construct year-to-year transitions between consecutive survey years and compute survey-weighted flows between every origin and destination market, averaging these flows over the sample period. Dividing each flow by the total mass of workers from its origin market, including those who stay, yields the empirical migration shares $\hat{m}_{nm}$.

The mobility shares of Section 2 imply that, for a worker in market n,

$$X_{nm} \equiv \log \hat{m}_{nm} - \log \hat{m}_{nn} = \nu \, (V_m - V_n) - \nu \, \xi_{nm}, \tag{B10}$$

where we impose the normalization $\xi_{nn} = 0$. Taking the corresponding expression for the reverse flow and adding the two equations eliminates the value terms:

$$X_{nm} + X_{mn} = -\nu \, (\xi_{nm} + \xi_{mn}).$$

Assuming symmetric migration costs, $\xi_{nm} = \xi_{mn}$, we recover

$$\xi_{nm} = -\frac{X_{nm} + X_{mn}}{2\nu}, \tag{B11}$$

using the calibrated taste dispersion ν.

Two qualifications are worth noting. First, we do not recover skill-specific migration costs. Although mobility shares in the model depend on skill through workers' values, we pool workers across skill levels when constructing the empirical shares and recover a single migration cost for each market pair. Second, market pairs with no observed flows in either direction are assigned an infinite migration cost, ruling out migration between them in the model.

B.5 Construction of Model Moments

The estimation in Section 4 matches ten moments, all computed directly from the stationary equilibrium, so no simulation is required. Two of these moments are cross-sectional. The modern-sector employment share is given by the mass of workers above the sectoral cutoffs, $\sum_n \int N_n^M(z)\, \mathrm{d}z$, where total population is normalized to one. The variance of modern log wages is the variance of $\log w_n^M(z)$ under the modern-sector distribution $N_n^M(z)$. Both moments are computed directly from the stationary distribution obtained in Appendix B.2.

The transition moments track individual workers over time and therefore require the law of motion of a worker's state, (n, x). On the NJ-state grid, this law of motion is summarized by the generator

$$\mathcal{Q} \;=\; \frac{Q - I}{\Delta t} \;+\; \lambda\big(\boldsymbol{M}(V) - I\big) \;-\; \delta I, \tag{B12}$$

where the three terms capture the per-year rates of learning through drift and diffusion, migration across labor markets, and death, respectively. The generator is evaluated at the stationary equilibrium using the operators defined in Appendix B.2.

Death is absorbing in (B12). The newborn who replaces an exiting worker is a different individual in the panel, so the rebirth inflow that appears in the KFE is excluded here. For any mass vector $\boldsymbol{\nu}$ of workers, $e^{h\mathcal{Q}}\boldsymbol{\nu}$ gives their distribution across states h years later, scaled by the surviving share $e^{-\delta h}$. Because the sector is a property of the grid cell, with modern cells above the stationary cutoff and traditional cells below it, sector and wage classifications at any two dates use the same time-invariant labels.

The sector transition probabilities condition on survival. Let $\mathbf{1}_S$ indicate the cells assigned to sector S and let $\boldsymbol{N}^S$ denote the stationary mass in those cells. The probability that a worker who starts in sector S is in sector S' after h years, conditional on surviving, is

$$P^h_{SS'} \;=\; \frac{\mathbf{1}'_{S'}\, e^{h\mathcal{Q}}\, \boldsymbol{N}^S}{\mathbf{1}'\, e^{h\mathcal{Q}}\, \boldsymbol{N}^S}, \tag{B13}$$

where $\mathbf{1}$ sums over all cells, so the denominator removes the uniform survival factor. The moment vector includes the one-year staying probabilities, P^1_{MM} and P^1_{TT}.

The within-sector wage-quintile transitions mirror their construction in the data. Quintiles are defined separately within each sector's wage distribution, and the sample is restricted to workers observed in the same sector at both endpoints. Let $\mathbf{1}^S_i$ indicate the sector-S cells in the i-th quintile of the sector's stationary wage distribution, and let $\boldsymbol{N}^S_i$ denote the stationary mass in those cells. The five-year transition probability from quintile i to quintile j, conditional on surviving and being in sector S at both endpoints, is

$$P^{5,S}_{ij} \;=\; \frac{(\mathbf{1}^S_j)'\, e^{5\mathcal{Q}}\, \boldsymbol{N}^S_i}{\sum_{j'} (\mathbf{1}^S_{j'})'\, e^{5\mathcal{Q}}\, \boldsymbol{N}^S_i}, \tag{B14}$$

where the denominator retains only the mass of workers who remain in sector-S cells after five years. The moment vector uses the interior diagonal elements, $P^{5,S}_{22}$, $P^{5,S}_{33}$, and $P^{5,S}_{44}$, for each sector, giving six moments. We exclude the corner elements for the reasons discussed in Section 4.

B.6 Transitional Dynamics via the Sequence-Space Jacobian

We compute the perfect-foresight transitions generated by the two permanent shocks considered in Section 5: eliminating the payroll tax and equalizing the learning technology for traditional workers. We solve for these transitions to first order in sequence space, following the continuous-time formulation of Bilal and Goyal (2026). Both reforms are unanticipated and permanent.

Following the convention in Auclert et al. (2021) for permanent shocks, we therefore linearize the equilibrium conditions around the terminal steady state, whose objects we denote with an asterisk. The transition is then expressed as a path of deviations from this terminal equilibrium. At the time of the reform, the economy is still distributed according to the baseline stationary equilibrium, so the initial deviation is $\Delta N_0 = N^0 - N^*$. These deviations converge to zero along the transition path. We truncate the computation at $T = 300$ years and use the remaining deviation at the terminal date as a diagnostic for the accuracy of the truncation.

