# Labor Market Policy, Worker Careers, and Aggregate Productivity

Niklas Engbom*
NYU, CEPR & NBER

September, 2026

**Abstract**

Policies that impede worker reallocation reduce productivity by altering the skills workers acquire over their careers. Using worker-level panel data from 15 advanced economies, I estimate a life-cycle equilibrium search model with firm entry, vacancy creation, and firms that differ in productivity and learning environment. Entry and hiring frictions delay workers' access to learning opportunities and reduce their expected returns to acquiring skills. Moving from labor-market wedges inferred for France to those inferred for the United States raises output per worker by 28.3 log points, of which higher human capital contributes 21.5. In a separately calibrated model with policy-invariant skill accumulation, the corresponding increase in output per worker is 13.3 log points. Allowing employment histories to shape skills thus more than doubles the estimated output effect of labor-market wedges.

*Niklas Engbom: ne466@nyu.edu. This paper builds on and replaces an earlier draft circulated as "Labor Market Fluidity and Human Capital Accumulation." I thank Victoria Gregory, Gregor Jarosch, Kyle Herkenhoff, Greg Kaplan, Guido Menzio, Claudio Michelacci, Ben Moll, Chris Moser, Diego Restuccia, Richard Rogerson, Todd Schoellman, Venky Venkateswaran, Gianluca Violante, and many seminar participants. I also thank Eurostat for granting me access to the ECHP and EU-SILC data sets. The results and conclusions in this paper are mine and do not represent Eurostat, the European Commission or any of the national statistical agencies whose data are used. All errors are my own.

# 1 Introduction

A large literature studies how labor market institutions and frictions shape the allocation of workers across firms and, in turn, aggregate productivity (Bentolila and Bertola, 1990; Hopenhayn and Rogerson, 1993; Ljungqvist and Sargent, 1998). This work has emphasized firms' job-creation and job-destruction decisions, as well as the static misallocation of labor across producers induced by these institutions. Yet labor markets do more than allocate a fixed stock of skills across firms. They also assign workers—especially young workers—to jobs that shape the experience and skills they accumulate over their careers. Through this dynamic margin, less reallocation may depress productivity not only by misallocating workers across firms at a point in time, but also by changing the human capital workers carry into the future.

This paper quantifies this dynamic margin in a life-cycle equilibrium search model in which firms differ in both productivity and the learning opportunities they offer. Policies and frictions that discourage firm entry and hiring reduce workers' access to jobs where they acquire skills and their subsequent opportunities to use those skills. I discipline the model with worker-level panel data from 15 advanced economies. Moving from the labor-market wedges inferred for France to those inferred for the United States raises output per worker by 28.3 log points. The corresponding increase is 13.3 log points in a separately calibrated model in which skill accumulation does not respond to these wedges. Allowing skill accumulation to respond to labor market policy thus more than doubles the estimated output effect of labor market policy that reduce worker flows.

The empirical analysis establishes three motivating facts. First, direct employer-to-employer mobility differs sharply across countries. I measure the annual *poaching rate* as the share of employed workers who joined their current employer in the preceding 12 months directly from another employer, without an intervening month of nonemployment. This rate ranges from 3.6 percent in Greece to 14.4 percent in Sweden. Second, poaching is lower in countries with higher labor taxes and higher costs of starting a business. Higher startup costs are also associated with less firm entry, consistent with an important role for young firms in hiring workers from other employers (Haltiwanger et al., 2018; Bilal et al., 2022). Third, workers experience greater life-cycle wage growth in countries with more poaching. The estimated wage profiles, net of individual and country-year fixed effects, rise by 10.9 log points between ages 20–29 and 50–59 in Greece, compared with 51.7 log points in Sweden.

The wage-growth relationship extends beyond gains at employer changes. Even workers who remain with their employer experience faster wage growth in high-poaching countries. Outside offers could generate this pattern by allowing stayers to renegotiate their pay (Postel-Vinay and Robin, 2002). Yet the relationship remains positive among workers recently hired from nonemployment, whose accumulated bargaining position is reset in standard counteroffer models.

To understand these patterns and assess the impact of labor market policy on workers' careers, I develop an equilibrium search model in which firms differ in two attributes: current productivity and their *learning environment*. The latter is motivated by growing evidence of important firm-level heterogeneity in the speed at which workers accumulate skills (Jarosch, Oberfield and Rossi-Hansberg, 2021; Herkenhoff et al., 2024; Gregory, 2026). Firms choose whether to enter and how many vacancies to create, taking into account the human capital of potential recruits and their willingness to accept an offer. Workers enter the labor market as nonemployed and search for jobs throughout their careers. Wages respond to outside offers, and changing employers entails a loss of specific skills. Young workers attach a high value to a good learning environment because they learn relatively quickly and have a long remaining career over which to use their skills. As they age, they place greater weight on current productivity. A worker's career therefore involves both finding a job in which to learn and subsequently finding jobs in which those skills are productive.

Labor taxes, entry costs, incumbent recruitment costs, and aggregate matching efficiency affect the equilibrium worker allocation through a common cost composite. An increase in this cost composite reduces poaching. Less poaching delays young workers' arrival at good learning environments. Poorer prospects for reaching productive employers later in life also reduce the value of acquiring skills today. Consequently, nonemployed workers are less willing to sacrifice current output for a better learning environment in a low-poaching economy.

I estimate the model in two steps. First, I choose common parameters to match worker-flow and wage profiles, wage changes around job transitions and re-employment, the firm-size distribution, and employment and wage growth at training firms in a hypothetical economy with average poaching. A central parameter is the dispersion in firms' learning environments, which governs how much workers' employment histories can affect their skills. I discipline this parameter using within-worker variation in wage growth across training and nontraining employers, where the latter is based on workers' self-reports of whether their employer provides any type of on-the-job training. A given worker experiences faster wage growth at training employers, particularly early in the career, and these wage gains remain with the worker even after she has left the training employer. Higher wage growth at training employers relative to nontraining employers implies greater dispersion in learning environments and greater scope for dynamic misallocation.

Second, I recover four country-specific wedges: the labor tax, the cost of firm entry, the cost of incumbent recruitment, and the relative efficiency of employed search. I take labor taxes from the OECD and infer the remaining wedges jointly from poaching, transitions from nonemployment to employment, employment at large firms, and reported search by employed workers. The inferred wedges to firm entry and incumbent expansion vary substantially across countries, yet they are broadly comparable in magnitude to that in the monetary cost of starting a business measured by the World Bank, using the methodology developed by Djankov et al. (2002).

I use the estimated model as a laboratory to isolate and quantify the impact of these four policy wedges on other cross-country outcomes. Combined, these four wedges generate substantial differences in career wage growth, accounting for about 68 percent of the variation across these 15 advanced economies. Individually, the entry and incumbent recruitment wedges are the most important. The wedges also produce economically significant differences in output per worker. For instance, moving the policy wedges from France, a typical low-fluidity country, to the United States, a typical high-fluidity country, raises output per worker by 28.3 log points. About 76 percent of the aggregate response is accounted for by a change in the stock of skills, as young workers move more into high-learning jobs in high-poaching countries, a prediction for which I find empirical support. At the same time, the data display substantial variation in output per worker conditional on poaching, which these four policy wedges cannot explain. Overall, these four policy wedges explain only about 7 percent of the cross-country variation in output per worker.

To assess how much endogenous skill accumulation changes the inferred output effects, I compare the full model with an alternative in which skills evolve exogenously and independently of workers' employment histories. I re-calibrate the alternative using many of the same moments as in the baseline estimation. Subsequently, I move the estimated policy wedges from those in France to those in the U.S. The resulting output gain is 47 percent of that in the baseline model, as skills by construction do not respond. A model with policy-invariant skill accumulation therefore substantially understates the output effects of labor-market policy.

Most of the aggregate effect of the policy wedges arises because workers reach learning opportunities sooner. Holding workers' acceptance and mobility rules and firms' vacancy choices fixed while moving contact rates from their French to their U.S. values generates 19.5 of the 28.3 log-point increase in output per worker and 14.0 of the 21.5 log-point increase in human capital. Behavioral responses account for the remaining 31 percent of the output effect and 35 percent of the human-capital effect: young workers place greater value on good learning environments in a high-poaching economy, and firms adjust vacancy creation in response. Consistent with this mechanism, I find that, among workers hired from nonemployment, the difference between younger and older workers in the share joining training firms rises with aggregate poaching.

The size of these dynamic effects is disciplined by the observed wage-growth advantage of training employers. I demonstrate this by varying the training wage-growth targets and re-estimating all internally calibrated parameters, holding the other empirical targets fixed. Raising the training wage-growth advantage by 50 percent increases the estimated dispersion of learning environments, $\sigma_x$, and the human-capital gain from the France-to-U.S. policy change from 21.5 to 32.8 log points.

One concern is that the model treats each firm's learning environment as a fixed technological attribute. If firms choose how much training to provide and workers cannot compensate firms for training, greater poaching can weaken their incentives to invest because some returns accrue after workers leave the firm that financed their training (Acemoglu and Pischke, 1999). Two features of the data do not support lower training provision in high-poaching countries. First, the wage-growth advantage of training over nontraining firms tends to be greater where poaching is higher, consistent with a larger gap in skill accumulation. Second, reported time spent in training among workers at training firms also increases with aggregate poaching.

Finally, I examine differences by education and gender. Allowing the mean learning environment to differ across groups, I infer that college-educated workers and men interact with better learning environments (or isomorphically that college educated workers and men have higher learning ability). Feeding in the same country variation in wedges reproduces the stronger relationship between poaching and wage growth for college graduates and for men in the data.

This paper connects a literature on labor-market institutions, firm dynamics, and aggregate productivity to research on firms as sources of human capital. Models of job creation and labor reallocation have established how policies affect employment and the allocation of workers across producers (Hopenhayn and Rogerson, 1993; Pries and Rogerson, 2005). Ljungqvist and Sargent (1998) connect labor-market institutions to skill accumulation during employment and skill loss during nonemployment. More recent research shows how coworkers and employers shape the skills workers carry to subsequent jobs (Jarosch, Oberfield and Rossi-Hansberg, 2021; Arellano-Bover, 2024; Herkenhoff et al., 2024; Arellano-Bover and Saltiel, 2026). Gregory (2026) quantifies the implications of heterogeneous learning environments for life-cycle earnings inequality, while Acabbi, Alati and Mazzone (2026) study how recessions disrupt sorting and human capital accumulation. I quantify the long-run effects of cross-country labor-market wedges when firm entry, incumbent recruitment, and workers' career choices jointly determine access to learning opportunities.

The cross-country analysis also builds on evidence of international differences in worker flows and life-cycle wage growth (Donovan, Lu and Schoellman, 2023; Lagakos et al., 2018). I use these patterns, together with microeconomic evidence on training and wages, to evaluate a common equilibrium mechanism.

The analysis also complements theories of training investment in frictional labor markets. Acemoglu (1997) shows that training is inefficiently low because future employers capture part of its return, even when workers can compensate their current employer. In his baseline model, a higher probability of separation following an adverse match shock reduces training, holding other parameters fixed. Acemoglu and Pischke (1999) show how wage compression gives firms an incentive to finance general training; when workers cannot contribute to its cost, higher exogenous turnover reduces firms' investment, holding the wage structure fixed. These separation comparative statics do not directly establish how training responds to more frequent outside offers. Here, firms' learning environments are fixed attributes, and policy changes which jobs are created and how quickly workers reach them through firms' entry and recruitment decisions and workers' career choices.

The paper proceeds as follows. Section 2 presents the cross-country evidence. Section 3 develops the model. Section 4 describes the estimation, and Section 5 presents the quantitative results. Section 6 concludes.

## 2 Motivating Cross-Country Evidence

I document three facts that motivate the analysis: employer-to-employer mobility differs sharply across countries; mobility is lower where labor taxes and business startup costs are higher; and workers experience greater career wage growth in countries with more mobility. I then examine whether the wage-growth relationship reflects only wage gains at job changes or bargaining leverage accumulated through outside offers.

### 2.1 Data and Measurement

The analysis combines the European Community Household Panel (ECHP, 1994–2001), the European Union Statistics on Income and Living Conditions (EU-SILC, 2003–2024), and the Panel Study of Income Dynamics (PSID, 1993–2023). These surveys provide worker-level panel data for 15 advanced economies. Coverage varies by country and outcome; the poaching measure is constructed over 1994–2020. Appendix A.1 describes the country-specific gaps, harmonization, and construction of the country averages.

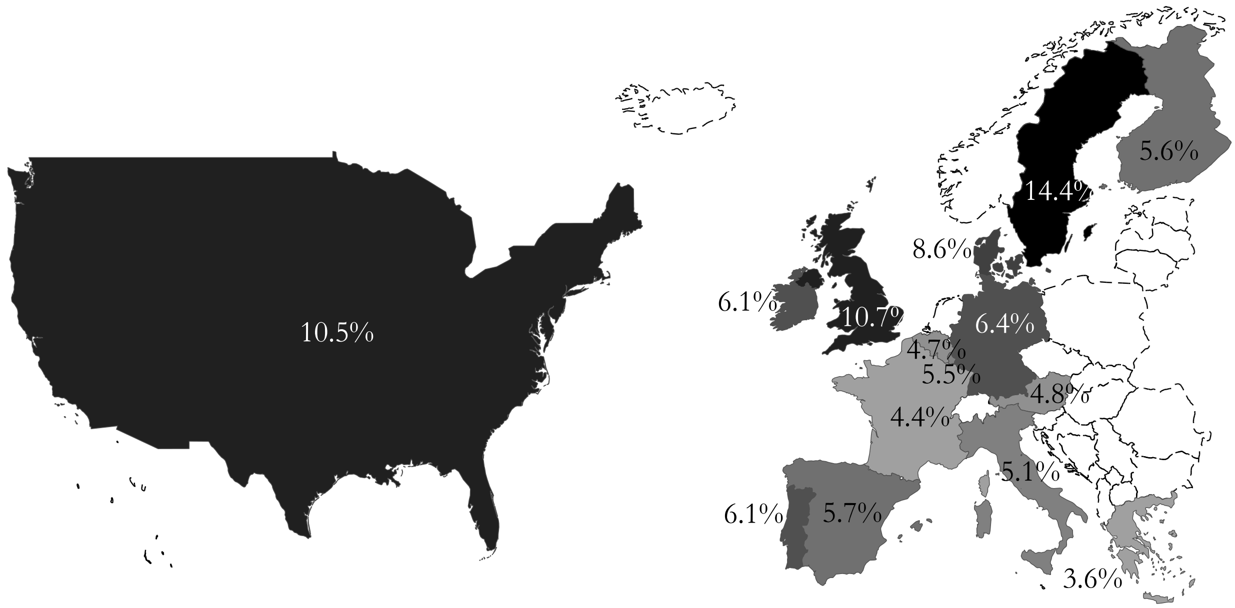
The sample covers men and women aged 20–59, with a minimum age of 22 for workers with upper-secondary education and 24 for those with tertiary education. Employment includes private- and public-sector wage work and self-employment. I measure hourly wages as annual labor income divided by annual hours and express them in 2007 PPP-adjusted U.S. dollars. The annual *poaching rate* is the share of employed workers who joined their current employer during the preceding 12 months directly from another employer, without a recorded month of nonemployment during that interval. This measure records recent hires from other employers rather than all job transitions within a year. Country averages give equal weight to ten-year age groups; the treatment of missing years is described in the appendix. Appendix A.2 compares it with the quarterly JJ transition rate of Donovan, Lu and Schoellman (2023).

### 2.2 Three Facts on Poaching across Countries

**Fact 1. Poaching differs sharply across countries.** Figure 1 shows that annual poaching rates vary nearly fourfold, from 3.6 percent in Greece to 14.4 percent in Sweden. France also has a low rate, at 4.4 percent, while the U.K. and the U.S. are among the high-poaching economies. High-poaching countries also have higher nonemployment-to-employment (NE) rates and employment rates. Employment-to-nonemployment (EN) rates, by contrast, show no systematic relationship with poaching (Appendix A.3). Thus, differences in poaching are associated with differences in workers' access to jobs rather than simply with more frequent job loss.

**Fact 2. Poaching is lower where labor taxes and startup costs are higher.** Figure 2 relates poaching to two measures of the costs facing employers and workers. Countries with higher total labor

Figure 1: The Poaching Rate across 15 Advanced Economies



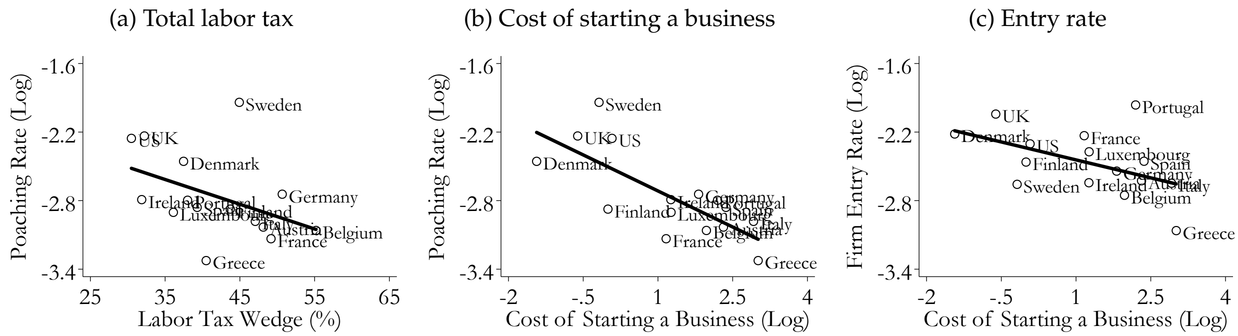

*Notes:* The poaching rate is the share of currently employed workers who began working for their current employer in the past 12 months directly from another employer, without an intervening month of nonemployment. Darker shading denotes a higher rate; countries shown in white are not in the sample. Alaska and Hawaii are omitted. *Source:* ECHP 1994–2001; EU-SILC 2003–2024; PSID 1993–2023. Poaching is measured over 1994–2020, with country-specific gaps.

taxes have lower poaching rates (panel (a)), as do countries where registering a business is more expensive (panel (b)). Higher startup costs are also associated with lower firm entry (panel (c)). Together with evidence that young firms are an important source of JJ hiring (Haltiwanger et al., 2018; Bilal et al., 2022), the latter pattern motivates the model's entry margin. Wage floors, dismissal costs, and unemployment-insurance generosity have weaker associations with poaching. Appendix A.4 defines these measures and considers alternative measures of entry costs and labor-market institutions. These comparisons are descriptive: correlated institutions and other country differences preclude identifying separate causal policy effects.

**Fact 3. Higher poaching is associated with greater career wage growth.** I estimate country-specific life-cycle wage profiles from regressions of log hourly wages on individual, country-year, and country-age fixed effects. To resolve the age–period–cohort collinearity, I restrict the age profiles to be flat from age 45 onward. I summarize each profile by the difference in estimated log wages between the 20–29 and 50–59 age groups. Appendix A.5.1 gives the specification and the country profiles. The estimates describe wage growth among workers observed with wages; they do not follow an entire birth cohort from labor-market entry to retirement.

Panel (a) of Figure 3 shows that wage profiles are steeper in countries with more poaching. The estimated increase between workers' twenties and fifties is about 11 log points in Greece and 52 log points in Sweden. Several additional patterns support the relevance of career dynamics. Wages in workers' twenties are also higher in high-poaching countries, so the steeper profiles do

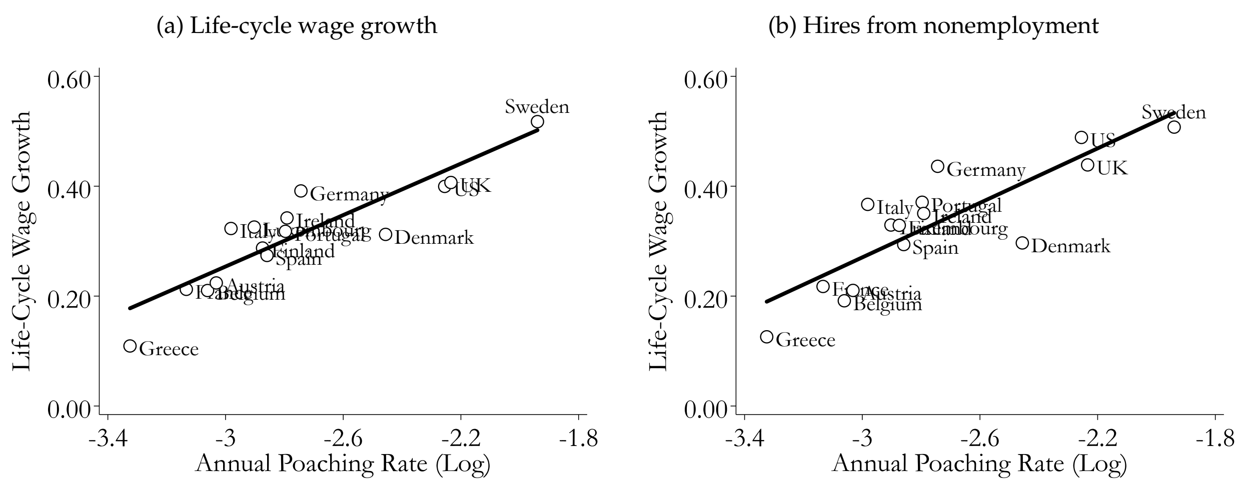
Figure 2: Poaching, Labor Taxes, Startup Costs, and Firm Entry




*Notes:* The total labor tax is personal income tax plus employee and employer social-security contributions minus cash transfers, expressed as a share of total labor cost. Startup cost is the cost of registering a limited-liability company relative to GNI per capita. The entry rate is the number of newly founded firms in a year divided by the stock of firms. Poaching and firm entry are plotted in logs, as are startup costs; the tax wedge is expressed as a percentage of labor cost. Each marker is a country; lines are linear fits. Appendix A.4 provides definitions and sample periods. *Source:* OECD; World Bank Doing Business; Eurostat Business Demography; U.S. Census Business Dynamics Statistics; ECHP; EU-SILC; PSID.

not simply reflect convergence from lower entry wages. The association with annual wage growth is strongest early in the career and persists when allowing for differences by education, gender, and occupation. Appendix A.5 presents these results.

Figure 3: Poaching and Wage Growth




*Notes:* Panel (a) plots the increase in log hourly wages between the 20–29 and 50–59 age groups based on the profiles described in Appendix A.5.1. Panel (b) reports the corresponding age-bin wage difference among recent hires from nonemployment, net of individual and country-year effects estimated in the full sample. Appendix A.5.1 describes this construction. *Source:* ECHP 1994–2001; EU-SILC 2003–2024; PSID 1993–2023. Poaching is measured over 1994–2020, with country-specific gaps.

One explanation is that workers in high-poaching countries change employers more often and receive a wage increase at each transition. Table 1 examines this possibility by relating annual wage growth to log poaching interacted with an indicator for ages 20–39. The regressions absorb country-year effects and allow age, education, gender, and occupation effects to vary by year. Adding an indicator for the worker's own JJ transition changes the coefficient on young age interacted with log poaching only from 0.018 to 0.017 (columns 1–2). Column 3 allows the relationship to differ by stayer status. Among workers with neither a JJ transition nor a hire from nonemployment in the preceding year, the coefficient is 0.015 and statistically significant. The estimate for the remaining sample is smaller and imprecise. Direct wage gains at employer changes therefore do not fully account for the wage-growth relationship. Appendix A.5.4 gives the specifications and the definition of stayer status.