The only aggregate sequences we need to solve for are the $2N$ vibrancy paths, $\Delta Z = \{\Delta Z_n^S(t)\}$. Vibrancies play a dual role in the equilibrium. On the one hand, they enter the individual worker problem through the learning drift in (B1). On the other hand, they are determined by the cross-sectional distribution of workers through the CES aggregator in (4). In this sense, vibrancies play the role of the price sequences in Auclert et al. (2021). No other aggregate sequence needs to be solved for. Wages depend only on parameters and market productivities, while the subsidy rate shifts flow utility uniformly across states and therefore does not affect workers' decisions.

Writing $\Delta V(\Delta Z)$ for the value deviations implied by the backward block and $\Delta N(\Delta Z; \Delta N_0)$ for the distribution deviations implied by the forward block, equilibrium requires the vibrancies generated by the resulting cross-sectional distribution to coincide with the conjectured paths:

$$\Delta Z_t^{\text{impl}}\big(\Delta V, \Delta N\big) - \Delta Z_t = 0, \qquad t = 0, \ldots, T-1. \tag{B15}$$

We construct the linearized blocks using exact derivatives of the discretized steady-state operators described in Appendices B.1 and B.2. This approach preserves the adjoint relation between the backward and forward blocks at the discrete level. Let Φ denote the one-year worker-level transition operator at the terminal equilibrium, incorporating learning, migration, and death—that is, $\Phi = e^{\mathcal{Q}}$, with the generator (B12) evaluated at the terminal equilibrium—and let Φ' denote its adjoint. Throughout this appendix, primes denote adjoint operators. Newborn entry remains at its terminal value, except for the behavioral response of location choice that we introduce below. We also define Φ_V as the discounted counterpart of Φ under the discount rate ρ. Because death is already incorporated into Φ, together these two components imply that workers' values discount at rate $\rho + \delta$.

We first consider the backward block. Changes in vibrancy affect the Bellman equation through the learning drift. Hence, the value response to a change in vibrancy is given by the discounted expected marginal value of the resulting increase in skill growth:

$$\Delta V_t = \sum_{a \geq 0} \Phi_V^a \, G_V \, \Delta Z_{t+a}, \qquad \big(G_V \, \Delta Z\big)_n(x) \equiv \gamma \, \partial_x V_n^*(x) \, \Delta Z_n^{S_n^*(x)}, \tag{B16}$$

where $S_n^*(x)$ denotes the terminal sector assignment. We compute the full set of kernels with a single linear backward pass. We next linearize the sector-indifference condition of Appendix B.1 around the terminal cutoffs. The resulting cutoff response is

$$\Delta \bar{x}_t = C_V \, \Delta V_t + C_Z \, \Delta Z_t, \tag{B17}$$

where the first term captures how changes in workers' values affect the sector choice, while the second captures the direct effect of changes in vibrancy on the dynamic gains from each sector.

We then turn to the forward block. The distribution deviation evolves according to the linearized Kolmogorov forward equation,

$$\Delta N_{t+1} = \Phi\, \Delta N_t + G_N\, \Delta Z_t + B\, \Delta V_t, \tag{B18}$$

where G_N and B capture two forces affecting the evolution of the distribution. The operator G_N moves the terminal distribution in response to changes in the learning drift through the derivative of the conservative flux in (B7). The operator B captures behavioral reallocation through changes in migration-destination and newborn-location choice probabilities induced by changes in workers' values. Both operators also incorporate the cutoff response in (B17), which changes the sector assignment, and therefore the learning technology, of workers at the cutoff.

The final block maps the distribution back into implied vibrancies. Changes in the distribution affect vibrancy through the CES gradients of (4), which we collect in the matrix E, as well as through the reassignment of workers around the sectoral cutoff:

$$\Delta Z_t^{\text{impl}} = E'\, \Delta N_t + G_{\bar{x}}\, \Delta \bar{x}_t. \tag{B19}$$

Combining (B16)–(B18) with (B19) turns the equilibrium condition in (B15) into a linear system in ΔZ. We construct its Jacobian, $\mathcal{J} = \partial \Delta Z^{\text{impl}} / \partial \Delta Z$, using the fake-news algorithm of Auclert et al. (2021).

In particular, the expectation vectors $E_a = (\Phi')^a E$ propagate the output functionals forward, while the distribution impulses $D_q = B\, \Phi_V^q\, G_V + \mathbb{1}\{q = 0\}\, G_N$ collect the forward response to vibrancy news arriving at horizon q. Their inner products give the fake-news matrix, $\mathcal{F}_{a,q} = E_a'\, D_q$, from which the standard recursion delivers $\mathcal{J}$. We additionally include the contemporaneous response operating through the within-period cutoff, given by $G_{\bar{x}} C_V \Phi_V^q G_V$ and $G_{\bar{x}} C_Z$. The equilibrium path then solves

$$(I - \mathcal{J})\, \Delta Z = R, \qquad R_a = E_a'\, \Delta N_0, \tag{B20}$$

where R captures the autonomous propagation of the initial distribution gap.

With $N = 27$ markets, two sectors, and a horizon of $T = 300$ years, (B20) is a single dense linear system of size $2NT = 16{,}200$. We aggregate time into one-year periods and construct the annual transition operators using twelve implicit Euler substeps, with the required factorization computed only once.

C Additional Figures

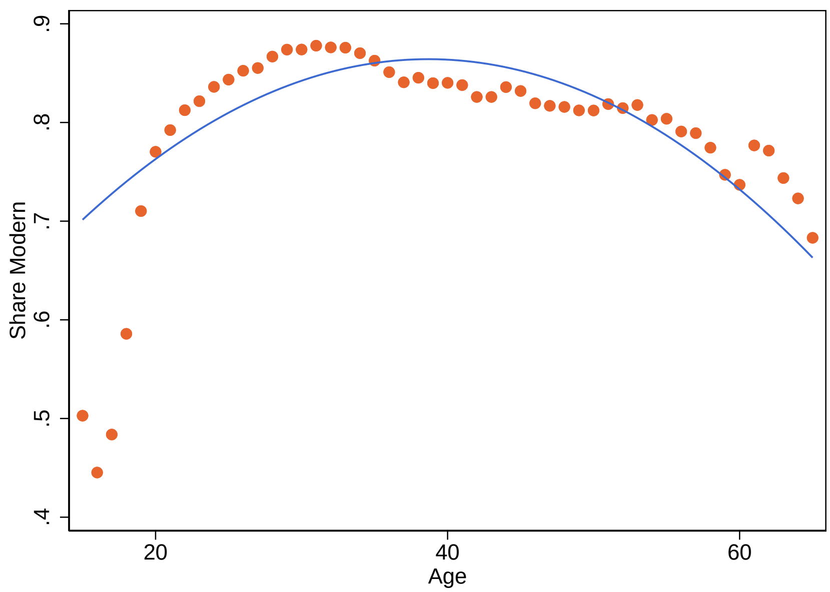
C.1 Descriptive Statistics

This section reports additional descriptive statistics on the share of modern workers in Chile. We confirm previously well-established patterns for the Chilean economy:

1. The share of modern workers follows an inverse U-shaped pattern over the life cycle.
2. The share of modern workers increases with educational attainment.
3. The share of modern workers increases with firm size.

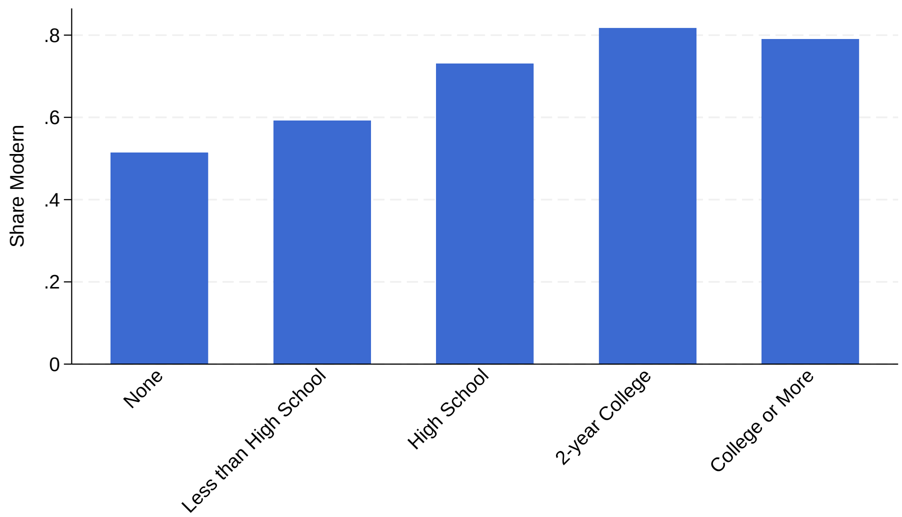
Then, we report the share of modern workers by occupation and economic sector.

Figure C1: Share of Modern Workers by Age



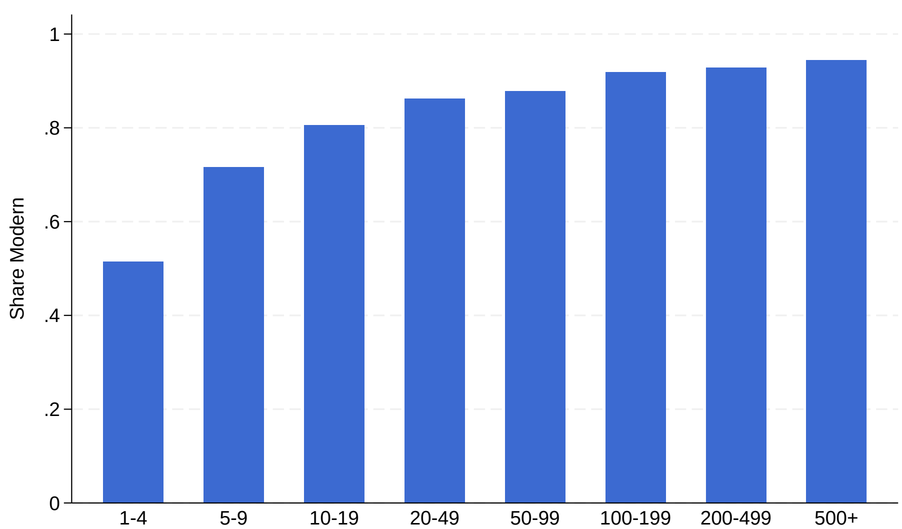
Notes: Figure C1 displays the average share of modern workers at each age, together with a quadratic fit. We restrict our sample to salaried workers, excluding self-employed workers, entrepreneurs, workers without remuneration, and public-sector and armed forces employees.