Table 1: Wage Growth Within and Between Jobs and the Poaching Rate

| | Baseline | With JJ control | By stayer status |
|---|---|---|---|
| Age 20–39 × log poaching | 0.018*** | 0.017*** | |
| | (0.003) | (0.003) | |
| × stayer | | | 0.015*** |
| | | | (0.004) |
| × nonstayer | | | 0.007 |
| | | | (0.005) |
| JJ transition in last 12 months | | 0.046*** | |
| | | (0.009) | |
| Observations | 1,226,517 | 1,226,517 | 1,226,517 |
| Years | 28 | 28 | 28 |
| Countries | 15 | 15 | 15 |

*Notes:* OLS estimates of annual log-wage growth. Poaching enters in logs and is interacted with an indicator for ages 20–39. Column (2) adds a JJ-transition indicator. Column (3) estimates separate young-age interactions for stayers and nonstayers and allows country-year effects to differ by stayer status. Stayers have neither a JJ transition nor a hire from nonemployment in the preceding year; nonstayers include workers with transitions or missing transition information. All columns include age-year, education-year, gender-year, and occupation-year fixed effects. Appendix A.5.4 gives the specifications. Standard errors clustered by country in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Source:* ECHP 1994–2001; EU-SILC 2003–2024; PSID 1993–2023. Poaching is measured over 1994–2020, with country-specific gaps.

Faster wage growth among stayers need not reflect faster skill accumulation: outside offers can also allow workers to renegotiate their pay. In canonical counteroffer models, a spell of nonemployment resets the bargaining leverage accumulated through previous offers (Cahuc, Postel-Vinay and Robin, 2006). Panel (b) therefore examines the age profile of wages among workers recently hired from nonemployment, after removing the individual and country-year effects estimated in the full sample. The relationship between poaching and career wage growth remains positive for these workers. This pattern is inconsistent with inherited bargaining leverage being the sole explanation. It does not eliminate differences in selection into nonemployment or in the jobs and wages available upon re-entry.

Taken together, the evidence connects worker mobility to firm entry, access to employment, and wage growth throughout the career. The association with wage growth extends beyond contemporaneous switching gains and the bargaining leverage carried from previous jobs. These facts motivate a model in which employment histories shape human capital, alongside match productivity and bargaining. The next section develops that model to quantify the effects of labor-market wedges on careers and output.

# 3 Model

I develop a life-cycle search model in which firms differ in current productivity and the learning opportunities they offer. Workers choose jobs with both attributes in mind. Firm entry and recruitment determine how quickly workers reach those jobs, while workers' accumulated skills and willingness to move determine the return to hiring. The model links labor-market policy to output through both the allocation of a given stock of skills and the stock itself.

## 3.1 Environment

Time is continuous and infinite, and the economy is in its long-run steady state.

**Workers.** Workers differ in age $a \in [0, A]$ and log human capital $h$. A flow $1/A$ enters nonemployment with $a = h = 0$, and workers retire at age $A$. The stationary worker population therefore has unit mass and a uniform age distribution.

Workers are risk neutral and discount at rate $\rho > 0$. As nonemployed, a worker of age $a$ with log human capital $h$ enjoys flow value of leisure $e^{b(a)+h}$.

A worker of age $a$ has learning ability $\psi(a)$. For the continuous-age analysis, $\psi$ is continuously differentiable on $[0, A]$ and $\psi(a) > 0$ for $a < A$. Section 4 uses a discrete-age approximation that allows learning to cease in the oldest group. Whenever an employment relationship ends—because the worker separates to nonemployment or moves to another employer—her log human capital falls by $\epsilon \geq 0$,

$$h' = h - \epsilon.$$

This is a parsimonious way to capture the notion that some share of a worker's skills are specific to her particular employer, and hence lost when the relationship ends.

Nonemployed workers receive job offers at rate $p$. Employed workers search with relative efficiency $\phi > 0$, so they receive offers at rate $\phi p$. Offers are drawn from a distribution with density $f$ and cumulative distribution function (CDF) $F$. While workers treat both $p$ and $f$ as exogenous, I outline below how they are determined in equilibrium.

Workers separate to nonemployment for worker-idiosyncratic reasons at rate $\delta(a)$, where $\delta(a)$ is a continuous function of age satisfying $0 \leq \delta(a) \leq \bar{\delta}$. They also become nonemployed due to exogenous firm exit at rate $\iota$. Finally, they are free to quit to nonemployment at any point.

**Firms.** A mass $M$ of firms hire workers to produce a homogeneous final consumption good. Profits are distributed lump-sum to households and do not affect job choices under linear preferences. The values below exclude these common transfers.

A firm is characterized by two permanent attributes—its productivity $y$ and its *learning environment* $x$. A worker with log human capital $h$ employed at a type-$(y, x)$ firm produces

$$e^{y+h},$$

and accumulates human capital according to

$$\dot{h}(a) \quad = \quad \psi(a)x.$$

To recruit workers, firms advertise $v$ vacancies subject to an isoelastic recruitment cost

$$C_v\big(v(y,x)\big) \quad = \quad \frac{c_v}{1+\eta}v(y,x)^{1+\eta}, \qquad c_v > 0, \quad \eta > 0.$$

Each vacancy contacts a potential hire at rate $q$.

Potential entrants pay $c_e > 0$ in return for a random draw of $(y, x)$ from an entry distribution with continuous, bounded density $\gamma$ over compact support $\mathcal{E} \subseteq [\underline{y}, \overline{y}] \times [\underline{x}, \overline{x}]$. Let $\Gamma$ denote the associated CDF. I maintain a positive effective discount margin

$$\rho + \delta(a) + \iota - \psi(a)x \quad > \quad 0 \qquad \text{for every } (y, x, a) \in \mathcal{E} \times [0, A].$$

Firms begin with no workers and exit exogenously at rate $\iota > 0$. Because exit is independent of type, $\Gamma$ is also the cross-sectional distribution of incumbent firms.[1] Meanwhile, the offer distribution $f$ is the vacancy-weighted density of firms

$$f(y,x) \quad = \quad \frac{M}{V}v(y,x)\gamma(y,x),$$

where $V$ is the aggregate number of advertised jobs

$$V \quad = \quad M\int_{\mathcal{E}} v(y,x)d\Gamma(y,x).$$

**The labor market.** Let $U$ be the nonemployment rate and $S$ aggregate search effort

$$S \quad = \quad U + \phi(1-U).$$

[1]In the equilibrium considered below, some firms are unable to recruit any workers. In the theoretical analysis, I assume that such firms remain until they exit at rate $\iota$. Meanwhile, the quantitative analysis computes firm size outcomes only among firms with positive employment, consistent with the data.

Search is random. The number of contacts is given by the Cobb-Douglas matching function

$$m \;=\; \chi V^{\alpha} S^{1-\alpha}, \qquad \chi > 0, \quad \alpha \in (0,1).$$

The contact rates per unit of worker search and per vacancy are

$$p \;=\; \frac{m}{S}, \qquad q \;=\; \frac{m}{V} = \chi^{1/\alpha} p^{(\alpha-1)/\alpha}.$$

Workers and firms split the surplus of a match following the counteroffer protocol of Cahuc, Postel-Vinay and Robin (2006), with worker bargaining power $\beta \in (0,1)$. Value is delivered to the worker via a piece rate $w \in [0,1]$. Let $\mathcal{N}(a,h)$ denote the value of nonemployment, and let $\mathcal{V}^w(y,x,a,h,w)$ and $\mathcal{V}^f(y,x,a,h,w)$ denote, respectively, the value of a worker employed at a type-$(y,x)$ firm at piece rate $w$ and the value of that employment relationship to the incumbent firm. Define the sum of the worker's and firm's value

$$\mathcal{J}(y,x,a,h) \;\equiv\; \mathcal{V}^w(y,x,a,h,w) + \mathcal{V}^f(y,x,a,h,w).$$

Transferable utility makes this joint value independent of the piece rate, as established in Appendix B.1.

Consider a worker who receives an outside offer. Let $\mathcal{J}_i$ be the incumbent joint value and $\mathcal{J}_o$ the destination joint value evaluated after the skill loss. If $\mathcal{J}_o > \mathcal{J}_i$, the worker moves and receives $\mathcal{J}_i + \beta(\mathcal{J}_o - \mathcal{J}_i)$; the destination firm receives the remaining surplus. If the incumbent has the higher value, a credible offer instead supports the worker promise $\mathcal{J}_o + \beta(\mathcal{J}_i - \mathcal{J}_o)$. The worker stays and renegotiates only when that promise improves on her current value. Offers that cannot do so are discarded. Appendix B.1 states the exact renegotiation region and individual recursions.

Contracts are subject to limited commitment, so they must satisfy $\mathcal{V}^w(y,x,a,h,w) \geq \mathcal{N}(a,h-\epsilon)$: quitting ends the relationship and destroys specific skills.[2] If deterministic aging or learning brings a continuing contract to the worker's participation boundary, the piece rate is reset just enough to keep her indifferent, as stated in the wage-policy system below.

### 3.2 Worker values and job choices

Human capital raises current output and every subsequent payoff in the same proportion. It therefore scales values without changing job rankings. This property makes it possible to solve for job choices without keeping track of each worker's human capital.

**Lemma 1.** Fix a finite contact rate $p$, a probability offer distribution $F$, and a bounded continuous leisure schedule $b$. Defining $N(a) \equiv \mathcal{N}(a,0)$ and $J(y,x,a) \equiv \mathcal{J}(y,x,a,0)$, the joint value and the

[2] Note that the firm's counterpart, $\mathcal{V}^f \geq 0$, holds automatically since the piece rate $w \leq 1$.

value of nonemployment satisfy

$$\mathcal{N}(a,h) \quad = \quad e^h N(a), \qquad \mathcal{J}(y,x,a,h) = e^h J(y,x,a). \tag{1}$$

Writing $z^+ \equiv \max\{z, 0\}$, the scaled value of nonemployment satisfies

$$\rho N(a) \quad = \quad e^{b(a)} + N'(a) + p\beta \int_{\mathcal{A}(a)} \left(J(y,x,a) - N(a)\right) dF(y,x), \tag{2}$$

with $N(A) = 0$. In the continuation region $\mathcal{C}(a)$, the scaled joint value satisfies

$$\begin{aligned}\big(\rho + \delta(a) + \iota\big) J(y,x,a) \quad &= \quad e^y + \psi(a) x J(y,x,a) + J_a(y,x,a) + \big(\delta(a) + \iota\big) e^{-\epsilon} N(a) \\ &+ \quad \phi p \beta \int_{\mathcal{M}(y,x,a)} \big(e^{-\epsilon} J(\tilde{y}, \tilde{x}, a) - J(y,x,a)\big)\, dF(\tilde{y}, \tilde{x}),\end{aligned} \tag{3}$$

with $J(y,x,A) = 0$. The four policy regions are

$$\begin{aligned}\mathcal{C}(a) \quad &\equiv \quad \left\{(y,x) : J(y,x,a) > e^{-\epsilon} N(a)\right\}, \\ \mathcal{M}(y,x,a) \quad &\equiv \quad \left\{(\tilde{y},\tilde{x}) : e^{-\epsilon} J(\tilde{y},\tilde{x},a) > J(y,x,a)\right\}, \\ \mathcal{A}(a) \quad &\equiv \quad \left\{(y,x) : J(y,x,a) > N(a)\right\}, \\ \mathcal{P}(y,x,a) \quad &\equiv \quad \left\{(\tilde{y},\tilde{x}) : e^{-\epsilon} J(y,x,a) > J(\tilde{y},\tilde{x},a)\right\}.\end{aligned} \tag{4}$$

*Proof.* See Appendix B.2. □

The continuation equation is supplemented by the option to quit: $J \geq e^{-\epsilon} N$, with equality when the match ends. Appendix B.3 states the resulting variational inequality. The distinction between hiring and retention is consequential. A nonemployed worker accepts a job only if $J > N$, whereas an incumbent match continues if $J > e^{-\epsilon} N$. The loss of specific skills creates a band of jobs that workers retain but would decline from nonemployment. The actual transition rate into employment is $pF(\mathcal{A}(a))$; $p$ itself is the offer-arrival rate.

**Wage policies.** Let $V^w$ and $V^f$ denote the corresponding individual values at $h = 0$

$$\begin{aligned}\mathcal{V}^w(y,x,a,h,w) \quad &= \quad e^h V^w(y,x,a,w), \\ \mathcal{V}^f(y,x,a,h,w) \quad &= \quad e^h V^f(y,x,a,w).\end{aligned}$$

Once $(N, J)$ and the allocation are known, piece rates implement the bargaining promises. I maintain a contract-spanning condition: each required promise lies in the range of feasible worker values generated by $w \in [0,1]$. On the interior of the continuing contract region, worker value is continuous and strictly increasing in $w$, so the implementing piece rate is unique. At a participation reset I use the smallest feasible rate. Appendix B.1 gives the individual recursions and the implementation argument. In scaled units, a hire from nonemployment, a move from $(y, x)$ to $(\widetilde{y}, \widetilde{x})$, and a renegotiation at the incumbent firm satisfy, respectively,

$$
\begin{aligned}
V^w\left(y, x, a, w^N\right) &= N(a) + \beta\left(J(y, x, a) - N(a)\right), \\
e^{-\epsilon} V^w\left(\widetilde{y}, \widetilde{x}, a, w^M\right) &= J(y, x, a) + \beta\left(e^{-\epsilon} J(\widetilde{y}, \widetilde{x}, a) - J(y, x, a)\right), \\
V^w\left(y, x, a, w^R\right) &= e^{-\epsilon} J(\widetilde{y}, \widetilde{x}, a) + \beta\left(J(y, x, a) - e^{-\epsilon} J(\widetilde{y}, \widetilde{x}, a)\right), \\
V^w(y, x, a, \underline{w}(y, x, a)) &= e^{-\epsilon} N(a).
\end{aligned}
$$

Whenever a continuing contract reaches this boundary, the piece rate is reset upward to $\underline{w}(y, x, a)$.

**The reference job.** To parameterize the value of nonemployment, I use a reference job $(y_r, x_r)$ on the acceptance frontier. This normalization is also used in the quantitative model. It locates the job at which a nonemployed worker is indifferent to accepting an offer; when $\epsilon > 0$, a worker already in that job strictly prefers retention to quitting.

**Assumption 1.** The leisure schedule is chosen so that

$$
\mathcal{N}(a, h) = \mathcal{J}(y_r, x_r, a, h), \qquad N(a) = J(y_r, x_r, a).
$$

The reference job satisfies its continuation equation, including the continuation-side equality when $\epsilon = 0$. The implied leisure flow is continuous and strictly positive on the candidate domain used in the equilibrium argument.

Let $\lambda(a) \equiv \delta(a) + \iota$. Substituting the reference condition into the match equation gives

$$
\begin{aligned}
\left[\rho + \lambda(a)(1 - e^{-\epsilon}) - \psi(a) x_r\right] N(a) &= e^{y_r} + N'(a) \\
&+ \phi p \beta \int_{\mathcal{E}} \left(e^{-\epsilon} J(y, x, a) - N(a)\right)^+ dF(y, x).
\end{aligned} \tag{5}
$$

Combining this equation with (2) recovers leisure:

$$
\begin{aligned}
e^{b(a)} &= e^{y_r} + \left[\psi(a) x_r - \lambda(a)(1 - e^{-\epsilon})\right] N(a) \\
&+ \phi p \beta \int_{\mathcal{E}} \left(e^{-\epsilon} J - N\right)^+ dF - p\beta \int_{\mathcal{E}} \left(J - N\right)^+ dF.
\end{aligned} \tag{6}
$$

The two option values differ because leaving the reference employer destroys specific skills, whereas accepting a job from nonemployment does not. The leisure schedule is recalibrated when equilibrium conditions change. The policy results below are conditional on this normalization; the calibration section also discusses the exercise that holds leisure fixed.

### 3.3 Career dynamics

Define the iso-value locus

$$\mathcal{I}(a;\nu) \quad \equiv \quad \{(y,x) : J(y,x,a) = \nu\}\,.$$

This definition nests the separation frontier when $\nu = e^{-\epsilon}N$, the acceptance frontier when $\nu = N$, and, for an incumbent $(y_i, x_i)$, the poaching frontier when $\nu = e^{\epsilon}J(y_i, x_i, a)$.

**Lemma 2.** Fix the contact rate and an offer distribution with a density. At each age $a < A$, the joint value is continuous and weakly increasing in productivity and learning quality. Raising either attribute strictly increases value whenever the lower-type match is in the continuation region. Thus an iso-value locus through an interior continuation point is locally the graph of a continuous, strictly decreasing function of learning quality.

For types with $y > y_r$ and $x \geq x_r$, the joint value is continuously differentiable in the firm attributes and

$$J_y(y,x,a) \quad > \quad 0, \qquad J_x(y,x,a) > 0.$$

At support boundaries, derivatives refer to the value of a hypothetical current type in a neighborhood, with $F$ held fixed. At such types, the local slope of an iso-value locus is $-J_x/J_y$. If

$$-\frac{\psi'(a)}{\psi(a)} \quad \geq \quad \rho + \delta(a) + \iota - \psi(a)x_r, \qquad a < A, \tag{7}$$

then the absolute slope at a fixed firm type strictly decreases with age:

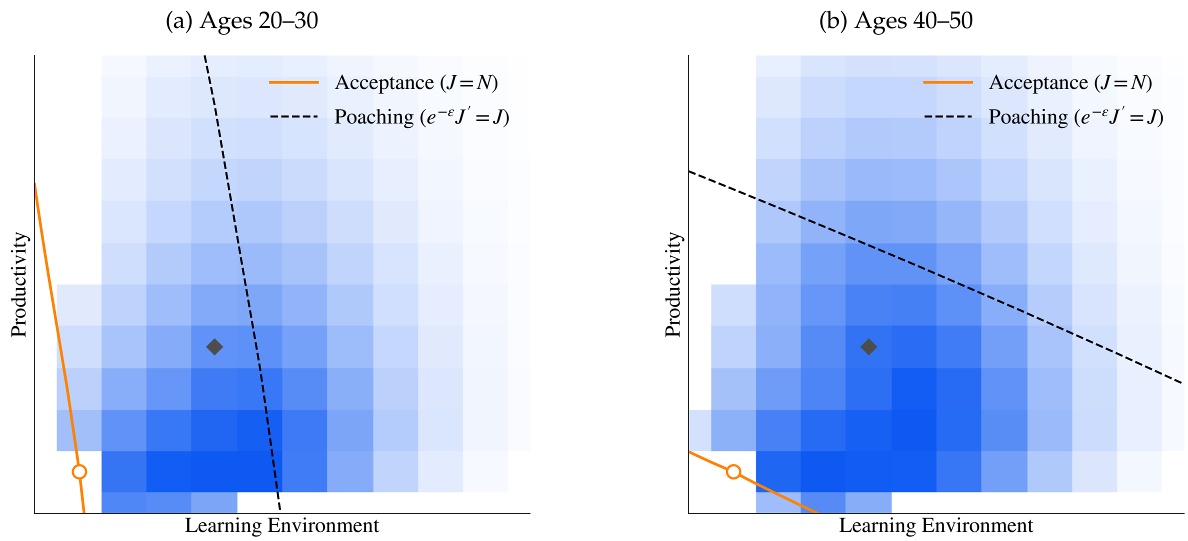
$$\frac{J_x(y,x,a_1)}{J_y(y,x,a_1)} \quad > \quad \frac{J_x(y,x,a_2)}{J_y(y,x,a_2)}, \qquad 0 \leq a_1 < a_2 < A.$$

*Proof.* See Appendix B.8. □

The derivative result applies to jobs that are more productive than the reference job and provide at least as much learning. Such jobs remain strictly above the quitting boundary until retirement. The sufficient condition bounds the decline in learning ability relative to discounting and job destruction. Under that condition, an older worker places less relative value on learning at any fixed job in this region. The numerical career frontiers below describe the calibrated model, including jobs outside this sufficient-condition region.

Figure 4 displays career dynamics. The poaching frontier is steep for young workers, who are willing to trade substantial contemporaneous productivity for a better learning environment. The reason is twofold. First, young workers have a high estimated learning ability. Second, they have a long remaining time in the labor market to benefit from higher skills. Older workers, on the other

Figure 4: Career Dynamics over the Life Cycle




*Notes:* The panels use the estimated baseline parameters, including $\epsilon = 0.034$. Darker shading indicates more employment at a firm type in the indicated age group; shading is normalized within each panel and transformed by a power of 0.3 for visibility. The orange solid locus is the acceptance frontier ($J = N$). The black dashed locus is the poaching frontier relative to the incumbent marked by the diamond ($e^{-\epsilon}J(\tilde{y}, \tilde{x}, a) = J(y_i, x_i, a)$). The circle marks the reference job. All displayed types satisfy $J > e^{-\epsilon}N$ in these age groups, so the quitting frontier lies outside the plotted support. Curves interpolate the model's type grid.

hand, are less willing to give up current productivity in favor of a better learning environment.

## 3.4 Equilibrium

The stationary worker distribution records the histories generated by these job choices. Let $\tilde{n}(a, dh)$ and $\tilde{g}(dy, dx, a, dh)$ denote the age-specific measures of nonemployed and employed workers. Entrants start at $h = 0$; employment raises $h$ at rate $\psi(a)x$; and every departure from an employer lowers $h$ by $\epsilon$. Workers return to nonemployment after an exogenous separation or when their match first reaches the quitting boundary. Appendix B.3 gives the forward equations, including this boundary flow.

Two moments of the distribution suffice for recruitment values:

$$\hat{n}(a) = \int e^h \tilde{n}(a, dh), \qquad \hat{g}(dy, dx, a) = \int e^h \tilde{g}(dy, dx, a, dh). \tag{8}$$

Where a type density exists, I write it as $\hat{g}(y, x, a)$. These measures weight workers by their human capital. In contrast, matching uses their numbers:

$$U = \int_0^A \int \tilde{n}(a, dh)\, da, \qquad S = U + \phi(1 - U). \tag{9}$$

**Job creation.** A type-$(y, x)$ incumbent chooses how many vacancies to advertise

$$\max_{v(y,x)\geq 0} \left\{ qv(y,x)Q(y,x) - \frac{c_v}{1+\eta} v(y,x)^{1+\eta} \right\}.$$

The corresponding first-order condition is

$$c_v v(y,x)^{\eta} \;\;=\;\; qQ(y,x), \tag{10}$$

where the value of a contact between a type-$(y, x)$ firm and a random prospective hire is

$$\begin{aligned} Q(y,x) \;\;&=\;\; (1-\beta) \int_0^A \frac{\widehat{n}(a)}{S} \mathbf{1}_{\mathcal{A}}(y,x) \left( J(y,x,a) - N(a) \right) da \\ &\quad + (1-\beta) \int_0^A \frac{\phi}{S} \int_{\mathcal{P}} \left( e^{-\epsilon} J(y,x,a) - J(\tilde{y}, \tilde{x}, a) \right) \widehat{g}(\tilde{y}, \tilde{x}, a) d\tilde{y}\, d\tilde{x}\, da. \end{aligned} \tag{11}$$

The first term is the expected surplus from meeting a nonemployed worker; the second is the expected incremental surplus from poaching an employed worker.

**Entry.** After drawing type $(y, x)$, an entrant implements the optimal vacancy policy (10) until it exits at rate $\iota$. Its implied discounted net return from hiring is

$$\mathcal{V}^e(y,x) \;\;\equiv\;\; \frac{1}{\rho+\iota} \max_{v\geq 0} \left\{ qvQ(y,x) - \frac{c_v}{1+\eta} v^{1+\eta} \right\} \;\;=\;\; \frac{1}{\rho+\iota} \frac{\eta}{1+\eta} c_v^{-\frac{1}{\eta}} \left( \chi^{\frac{1}{\alpha}} p^{\frac{\alpha-1}{\alpha}} Q(y,x) \right)^{\frac{1+\eta}{\eta}}.$$

Positive entry requires the entry cost to equal the expected return from the draw

$$c_e \;\;=\;\; \int_{\mathcal{E}} \mathcal{V}^e(y,x) d\Gamma(y,x) \;\;=\;\; \frac{1}{\rho+\iota} \frac{\eta}{1+\eta} c_v^{-\frac{1}{\eta}} \int_{\mathcal{E}} \left( \chi^{\frac{1}{\alpha}} p^{\frac{\alpha-1}{\alpha}} Q(y,x) \right)^{\frac{1+\eta}{\eta}} d\Gamma(y,x).$$

To study the labor tax in Section 2, let $\tau$ be the fraction of gross match output paid in tax. Bargaining divides the remaining output. The following normalization puts the tax and the two recruitment costs on a common footing.