Figure C2: Share of Modern Workers by Educational Attainment



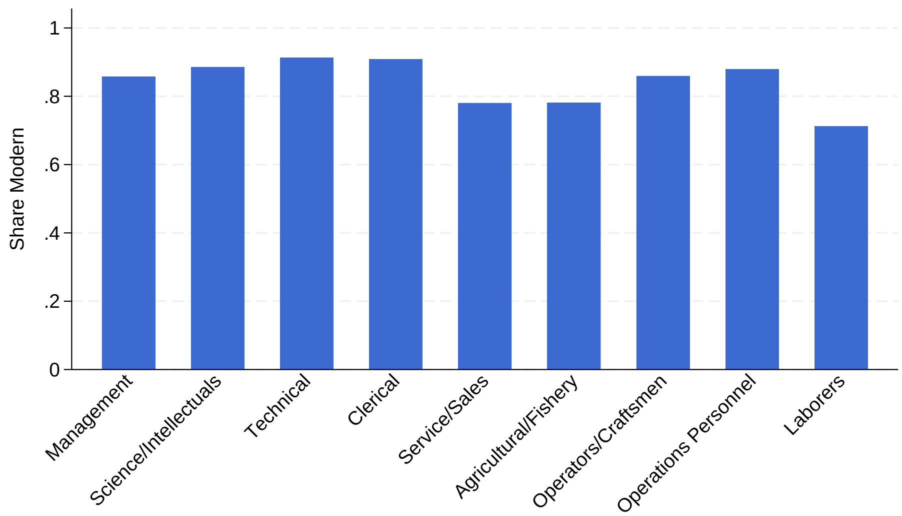
Notes: Figure C2 displays the share of modern workers by education group. We consider all workers in our baseline sample, including self-employed workers and entrepreneurs.

Figure C3: Share of Modern Workers by Firm Size



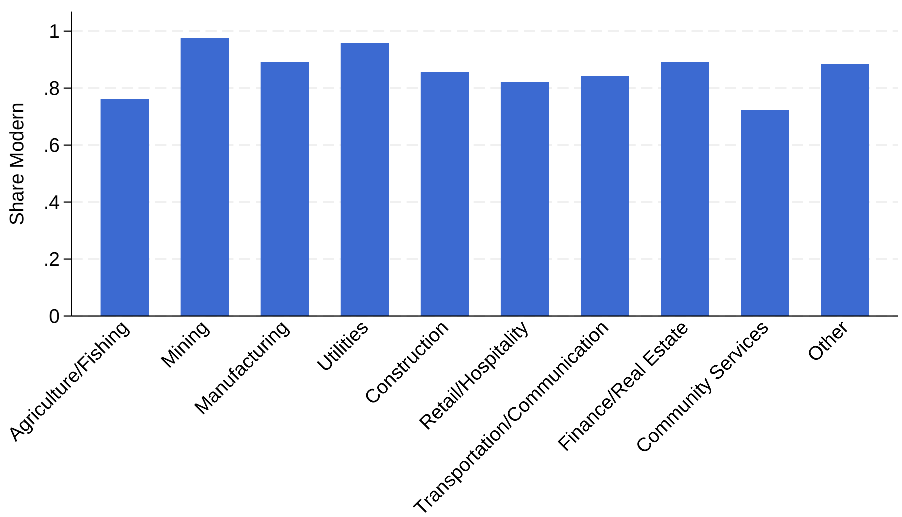
Notes: Figure C3 displays the share of modern workers by firm size. Bars denote firm size groups, and labels denote the range of employees in each group. We restrict our sample to salaried workers, excluding self-employed workers, entrepreneurs, workers without remuneration, and public-sector and armed forces employees.

Figure C4: Share of Modern Workers by Occupation



Notes: Figure C4 displays the share of modern workers by occupation. We restrict our sample to salaried workers, excluding self-employed workers, entrepreneurs, workers without remuneration, and public-sector and armed forces employees.

Figure C5: Share of Modern Workers by Sector



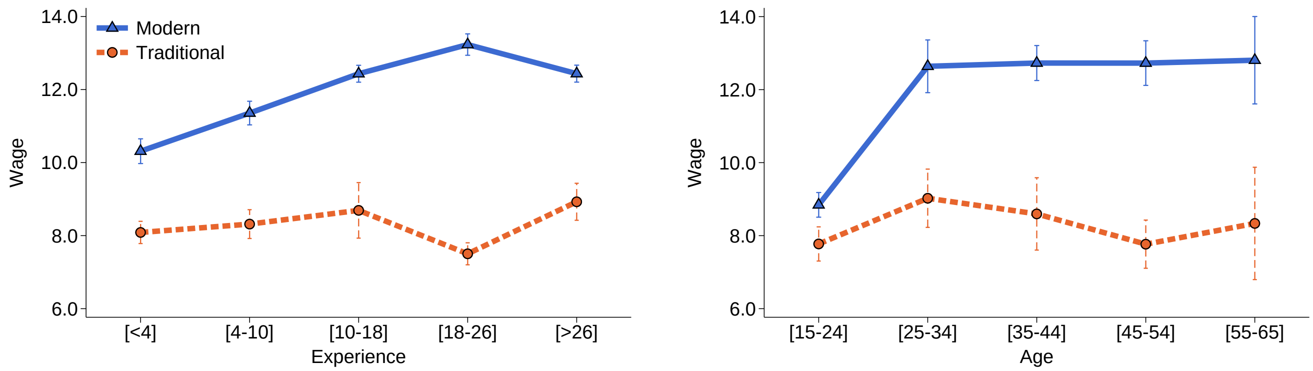
Notes: Figure C5 displays the share of modern workers by economic sector. We restrict our sample to salaried workers, excluding self-employed workers, entrepreneurs, workers without remuneration, and public-sector and armed forces employees.

C.2 Modern and Traditional Wage Dynamics

In this section, we report wage-experience and wage-age profiles for modern and traditional workers, restricting the sample to salaried workers.

Figure C6: Wages Over the Life Cycle for Modern and Traditional Workers, Salaried Workers

(a) Wage-experience profiles **(b)** Wage-age profiles

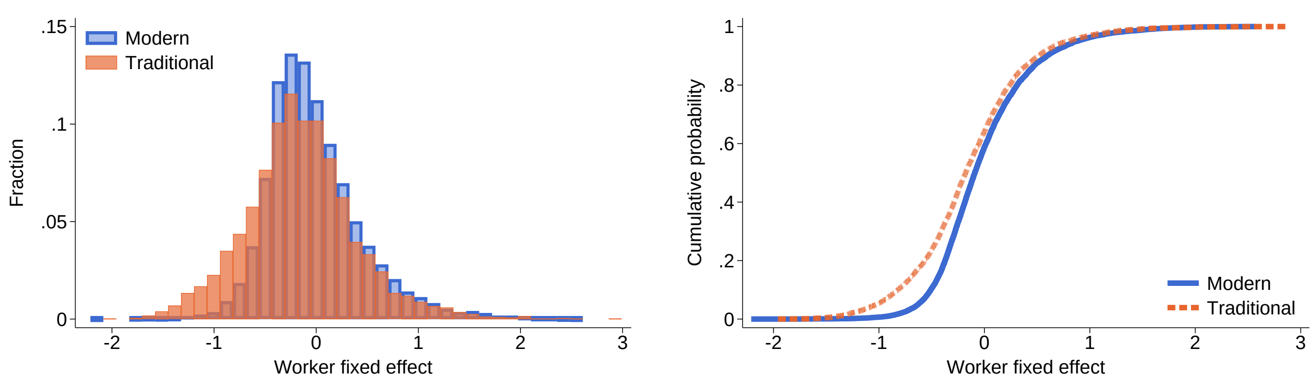


Notes: Figure C6 displays average wages by experience quintile (panel (a)) and by 10-year age bin (panel (b)) for modern and traditional workers. We restrict our sample to salaried workers, excluding self-employed workers and entrepreneurs. Vertical dashed lines denote 95 percent confidence intervals.

C.3 Robustness: Omitting Firm Size

In this section, we show that the main results of our empirical analysis are robust to omitting firm size. We re-estimate the specifications underlying Figures 2, 3, 4, and 5, excluding the firm-size bracket both from the set of controls and from the definition of workers' labor markets, which become Region × Year × Industry × Occupation cells.

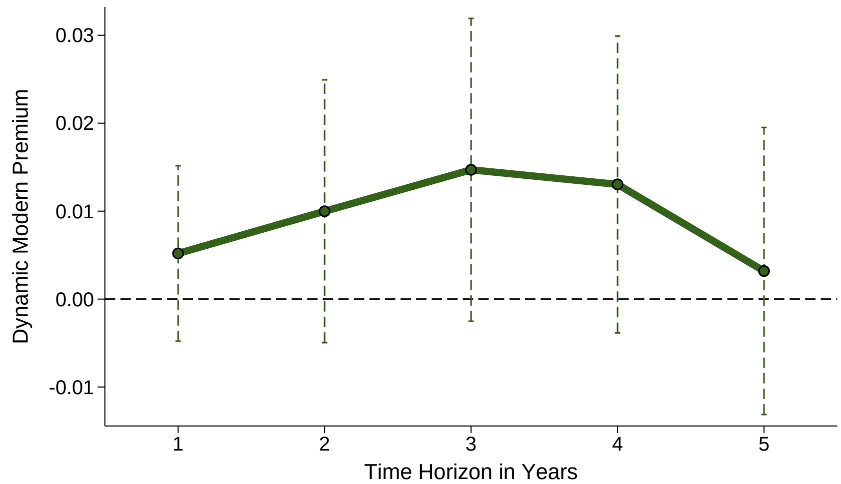
Figure C7: Worker Fixed Effects for Modern and Traditional Workers, Omitting Firm Size



(a) Histogram **(b)** Cumulative distribution

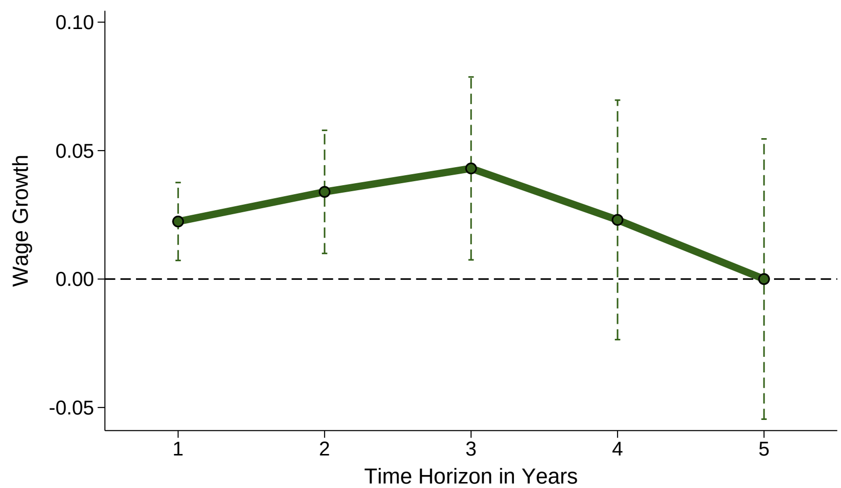
Notes: Figure C7 replicates Figure 2 omitting firm size from the set of controls. Worker fixed effects α_i are estimated from equation (14), controlling for age, education, occupation, industry, region, experience in each sector, and time fixed effects. We keep one observation per worker and assign workers to the sector in which we observe them most often.