**Lemma 3.** Let $\tau \in [0, 1)$ be a proportional tax on gross match output, and suppose worker value is delivered as a piece rate of post-tax match output. Under the reference-job normalization in Assumption 1, the taxed economy in units of retained output is isomorphic to the worker problem above and the firm problem with effective costs

$$\tilde{c}_v \;\;\equiv\;\; \frac{c_v}{1-\tau}, \qquad \tilde{c}_e \equiv \frac{c_e}{1-\tau}.$$

*Proof.* See Appendix B.6. □

**Definition.** The *cost composite* $c$ is

$$c \quad \equiv \quad (1-\tau)^{-\frac{1+\eta}{\eta}} c_v^{\frac{1}{\eta}} c_e \chi^{-\frac{1+\eta}{\alpha\eta}}. \tag{12}$$

**Definition.** A stationary equilibrium is a continuous contact-value function $Q \geq 0$ with $\int Q^{1/\eta} d\Gamma > 0$ satisfying $Q = \mathcal{T}_c(Q)$. The map consists of the following steps:

(1) Given $Q$, the contact rate $p$ and offer distribution $F$ are given by

$$p \quad = \quad \left( \frac{\eta}{1+\eta} \frac{1}{\rho+\iota} \frac{1}{c} \int_{\mathcal{E}} Q(y,x)^{(1+\eta)/\eta} d\Gamma(y,x) \right)^{\kappa}, \qquad \kappa \equiv \frac{\eta\alpha}{(1-\alpha)(1+\eta)}, \tag{13}$$

$$dF(y,x) \quad = \quad \frac{Q(y,x)^{1/\eta} d\Gamma(y,x)}{\int_{\mathcal{E}} Q(\tilde{y},\tilde{x})^{1/\eta} d\Gamma(\tilde{y},\tilde{x})}. \tag{14}$$

(2) Given $(p, F)$, the joint-value obstacle problem associated with (3) and (5) determine $(J, N)$ and the policy regions $(\mathcal{A}, \mathcal{M}, \mathcal{P}, \mathcal{C})$ in (4).

(3) Given $(p, F, \mathcal{A}, \mathcal{M}, \mathcal{P}, \mathcal{C})$, the forward equations in Appendix B.3 determine $(\tilde{n}, \tilde{g})$. Equations (8) and (9) then determine $(\hat{n}, \hat{g}, U, S)$.

(4) Given $(J, N, \mathcal{A}, \mathcal{P}, \hat{n}, \hat{g}, S)$, the right-hand side of (11) defines $\mathcal{T}_c(Q)$. Equilibrium requires $\mathcal{T}_c(Q) = Q$.

**Proposition 1.** Holding the remaining primitives and the reference-job convention fixed, any two configurations of $(\tau, c_v, c_e, \chi)$ with the same $c$ have the same equilibrium correspondence for $(Q, p, F)$ and the same worker allocation, with values measured in units of retained output.

*Proof.* See Appendix B.6. □

Proposition 1 allows policy to be summarized by the cost composite $c$.

Conditional on the remaining primitives, worker flows can at most identify the composite. They cannot distinguish its separate components. Firm size provides additional information. Across all incumbent firms, including those with no workers,

$$\frac{1-U}{M} \quad = \quad \tilde{c}_e \frac{1-U}{S} \frac{(1+\eta)(\rho+\iota)}{\eta p} \frac{\int Q^{1/\eta} d\Gamma}{\int Q^{(1+\eta)/\eta} d\Gamma}.$$

For a given worker allocation, a higher entry cost raises average firm size. For the quantitative size moments, I retain firm types with positive mean employment. Their conditional average size equals the right-hand side divided by their share in $\Gamma$. I take taxes from the data, normalize matching efficiency in the absence of vacancy data, and use firm-size information to separate entry from recruitment costs. The baseline calibration uses average size; the country-specific calibration uses the employment share at firms with at least 50 workers.

**Existence and uniqueness.** The equilibrium combines two feedbacks. More valuable worker contacts induce vacancy creation, while more frequent offers change job choices and the human capital of future recruits. The following results give sufficient conditions for this system to have an equilibrium and for that equilibrium to be unique.

**Assumption 2.** There is $\Delta_y > 0$ such that

$$\Gamma\{(y,x) \in \mathcal{E} : y \geq y_r + \Delta_y\} \quad > \quad 0.$$

This condition supplies a positive mass of firms with a productivity advantage over the reference job and hence potential gains from hiring.

**Proposition 2.** Maintain Assumptions 1 and 2 and the standard technical assumptions stated in Appendix B.4. If every type in $\mathcal{E}$ satisfies $y \geq y_r$ and $x \geq x_r$, a stationary equilibrium exists for every $\eta > 0$.

*Proof.* See Appendix B.4. □

The support restriction places the reference job weakly below every offered job in both attributes. Under this restriction, realized matches continue until an exogenous separation, an accepted outside offer, or retirement. This eliminates discontinuities associated with endogenous quitting from the fixed-point argument. The general model and its numerical implementation allow types on either side of the reference job.

**Proposition 3.** Maintain the hypotheses of Proposition 2 and suppose $0 < \eta \leq 1$. Fix all primitives other than the cost composite, maintaining the standard technical assumptions stated in Appendix B.5. There is $c_* > 0$ such that the stationary equilibrium is unique for every $c > c_*$.

*Proof.* See Appendix B.5. □

For a sufficiently high cost composite, contacts are infrequent and the feedback from contact values through hiring and worker careers is a contraction. This is a sufficient result for a restricted parameter region. It does not establish uniqueness of the calibrated economy, which is solved numerically.

### 3.5 The impact of policy

I next turn to the comparative statics of policy. Appendix B.7 defines local regularity and stability.

**Proposition 4.** Holding the remaining primitives fixed, at a positive, locally regular equilibrium that is linearly stable under the adjustment rule defined in Appendix B.7, an increase in the cost composite reduces the contact rate

$$\frac{dp}{dc} \quad < \quad 0.$$

*Proof.* See Appendix B.7. $\square$

For the next result, allow the support to include jobs below the reference productivity. The support restriction in Propositions 2–3 applies only to those sufficient existence and uniqueness results. Maintain (7) and fix an offer distribution $F$ satisfying

$$F\{(y,x) : y > y_r\} \quad > \quad 0.$$

Compare two contact rates $p_H > p_L$, with all other primitives fixed and the hiring-reference normalization maintained. Assume that both worker problems remain admissible for all sufficiently small nonnegative portability losses. Write $(J_k^\epsilon, N_k^\epsilon)$ for their values, where $k \in \{L, H\}$ indexes the contact rate. For any learning quality whose feasible output section is nonempty, define

$$z_k^{A,\epsilon}(x,a) \quad = \quad \inf\{e^y : (y,x) \in \mathcal{E},\ J_k^\epsilon(y,x,a) > N_k^\epsilon(a)\}.$$

**Lemma 4.** Fix $a < A$ and $x_r \leq x_1 < x_2$. Suppose that each of the four zero-loss acceptance cutoffs $z_k^{A,0}(x_i, a)$, $k \in \{L, H\}$ and $i \in \{1, 2\}$, lies in the interior of its feasible output section. There is $\bar{\epsilon} > 0$ such that, for $0 \leq \epsilon < \bar{\epsilon}$,

$$z_H^{A,\epsilon}(x_1,a) - z_H^{A,\epsilon}(x_2,a) \quad > \quad z_L^{A,\epsilon}(x_1,a) - z_L^{A,\epsilon}(x_2,a) > 0.$$

Each of these interior cutoffs satisfies

$$J_k^\epsilon(\log z_k^{A,\epsilon}(x_i,a), x_i, a) \quad = \quad N_k^\epsilon(a).$$

*Proof.* See Appendix B.9. $\square$

At fixed offer composition, more frequent contacts increase the value of future opportunities. With fully portable skills, this raises the current-output sacrifice that a nonemployed worker accepts in exchange for a better learning environment. The strict comparison persists for sufficiently small portability losses. The result does not sign the response to changes in the offer distribution; that equilibrium response is included in the quantitative analysis.

# 4 Calibration

I use a calibrated version of the model to isolate and quantify how four policy wedges contribute to the observed cross-country variation in life-cycle wage growth and aggregate output per worker:

the labor tax $\tau$, the cost of expanding incumbent firms $c_v$, the entry cost $c_e$, and on-the-job search efficiency $\phi$.

For computation, I approximate the continuous age dimension with four 10-year age groups: 20–29, 30–39, 40–49, and 50–59. Let $j = 1, \ldots, 4$ index these groups. Workers move stochastically between adjacent groups at a monthly rate of $1/120$, implying an average duration of 10 years in each group. I discretize firm productivity, the learning environment, and the piece rate using twelve grid points each. I do not discretize human capital because it scales out by Lemma 1. I first calibrate the common parameters to a hypothetical average-poaching economy and then recover the four country-specific policy wedges using a limited set of cross-country moments. The next section varies these wedges while holding the remaining structural parameters at their baseline values and compare the model's predictions for untargeted moments with the data.

### 4.1 Externally calibrated parameters

I fix nine parameters using either standard values from the literature or direct empirical counterparts. I set the monthly discount rate to $\rho = 0.04/12$, corresponding to an annual real interest rate of approximately four percent. Because aggregate matching efficiency is not separately identified from the cost parameters without vacancy data, I normalize $\chi = 1$ and hold it fixed in the cross-country exercises below. Variation in matching efficiency could therefore account for some of the cross-country variation left unexplained by the four policy wedges $(\tau, c_v, c_e, \phi)$. I set the matching-function elasticity to $\alpha = 0.63$, based on the estimate in Moscarini and Postel-Vinay (2025) from a model with on-the-job search.

I set the tax rate to $\tau = 0.417$, the total labor tax implied for the hypothetical average-poaching economy by the OECD data described in Section 2. I construct this and all other baseline outcomes below by regressing each country-level outcome on a constant and the log poaching rate across the 15 sample countries, then evaluating the fitted relationship at the sample mean of log poaching.[3]

Because I do not have cross-country data on worker separations caused by firm exit, I set $\iota = 0.05/12$, corresponding to a five-percent annual employment-weighted firm exit rate, consistent with the job destruction attributable to establishment deaths in the U.S. *Business Dynamics Statistics*. For each age group, I then set $\delta_j$ equal to the observed monthly EN rate minus $\iota$.[4] Panel A of Table 2 summarizes the externally calibrated parameters.

[3] When an outcome is unavailable for some countries, I estimate the regression using the available subsample but evaluate the fitted relationship at mean log poaching in the full 15-country sample.

[4] The model also allows endogenous separations as workers age into the separation region. These separations are rare in the calibrated model because workers accept only jobs satisfying $J > N > e^{-\epsilon}N$. Appendix C.1 reports the empirical EN profile used to set the separation rates.

### 4.2 Internally calibrated parameters

I parameterize age-specific learning ability as

$$\psi_j = \left(\frac{4-j}{3}\right)^{\psi}, \qquad j = 1, \ldots, 4,$$

so $\psi_1 = 1$ and $\psi_4 = 0$, with $\psi$ governing how quickly learning ability declines with age. Log firm productivity $y$ has a normalized exponential marginal with standard deviation $\sigma_y$, while the learning environment $x$ has a Gaussian marginal with standard deviation $\sigma_x$. A Gaussian copula links the two marginals, and $\varrho$ denotes its latent-normal dependence parameter.

These assumptions leave the following 10 parameters to be estimated:

$$\Big( \underbrace{p}_{\substack{\text{job finding}\\\text{rate}}}, \underbrace{\phi}_{\substack{\text{OTJ search}\\\text{efficiency}}}, \underbrace{\mu}_{\substack{\text{mean}\\\text{learning}}}, \underbrace{\psi}_{\substack{\text{learning}\\\text{decay}}}, \underbrace{\sigma_y}_{\substack{\text{productivity}\\\text{dispersion}}}, \underbrace{\epsilon}_{\substack{\text{human capital}\\\text{specificity}}}, \underbrace{\beta}_{\substack{\text{bargaining}\\\text{power}}}, \underbrace{\sigma_x}_{\substack{\text{learning}\\\text{dispersion}}}, \underbrace{\varrho}_{\substack{y\text{–}x\\\text{copula dep.}}}, \underbrace{\eta}_{\substack{\text{vacancy}\\\text{curvature}}} \Big).$$

I estimate these 10 parameters using 29 moments, not all of which are independent. Although identification is joint, the following discussion provides a heuristic mapping from moments to parameters.

The age profiles of the NE and JJ rates jointly inform the job-finding rate $p$ and on-the-job search efficiency $\phi$. Although $p$ is an equilibrium object, I can recover cost parameters $(c_v, c_e)$ that rationalize it; I explain this recovery below. A higher $p$ raises both the NE and JJ rates, whereas a higher $\phi$ raises the JJ rate relative to the NE rate. As before, the JJ mobility rate is the share of currently employed workers who started a new job during the previous 12 months without experiencing nonemployment over that interval. I construct this rate identically in the model and data and evaluate both age-specific flow rates at the midpoint of each age bin.

The life-cycle wage profile informs the mean learning environment $\mu$ and the rate at which learning ability declines with age, $\psi$. A higher $\mu$ generates more human capital and therefore more wage growth over the life cycle. A larger $\psi$ reduces human-capital accumulation and wage growth, especially later in workers' careers, making wage growth more front-loaded.

I use the wage change of job switchers relative to stayers to inform productivity dispersion $\sigma_y$. Greater productivity dispersion generates larger productivity differences and, consequently, larger wage gains for movers. Because the age distribution is uniform in the model but not in the data, I construct this and all other aggregate moments below as equally weighted means across the 20–29, 30–39, 40–49, and 50–59 age groups in both the model and data. To limit the influence of extreme outliers, I use the median wage change in the data. Computing the corresponding median in the model is difficult without simulation, so I use the model mean during estimation. Simulated model data confirm that the median and mean wage changes are very similar.

The wage loss of job losers informs the human-capital loss upon separation, $\epsilon$. I target the wage change between years $t-1$ and $t$ among workers who make an ENE transition during that interval. A larger $\epsilon$ generates a larger and more persistent wage loss.[5]

Two moments, ordered as in Figure 5, inform workers' bargaining power $\beta$. First, I use annual wage growth from $t+1$ to $t+2$ among workers hired from nonemployment between $t$ and $t+1$ who do not subsequently switch employers. If either $\epsilon$ is high or $\beta$ is low, workers start at low wages after a spell of nonemployment. The effect of low bargaining power is temporary, however, because workers' wages subsequently rise as they accumulate outside offers. Second, I use the wage difference between ENE movers who re-enter at large rather than small firms. Large firms tend to be highly productive in the estimated model, so a higher $\beta$ implies a larger large–small wage gap among workers re-entering from nonemployment.

The employment shares in the firm-size bins 1–4, 5–19, 20–49, and 50 or more discipline the vacancy-cost curvature $\eta$. A larger $\eta$ dampens differences in vacancy creation across firm types and compresses the firm-size distribution.

Excess wage growth at training relative to nontraining firms disciplines the dispersion of learning environments, $\sigma_x$.[6] The ECHP records whether a worker's employer provides training. Appendix C.2 validates this measure using within-worker variation across employers. In the model, I classify training firms using a smoothed cutoff in $x$, chosen so that approximately 45 percent of equal-age-weighted employment is at training firms in the hypothetical average-poaching economy. A larger $\sigma_x$ raises the training wage-growth premium.

Finally, four age-specific employment shares at training firms discipline the Gaussian-copula dependence parameter $\varrho$. Each share pools observations within a 10-year age bin. A positive dependence parameter makes high-productivity firms more likely to offer strong learning environments, so age-related sorting toward productive firms shapes the training-share profile.

I choose the 10 parameters to minimize a weighted sum of squared model–data deviations across the 29 targeted moments. Each moment family receives unit weight, divided equally among its constituent moments. I normalize each squared deviation within a family by the square of that family's mean data moment, treating the age profile of training shares as a separate family. For example, the contribution of the four age-specific NE rates to the minimum-distance objective is

$$Q_{\mathrm{NE}}(\theta) \quad = \quad \frac{1}{4}\sum_{j=1}^{4}\frac{\left(NE_j^{\text{model}} - NE_j^{\text{data}}\right)^2}{\left(\overline{NE}^{\text{data}}\right)^2}.$$

Normalizing by the family-level mean keeps the different moment families comparable in scale without placing excessive weight on age-specific outcomes that are close to zero.

[5] I measure this wage change among workers with a brief spell of nonemployment and assume that they lose no general skills. If they also lose some general skills, I overstate the importance of firm-specific skills, meaning that the true $\epsilon$ is lower than my estimate. Because a lower $\epsilon$ would imply even larger effects of policies that reduce worker flows, the estimates below are conservative.

[6] I exclude the oldest group because selection close to retirement makes this comparison difficult to interpret.

Panel B of Table 2 reports the internally estimated parameters. The monthly job-finding rate is $p = 0.044$. Matching the observed JJ rate requires on-the-job search efficiency of $\phi = 2.405$ relative to nonemployed search. This value is substantially higher than conventional estimates because I target the NE rate rather than the unemployment-to-employment (UE) rate. The lower NE rate implies a lower $p$ and therefore a higher $\phi$ to match observed JJ mobility. I have verified that the cross-country results below are robust to calibrating $p$ to the higher UE rate instead.

Learning ability falls rapidly with age: the estimate $\psi = 1.973$ implies that workers aged 30–39 have less than half the learning ability of workers aged 20–29. Match-productivity dispersion is $\sigma_y = 0.021$, much lower than the dispersion in value added or TFP across firms. Workers lose slightly more than three percent of their skills when switching employers, $\epsilon = 0.034$, consistent with evidence in Lazear (2003) and Gregory (2026) that most skills are portable across employers.

Workers' bargaining power is $\beta = 0.619$. For comparison, Gregory (2026) estimates $\beta = 0.66$.

The estimated Gaussian-copula dependence parameter is positive but below one, $\varrho = 0.298$.

The estimated vacancy-cost function is somewhat more convex than quadratic, $\eta = 1.308$.[7]

Table 2 also reports the implied values of $c_e$ and $c_v$. Because the normalization $\chi = 1$ is arbitrary, their levels are not economically meaningful, even when expressed relative to output per worker. As (12) makes clear, any alternative normalization of $\chi$ would imply different values of $c_e$ and $c_v$ while fitting the data equally well.

Conditional on the calibrated parameter vector, I infer the age-specific component $b_j$ of the flow value of leisure so that workers are indifferent between nonemployment and employment at the fixed reference match $(y_r, x_r)$. The implied schedule $b(a)$ rises with age because younger workers have a stronger incentive to work and accumulate skills. The average flow value of leisure $e^{b(a)+h}$ exceeds 80 percent of the average wage for three reasons. First, heterogeneity in $(y, x)$ is relatively modest: as shown below, the model matches within-worker wage dispersion over time, which is substantially smaller than wage dispersion across workers even after removing observable characteristics. Second, workers search more efficiently while employed, making nonemployment less attractive. Third, employment allows workers to accumulate skills.

### 4.3 Model fit

Because the number of moments exceeds the number of estimated parameters, the model does not fit every target exactly. Figure 5 nevertheless shows that it reproduces the main life-cycle and firm-size patterns. The model captures the decline in the NE rate with age reasonably well in panel (a). The NE rate declines because young nonemployed workers are willing to accept some low-productivity, high-learning jobs that their older counterparts reject.

[7] The sufficient uniqueness result in Proposition 3 requires $\eta \leq 1$ and a reference job below the support in both attributes. These conditions are not imposed in the calibration. Numerical convergence does not establish global uniqueness.

Table 2: Parameters

| Parameter | Description | Value | Source / identifying moment |
|---|---|---|---|
| **Panel A. Externally Calibrated** | | | |
| $\rho$ | Discount rate | 0.003 | 4% annual real rate |
| $\chi$ | Matching efficiency | 1.000 | Normalized |
| $\alpha$ | Matching elasticity | 0.630 | Moscarini and Postel-Vinay (2025) |
| $\tau$ | Labor tax wedge | 0.417 | OECD Taxing Wages, 2000–2020 |
| $\iota$ | Firm exit hazard | 0.004 | 5% annual, employment-weighted |
| $M$ | Mass of firms | 0.245 | Eurostat, 2004–2020 |
| $\delta(a)$ | Separation hazard by age | | EN rate |
| | Age 20–30 | 0.017 | |
| | Age 30–40 | 0.006 | |
| | Age 40–50 | 0.003 | |
| | Age 50–60 | 0.004 | |
| **Panel B. Internally Calibrated (Figure 5)** | | | |
| $p$ | Job finding rate | 0.044 | NE rate (panel (a)) |
| $\phi$ | On-the-job search efficiency | 2.405 | JJ rate (panel (b)) |
| $\mu$ | Mean learning environment ($\times$100) | -0.098 | Wages by age (panel (c)) |
| $\psi$ | Decline in learning ability | 1.973 | Wage growth shares (panel (d)) |
| $\sigma_y$ | Std. dev. of productivity | 0.021 | Job switchers (panel (e)) |
| $\epsilon$ | HC loss on accepting a new job | 0.034 | Job losers (panel (e)) |
| $\beta$ | Worker bargaining power | 0.619 | Post re-entry, large/small (panel (f)) |
| $\sigma_x$ | Std. dev. of learning environment ($\times$100) | 0.176 | Training wage premium (App. C.2) |
| $\varrho$ | Gaussian-copula dependence of $y$ and $x$ | 0.298 | Training share by age (panel (h)) |
| $\eta$ | Vacancy cost curvature | 1.308 | Employment by size (panel (i)) |
| **Panel C. Implied** | | | |
| $c_e$ | Cost of starting a business | 8.157 | Average firm size of 6.06 |
| $c_v$ | Vacancy cost scale | 13.78 | Rationalize equilibrium $p$ |
| $e^{b(a)}$ | Flow value of leisure | | |
| | Age 20–30 | 0.681 | |
| | Age 30–40 | 0.828 | |
| | Age 40–50 | 0.922 | |
| | Age 50–60 | 0.986 | |

The JJ rate in panel (b) declines with age by a similar amount in the model and data for two reasons. First, young nonemployed workers accept some stepping-stone jobs with low productivity but strong learning environments that their older counterparts reject. Their subsequent movement toward more productive jobs raises their JJ rate. Second, young workers have had less time to find a good match and therefore remain more mobile. JJ mobility also declines more gradually with age than is typical in one-dimensional job-ladder models. The reason is twofold. First, a two-dimensional ladder creates more scope for mismatch than a one-dimensional model with the same productivity dispersion, slowing convergence toward the ergodic distribution. Second, workers' relative valuations of job attributes change over the life cycle, sustaining mobility later in life as workers re-sort.

The model tracks the life-cycle wage profile in panel (c) and captures the concentration of wage growth early in workers' careers in panel (d). Although these two sets of moments are linearly dependent, including both is useful because the latter better highlights the front-loading of wage growth that informs $\psi$. Conditional on adjusting the mean learning environment $\mu$ to match the same overall life-cycle wage growth, a higher $\psi$ shifts more wage growth toward the beginning of the life cycle.

The model also comes close to the excess wage-change measures for job switchers and job losers in panel (e). In both the model and data, workers make only small average wage gains around a JJ transition. Several factors explain these small gains.[8] First, in both the model and data, a JJ switcher is a worker who changed employers during the previous 12 months without a recorded month of nonemployment during that 12-month interval. This group includes workers who made an ENE transition with a nonemployment spell lasting less than one month, and these workers tend to experience wage losses. Second, I measure the wage change between months 1–12 and 13–24 for workers who switched employers between months 6–17. Additional mobility before or after the recorded transition mutes measured wage-growth differences between job stayers and switchers. Third, some job switchers move to less productive firms with better learning environments, typically accepting a wage cut.

Panel (f) shows that the model closely matches workers' wage growth in the year after re-entry from nonemployment and the wage difference between ENE movers who re-enter at large rather than small firms. As discussed above, these moments are particularly informative about workers' bargaining power $\beta$. When $\beta$ is low, workers re-entering from nonemployment start at low wages but subsequently experience large wage gains as they accumulate renegotiation capital. A low $\beta$ also implies a smaller wage difference between workers re-entering at high-productivity, typically large firms and those re-entering at low-productivity, typically small firms.

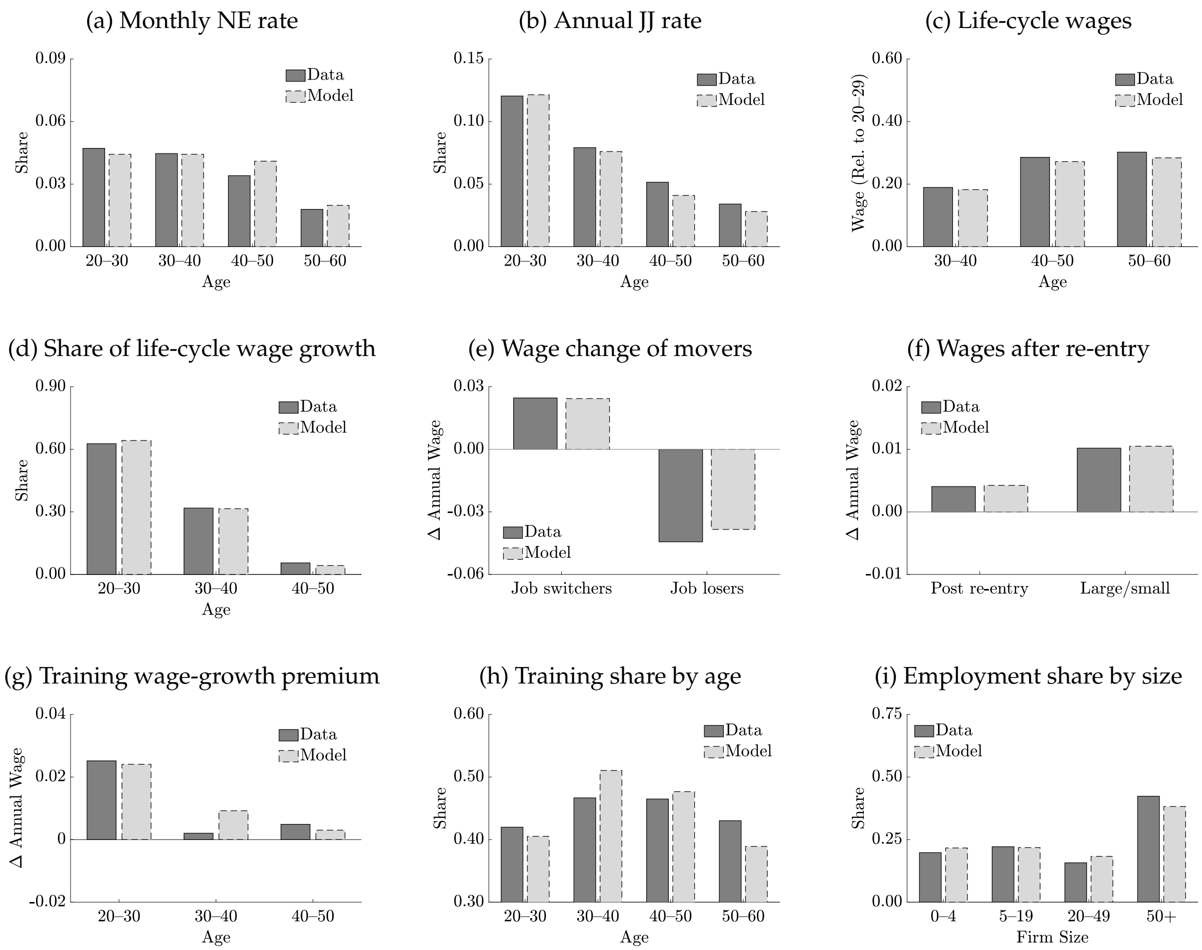
The model also captures that the training wage-growth premium is concentrated among workers aged 20–29 in panel (g). As discussed above, this age profile is particularly informative about the dispersion of learning environments, $\sigma_x$.

The model closely matches the hump-shaped age profile of employment at training firms in panel (h). Although young workers value training firms most, finding these jobs takes time. Moreover, the positive dependence between productivity and the learning environment causes workers initially to move toward training firms. If the dependence between the two attributes were perfect—so that the two-dimensional job ladder collapsed to a one-dimensional ladder—workers would move monotonically toward high-productivity training firms as they aged.

The model also matches well the distribution of employment across firm-size bins in panel (i).

[8] As noted above, I use the median gain in the data and the mean gain in the model to limit the influence of outliers that the model cannot match. Simulated model data confirm that the median and mean wage gains of JJ switchers are very similar.

Figure 5: Targeted Moments




*Notes:* Each data moment is the fitted value from a regression of the country-level moment on log aggregate poaching, evaluated at the average log poaching rate across countries. Panel (a) reports the monthly NE rate by age; panel (b) the annual JJ mobility rate by age; panel (c) log hourly wages by age relative to ages 20–29; and panel (d) the share of total life-cycle log-wage growth realized in each subsequent ten-year interval. Panel (e) reports two pooled excess log-wage-change measures from $t$ to $t+1$, each relative to job stayers: job switchers and job losers. Panel (f) reports two further measures for workers re-entering from nonemployment: the change from $t+1$ to $t+2$ for those who stay with the same employer, and the difference between re-entry at large (50+) and small employers. Panel (g) reports excess log-wage growth at training relative to nontraining firms by age for the three targeted age groups; panel (h) the employment share at training firms by age; and panel (i) the employment share by firm-size bin. *Source:* ECHP 1993–2001; EU-SILC 2003–2024; PSID 1993–2023; author's calculations.

## 4.4 Cross-country wedges

I next recalibrate four wedges separately for each country: the tax on match output $\tau$, the entry cost $c_e$, the cost of expanding incumbent firms $c_v$, and on-the-job search efficiency $\phi$. The on-the-job search wedge is a reduced-form representation of employed workers increasing their search effort when the job-finding rate is higher, as observed empirically. Some factor that differentially affects employed search is necessary to match the cross-country patterns in the NE and JJ rates jointly.

I take the tax rate $\tau$ from the OECD measure of the total labor tax, which includes taxes levied on both the worker and the firm. Conditional on matching each country's aggregate poaching rate exactly, I calibrate the remaining three wedges $(c_e, c_v, \phi)$ to match as closely as possible the employment share at firms with at least 50 employees, the NE rate, and on-the-job search, with the former expressed as percentage points deviations and the latter as log deviations from the hypothetical average-poaching economy in both the data and model. This is an overidentified joint exercise, with the following identification logic.

Given a tax rate $\tau$ from the data, $c_e$ and $c_v$ governs both the aggregate poaching rate and the NE rate via the cost composite $c$ and its effect on the job finding rate $p$. The two, however, have different effects on the firm-size distribution. Specifically, a higher $c_v$ makes it more expensive for incumbent firms to scale up, which in equilibrium raises the mass of active firms and hence lowers firm size. A higher $c_e$, on the other hand, reduces entry and the mass of active firms, raising average firm size. I hence use the employment share at large firms to inform the entry wedge $c_e$, and the NE rate to infer $c_v$. Finally, relative on-the-job search efficiency $\phi$ affects both aggregate poaching and on-the-job search. In the data, I measure the latter as the fraction of employed workers who looked for another job during the previous four weeks.[9]

Figure 6 reports the resulting wedges and calibration fit for all 15 countries. By construction, the model exactly matches the tax rate, aggregate poaching rate, and employment share at large firms. It trades off the NE rate against on-the-job search but broadly reproduces both outcomes. The calibration implies substantial cross-country variation in the entry, incumbent, and on-the-job search wedges. Panel (d) shows that the variation in the entry wedge is similar in magnitude to the World Bank's cost-of-starting-a-business measure. Although no comparable measure exists for the cost of expanding incumbent firms, panel (e) shows that variation in the incumbent wedge is also comparable in magnitude to the World Bank series.

In the next section, I vary these wedges either jointly or individually, hold all other structural parameters at their baseline values, and compare the model's predictions with the data along untargeted dimensions.[10] This exercise isolates and quantifies the effects of the wedges on workers' careers and aggregate outcomes.

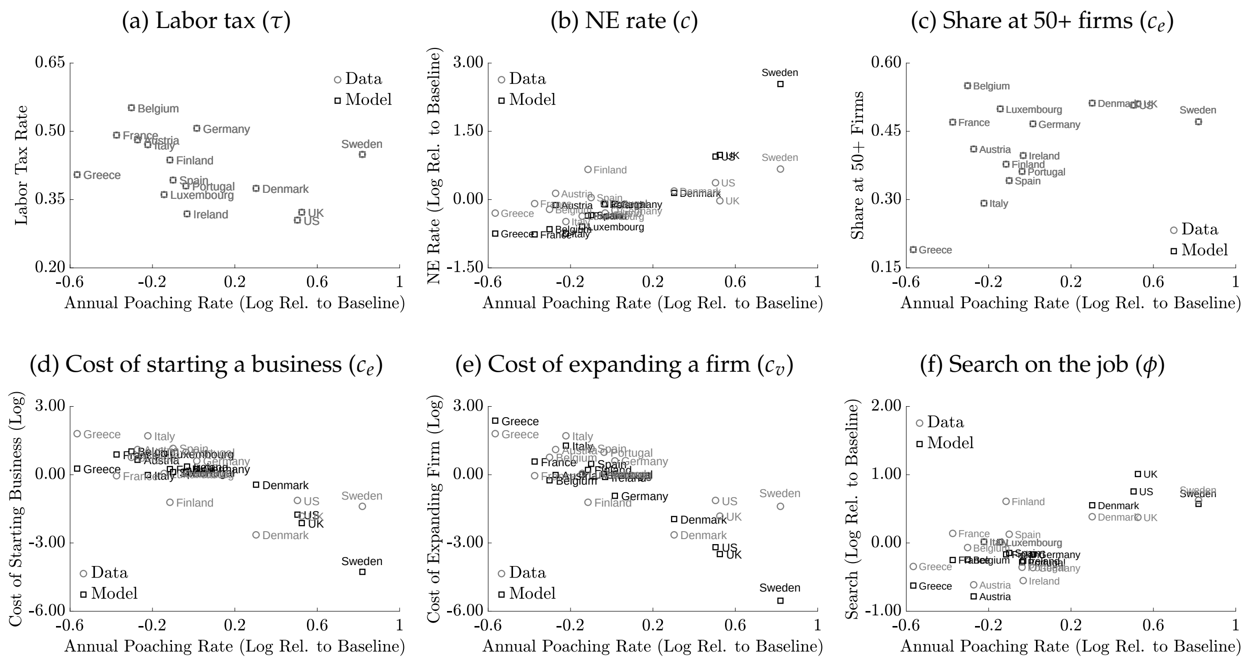
# 5 Dynamic Misallocation

I use the estimated model to quantify how the four labor-market wedges shape worker careers and output. I begin by comparing their predicted effects—jointly and individually—with untargeted

[9]This measure is available for every sample country except the United States. I impute the U.S. value by regressing log on-the-job search on a constant and log poaching across countries with valid data.

[10]I also recalibrate the age-specific leisure values $b_j$ for each country to preserve the reference-job hiring indifference condition in Assumption 1. The implied flow value of leisure rises modestly with poaching, which generally dampens the effect of poaching on wage growth. In practice, however, holding $b_j$ fixed at their baseline values produces very similar results.

Figure 6: Cross-Country Wedges and Calibration Fit




*Notes:* Each panel is plotted against the log annual poaching rate; squares show the model and circles the data. No fitted lines are shown. Panels (f) and (b) report log deviations of on-the-job search and the NE rate from their cross-country geometric means. Panel (a) reports the total labor tax in levels. Panels (d) and (e) report log deviations of the recovered entry and incumbent cost wedges, alongside the World Bank cost of starting a business as an external validation series rather than a calibration target. Panel (c) reports the employment share at firms with at least 50 employees in levels; the model matches this share exactly. Aggregate poaching, which is not shown as a separate panel, is also matched exactly in every country. *Source:* ECHP, EU-SILC, PSID, EU-LFS, OECD Taxing Wages, World Bank; author's calculations.

cross-country patterns in the data. This exercise quantifies the contribution of each wedge to the observed differences across countries. I then conduct three model-based exercises comparing representative low- and high-flow economies designed to quantify the importance of dynamic misallocation. Finally, I provide additional supportive evidence and robustness exercises.

## 5.1 Cross-country implications of the wedges

I quantify the contribution of the estimated policy variation toward observed cross-country differences in two untargeted outcomes: life-cycle wage growth and output per worker. To that end, I feed the four country-specific policy wedges from Section 4.4 into the model, holding all other parameters fixed at their baseline values.[11] I discuss the two outcomes in turn.

Table 3 decomposes the joint and individual contributions of the four wedges to cross-country

[11]The moments in this section differ slightly from the previous section for the following reason. In order to speed up the estimation, I target moments that can be directly computed in the model without simulation, such as the log of average annual wages, identically computed in model and data. Having estimated the model, I use a simulated model to compute average log annual wages, consistent with the treatment of the data in Section 2. It is not possible to derive analytical expressions for these outcomes. In practice, the two sets of moments correspond closely.

variation in life-cycle wage growth. Combined, the wedges generate cross-country variation in life-cycle wage growth that matches well the magnitude of the empirical variation—the cross-country standard deviation of life-cycle wage growth in the model is 1.16 times its empirical value. Projecting wage growth on poaching, the four wedges somewhat overstate the systematically higher wage growth in high poaching countries—the slope coefficient induced by these four policy wedges is 1.23 times the empirical counterpart. These four wedges explain 66 percent of the empirical variation.

Table 3: The Joint and Individual Effects of Policy Wedges on Life-Cycle Wage Growth

| | Joint Fit | | | Individual Wedges | | | |
|---|---|---|---|---|---|---|---|
| Country | Data | Model | Gap | $\tau$ | $c_e$ | $c_v$ | $\phi$ |
| Austria | -0.086 | -0.077 | 0.009 | 0.000 | -0.009 | -0.014 | -0.026 |
| Belgium | -0.100 | -0.058 | 0.042 | -0.007 | -0.016 | -0.008 | -0.006 |
| Denmark | 0.002 | 0.048 | 0.046 | 0.002 | 0.003 | 0.024 | 0.026 |
| Finland | -0.023 | -0.022 | 0.001 | 0.013 | 0.007 | -0.007 | 0.003 |
| France | -0.098 | -0.087 | 0.011 | -0.018 | -0.037 | -0.040 | -0.018 |
| Germany | 0.081 | -0.043 | -0.124 | -0.015 | -0.021 | -0.007 | -0.016 |
| Greece | -0.201 | -0.117 | 0.084 | 0.010 | -0.003 | -0.049 | -0.025 |
| Ireland | 0.032 | -0.039 | -0.071 | 0.008 | -0.031 | -0.015 | -0.012 |
| Italy | 0.013 | -0.043 | -0.055 | -0.006 | -0.016 | -0.030 | -0.007 |
| Luxembourg | 0.015 | -0.059 | -0.074 | 0.000 | -0.040 | -0.032 | -0.017 |
| Portugal | 0.007 | -0.047 | -0.054 | 0.018 | -0.004 | -0.006 | -0.010 |
| Spain | -0.036 | -0.046 | -0.009 | 0.007 | -0.016 | -0.025 | -0.009 |
| Sweden | 0.207 | 0.288 | 0.081 | -0.002 | 0.116 | 0.104 | 0.044 |
| UK | 0.097 | 0.161 | 0.064 | -0.009 | 0.031 | 0.055 | 0.063 |
| US | 0.089 | 0.141 | 0.051 | -0.002 | 0.035 | 0.050 | 0.010 |
| Std. dev. | 0.096 | 0.108 | | 0.010 | 0.037 | 0.040 | 0.025 |
| Std. dev. ratio (model/data) | | 1.13 | | 0.10 | 0.39 | 0.42 | 0.26 |
| Correlation | | 0.82 | | -0.18 | 0.62 | 0.80 | 0.67 |
| Slope ratio (model/data) | | 1.21 | | -0.01 | 0.35 | 0.45 | 0.25 |
| $R^2$ | | 0.68 | | 0.03 | 0.39 | 0.64 | 0.45 |

*Notes:* Entries are deviations from the fitted value in a hypothetical economy at mean log poaching. For each outcome and each data or model series, I regress the country value on log poaching and subtract the fitted value evaluated at mean log poaching. Under "Joint Fit," "Model" feeds all four country-specific wedges jointly and "Gap" is Model minus Data. Under "Individual Wedges," each column feeds the indicated wedge at its country-specific value and holds the other three at their baseline values. These are separate equilibrium counterfactuals, so the individual wedge effects need not sum to the joint model effect. The summary rows report the cross-country standard deviation, the model–data standard-deviation ratio, correlation, slope ratio, and squared correlation. *Source:* ECHP, EU-SILC, PSID, EU-LFS, OECD Taxing Wages, World Bank; author's calculations.

Individually, the tax wedge generates dispersion in life-cycle wage growth of 0.01 or about 10 percent of the empirical dispersion. Projected on the aggregate poaching rate, its slope coefficient is only two percent of its empirical counterpart. Hence, alone it accounts for little of the cross-country patterns. The entry- and incumbent-cost wedges each generate dispersion in life-cycle wage growth of about 40 percent of its empirical counterpart, and a slope coefficient with respect to poaching that is 35–43 percent of its empirical counterpart. Hence, these policy wedges are the most important quantitatively. Finally, the wedge to search on-the-job generate dispersion of life-cycle wage growth equal to about a quarter of the empirical dispersion, and a slope coefficient on poaching that is also about a quarter of its empirical counterpart.

Table 4: The Joint and Individual Effects of Policy Wedges on Output

| | Joint Fit | | | Individual Wedges | | | |
|---|---|---|---|---|---|---|---|
| Country | Data | Model | Gap | $\tau$ | $c_e$ | $c_v$ | $\phi$ |
| Austria | 0.077 | -0.079 | -0.156 | -0.009 | -0.041 | -0.028 | 0.001 |
| Belgium | 0.125 | -0.096 | -0.220 | -0.019 | -0.054 | -0.020 | 0.001 |
| Denmark | 0.028 | 0.047 | 0.019 | 0.007 | 0.006 | 0.042 | -0.002 |
| Finland | -0.084 | -0.065 | 0.020 | -0.001 | -0.023 | -0.032 | 0.003 |
| France | 0.071 | -0.102 | -0.173 | -0.007 | -0.046 | -0.042 | 0.005 |
| Germany | 0.087 | -0.041 | -0.128 | -0.012 | -0.025 | 0.003 | 0.002 |
| Greece | -0.591 | -0.116 | 0.474 | 0.003 | -0.024 | -0.091 | 0.004 |
| Ireland | 0.250 | -0.049 | -0.299 | 0.010 | -0.030 | -0.026 | 0.000 |
| Italy | -0.120 | -0.089 | 0.030 | -0.004 | -0.011 | -0.063 | 0.004 |
| Luxembourg | 0.468 | -0.079 | -0.547 | 0.005 | -0.051 | -0.029 | -0.001 |
| Portugal | -0.533 | -0.051 | 0.481 | 0.001 | -0.022 | -0.032 | -0.001 |
| Spain | -0.196 | -0.064 | 0.132 | 0.002 | -0.020 | -0.043 | 0.001 |
| Sweden | 0.148 | 0.393 | 0.245 | -0.002 | 0.195 | 0.180 | -0.002 |
| UK | -0.052 | 0.210 | 0.261 | 0.010 | 0.081 | 0.095 | -0.010 |
| US | 0.322 | 0.182 | -0.140 | 0.015 | 0.067 | 0.087 | -0.005 |
| Std. dev. | 0.275 | 0.142 | | 0.009 | 0.064 | 0.069 | 0.004 |
| Std. dev. ratio (model/data) | | 0.52 | | 0.03 | 0.23 | 0.25 | 0.01 |
| Correlation | | 0.26 | | 0.08 | 0.13 | 0.40 | -0.20 |
| Slope ratio (model/data) | | 1.43 | | 0.04 | 0.60 | 0.70 | -0.03 |
| $R^2$ | | 0.07 | | 0.01 | 0.02 | 0.16 | 0.04 |

*Notes:* Entries are deviations from the fitted value in a hypothetical economy at mean log poaching. For each outcome and each data or model series, I regress the country value on log poaching and subtract the fitted value evaluated at mean log poaching. Under "Joint Fit," "Model" feeds all four country-specific wedges jointly and "Gap" is Model minus Data. Under "Individual Wedges," each column feeds the indicated wedge at its country-specific value and holds the other three at their baseline values. These are separate equilibrium counterfactuals, so the individual wedge effects need not sum to the joint model effect. The summary rows report the cross-country standard deviation, the model–data standard-deviation ratio, correlation, slope ratio, and squared correlation. *Source:* ECHP, EU-SILC, PSID, EU-LFS, OECD Taxing Wages, World Bank; author's calculations.

Table 4 provides the corresponding decomposition for output per worker, measured in PPP-adjusted USD. Combined, the wedges generate dispersion in output per worker that corresponds to roughly half the empirical dispersion across these countries. They somewhat overstate the systematic relationship between output per worker and poaching in the data. However, these wedges explain only a small share of the empirical variation in output per worker, with an $R^2$ of 0.07. Hence, other factors account for the majority of the dispersion in output per worker across these countries. Individually, the entry- and vacancy-cost wedges again generate the largest effects, while the tax wedge and wedge to on-the-job search are less important.

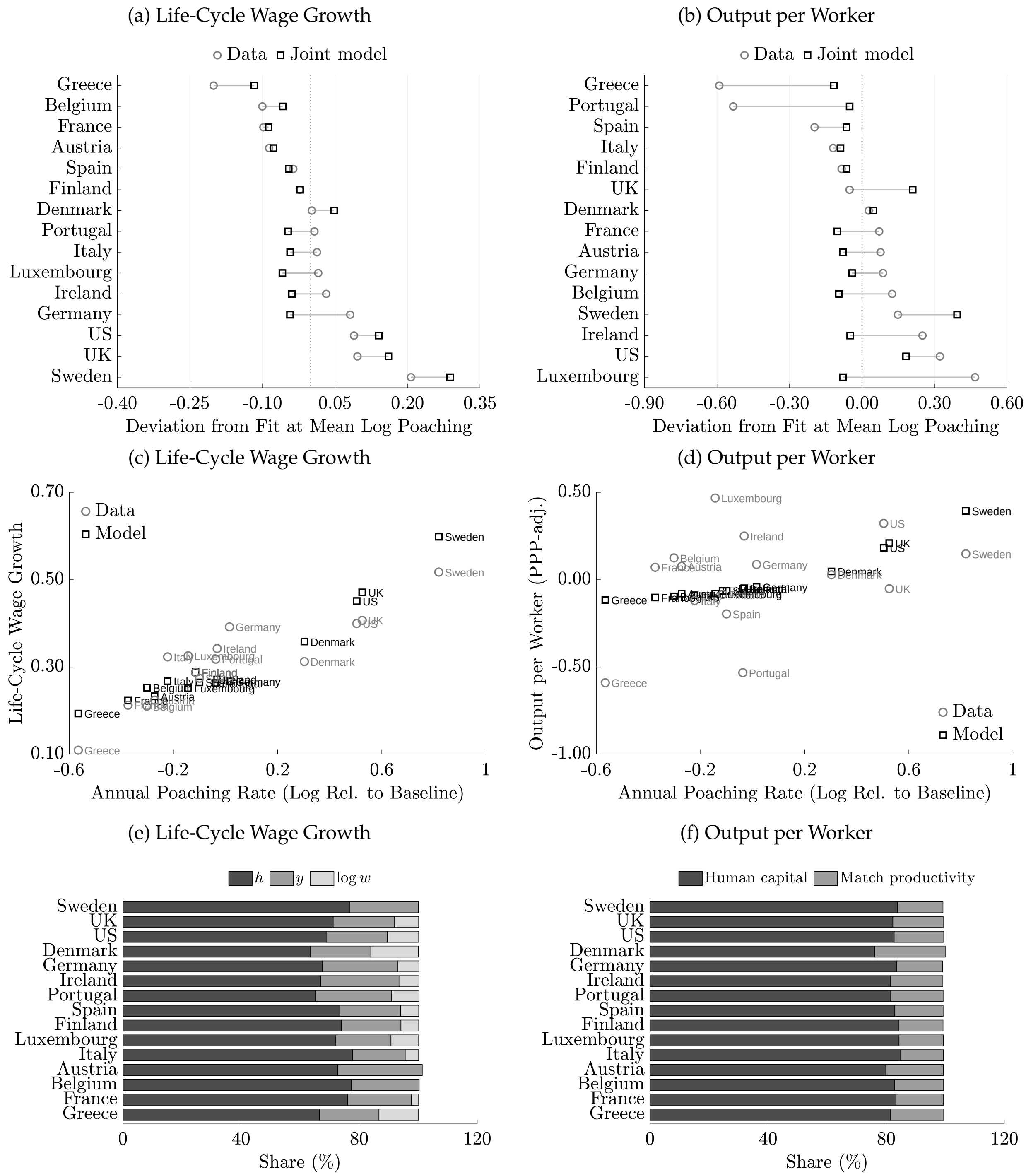
Figure 7 visualizes the combined impact of the four policy wedges on life-cycle wage growth (left panels) and output per worker (right panels), as well as the raw data. The top panels show the gap between country $c$ and the hypothetical baseline average-poaching country in the data and model, while the middle panels plot the outcome of interest against the aggregate poaching rate. Contrasting the effect of these wedges on life-cycle wage growth with the data country-by-country, the wedges particularly fail to explain the experiences of Germany and Sweden. These wedges result in lower life-cycle wage growth in Germany relative to the hypothetical average poaching country, whereas in the data wage growth is higher in Germany than the average poaching country. One possibility is that its well-developed apprenticeship program raises wage growth, countering the forces emphasized here. The wedges predict even larger wage growth in Sweden relative to the hypothetical average poaching country than what is observed in the data. One possibility is that its high degree of progressive taxation dampens wage growth (Guvenen, Kuruscu and Ozkan, 2014). For output per worker, these four wedges particularly fail to explain the much lower output per worker in Greece and Portugal than the hypothetical average poaching country, as well as the much higher output per worker in Luxembourg in the data.