Figure C8: Modern Sector Dynamic Wage Premium, Omitting Firm Size



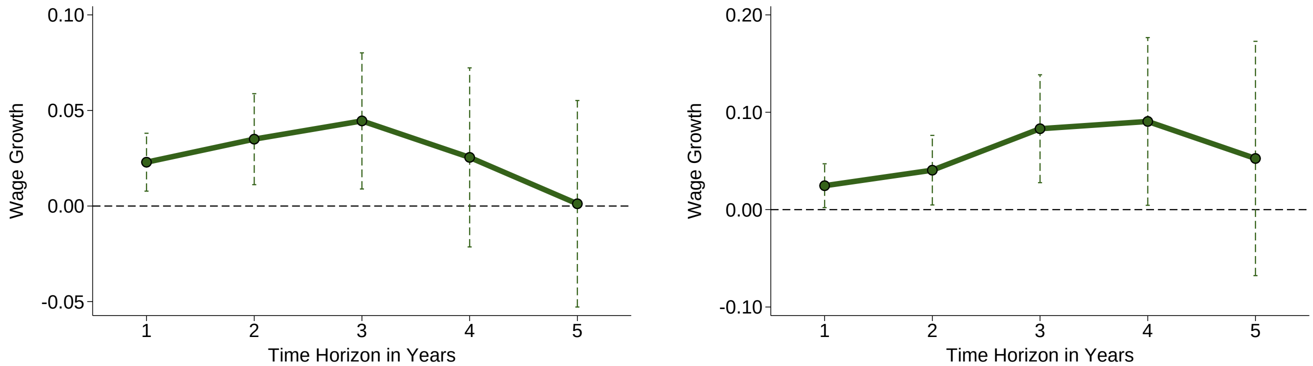
Notes: Figure C8 replicates Figure 3 omitting firm size from the set of controls. Each dot represents the estimated wage premium at horizon $t+h$: $\exp(\hat{\beta}_h)-1$, for $h=1,\ldots,5$, from equation (15) including worker fixed effects. Controls include age, education, gender, occupation, industry, region, experience, wage decile, and time and worker fixed effects. Vertical dotted lines denote 95 percent confidence intervals.

Figure C9: Peer Effects on Wage Growth, Omitting Firm Size



Notes: Figure C9 replicates Figure 4 omitting the firm-size bracket from the definition of peer groups. Peers are defined as all other workers in the same Region × Year × Industry × Occupation cell. Each dot represents the estimated effect of peer wages on workers' future wage growth at the time horizon $t+h$, based on the estimates of β_h from equation (16). Controls include age, education, gender, occupation, industry, region, experience, wage decile, and time fixed effects. Vertical dotted lines denote 95 percent confidence intervals.

Figure C10: Peer Effects on Wage Growth Robustness, Omitting Firm Size



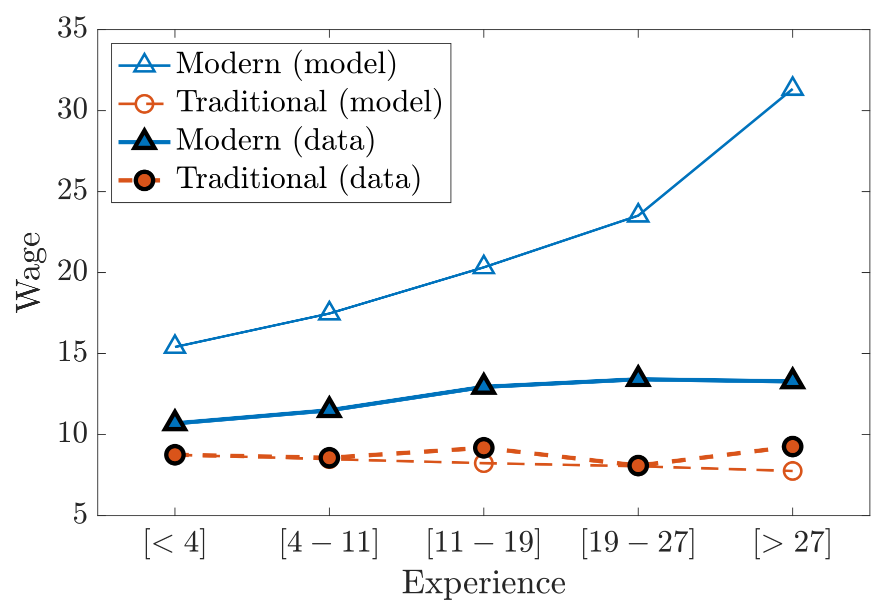
(a) Modern-sector status

(b) Wage-bill growth

Notes: Figure C10 replicates Figure 5 omitting the firm-size bracket from the definition of peer groups. In Panel C10(a) we control for an indicator for modern-sector employment, Modern_{it}. In Panel C10(b) we control for the growth of the labor market's total wage bill between $t-2$ and $t-1$, $t-1$ and t, and t and $t+1$. The remaining controls are the same as in Figure C9. Vertical dotted lines denote 95 percent confidence intervals.