The middle panels plot life-cycle wage growth and output per worker against the aggregate poaching rate. Combined, the four policy wedges explain most of the systematically higher life-cycle wage growth and output per worker in high-poaching countries. Yet as is clear from panel (d), the data display substantial variation in output per worker conditional on poaching, which these four policy wedges cannot explain. Taken together, the results show that these four wedges explain a substantial share of the cross-country variation in life-cycle wage growth, but much less of that in output per worker. Put another way, other factors are clearly first-order to explain cross-country differences in productivity. Yet these four policy wedges generate economically meaningful variation in output per worker. The model also predicts higher entry wages and employment in high-flow economies, as shown in Appendix C.4.

The bottom panels of Figure 7 decompose the aggregate effects of the four policy wedges into their effect on human capital accumulation, match productivity, and in the case of life-cycle wage growth, the piece rate.[12] The aggregate effect of these wedges on both life-cycle wage growth and output per worker are predominantly driven by their effect on human capital accumulation, which accounts for roughly 65 percent of the rise in life-cycle wage growth and 80 percent output per worker with poaching. Greater match productivity accounts for most of the remaining effects.

[12]To be consistent with the model-data comparison above, I decompose average log annual wages into average log annual human capital, average log annual match productivity, average log annual piece rates, and a residual arising from the fact that while log wages are additive in log human capital, log match productivity and the log piece rate, log annual wages are not (although in practice the decomposition is close to exact). I decompose log average output per worker into log average human capital per worker, log average match productivity per worker and a residual that arises from the fact that while log output is additive in log human capital and log match productivity, log average output per worker is not.

Figure 7: Country-Level Fit Under All Four Policy Wedges




*Notes:* Life-cycle wage growth is growth between ages 20–29 and 50–59 estimated from a regression with worker fixed effects, year fixed effects, and age effects restricted to be the same after age 45 (in both the model and data). The model adds a small aggregate proportional shifter to all wages in that year. Output per worker is log PPP-adjusted output per worker in the data and log gross output in the model. Model wage-growth estimates are based on monthly simulations of 50,000 workers per country, aggregated to annual observations. For each worker, I retain a randomly located consecutive block whose length is drawn from the country-specific empirical distribution of observations per worker, capped at eight years. *Source:* ECHP, EU-SILC, PSID, EU-LFS, OECD Taxing Wages, World Bank; author's calculations.

Variation in the piece rate plays a relatively small role for life-cycle wage growth, and does by construction not contribute to variation in output per worker.

### 5.2 The role of dynamic misallocation

To highlight the importance of dynamic misallocation, I contrast the impact of the estimated policy wedges in the baseline model with that in an alternative model in which firms do not differ in learning environments and skills accumulate exogenously, independently of labor-market dynamics. I calibrate this alternative model to fit many of the same moments as the baseline model in the hypothetical average poaching country—see Appendix C.3 for details. I subsequently analyze the impact of variation in the four policy wedges above, moving from France as a typical low poaching country to the U.S. as a typical high poaching country. In the baseline model, I simply set $(\tau, c_e, c_v, \phi)$ to their calibrated values in these two countries. The alternative model has the same tax rate $\tau$ in the hypothetical average poaching country as the baseline model, so I vary $\tau$ by the same amount in both models. The calibrated values of $(c_e, c_v, \phi)$ in the hypothetical average poaching country however differ across the two models. To make the results comparable, I move $(c_e, c_v, \phi)$ by the same proportional amount in the alternative model as in the baseline model.

Table 5 reports the results from this exercise. Life-cycle wage growth increases by 15.7 log points in the baseline model, including 7.0 log points from human capital. In contrast, life-cycle wage growth rises by only 5.9 log points when skills are policy invariant, as by construction human capital does not respond.

The key comparison is output per worker. The effect is 28.3 log points in the baseline model and 13.3 log points in the restricted model. Allowing policy to affect skill accumulation therefore more than doubles the estimated output effect. Within the baseline model, higher human capital contributes 21.5 log points and higher match productivity contributes 4.3 log points. By construction, human capital does not change in the alternative model. The policy wedges have a larger effect on match productivity in the alternative model, for two reasons. First, I calibrate higher productivity dispersion $\sigma_y$ in the alternative model to match the same moments. Second, workers in the baseline model trade off productivity with learning environment, so that the greater mobility into high-learning environments in high-poaching countries comes at the cost of a smaller effect of poaching on match productivity.

Figure 8 illustrates the reallocation toward better learning environments that accompanies the move from France to the United States. Panel (a) plots the marginal distribution of employed workers over firms' learning environments, pooling all ages. Relative to France, the United States places more employment mass at high-$x$ firms and above the common training-firm cutoff. Panel (b) shows the corresponding pattern across all countries. The training-firm employment share generally rises with poaching in both the model and the data, although the empirical relationship is more dispersed. Panel (c) examines age sorting among NE hires after netting out common

Table 5: Impact of Moving Policy Wedges from France to the U.S.

| | Baseline | Policy-invariant skills |
|---|---|---|
| Life-cycle wage growth (log points) | 15.7 | 5.9 |
| Match productivity | 5.1 | 5.2 |
| Human capital | 7.0 | 0.0 |
| Piece rate | 3.5 | 0.7 |
| Output per worker (log points) | 28.3 | 13.3 |
| Match productivity | 4.3 | 12.7 |
| Human capital | 21.5 | 0.0 |

*Notes:* Entries report the change from France to the United States. The low (high) economy imposes $\phi$, the labor tax $\tau$, the entry cost $c_e$, and the vacancy cost $c_v$ at that country's own calibrated value from the cross-country re-estimation; no projection or quartile average is involved. The job-finding rate $p$ adjusts so that free entry supports the imposed $c_e$. The tax changes by the same amount in the two models; $\phi$, $c_e$, and $c_v$ change by the same proportions. The job-finding and employment rates are in percentage points; all other entries are in log points. The baseline model is the full model with endogenous learning environments. In the policy-invariant-skills model, human-capital accumulation does not respond to the wedges. Components can differ slightly from their totals because of rounding and aggregation. *Source*: author's calculations.

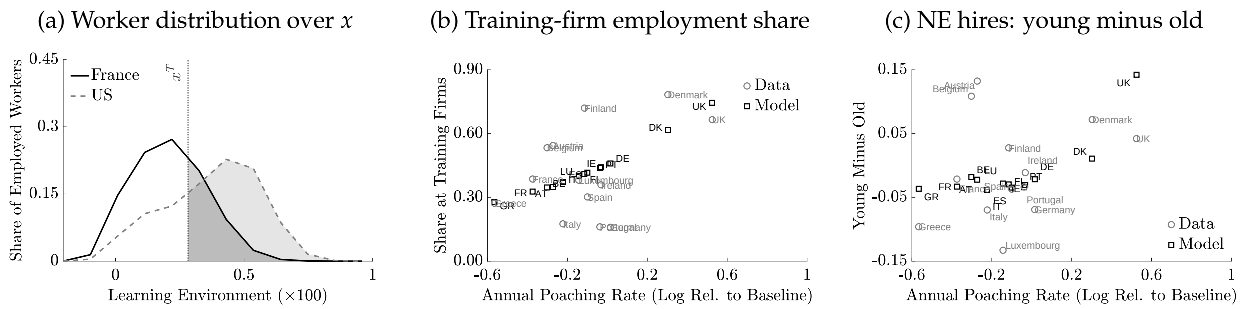
employer-size effects. It compares the training-firm share of workers aged 20–34 with that of workers aged 40–59. This young–old gap rises with poaching in both the model and the data, with a steeper response in the model.

### 5.3 The role of changes in behavior

Higher contact rates affect skill accumulation through both faster access to learning opportunities and changes in behavior. Even with decision rules held fixed, more frequent contacts help workers reach firms with better learning environments sooner. Better prospects for future mobility also raise workers' valuation of learning, inducing changes in their acceptance and mobility decisions and in firms' vacancy creation. To separate these channels, I conduct a partial-equilibrium exercise that moves $p$ and $\phi$ from their French to their U.S. values while holding workers' acceptance and mobility rules and firms' vacancy-creation rules fixed at their baseline values.

Table 6 shows that the fixed-behavior exercise generates 19.5 of the 28.3 log-point increase in output per worker. It also generates 14.0 of the 21.5 log-point increase in human capital and 3.7 of the 4.3 log-point increase in match productivity. Most of the output effect therefore comes from faster access to learning opportunities, with behavioral responses accounting for 31 percent of the aggregate effect. The final column varies the contact rate of nonemployed workers, $p$, while holding the contact rate of employed workers, $\phi p$, at its baseline value and keeping behavior fixed. Faster job finding reduces time spent in nonemployment, where workers do not accumulate skills. The resulting human-capital gain is about one-quarter of the full-equilibrium effect and just over one-third of the fixed-behavior effect. Reducing time in nonemployment therefore explains only

Figure 8: Employment across Learning Environments




*Notes:* Panel (a) plots the model-implied marginal distribution of employed workers over the learning environment $x$, pooling workers aged 20–59. France and the United States use their respective recovered policy wedges, while all common parameters remain at their baseline estimates. The vertical line marks the common center $x^T$ of the smoothed training-firm cutoff, and the shaded tails highlight the region $x > x^T$. Panel (b) plots the training-firm employment share in the model and data against log annual poaching relative to the hypothetical average-poaching economy. Panel (c) plots the country-specific difference in training-firm shares between NE hires aged 20–34 and 40–59. The empirical estimates come from one pooled regression of the training indicator on country-specific indicators for nine mutually exclusive age-by-transition cells, common fixed effects for the six ECHP employer-size categories and education, and country-year fixed effects. The nine cells cross ages 20–34, 35–39, and 40–59 with nonhires, NE hires, and JJ hires. Survey weights are normalized to sum to one within country, giving each country equal influence on the common effects. The model uses the corresponding population masses and the pooled specification, with country and common employer-size effects; education and year effects have no model counterpart. In Panel (c), each series is normalized to zero at its fitted value in the hypothetical average-poaching economy. The model applies the same training-firm cutoff in every country. Squares show the model and circles the data; two-letter country codes label the model dots and full country names label the data dots. No fitted lines are shown. *Source:* ECHP; author's calculations.

part of the dynamic gains, with faster movement toward better learning environments being key.

Panel (c) of Figure 8 provides empirical evidence consistent with young workers placing greater value on learning opportunities in high-poaching countries. Among workers hired from nonemployment, it plots the difference between young and older workers in the share joining training firms. Under the assumption that both age groups sample the same offer distribution, comparing them helps address cross-country differences in the prevalence of training vacancies. Consistent with the prediction of the model, this young–old gap rises with aggregate poaching in the data. However, the empirical relationship is very noisy.

Table 6: Mechanical Access and Behavioral Responses

| | Baseline | Fixed behavior | Nonemployment |
|---|---|---|---|
| Output per worker | 28.3 | 19.5 | 6.0 |
| Match productivity | 4.3 | 3.7 | 0.0 |
| Human capital | 21.5 | 14.0 | 5.4 |

*Notes:* All entries are changes from France to the United States, in log points. "Baseline" is the full-equilibrium change. "Fixed behavior" changes only $p$ and $\phi$, holding workers' and firms' decision rules at their baseline values; the gap to Baseline is the contribution of changes in behavior. "Nonemployment" is the $p$-only contact-rate experiment: it changes $p$ while holding the job-to-job contact rate $\phi p$ fixed. These counterfactual channels need not add because the contact-rate margins interact in equilibrium. *Source*: author's calculations.

### 5.4 The features of the data that imply substantial dynamic misallocation

The wage-growth advantage of training firms is central to the estimated importance of dynamic misallocation. As discussed in Section 4, this moment disciplines the dispersion of learning environments, $\sigma_x$. Higher wage growth at training firms points to differences in learning environments, making the firms at which workers spend their careers consequential for their skills.

To further highlight this, I counter-factually vary the wage-growth advantage of training firms relative to its true value in the data, and re-estimate the model to target this augmented moment. Specifically, I multiply the training-firm wage-growth differentials for workers aged 20–29, 30–39, and 40–49 by a common factor ranging from 0.25 to 1.75, preserving their relative age profile. For each factor, I re-estimate all 10 internally calibrated parameters using the same minimum-distance criterion, leaving the other empirical targets and externally calibrated parameters unchanged. Thus, the mean learning environment, productivity dispersion, and the other parameters can adjust along with $\sigma_x$. I recover the associated entry and vacancy costs and apply the same France-to-U.S. policy change as above: I set $\tau$ to its respective country values and move $(c_e, c_v, \phi)$ by the same proportional amounts around each re-estimated baseline as in the original calibration.

Table 7 reports the results. Raising the training wage-growth targets by 50 percent increases estimated learning dispersion from $\sigma_x = 0.00176$ to 0.00235. The human-capital effect of the policy change rises from 21.5 to 32.8 log points, and the output-per-worker effect rises from 28.3 to 41.8 log points. The comparison shows that the magnitude of dynamic misallocation is disciplined by the observed wage-growth advantage of training firms. Substantially smaller training gains would lead to smaller estimated differences in learning environments and smaller dynamic effects of policy, even after allowing the remaining parameters to adjust to the other features of the data. This interpretation makes it important that the measured wage gains reflect persistent and portable skills, which I examine next.

### 5.5 Supporting evidence

A first concern is that the empirical wage gains at training firms used to discipline dispersion in learning environments may be temporary rather than reflect skill accumulation. To assess this possibility, I use the eight-year ECHP panel to examine how wages evolve after employment at a training firm. Let $T_{it}$ indicate that worker $i$ is employed at a training firm in year $t$, and let $M_{it}^{\tau}$ indicate that the worker's current job began after year $t - \tau$. Conditional on $T_{i,t-\tau} = 1$, $M_{it}^{\tau}$ therefore identifies a worker who has changed employers since the training-firm year. Pooling countries, I estimate

$$\log w_{ict} \quad = \quad \xi_i + \alpha_{ct} + \sum_{\tau=-2}^{2} \beta_\tau T_{i,t-\tau} + \sum_{\tau=1}^{2} \left(\kappa_\tau M_{it}^{\tau} + \gamma_\tau T_{i,t-\tau} M_{it}^{\tau}\right) + \varepsilon_{ict}, \tag{15}$$

Table 7: Training Wage Growth and Dynamic Misallocation

| Training wage-growth target | 0.25× | 0.50× | 0.75× | Baseline | 1.25× | 1.50× | 1.75× |
|---|---|---|---|---|---|---|---|
| Estimated $\sigma_x$ (×100) | 0.045 | 0.095 | 0.125 | 0.176 | 0.195 | 0.235 | 0.278 |
| Effect of France-to-U.S. policy change (log points) | | | | | | | |
| Output per worker | 22.5 | 21.7 | 27.0 | 28.3 | 33.1 | 41.8 | 52.1 |
| Human capital | 16.2 | 15.7 | 20.3 | 21.5 | 25.6 | 32.8 | 41.7 |
| Match productivity | 4.5 | 4.1 | 4.4 | 4.3 | 4.6 | 5.5 | 6.0 |

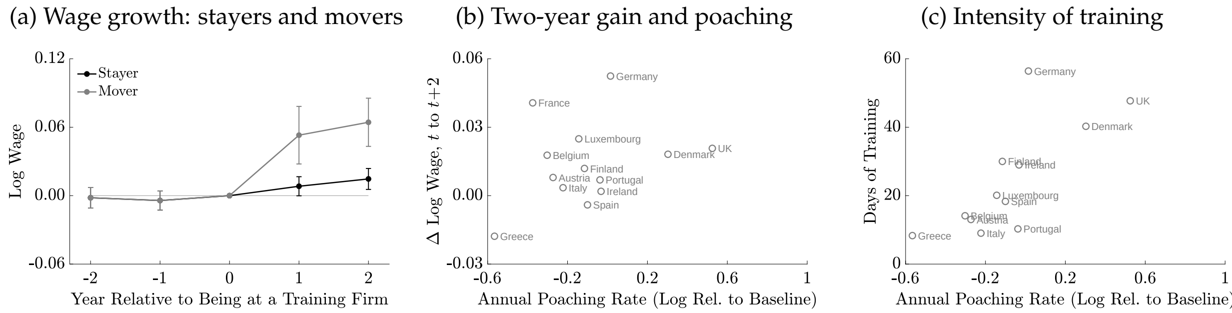
*Notes:* Column headings give the common multiplier applied to the three age-specific training-firm wage-growth targets; Baseline uses the empirical targets. All internally calibrated parameters are re-estimated in each column, while other empirical targets and externally calibrated parameters are unchanged. Entry and vacancy costs are recovered for each re-estimated baseline. The policy experiment uses the France and U.S. tax rates and applies the original country-specific proportional changes in $(c_e, c_v, \phi)$ around that baseline. Outcome entries are differences between the resulting U.S.-policy and France-policy economies, in log points. Human capital and match productivity are measured per employed worker, holding age-composition weights fixed. They are reported as separate outcomes and do not form an exact additive decomposition of output. These are sensitivity exercises with alternative calibration targets, not confidence intervals. Missing results are shown as dashes. *Source:* Author's calculations.

where $\xi_i$ and $\alpha_{ct}$ are worker and country-year fixed effects. I interact the training terms and their interactions with $M_{it}^{\tau}$ with centered log poaching, so the coefficients in equation (15) are evaluated at the hypothetical average-poaching country. Panel (a) plots the training-associated wage coefficients relative to the training-firm year. At horizons $\tau = 1, 2$, these are $\widehat{\beta}_\tau - \widehat{\beta}_0$ for workers who remain with the training employer and $\widehat{\beta}_\tau + \widehat{\gamma}_\tau - \widehat{\beta}_0$ for workers who change employers. Wage gains remain positive after two years for both groups and are larger among movers.

A second concern is that firms may choose training intensity within a match, whereas the model treats a firm's learning environment as a fixed attribute. If workers cannot compensate firms for received training through a lower wage, firms may invest more in training when workers are less likely to leave (Acemoglu, 1997; Acemoglu and Pischke, 1999). Fully assessing this mechanism would require introducing a contracting friction within the match and is beyond the scope of the paper. Nevertheless, I document here two patterns that appear at odds with this predictions.

Panel (b) plots $\widehat{\beta}_{2,c} - \widehat{\beta}_{0,c}$, the change in the training-associated log-wage coefficient from the training-firm year to two years later, against the country's log poaching rate. This is based on a re-estimation of equation (15) separately for each country $c$. The wage gain remains positive after two years in most countries and, if anything, rises with poaching, although the relationship is not statistically significant. If instead training firms had offered less training in high-poaching countries, one might have expected the wage gains at training firms to be larger in high-poaching countries. The ECHP records the number of days workers spent in employer-provided training during the preceding year. Panel (c) plots the average among workers at training firms against aggregate poaching. Workers report more days of training in high-poaching countries.

Figure 9: Training-Firm Outcomes and Poaching




*Notes:* Panel (a) reports the pooled event-study coefficients from equation (15), evaluated at the hypothetical average-poaching country and normalized to zero in the training-firm year. The lines distinguish workers who remain with the training employer from those who subsequently change employers; vertical bars are 95 percent confidence intervals. Panel (b) plots the country-specific change in the training-associated log-wage coefficient from year $t$ to year $t+2$, $\widehat{\beta}_{2,c} - \widehat{\beta}_{0,c}$, from regressions with worker and year fixed effects. Panel (c) plots average days of employer-provided training during the preceding year among workers at training firms. Panels (b) and (c) plot country observations against the log annual poaching rate without fitted lines. All empirical estimates use ECHP survey weights; no panel contains a model series. *Source:* ECHP; author's calculations.

## 5.6 Heterogeneity

The aggregate calibration imposes a common distribution of learning environments across workers. I allow the mean learning environment $\mu$ to differ separately by education and gender, holding the remaining calibration inputs, including the dispersion $\sigma_x$, at their baseline values. Within each partition, I estimate the two groups' wage profiles jointly with worker fixed effects and common country-year fixed effects. I then project each country-group-age wage cell on log poaching centered at its cross-country mean. The intercepts give the group-specific profiles in the hypothetical average-poaching economy.

Table 8 reports the resulting estimates of $\mu$ and the targeted life-cycle wage growth, measured as the change in log average wages between ages 20–29 and 50–59 in the adjusted wage profiles. The model closely matches each group's target. The steeper life-cycle wage profile of college graduates requires a higher mean learning environment than that of workers with less than college. The difference between the estimated means for men and women is smaller, consistent with the smaller gender difference in the wage-growth targets.

I next ask whether the four policy wedges account for cross-country differences in life-cycle wage growth within these groups. For each country and group, I retain the corresponding value of $\mu$ and vary $(\tau, c_e, c_v, \phi)$ jointly. Because changing $\mu$ changes the costs recovered in the baseline economy, I apply each country's proportional changes in $c_e$, $c_v$, and $\phi$ around the relevant group-specific baseline and set $\tau$ to its country value.

Table 9 applies the fully interacted specification in (18) to the data and the model-generated worker panels. Columns (1)–(2) use workers with less than a college degree as the omitted group, while columns (3)–(4) use women as the omitted group. The data columns reproduce columns (3) and (5) of Table 10. The model closely matches the group-specific age slopes and reproduces

Table 8: Mean Learning Environments by Education and Gender

| | Less than college | College or more | Women | Men |
|---|---|---|---|---|
| Mean learning environment ($\mu \times 100$) | -0.100 | 0.119 | -0.019 | 0.028 |
| Life-cycle wage growth, data | 19.8 | 42.7 | 27.7 | 32.2 |
| Life-cycle wage growth, model | 19.8 | 42.6 | 27.7 | 32.2 |

*Notes:* For each group, only the mean learning environment $\mu$ is recalibrated to its life-cycle wage growth in the hypothetical average-poaching economy. The parameter row reports $100\mu$, with $\mu$ measured at the model's monthly frequency. The targeted moment is the change in log average wages between ages 20–29 and 50–59 in the adjusted wage profiles; Data and Model report this change in log points. Within each education or gender partition, the data wage profiles are estimated jointly with worker and common country-year fixed effects. The data age effects are constrained to be flat from age 45 onward; model profiles remove mean wage growth over ages 45–59. Country-group-age cells are projected on log poaching centered at its cross-country mean; the intercepts define the average-poaching profiles. The remaining calibration inputs, including learning dispersion $\sigma_x$, are held at their baseline values. *Source:* ECHP, EU-SILC, PSID; model calculations.

the positive relationship between poaching and wage growth, although it predicts a somewhat steeper poaching–age gradient for the omitted groups. The triple interactions are positive in both the data and the model: wage growth responds more strongly to poaching for college graduates than for workers with less than college and for men than for women. The model generates a slightly smaller education differential and a more noticeably smaller gender differential than the data.

## 6 Conclusion

Labor-market policy affects productivity through both the allocation of existing skills and the jobs in which workers acquire new ones. In a model disciplined by panel data from 15 advanced economies, moving from the labor-market wedges inferred for France to those inferred for the United States raises output per worker by 28.3 log points, compared with 13.3 log points in a separately calibrated model with policy-invariant skill accumulation. Allowing workers' careers to shape their skills thus makes the estimated output effect 2.1 times as large.

Two extensions merit further study. First, an intensive margin of training would allow firms and workers to choose how much training takes place within a match, conditional on the firm's learning environment. This would capture how mobility affects both the returns to training and the time available to recover its costs.

Second, firms could choose their learning environments at a cost, rather than draw them from an exogenous distribution at entry. Firms might then choose better learning environments in high-poaching economies: workers with better prospects for using skills later may be willing to sacrifice more current pay for learning. This could raise firms' returns to providing such environments and amplify the dynamic effects quantified here. Whether it does so would depend on investment

Table 9: Policy Wedges and Life-Cycle Wage Growth by Education and Gender

| | Education | | Gender | |
|---|---|---|---|---|
| | Data | Model | Data | Model |
| Poaching × age | 0.010*** | 0.013 | 0.010*** | 0.014 |
| | (0.002) | | (0.002) | |
| Group × age | 0.014*** | 0.016 | 0.003*** | 0.003 |
| | (0.001) | | (0.001) | |
| Poaching × age × group | 0.005** | 0.004 | 0.004*** | 0.002 |
| | (0.002) | | (0.001) | |
| Observations | 1,854,494 | 3,123,821 | 1,854,494 | 3,097,603 |
| Years | 30 | 30 | 30 | 30 |
| Countries | 15 | 15 | 15 | 15 |

*Notes:* OLS estimates of (18). The education and gender data columns reproduce the fully interacted specifications in columns (3) and (5), respectively, of Table 10. The group indicator denotes college or more in columns (1)–(2) and male in columns (3)–(4); the omitted groups are less than college and female. Every regression uses log wages and includes the group–age slope, the log-poaching–age slope, and their triple interaction, absorbing individual, country×year, and restricted-age fixed effects. Age is capped at 45. The model columns apply the same estimator to simulated worker panels, with group-specific $\mu$ held fixed and $(\tau, c_e, c_v, \phi)$ varied jointly. Both groups use the same aggregate model poaching rate within a country. Model panels are thinned to the corresponding country–group distribution of observations per worker in the data. For the data columns, standard errors clustered by country are in parentheses, and * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Source:* ECHP 1993–2001; EU-SILC 2003–2024; PSID 1993–2023; model simulations; author's calculations.

costs and the shorter expected duration of employment relationships.

## References

**Acabbi, Edoardo Maria, Andrea Alati, and Luca Mazzone.** 2026. "The Collapse of Human Capital Ladders in Recessions." Working paper, available at SSRN 4068858.

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

**Acemoglu, Daron, and Jörn-Steffen Pischke.** 1999. "The Structure of Wages and Investment in General Training." *Journal of Political Economy*, 107(3): 539–572.

**Arellano-Bover, Jaime.** 2024. "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size." *Journal of Labor Economics*, 42(2): 549–589.

**Arellano-Bover, Jaime, and Fernando Saltiel.** 2026. "Differences in On-the-Job Learning across Firms." *Journal of Labor Economics*, 44(1): 149–188.

**Bentolila, Samuel, and Giuseppe Bertola.** 1990. "Firing Costs and Labour Demand: How Bad Is Eurosclerosis?" *Review of Economic Studies*, 57(3): 381–402.

**Bilal, Adrien, Niklas Engbom, Simon Mongey, and Giovanni L Violante.** 2022. "Firm and Worker Dynamics in a Frictional Labor Market." *Econometrica*, 90: 1425–1462.

**Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin.** 2006. "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica*, 74(2): 323–364.

**Djankov, Simeon, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer.** 2002. "The Regulation of Entry." *The Quarterly Journal of Economics*, 117(1): 1–37.

**Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman.** 2023. "Labor Market Dynamics and Development." *Quarterly Journal of Economics*, 138(4): 2287–2325.

**Gregory, Victoria.** 2026. "Firms as Learning Environments." Working paper.

**Guvenen, Fatih, Burhanettin Kuruscu, and Serdar Ozkan.** 2014. "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis." *Review of Economic Studies*, 81(2): 818–850.

**Haltiwanger, John C., Henry R. Hyatt, Lisa B. Kahn, and Erika McEntarfer.** 2018. "Cyclical Job Ladders by Firm Size and Firm Wage." *American Economic Journal: Macroeconomics*, 10(2): 52–85.

**Herkenhoff, Kyle, Jeremy Lise, Guido Menzio, and Gordon M. Phillips.** 2024. "Production and Learning in Teams." *Econometrica*, 92(2): 467–504.

**Hopenhayn, Hugo, and Richard Rogerson.** 1993. "Job Turnover and Policy Evaluation: A General Equilibrium Analysis." *Journal of Political Economy*, 101(5): 915–938.

**Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg.** 2021. "Learning from Coworkers." *Econometrica*, 89(2): 647–676.

**Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman.** 2018. "Life Cycle Wage Growth across Countries." *Journal of Political Economy*, 126(2): 797–849.

**Lazear, Edward P.** 2003. "Firm-Specific Human Capital: A Skill-Weights Approach." National Bureau of Economic Research Working Paper 9679.

**Ljungqvist, Lars, and Thomas J. Sargent.** 1998. "The European Unemployment Dilemma." *Journal of Political Economy*, 106(3): 514–550.

**Moscarini, Giuseppe, and Fabien Postel-Vinay.** 2025. "The job ladder: Inflation vs. reallocation." *Working Paper, Yale University and University College London*.

**Postel-Vinay, Fabien, and Jean-Marc Robin.** 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica*, 70(6): 2295–2350.

**Pries, Michael, and Richard Rogerson.** 2005. "Hiring Policies, Labor Market Institutions, and Labor Market Flows." *Journal of Political Economy*, 113(4): 811–839.

# A Data and Supporting Evidence

This appendix describes the sample and measurement choices underlying Section 2, compares the poaching measure with an independent measure of worker mobility, and presents additional evidence on institutions and career wage growth.

## A.1 Data Sources, Sample, and Measurement

The analysis combines worker-level panel data from the European Community Household Panel (ECHP), the European Union Statistics on Income and Living Conditions (EU-SILC), and the Panel Study of Income Dynamics (PSID). The survey observations span 1994–2001 for the ECHP, 2003–2024 for EU-SILC, and 1993–2023 for the PSID, with coverage varying by country and outcome. I harmonize employment histories, earnings, worker characteristics, and available employer characteristics across these sources to construct the sample of 15 advanced economies.

**ECHP.** The ECHP survey waves run from 1994 to 2001, with some retrospective information referring to the preceding year. Its release covers 15 Western European countries: Austria (from 1995), Belgium, Denmark, Finland (from 1996), France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden (from 1997), and the U.K. Five features of these files affect the analysis.

First, Germany participated in the original ECHP only from 1994 to 1996, but partially overlapping data from the German Socio-Economic Panel (SOEP) for 1994–2001 are included in the ECHP files. The German ECHP component suppresses the interview month, so its poaching rate cannot be computed. Because the ECHP and SOEP samples contain different respondents, I use both sources for all other outcomes to maximize sample size.

Second, Luxembourg participated in the original ECHP only from 1994 to 1996. For 1995–2001, the files include complementary data from the Panel Socio-Economique Liewen zu Lëtzebuerg (PSELL). PSELL does not report the start date of the current job for 1995–1997 and never records the survey month. Its poaching rate therefore cannot be computed, nor can Luxembourg's for 1997–2001. I use both the ECHP and PSELL samples for outcomes that can be constructed in both.

Third, the Netherlands is excluded because its national components are not available under the project's data-access permissions. Fourth, Sweden supplied a national cross-section in place of an ECHP panel; these observations cannot be used to estimate within-worker changes. Fifth, the U.K. participated in the original ECHP from 1994 to 1996, with overlapping British Household Panel Survey (BHPS) data for 1994–2001. I exclude the BHPS from the poaching calculation because the series constructed from it exhibits unusually large year-to-year fluctuations, while retaining it for other outcomes that can be measured consistently.

**EU-SILC.** EU-SILC succeeded the ECHP. The files used here cover 2004–2024, with several exceptions: Denmark and Luxembourg also provide data for 2003; Denmark and Ireland end in 2023; Greece starts in 2006; Germany starts in 2015 and is unavailable in 2020; and the U.K. covers 2005–2018. Eurostat releases separate cross-sectional and longitudinal files. Firm size and sector are available only in the cross-sectional files used here. I impute them in the longitudinal files by matching observations on country, year, gender, age, education, occupation, usual hours, and income. The variables required for the poaching measure are not consistently available after 2020.

**PSID.** The PSID is annual from 1993 to 1997 and biennial thereafter. I compute poaching using retrospective questions on employment status in each month of the calendar year before the survey. In the biennial sample, these calendars recover employment around the previous survey only from the early 2000s onward. The 1999 and 2001 surveys cover only the previous calendar year, so poaching cannot be computed for 1997–2000. Firm size is recorded only from 2005 onward, apart from additional measures in 1993 and 1999. Because earnings questions change over time, I harmonize earnings as total wage, salary, and self-employment income.

**Sample and variables.** The baseline sample consists of men and women aged 20–59. To limit the inclusion of workers still completing their initial education, the minimum age is 20 for respondents with less than an upper-secondary degree, 22 for those with an upper-secondary degree, and 24 for those with a tertiary degree.[13]

Employment includes private- and public-sector wage work and self-employment; nonemployment includes all other labor-market states. This broad definition avoids imposing distinctions that are not consistently available across the surveys.

I group education into less than upper secondary, upper secondary, and tertiary, and harmonize occupational classifications across surveys. Annual income includes wages and salaries, overtime and bonuses, and self-employment income for the previous calendar year. I approximate annual hours by multiplying months worked by four and by usual weekly hours, and divide annual income by this measure. I deflate wages with national consumer price indices and convert them to 2007 U.S. dollars using purchasing-power-parity exchange rates.

The surveys do not support a consistently measured monthly employer-to-employer (JJ) transition rate. I therefore use the annual poaching indicator defined in Section 2: whether an employed respondent joined the current employer in the preceding 12 months directly from another employer, without an intervening month of nonemployment. The poaching calculation uses observations from 1994–2020 in all three surveys, subject to the gaps described above; later observations remain available for other outcomes.

To make country averages comparable, I first calculate survey-weighted poaching rates within

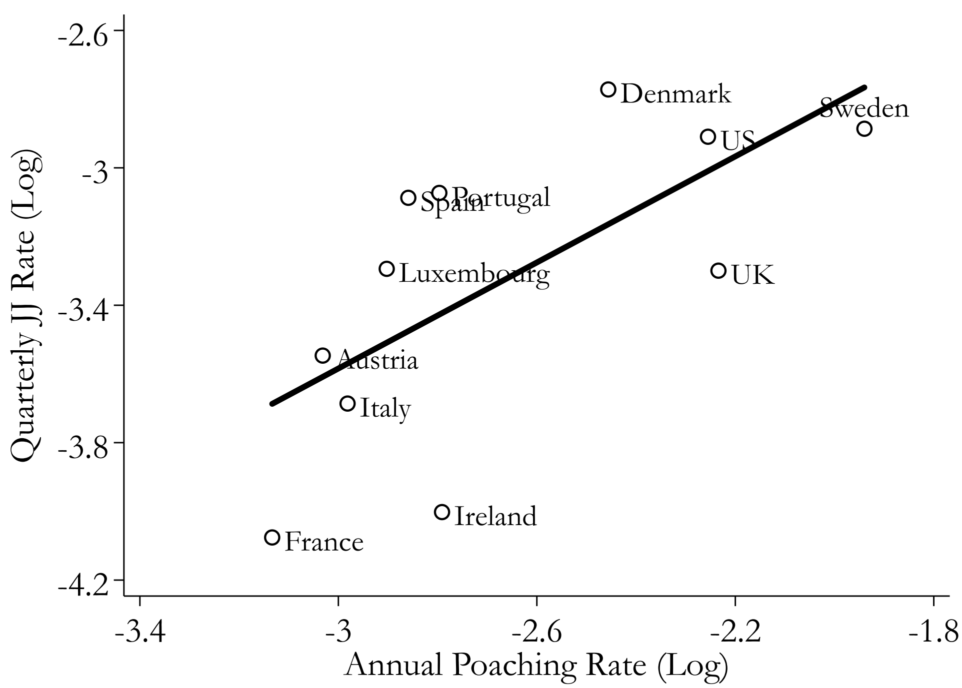
[13] The age at which respondents completed schooling is not consistently observed. These education-specific cutoffs are common across countries.

country-year cells for each of the four ten-year age groups. I then average across the age groups with equal weights. I construct a common 1994–2020 series for each country by linearly interpolating interior gaps and carrying the nearest observed rate backward or forward at the endpoints. For example, the observed U.S. poaching series ends in 2019, so its 2019 value is also used for 2020. The 1997 observation for Greece is treated as missing because of an unusually large fluctuation. A country's poaching rate is the mean of this completed annual series. This procedure limits the influence of differences in age composition and survey coverage; the resulting country rates include imputed years as well as observed years.

### A.2 Validation Against the Donovan-Lu-Schoellman Measure

Figure 10 compares the annual poaching rate with the quarterly JJ transition rate of Donovan, Lu and Schoellman (2023) for the 11 countries available in both data sets. The two measures differ in their underlying surveys, sampling frames, reference periods, and time horizons, so their levels need not coincide. Their strong positive association supports the use of the poaching rate to compare mobility across countries.

Figure 10: Annual Poaching Rate and Quarterly JJ Mobility

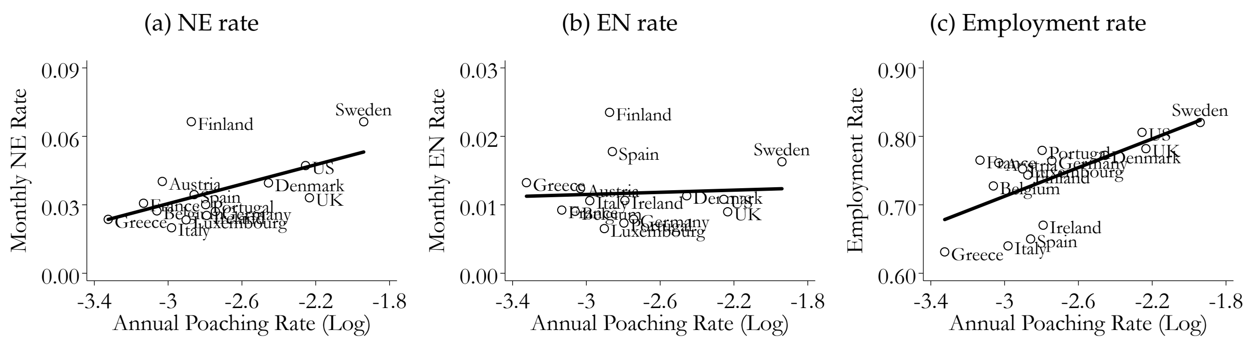



*Notes:* The horizontal axis is the log annual poaching rate constructed in Section 2. The vertical axis is the log of the quarterly JJ transition rate from Donovan, Lu and Schoellman (2023), who harmonize labor-force survey rotations across countries. Each marker is a country; the line is a linear fit. The sample is the 11 countries that appear in both data sets. The two measures are based on different time periods. *Source:* ECHP 1994–2001; EU-SILC 2003–2024 (poaching measure through 2020); PSID 1993–2023; Donovan, Lu and Schoellman (2023).

## A.3 Poaching and Other Worker Flows

Figure 11 relates poaching to transitions between employment and nonemployment and to the employment-to-population ratio. Countries with higher poaching have higher nonemployment-to-employment (NE) rates (panel (a)). Employment-to-nonemployment (EN) rates show no systematic relationship with poaching (panel (b)). Employment rates are also higher in high-poaching countries (panel (c)). These associations indicate that differences in poaching accompany broader differences in workers' access to employment.

Figure 11: Poaching, NE, EN, and Employment Rates in 15 Advanced Economies




*Notes:* The poaching rate is the share of currently employed workers who began working for their current employer in the past 12 months without an intervening month of nonemployment. The employed include private and public wage employees as well as the self-employed; the nonemployed are everyone else. *Source:* ECHP 1994–2001; EU-SILC 2003–2024 (poaching measure through 2020); PSID 1993–2023.

## A.4 Additional Policy Correlates

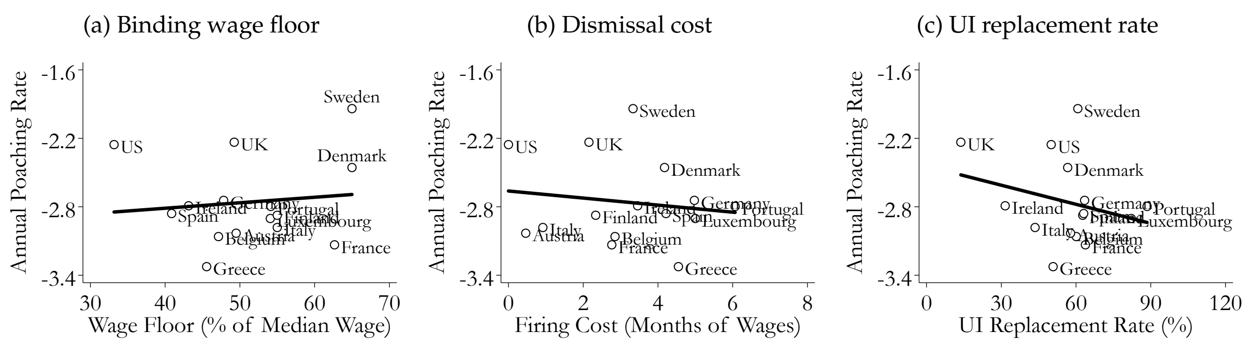
This subsection defines the institutional measures used in Section 2 and examines additional measures of entry costs, dismissal costs, and unemployment-insurance generosity. These country comparisons describe associations with poaching; they do not isolate the effects of individual policies.

**Policy measures.** Following Pries and Rogerson (2005), I consider wage floors, dismissal costs, unemployment insurance, labor taxes, and the cost of creating a firm. The binding wage floor is the maximum of the statutory minimum wage and prevailing union-negotiated wage floor, expressed relative to the median wage. Dismissal cost is the World Bank redundancy cost—statutory notice plus severance—in months of salary, averaged across workers with 1, 5, and 10 years of tenure. The UI measure is the SIED net replacement rate after 26 weeks of unemployment for a single worker earning the average manufacturing production wage. The OECD total labor tax is personal income tax plus employee and employer social-security contributions minus cash transfers, as a share of total labor cost for a single worker earning the average production wage, averaged over 2000–2020.

The World Bank startup-cost measure sums the official, professional, and publication fees required to register a standardized domestic limited-liability company, expressed as a percentage of annual GNI per capita. I combine the 1999 observation from Djankov, La Porta, Lopez-de-Silanes, and Shleifer (2002) with World Bank observations for 2003–2019 and linearly interpolate 2000–2002. Figure 2 uses country means over 1999–2019. The firm entry rate is the number of newly founded firms divided by the stock of firms. For European countries, I use Eurostat business-demography data on enterprise births as a share of active enterprises in the business economy; for the United States, I use the Census Bureau's Business Dynamics Statistics on firms of age zero as a share of all firms. Country entry rates are averages over the available years.

**Other institutions.** Figure 12 reports three additional institutional comparisons. The wage-floor measure is weakly positively associated with poaching (panel (a)); dismissal costs and replacement rates are negatively associated with it (panels (b) and (c)). These relationships are weaker than those for labor taxes and startup costs in the main text.

Figure 12: Poaching and Other Labor-Market Institutions




*Notes:* The binding wage floor is the maximum of the statutory and prevailing union-negotiated wage floors relative to the median wage. Dismissal cost is statutory notice plus severance in months of salary. The UI replacement rate is measured after 26 weeks of unemployment. Each marker is a country; lines are linear fits. *Source:* OECD; World Bank Employing Workers; SIED; ECHP; EU-SILC; PSID.

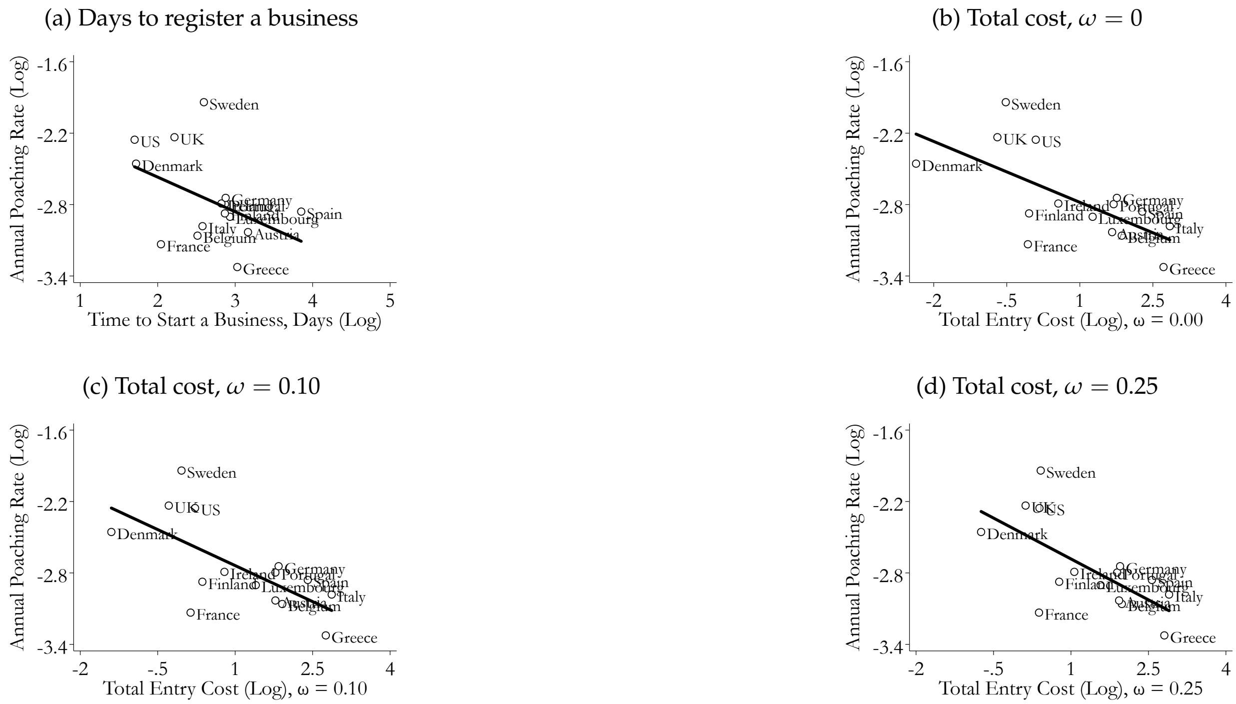
**The time cost of starting a business.** The World Bank *Starting a Business* cost indicator covers monetary fees, but registration also requires time. Panel (a) of Figure 13 relates the log poaching rate to the log number of calendar days required to register a business. To incorporate a possible income cost of this delay, I augment the fee measure as follows:

$$\text{total cost}_{ct} \quad = \quad \text{fees}_{ct} + 100\,\omega\,\frac{\text{days}_{ct}}{365}, \tag{16}$$

where fees and total cost are measured in percentage points of annual GNI per capita. The parameter $\omega$ scales forgone income relative to a full-time income loss valued at GNI per capita. Because registration days measure elapsed time rather than hours worked, I consider $\omega \in \{0, 0.10, 0.25\}$.