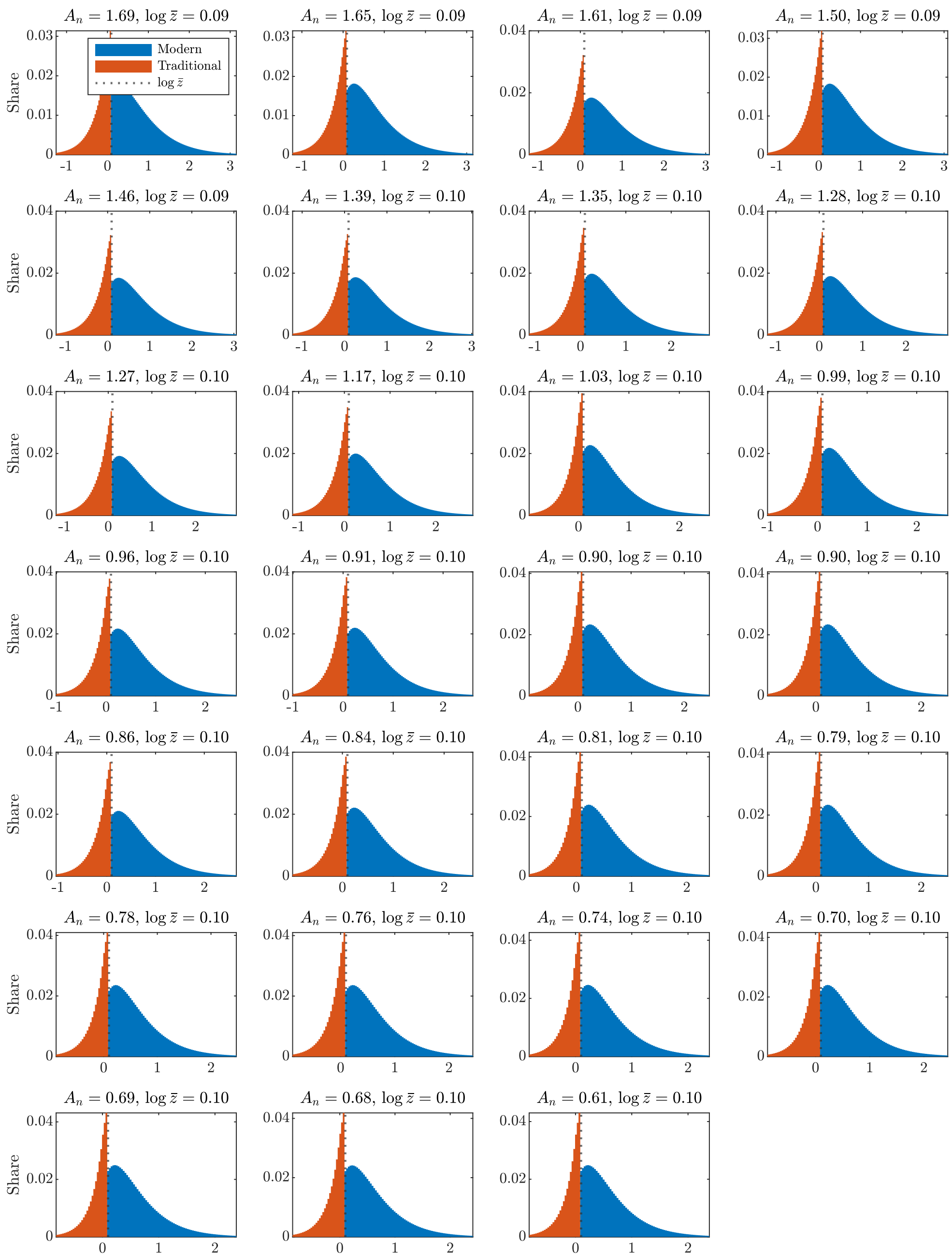
C.4 Life-Cycle Wages in Levels: Model vs Data

Figure C11: Life-Cycle Wages in Levels: Model vs Data



Notes: Figure C11 displays average wages by experience bin for modern and traditional workers in the model and in the data. Model averages are computed over the same experience-quintile bins as in Figure 1(a), weighting ages by surviving sector mass, and model wages are normalized so that the model matches the traditional average wage in the first experience bin. The data series coincide with those reported in Figure 1(a).

Figure C12: Skill Distributions Across Labor Markets

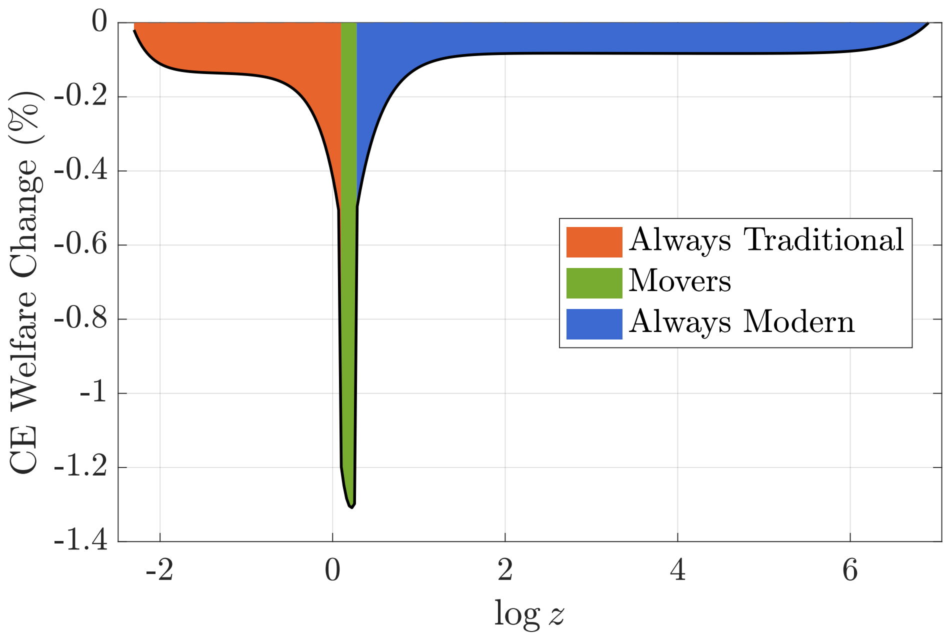


Notes: Figure C12 displays the stationary distribution of workers over log skill by sector in each of the 27 labor markets, expressed as a share of each market's workers. Markets are sorted by productivity A_n, with the most productive market first; each panel's title reports the market's productivity and modern-sector cutoff, and the dotted vertical line denotes the cutoff $\log \bar{z}$.

C.5 Learning-Channel Welfare Effects of Raising Payroll Taxes

Figure C13 replicates the left panel of Figure 9 for the counterfactual economy with a higher payroll tax, $\tau = 0.20$. The welfare effects are roughly the mirror image of those from eliminating payroll taxes, with losses concentrated among movers, who now switch from the modern to the traditional sector.

Figure C13: Welfare Losses from Peer Learning: Higher Payroll Taxes



Notes: Figure C13 decomposes learning welfare losses in the counterfactual economy with a payroll tax of $\tau = 0.20$. The x-axis denotes log skill, while the y-axis denotes consumption-equivalent welfare changes from the learning channel. The shaded areas split workers into always-traditional (orange), movers (green), and always-modern (blue) groups, where movers now switch from the modern to the traditional sector after the tax increase.

D Additional Tables

Table D1 reports estimates of the modern-traditional wage gap from equation (14), discussed in Section 3.

Table D1: Modern-Traditional Wage Gap

	Dependent variable: $\log w_{i,t}$		
	(1)	(2)	(3)
Modern	0.395***	0.093***	0.082***
	(0.015)	(0.014)	(0.018)
Corr(Modern$_{it}$, α_i)			0.095
			(0.011)
Observations	58,714	58,714	58,714
Adj R-squared	0.074	0.459	0.841
Controls		✓	✓
Worker Fixed Effects			✓

Notes: Table D1 reports estimates for β in equation (14). Column (1) displays the raw gap without any controls. Column (2) controls for worker characteristics, including experience in the modern and traditional sectors and firm size fixed effects. Column (3) adds worker fixed effects. Controls include age, education, gender, occupation, industry, region, firm size, experience in each sector, and time fixed effects. Corr(Modern$_{it}$, α_i) reports the correlation between the modern-sector indicator and the estimated worker fixed effects across worker-year observations, weighted by survey expansion factors, with standard errors clustered at the worker level. Standard errors in parentheses are clustered at the worker level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table D2 reports estimates from a pooled specification that allows the peer effect in equation (17) to differ for workers above and below the mean wage of their labor market.

Table D2: Peer Effects on Wage Growth: Workers Above and Below Average Wage

	Dependent variable: $\Delta \log w_{i,t+h}$				
	Time horizon h in years				
	1	2	3	4	5
Peer wages × Above	0.029***	0.081***	0.116***	0.144***	0.109***
	(0.010)	(0.014)	(0.019)	(0.026)	(0.036)
Peer wages × Below	0.014*	0.036***	0.040***	0.069***	0.074***
	(0.009)	(0.013)	(0.015)	(0.019)	(0.024)
Observations	39,338	29,997	21,577	15,124	9,704
Adj R-squared	0.086	0.163	0.224	0.267	0.266

Notes: Table D2 reports estimates of a pooled version of equation (16) in which log peer wages are interacted with indicators for the worker's wage being above or below the leave-one-out mean wage of her labor market, and the above-mean indicator enters directly (coefficient not reported). Controls are the same as in Figure 4, with coefficients common to the two groups. Standard errors in parentheses are clustered at the labor-market level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table D3 describes the composition of the region, industry, and occupation groups that define the labor markets used in the model quantification of Section 4.

Table D3: Labor-Market Groups

(1) Group	(2) Components
Panel A: Regions	
1	Tarapacá, Antofagasta, Atacama, Coquimbo
2	Metropolitana (Santiago)
3	Valparaíso, O'Higgins, Maule, Biobío, La Araucanía, Los Lagos, Aysén, Magallanes, Los Ríos
Panel B: Industries	
1	Mining, Utilities, Finance/Real Estate
2	Manufacturing, Construction, Transportation/Communication, Other
3	Agriculture/Fishing, Retail/Hospitality, Community Services
Panel C: Occupations	
1	Management, Science/Intellectuals, Technical
2	Clerical, Operators/Craftsmen, Operations Personnel, Other
3	Service/Sales, Agricultural/Fishery, Laborers

Notes: Table D3 lists the composition of the region, industry, and occupation groups that define the $N = 27$ labor markets used in Section 4. Each labor market is a Region group $\times$ Industry group $\times$ Occupation group cell. Industry and occupation groups are ordered by mean hourly wage in the baseline sample (group 1 highest, group 3 lowest). Industry and occupation categories follow the EPS classification.

Table D4 reports the full welfare decomposition behind the payroll-tax counterfactuals of Section 5.

Table D4: Payroll-Tax Counterfactuals: Welfare Decomposition

	Consumption-equivalent welfare change (%)							
	Total		Flow utility			Continuation value		
τ	Full	Fixed b	Full	Fixed b	Transfer	Learning	Volatility	Migration
0.20	-0.30	-5.98	1.29	-4.48	6.04	-0.36	-0.14	0.00
0.15	-0.01	-3.19	0.76	-2.44	3.28	-0.18	-0.05	0.00
0.10	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
0.05	-0.27	3.61	-1.07	2.77	-3.74	0.20	0.14	0.00
0.00	-0.78	7.66	-1.99	6.35	-7.84	0.37	0.03	0.00

Notes: Table D4 decomposes the consumption-equivalent welfare change in counterfactual economies with different payroll tax rates, τ, relative to the baseline stationary equilibrium ($\tau = 0.10$, shown with $\cdots$ where changes are zero by construction). All entries are utilitarian aggregates weighted by the baseline distribution of workers, in percentage terms. The first "Total" column reports the full welfare change; the second holds the lump-sum transfer b fixed at its baseline level. The "Flow utility" columns decompose the flow-utility component into its full change, the change at the fixed baseline transfer (the wage-wedge component), and the transfer component; because b is allocation-neutral and uniform across workers, the split is exact and multiplicative: $(1{+}\mathrm{CE}_u) = (1{+}\mathrm{CE}_u^{\text{fixed-}b})(1{+}\mathrm{CE}_{\text{transfer}})$. The "Continuation value" columns report the components of the change in workers' continuation value in (7): skill accumulation through learning (drift), skill volatility, and migration opportunities.