Panels (b)–(d) plot log poaching against the log augmented cost measure. The fitted slopes are approximately −0.17, −0.20, and −0.23, respectively. The negative association is therefore robust to these alternative valuations of registration time.

Figure 13: Poaching and the Time Cost of Starting a Business

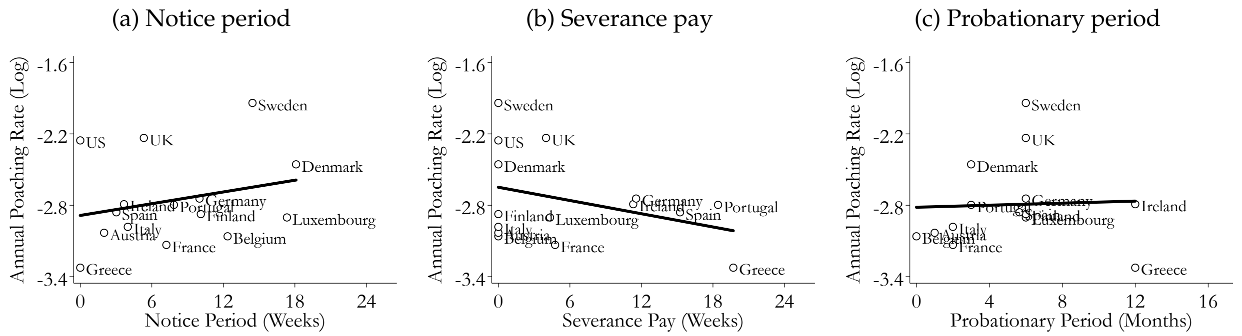



*Notes:* Panel (a) plots the log poaching rate against the log number of calendar days required to register a business, measured by the World Bank `IC.REG.DURS` indicator and averaged over 2003–2019. Panels (b)–(d) plot log poaching against the log cost measure defined in (16), for three values of $\omega$. Before taking logs, costs are expressed as a percentage of annual GNI per capita. Each marker is a country; lines are linear fits. *Source:* World Bank Doing Business; ECHP 1994–2001; EU-SILC 2003–2024 (poaching measure through 2020); PSID 1993–2023.

**Components of dismissal costs.** Figure 14 separates the notice and severance components of the World Bank redundancy-cost measure and also reports the maximum probationary period. Longer notice periods and larger severance payments increase the statutory cost of dismissal; a longer probationary period generally permits greater flexibility early in a match. None of the three measures displays a strong relationship with poaching. Examining these components separately thus gives a similar descriptive picture to the combined dismissal-cost measure.

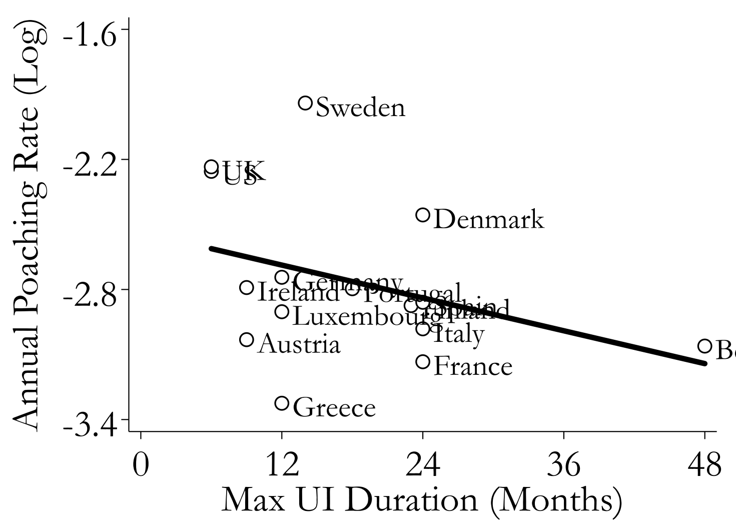
**Potential duration of unemployment benefits.** Unemployment-insurance generosity depends on benefit duration as well as the replacement rate. Figure 15 compares poaching with the maximum potential duration of benefits for a 40-year-old worker with a long contribution history. The association is weak, consistent with the limited relationship between poaching and replacement rates in Figure 12.

Figure 14: Poaching and the Components of Dismissal Costs




*Notes:* The notice period (panel (a)) and severance pay (panel (b)) are World Bank Employing Workers indicators for a standard redundancy dismissal, measured in weeks of pay and averaged across workers with 1, 5, and 10 years of tenure. The probationary period (panel (c)) is the maximum initial probation period, in months, reported by the same source. Country values are averaged over 2003–2019. Each marker is a country; lines are linear fits. *Source:* World Bank Employing Workers; ECHP 1994–2001; EU-SILC 2003–2024 (poaching measure through 2020); PSID 1993–2023.

Figure 15: Poaching and the Potential Duration of Unemployment Benefits




*Notes:* Potential duration is the maximum number of months for which benefits can be drawn by a 40-year-old worker with a long contribution history. The historical entitlement measure for Belgium is effectively indefinite and is represented by 48 months for plotting. Each marker is a country; the line is a linear fit. *Source:* SIED; OECD; ECHP 1994–2001; EU-SILC 2003–2024 (poaching measure through 2020); PSID 1993–2023.

## A.5 Wage-Growth Details and Robustness

### A.5.1 Life-Cycle Wage Profiles

I estimate country-specific life-cycle wage profiles by projecting log hourly wages on individual fixed effects, country-year fixed effects, and country-age effects:

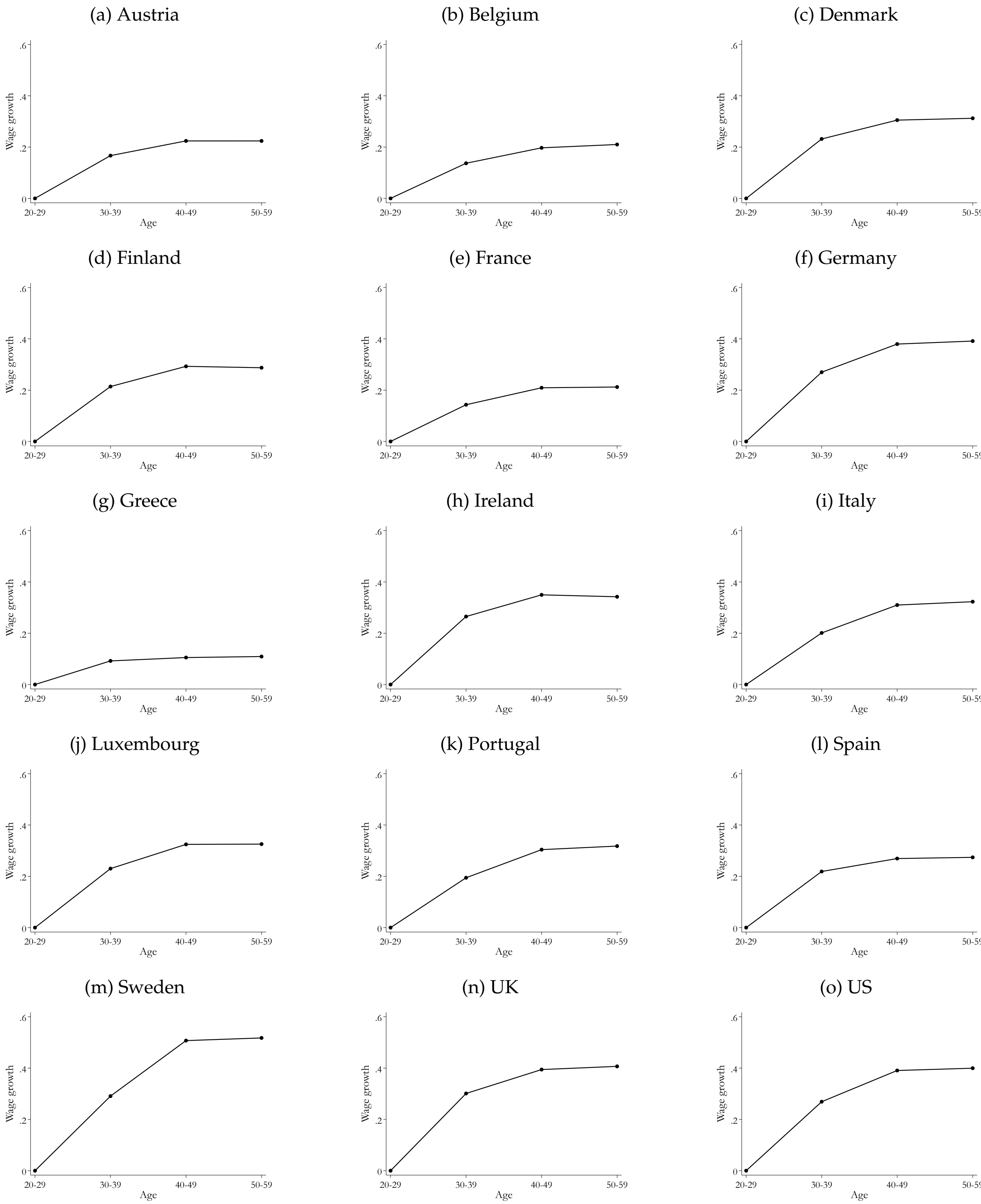
$$\log w_{it} \quad = \quad \alpha_i + \gamma_{ct} + \xi_{ca} + \varepsilon_{it}. \tag{17}$$

Individual effects absorb persistent worker characteristics, while country-year effects absorb wage movements common to workers in a country and year. To resolve the linear dependence between age and time within a worker's history, I pool ages 45–59 into a single age category. The profile is therefore constrained to be flat from age 45 onward. I estimate the regression with survey weights and average the estimated country-age effects within each ten-year age group. Career wage growth is the difference between the averages for ages 50–59 and 20–29. Figure 16 normalizes each profile to zero at ages 20–29. The estimated profiles rise most rapidly early in the career; their flatness at older ages is imposed rather than estimated.

For the comparison involving recent hires from nonemployment in Figure 3, I use the same full-sample regression. I add each worker's residual to the estimated country-age effect, equivalently retaining log wages net of the estimated worker and country-year effects. I then calculate age-group means of these adjusted wages among workers whose preceding-year record identifies a hire from nonemployment. The plotted measure is the difference between the oldest and youngest age groups in this selected sample. It is not a separate estimation of the age profile on recent hires.

A spell of nonemployment resets the bargaining position built through past outside offers in standard counteroffer models (Cahuc, Postel-Vinay and Robin, 2006). The positive association between poaching and adjusted wage growth among recent hires is therefore consistent with a role for accumulated skills. Selection into nonemployment and re-employment, and differences in the new jobs workers obtain, also affect this comparison. It does not provide a direct estimate of the human-capital channel.

Figure 16: Life-Cycle Wage Profile



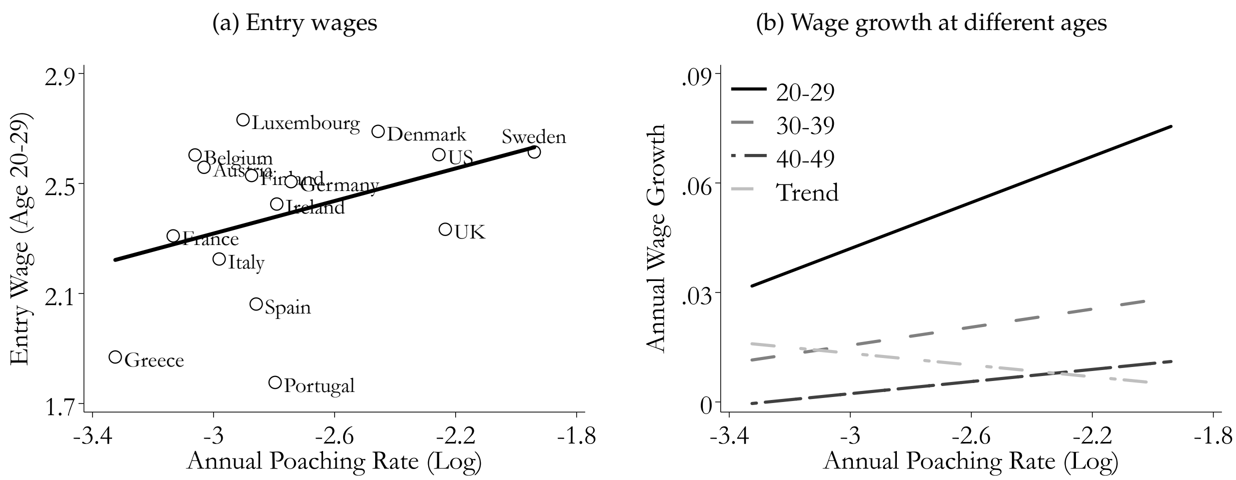

*Notes:* Survey-weighted country-age effects from (17), averaged within ten-year age groups and normalized to zero at ages 20–29. The regression includes individual and country-year fixed effects and constrains the age profile to be flat from age 45 onward. *Source:* ECHP 1994–2001; EU-SILC 2003–2024; PSID 1993–2023.

### A.5.2 Entry Wages and Wage Growth by Age

Panel (a) of Figure 17 relates mean PPP-adjusted log hourly wages at ages 20–29 to log poaching. Wages in this age group are higher in countries with more poaching. Thus, the steeper wage profiles in those countries do not simply reflect growth from a lower initial wage. These early-career wages need not be wages at the first job: the age group includes workers with different amounts of prior experience. The combination of higher wage levels and steeper profiles is consistent with the international evidence in Lagakos et al. (2018).

Panel (b) examines within-worker annual wage growth by age. For workers aged 20–29, 30–39, and 40–49, it subtracts the corresponding mean wage growth of workers aged 50–59. The oldest group's growth is also reported separately. Using that group as a benchmark removes a common wage trend under the assumption that its remaining life-cycle growth is small. The positive association with poaching is strongest among younger workers.[14] For workers aged 50–59, the estimated association is negative but statistically insignificant.

Figure 17: Additional Wage Evidence




*Notes:* Panel (a) plots mean PPP-adjusted log hourly wages among workers aged 20–29. Panel (b) shows fitted relationships between log poaching and annual within-worker log wage growth for workers aged 20–29, 30–39, and 40–49, each relative to workers aged 50–59. The line labeled "Trend" reports the relationship for the oldest group, which serves as the benchmark for broad wage trends. *Source:* ECHP 1994–2001; EU-SILC 2003–2024 (poaching measure through 2020); PSID 1993–2023.

[14] The annual wage-growth calculation conditions on observing valid wages in successive observations and therefore need not use the same sample as the life-cycle profile regression.

### A.5.3 Composition by Education and Gender

To assess whether the aggregate relationship reflects differences in workforce composition, I allow age slopes and their interactions with poaching to differ by education or gender:

$$\log w_{it} \quad = \quad \beta_0 \widetilde{p}_c \widetilde{a}_{it} + \mathbf{X}_i' \delta \, \widetilde{a}_{it} + \mathbf{X}_i' \eta \, \widetilde{p}_c \widetilde{a}_{it} + \alpha_i + \gamma_{ct} + \xi_{\widetilde{a}} + \varepsilon_{it}, \tag{18}$$

where $\mathbf{X}_i$ is empty in the baseline and contains an indicator for college completion or for being male in the other specifications. Here, $\widetilde{a}_{it} = \min\{a_{it}, 45\}$ and $\widetilde{p}_c$ is log poaching centered at its estimation-sample mean. The education- and gender-specific age slopes are therefore evaluated at the sample mean of log poaching. Individual, country-year, and restricted age fixed effects are included throughout, with survey weights and standard errors clustered by country.

Column (1) of Table 10 estimates a poaching–age coefficient of 0.012. A doubling of poaching is associated with approximately $0.012 \times \ln(2) \times 25 \simeq 0.21$ additional log wage growth over a 25-year age interval below the cap. Columns (2) and (4) add education- or gender-specific age slopes; the poaching coefficient remains close to the baseline. Columns (3) and (5) additionally allow the poaching gradient to differ across groups. The interaction is 0.005 larger for college graduates and 0.004 larger for men. Both differences are statistically significant under the reported inference procedure. These comparisons show that the aggregate association is not accounted for by education and gender composition alone, while also documenting heterogeneity in its magnitude.

Table 10: Life-Cycle Wage Growth and the Poaching Rate

| | | College | | Male | |
|---|---|---|---|---|---|
| Poaching × age | 0.012*** | 0.012*** | 0.010*** | 0.013*** | 0.010*** |
| | (0.001) | (0.002) | (0.002) | (0.001) | (0.002) |
| Control × age | | 0.016*** | 0.014*** | 0.005*** | 0.003*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) |
| Poaching × age × control | | | 0.005** | | 0.004*** |
| | | | (0.002) | | (0.001) |
| Observations | 1,854,494 | 1,854,494 | 1,854,494 | 1,854,494 | 1,854,494 |
| Years | 30 | 30 | 30 | 30 | 30 |
| Countries | 15 | 15 | 15 | 15 | 15 |

*Notes:* OLS estimates of (18). Each column is a separate regression of log hourly wages on the log poaching rate interacted with age, absorbing individual, country×year, and restricted age fixed effects. Age is capped at 45, and log poaching is centered at its estimation-sample mean. Column (1) is the pooled specification. Columns (2)–(3) add the age slope interacted with college, and column (3) further adds its triple interaction with the poaching rate; columns (4)–(5) repeat this with an indicator for male. The "Control" rows therefore refer to college in columns (2)–(3) and to male in columns (4)–(5). Standard errors clustered by country in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Source:* ECHP 1994–2001; EU-SILC 2003–2024; PSID 1993–2023.

### A.5.4 Job Stayers

Workers can obtain wage gains by changing employers. Greater wage growth in high-poaching countries could therefore reflect the greater frequency of JJ transitions. To assess this explanation, I regress annual log wage growth on an indicator for being younger than 40 interacted with log poaching, first without and then with a control for a JJ transition in the past 12 months:

$$\Delta \log w_{it} \quad = \quad \beta \, \mathbf{1}\{a_{it} < 40\} \ln(\text{poaching}_c) + \alpha \, \text{JJ}_{it} + \xi_{cy} + \xi_{ay} + \xi_{ey} + \xi_{gy} + \xi_{oy} + \varepsilon_{it}. \tag{19}$$

Here, $\Delta \log w_{it} = \log w_{i,t+1} - \log w_{it}$ uses wages observed one year apart. The fixed effects interact year with country, age, education, gender, and occupation. Country-year effects absorb the common wage-growth level within each country and year, so $\beta$ measures how the wage-growth difference between workers aged 20–39 and 40–59 varies with country-level poaching. The JJ indicator equals one for a recorded transition and zero otherwise. I use survey weights and cluster standard errors by country.

Column (1) of Table 1 estimates $\beta = 0.018$. A doubling of poaching is associated with $0.018 \times \ln(2) \simeq 0.0125$, or about 1.25 log points, greater annual wage growth for younger workers relative to older workers. In column (2), a JJ transition is associated with 4.6 log points of additional wage growth. Including this indicator changes the poaching coefficient only slightly, to 0.017.

Column (3) allows the interaction to differ by stayer status and replaces country-year effects with country-year-by-stayer effects. Stayers are workers recorded as having neither a JJ transition nor a hire from nonemployment. The complementary group includes recorded transitions and observations with missing transition information. The poaching coefficient is 0.015 for stayers, with a standard error of 0.004, and 0.007 for the complementary group, with a standard error of 0.005. The first estimate is statistically significant; the second is not. The evidence for stayers shows that the aggregate relationship extends beyond wage gains realized at employer changes. It does not distinguish skill accumulation from other sources of within-job wage growth.

### A.5.5 Sorting Across Occupations

I also allow the age slope and its interaction with poaching to differ across harmonized occupational groups:

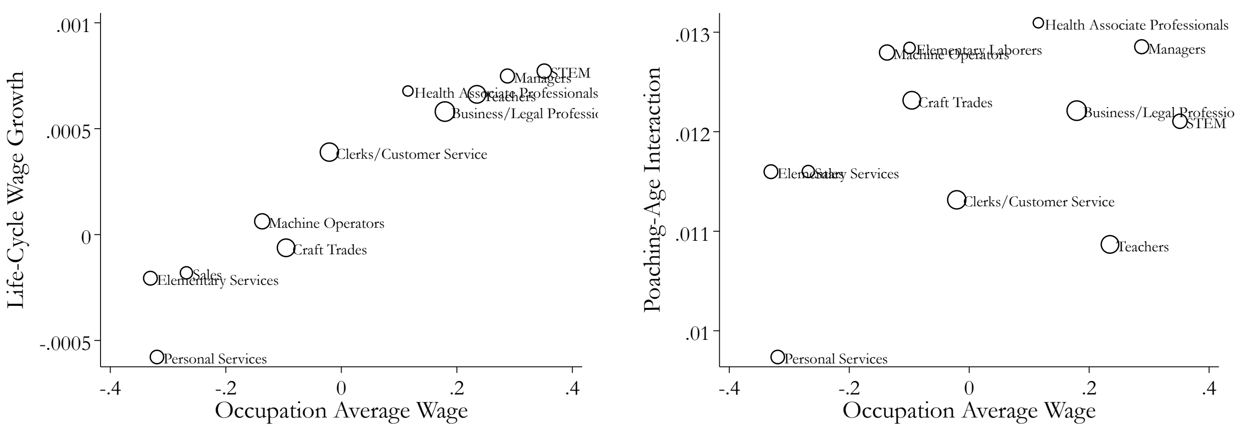
$$\log w_{it} \quad = \quad \sum_{o \in \mathcal{O}} (\alpha_o + \beta_o \widetilde{p}_c) \, \mathbf{1}\{o_{it} = o\} \widetilde{a}_{it} + \alpha_i + \gamma_{ct} + \xi_{\widetilde{a}} + \varepsilon_{it}. \tag{20}$$

Here, $\mathcal{O}$ is the set of occupational groups in the estimation sample. As in (18), age is capped at 45 and log poaching is centered at its estimation-sample mean. Figure 18 plots the occupation-specific age-slope coefficients, $\alpha_o$, and their interactions with poaching, $\beta_o$, against the occupation's mean log wage relative to its country-year mean. Higher-wage occupations tend to have steeper age slopes. The poaching interactions, however, show no systematic relationship with the occupation's average wage. This comparison describes heterogeneity across occupations; it does not by itself identify the contribution of occupational sorting to the aggregate relationship.

Figure 18: Life-Cycle Wage Growth and Poaching by Occupation

(a) Life-cycle wage growth

(b) Sensitivity to poaching




*Notes:* Each marker is a harmonized occupational group, plotted against its average log hourly wage net of the country-year mean; marker size is proportional to occupation employment. Both panels come from a single regression of log wages on occupation-specific age slopes and their interactions with centered log poaching, absorbing individual, country-year, and age fixed effects. Age is capped at 45. Panel (a) plots the occupation-specific age-slope coefficient; panel (b) plots its interaction with log poaching. The figure excludes skilled agricultural workers and coefficients omitted because of collinearity. *Source:* ECHP 1994–2001; EU-SILC 2003–2024 (poaching measure through 2020); PSID 1993–2023.

# B Model Appendix

This appendix gives the individual-value recursions, the worker-distribution equations, and proofs of the results in Section 3. Throughout, the reference job satisfies the hiring indifference condition $J(y_r, x_r, a) = N(a)$, while the quitting obstacle is $J \geq e^{-\epsilon}N$. These conditions coincide only when skills are fully portable.

## B.1 Individual values and counteroffer bargaining

Fix a finite contact rate $p$, an offer distribution $F$, and an admissible leisure schedule. Write $\mathcal{N}(a,h)$ for the value of nonemployment and $\mathcal{J}(y,x,a,h)$ for the candidate joint value. The individual worker and firm values are denoted by $\mathcal{V}^w(y,x,a,h,w)$ and $\mathcal{V}^f(y,x,a,h,w)$. The equations below apply when the current match continues and the worker's participation constraint is slack. The piece rate then remains constant between offers. At the participation boundary, it is adjusted upward as needed to maintain the worker's outside value. When the match reaches its stopping boundary, the worker receives $\mathcal{N}(a,h-\epsilon)$ and the incumbent receives zero. All values are zero at age $A$.

**The value of nonemployment.** The value of nonemployment $\mathcal{N}(a,h)$ satisfies

$$\rho\mathcal{N}(a,h) \quad = \quad e^{b(a)+h} + \mathcal{N}_a(a,h) + p\beta \int_{\mathcal{E}} \left(\mathcal{J}(y,x,a,h) - \mathcal{N}(a,h)\right)^+ dF(y,x)$$

with $\mathcal{N}(A,h) = 0$. As nonemployed, workers enjoy flow value of leisure $e^{b(a)+h}$, and age deterministically. They receive job offers at rate $p$ from the offer distribution $F$, which they accept if it provides positive surplus and receive a share $\beta$ of the surplus $\mathcal{J}(y,x,a,h) - \mathcal{N}(a,h)$.

**The value of a worker.** On the interior of the participation region, the worker value satisfies

$$\begin{aligned}
\rho\mathcal{V}^w(y,x,a,h,w) \quad &= \quad we^{y+h} + \mathcal{V}^w_a + \psi(a)x\mathcal{V}^w_h + \big(\delta(a)+\iota\big)\left(\mathcal{N}(a,h-\epsilon) - \mathcal{V}^w\right) \\
&+ \quad \phi p \int_{\mathcal{R}(y,x,a,h,w)} \Big(\mathcal{J}(\widetilde{y},\widetilde{x},a,h-\epsilon) + \beta\big(\mathcal{J} - \mathcal{J}(\widetilde{y},\widetilde{x},a,h-\epsilon)\big) - \mathcal{V}^w\Big) dF(\widetilde{y},\widetilde{x}) \\
&+ \quad \phi p \int_{\mathcal{M}(y,x,a,h)} \Big(\mathcal{J} + \beta\big(\mathcal{J}(\widetilde{y},\widetilde{x},a,h-\epsilon) - \mathcal{J}\big) - \mathcal{V}^w\Big) dF(\widetilde{y},\widetilde{x}), \qquad (21)
\end{aligned}$$

with $\mathcal{V}^w(y,x,A,h,w) = 0$, and where the move and renegotiation sets are

$$\begin{aligned}
\mathcal{M}(y,x,a,h) &\equiv \{(\tilde{y},\tilde{x}) : \mathcal{J}(\tilde{y},\tilde{x},a,h-\epsilon) > \mathcal{J}(y,x,a,h)\}, \\
\mathcal{R}(y,x,a,h,w) &\equiv \{(\tilde{y},\tilde{x}) : \mathcal{N}(a,h-\epsilon) < \mathcal{J}(\tilde{y},\tilde{x},a,h-\epsilon) \le \mathcal{J}(y,x,a,h), \\
&\qquad \mathcal{J}(\tilde{y},\tilde{x},a,h-\epsilon) + \beta\big(\mathcal{J}(y,x,a,h) - \mathcal{J}(\tilde{y},\tilde{x},a,h-\epsilon)\big) > \mathcal{V}^w(y,x,a,h,w)\}.
\end{aligned}$$

Workers get paid a piece rate of match output, grow older and accumulate skills in the match. They receive outside job offers at rate $\phi p$ from the offer distribution $F$, which they either discard, use to renegotiate their current pay, or accept.

**The value of a firm.** On the interior of the participation region, the firm value satisfies

$$\begin{aligned}
\rho\mathcal{V}^f(y,x,a,h,w) &= (1-w)e^{y+h} + \mathcal{V}^f_a + \psi(a)x\mathcal{V}^f_h - \big(\delta(a)+\iota\big)\mathcal{V}^f \\
&+ \phi p \int_{\mathcal{R}(y,x,a,h,w)} \Big((1-\beta)\big(\mathcal{J} - \mathcal{J}(\tilde{y},\tilde{x},a,h-\epsilon)\big) - \mathcal{V}^f\Big) dF(\tilde{y},\tilde{x}) \\
&- \phi p \int_{\mathcal{M}(y,x,a,h)} \mathcal{V}^f dF(\tilde{y},\tilde{x}),
\end{aligned} \tag{22}$$

with $\mathcal{V}^f(y,x,A,h,w) = 0$.

**The joint value.** Adding (21) and (22) yields in the participation region the joint value

$$\begin{aligned}
\rho\mathcal{J}(y,x,a,h) &= e^{y+h} + \partial_a\mathcal{J} + \psi(a)x\,\partial_h\mathcal{J} + \big(\delta(a)+\iota\big)\left(\mathcal{N}(a,h-\epsilon) - \mathcal{J}\right) \\
&+ \phi p\beta \int_{\mathcal{E}} \left(\mathcal{J}(\tilde{y},\tilde{x},a,h-\epsilon) - \mathcal{J}\right)^+ dF(\tilde{y},\tilde{x}).
\end{aligned}$$

The construction below verifies that this candidate joint value is the sum of the implemented worker and incumbent-firm values.

**Consistency of the joint and individual values.** First solve the nonemployment equation and the joint-value obstacle problem. Conditional on these values and the prescribed bargaining protocol, let $\mathcal{V}^w$ solve the worker recursion and participation conditions, and define

$$\mathcal{V}^f(y,x,a,h,w) \equiv \mathcal{J}(y,x,a,h) - \mathcal{V}^w(y,x,a,h,w).$$

Subtracting the worker equation from the joint equation verifies the incumbent-firm recursion directly. To see the cancellation at offers, abbreviate the current and outside joint values by $\mathcal{J}_i$ and $\mathcal{J}_o$. An accepted counteroffer gives the worker $(1-\beta)\mathcal{J}_o + \beta\mathcal{J}_i$ while leaving the incumbent $(1-\beta)(\mathcal{J}_i - \mathcal{J}_o)$. Their sum is $\mathcal{J}_i$, so renegotiation does not change the joint value. At a move, the worker receives $\mathcal{J}_i + \beta(\mathcal{J}_o - \mathcal{J}_i)$ and the old incumbent receives zero. Subtracting the worker's gain from the joint gain $\beta(\mathcal{J}_o - \mathcal{J}_i)$ therefore leaves exactly $-\mathcal{V}^f$. Exogenous separation likewise leaves the incumbent zero. The flow payoff obtained by subtraction is $(1-w)e^{y+h}$, and the derivative terms are those in (22).

The participation adjustments only alter the division of value within a continuing match. At its stopping boundary, $\mathcal{J} = \mathcal{N}(a, h-\epsilon)$ and worker participation implies $\mathcal{V}^f = 0$. The difference constructed above thus satisfies the same flow equation, continuation and reset rules, and terminal condition as the present value of incumbent profits. Finite-horizon backward uniqueness for this linear payoff recursion identifies the two. Consequently,

$$\mathcal{V}^w(y,x,a,h,w) + \mathcal{V}^f(y,x,a,h,w) \quad = \quad \mathcal{J}(y,x,a,h),$$

and the joint value is independent of the implementing piece rate. Because $w \leq 1$, the incumbent's realized profit flow is nonnegative, so its continuation value is nonnegative.

**Implementation of promises.** The maintained contract-spanning condition requires the bargaining promises to be attainable with feasible piece rates in $[0,1]$. Conditional on this condition and a well-defined participation-adjustment rule, the worker recursion is order preserving in the initial piece rate: current pay increases with $w$, discarded offers preserve the ordering, incumbent counteroffers replace the current promise by its maximum with a common offer-based promise, and a move implements a promise independent of the previous piece rate. Backward comparison therefore gives weak monotonicity of $\mathcal{V}^w$ in $w$.

Strict monotonicity holds at interior continuing contract states for $a < A$. For two distinct feasible rates $w_2 > w_1$ away from an instantaneous reset, continuity gives a nondegenerate interval during which the two rates remain distinct if no offer or separation occurs. Such an event has positive probability because all arrival rates are finite. Current pay is strictly higher for $w_2$ during that interval, and the subsequent value difference is nonnegative by comparison. Hence

$$\mathcal{V}^w(y,x,a,h,w_2) \quad > \quad \mathcal{V}^w(y,x,a,h,w_1).$$

Each promise in this attainable interior has a unique implementing piece rate. At $w = 1$, the incumbent earns no current profit, no improving incumbent counteroffer occurs, and a move or separation leaves it zero. Thus $\mathcal{V}^f = 0$ and $\mathcal{V}^w = \mathcal{J}$. At the participation boundary, the reset rule selects the smallest piece rate that implements $\mathcal{N}(a, h-\epsilon)$. No strict-monotonicity claim is needed at retirement or among nominal rates that are immediately reset to the same contract.

### B.2 Homogeneity in skills

*Proof of Lemma 1.* Fix $(p, F)$ and a bounded continuous leisure schedule. The coupled nonemployment and joint-value equations admit a finite-horizon auxiliary control representation. In the nonemployment state, the flow payoff is $e^{b(a)+h}$ and opportunities to switch to a match arrive at rate $\beta p$, with type drawn from $F$; the decision maker accepts only positive gains. In a match, the flow payoff is $e^{y+h}$, log human capital changes at rate $\psi(a)x$, exogenous separation to nonemployment occurs at rate $\delta(a) + \iota$, and opportunities to move to another match arrive at rate $\beta\phi p$. A move or separation reduces log human capital by $\epsilon$. The decision maker can also stop a match and take the nonemployment value after the same loss. These auxiliary arrival rates reproduce the positive-part terms of the joint and nonemployment equations; they are a representation of the value system, not the workers' physical contact rates.

The finite horizon, bounded firm support, bounded learning rate, and finite arrival rates bound these values by $Ce^h$ for a finite constant $C$. Downward jumps in $h$ do not violate this bound because $\epsilon \geq 0$. The dynamic programming recursion gives the unique bounded-growth solution of the coupled Bellman and obstacle system, with zero terminal values. Equivalently, this follows by backward comparison over sufficiently short time intervals and concatenation to the finite horizon.

Now increase initial log human capital from $h$ to $h + \log \lambda$, where $\lambda > 0$. Couple the offer types, jump times, and stopping and acceptance decisions in the two auxiliary problems. The two log-skill paths differ by $\log \lambda$ at every date, since their drift and portability losses coincide. Every flow payoff on the second path is therefore $\lambda$ times its counterpart on the first. Each admissible strategy in one problem has a corresponding strategy in the other. Taking suprema over strategies gives

$$\begin{aligned} \mathcal{N}(a, h + \log \lambda) &= \lambda \mathcal{N}(a, h), \\ \mathcal{J}(y, x, a, h + \log \lambda) &= \lambda \mathcal{J}(y, x, a, h). \end{aligned}$$

Setting the original skill level to zero and $\lambda = e^h$ proves (1). Dividing the raw equations by $e^h$ gives the scaled equations in the main text. Multiplication by $e^h > 0$ preserves the comparisons defining acceptance, continuation, mobility, and poachability, so these regions do not depend on $h$.

The same scaling applies to the individual values at a fixed feasible piece rate. Current wages and profits scale by $\lambda$; offer-based and participation promises scale by $\lambda$ because the joint and nonemployment values do; and the implementing piece-rate rule is unchanged. Backward comparison in the individual payoff recursions therefore yields

$$\begin{aligned} \mathcal{V}^w(y, x, a, h, w) &= e^h V^w(y, x, a, w), \\ \mathcal{V}^f(y, x, a, h, w) &= e^h V^f(y, x, a, w). \end{aligned}$$

□

**The hiring-reference normalization.** Let $d(a) \equiv \delta(a) + \iota$ and $k \equiv e^{-\epsilon}$. Under the normalization used in the main text, the reference job satisfies $J(y_r, x_r, a) = N(a)$. Define

$$I_H(a) \equiv \int_{\mathcal{E}} \left(J(y,x,a) - N(a)\right)^+ dF(y,x),$$
$$I_P(a) \equiv \int_{\mathcal{E}} \left(kJ(y,x,a) - N(a)\right)^+ dF(y,x).$$

The scaled nonemployment equation is

$$\rho N(a) = e^{b(a)} + N'(a) + p\beta I_H(a). \tag{23}$$

For $\epsilon > 0$ and $a < A$, positivity of $N$ implies $J(y_r, x_r, a) = N(a) > kN(a)$, so the reference match lies in the continuation region. Its joint-value equation can therefore be evaluated there. For $\epsilon = 0$, the normalization additionally imposes zero continuation residual at the reference job. In either case,

$$\left[\rho - \psi(a)x_r + d(a)(1-k)\right] N(a) = e^{y_r} + N'(a) + \phi p\beta I_P(a).$$

Combining this identity with (23) gives the residual leisure flow

$$e^{b(a)} = e^{y_r} + \left[\psi(a)x_r - d(a)(1-k)\right] N(a) + \phi p\beta I_P(a) - p\beta I_H(a).$$

The admissibility condition requires this expression to be continuous and strictly positive. The reference normalization does not hold the leisure schedule fixed when contact rates or the offer distribution change.

### B.3 Formal setup

The type space and its support bounds are defined in Section 3.1.

For a continuous candidate contact-value function $Q \geq 0$ with $\int Q^{1/\eta} d\Gamma > 0$, let $(p_Q, F_Q)$ be the firm-side objects in (13)–(14). Subscripts $Q$ identify the values, decisions, and distributions induced by this candidate. Write $k = e^{-\epsilon}$ and $\lambda(a) = \delta(a) + \iota$.

**Worker problem conditional on $(p_Q, F_Q)$.** The reference condition is $J_Q(y_r, x_r, a) = N_Q(a)$. Its reduced recursion is

$$[\rho + \lambda(a)(1-k) - \psi(a)x_r] N_Q(a) = e^{y_r} + N_Q'(a) + \phi p_Q \beta \int_{\mathcal{E}} (kJ_Q(y,x,a) - N_Q(a))^+ dF_Q(y,x),$$

with $N_Q(A) = 0$. The policy regions are those in (4) evaluated at $(J_Q, N_Q)$. For a pair $(J, N)$ that is differentiable in age on a continuation interval, define

$$\begin{aligned}\mathcal{H}_Q(J,N)(y,x,a) \quad &\equiv \quad (\rho + \delta(a) + \iota - \psi(a)x) J(y,x,a) - e^y - J_a(y,x,a) - \big(\delta(a)+\iota\big)e^{-\epsilon}N(a) \\ &\qquad -\phi p_Q \beta \int_{\mathcal{E}} \big(e^{-\epsilon} J(\tilde{y},\tilde{x},a) - J(y,x,a)\big)^+ \, dF_Q(\tilde{y},\tilde{x}).\end{aligned}$$

The match value solves the finite-horizon variational inequality (pointwise wherever differentiable and otherwise in the dynamic-programming sense)

$$\begin{aligned} J_Q(y,x,a) - e^{-\epsilon}N_Q(a) \quad &\geq \quad 0, \\ \mathcal{H}_Q(J_Q,N_Q)(y,x,a) \quad &\geq \quad 0, \\ \big(J_Q(y,x,a) - e^{-\epsilon}N_Q(a)\big)\mathcal{H}_Q(J_Q,N_Q)(y,x,a) \quad &= \quad 0. \end{aligned}$$

with $J_Q(y,x,A) = 0$. The reference condition is imposed together with its continuation equation. The implied leisure flow $B_Q = e^{b_Q}$ is given by (6). Its strict positivity is an admissibility restriction, not a consequence of value matching alone.

For finite $(p_Q, F_Q)$, the reduced reference equation and the match stopping problem have a unique finite-horizon solution. One construction fixes candidate continuation paths $(J, N)$ on a short terminal interval, integrates the reference equation, and solves the scalar match stopping problems with these paths in the offer gains and obstacle. The positive-part map is one-Lipschitz, and the optimal-stopping value is Lipschitz in flow and terminal rewards. In the norm for the pair of paths, the reference component has Lipschitz bound $C\ell$, whereas the match component has bound $k$ on the obstacle input plus $C\ell$ on the remaining inputs, where $\ell$ is the interval length. Weighting the reference component sufficiently heavily and then choosing $\ell$ small makes the combined map a contraction. Repeating on adjacent intervals establishes uniqueness on $[0, A]$. The same argument gives continuous dependence on $p_Q$, $F_Q$ in total variation, and $k$. This value argument does not by itself establish continuity of worker occupation measures at a moving stopping boundary. The existence proof below uses a support restriction that removes endogenous stopping from realized employment spells.

**Worker distributions.** Workers follow a piecewise-deterministic Markov process determined by $(p_Q, F_Q, J_Q, N_Q)$. A nonemployed worker accepts offers in $\mathcal{A}_Q(a)$. An employed worker at $(y, x)$ accumulates human capital at rate $\psi(a)x$, separates exogenously at rate $\delta(a) + \iota$, and accepts outside offers in $\mathcal{M}_Q(y, x, a)$. A match also terminates at the first hitting time of the stopping boundary $J_Q(y, x, a) = e^{-\epsilon}N_Q(a)$.

Let $\tilde{n}_Q(a, h)$ and $\tilde{g}_Q(y, x, a, h)$ denote the resulting stationary densities whenever they exist; otherwise, the equations below are understood in their measure-valued weak form. The nonemployment law and, on $\mathcal{C}_Q(a)$, the employment law are

$$\begin{aligned}
\partial_a \widetilde{n}_Q(a,h) &= -p_Q F_Q\big(\mathcal{A}_Q(a)\big)\widetilde{n}_Q(a,h) + \int\limits_{\mathcal{C}_Q(a)} \big(\delta(a)+\iota\big)\widetilde{g}_Q(y,x,a,h+\epsilon)dydx \\
&\quad + \int\limits_{\partial\mathcal{C}_Q(a)} \widetilde{g}_Q(y,x,a,h+\epsilon)v_{n,Q}(y,x,a)d\ell, \\
\partial_a \widetilde{g}_Q(y,x,a,h) &= -\psi(a)x\,\partial_h \widetilde{g}_Q(y,x,a,h) - \big(\delta(a)+\iota\big)\widetilde{g}_Q(y,x,a,h) \\
&\quad -\phi p_Q F_Q\big(\mathcal{M}_Q(y,x,a)\big)\widetilde{g}_Q(y,x,a,h) \\
&\quad +p_Q f_Q(y,x)\mathbb{1}_{\mathcal{A}_Q(a)}(y,x)\widetilde{n}_Q(a,h) \\
&\quad +\phi p_Q f_Q(y,x) \int\limits_{\mathcal{P}_Q(y,x,a)} \widetilde{g}_Q(\widetilde{y},\widetilde{x},a,h+\epsilon)d\widetilde{y}\,d\widetilde{x}.
\end{aligned}$$

Here $f_Q$ is the density of $F_Q$ and $v_{n,Q} \geq 0$ is the outward probability-flux speed at a regular moving stopping boundary. At irregular boundaries, the first-hitting construction defines the exit measure; the displayed surface integrals are its density notation. The boundary conditions are

$$\widetilde{n}_Q(0,h) = A^{-1}\delta_{\{h=0\}}, \qquad \widetilde{g}_Q(y,x,0,h) = 0.$$

Define the integrated measures

$$n_Q(a) \equiv \int \widetilde{n}_Q(a,h)dh, \qquad g_Q(y,x,a) \equiv \int \widetilde{g}_Q(y,x,a,h)dh,$$

and the efficiency-weighted measures

$$\widehat{n}_Q(a) \equiv \int e^h \widetilde{n}_Q(a,h)dh, \qquad \widehat{g}_Q(y,x,a) \equiv \int e^h \widetilde{g}_Q(y,x,a,h)dh.$$

Integrating the forward equations over $h$ gives the closed equations for $(n_Q, g_Q)$. Multiplying by $e^h$ and integrating gives

$$\begin{aligned}
\partial_a \widehat{n}_Q(a) &= -p_Q F_Q\big(\mathcal{A}_Q(a)\big)\widehat{n}_Q(a) + e^{-\epsilon}\int\limits_{\mathcal{C}_Q(a)} \big(\delta(a)+\iota\big)\widehat{g}_Q(y,x,a)dydx \\
&\quad +e^{-\epsilon}\int\limits_{\partial\mathcal{C}_Q(a)} \widehat{g}_Q(y,x,a)v_{n,Q}(y,x,a)d\ell, \\
\partial_a \widehat{g}_Q(y,x,a) &= \psi(a)x\widehat{g}_Q(y,x,a) - \Big(\delta(a)+\iota+\phi p_Q F_Q\big(\mathcal{M}_Q(y,x,a)\big)\Big)\,\widehat{g}_Q(y,x,a) \\
&\quad +p_Q f_Q(y,x)\mathbb{1}_{\mathcal{A}_Q(a)}(y,x)\widehat{n}_Q(a) + e^{-\epsilon}\phi p_Q f_Q(y,x)\int\limits_{\mathcal{P}_Q(y,x,a)} \widehat{g}_Q(\widetilde{y},\widetilde{x},a)d\widetilde{y}\,d\widetilde{x}.
\end{aligned}$$

Total effective search is

$$U_Q \equiv \int_0^A n_Q(a)da, \qquad S_Q \equiv U_Q + \phi(1-U_Q).$$

**Contact-value map.** The value of a worker contact to a type-$(y,x)$ firm is

$$\begin{aligned}\mathcal{T}_c(Q)(y,x) &= \frac{1-\beta}{S_Q}\int_0^A \widehat{n}_Q(a)\mathbb{1}_{\mathcal{A}_Q(a)}(y,x)\left(J_Q(y,x,a)-N_Q(a)\right)da \\ &\quad +\frac{(1-\beta)\phi}{S_Q}\int_0^A \int_{\mathcal{P}_Q(y,x,a)} \\ &\quad \left(e^{-\epsilon}J_Q(y,x,a)-J_Q(\widetilde{y},\widetilde{x},a)\right)\widehat{g}_Q(\widetilde{y},\widetilde{x},a)d\widetilde{y}\,d\widetilde{x}\,da. \end{aligned} \tag{24}$$

A stationary equilibrium satisfies $Q = \mathcal{T}_c(Q)$. The support restriction in Proposition 2 is used only for the equilibrium existence and uniqueness arguments. It is not imposed on the general worker problem or on the fixed-offer comparison in Lemma 4.

### B.4 Existence

This subsection uses the hiring-reference normalization $J(y_r, x_r, a) = N(a)$. Throughout, the reference lies weakly southwest of the support: $y \geq y_r$ and $x \geq x_r$ on $\mathcal{E}$. Write $k = e^{-\epsilon} \in (0,1]$, $\lambda(a) = \delta(a) + \iota$, and $z_r = e^{y_r}$. The primitive functions are continuous and bounded on $[0, A]$, $\psi(a) > 0$ for $a < A$, and $\Gamma$ has a bounded density on the compact type space. The matching and cost parameters satisfy the restrictions in the main text. The positivity of the leisure residual specified below and the spanning condition for wage promises are maintained separately; neither is implied merely by the southwest-reference restriction.

For a candidate $Q$, let $(p_Q, F_Q)$ be defined by (13)–(14). The worker equations are (5) and (3), with zero terminal values and stopping obstacle $J \geq kN$. The leisure residual $B_Q = e^{b_Q}$ is given by (6). For each fixed-point domain used below, maintain strict positivity of $B_Q$ and the contract-spanning condition for every candidate in that domain. Continuity then makes $b_Q(a) = \log B_Q(a)$ a bounded continuous schedule for each candidate. The restriction is imposed on the entire domain, so it does not alter the convex set used in the fixed-point argument.

**The worker block.** Use remaining lifetime $t = A - a$ and write a dot for its derivative. For finite $p$, equations (5) and (3) are an ODE in $\mathbb{R} \times C(\mathcal{E})$. Its right-hand side is globally Lipschitz in $(N, J)$, uniformly on the finite time interval, because the positive-part map is one-Lipschitz. Picard iteration therefore gives a unique continuous solution, continuously differentiable in time.

Nonnegativity follows by comparison: the derivative at any zero coordinate, when the remaining coordinates are nonnegative, is strictly positive. In particular, $N(a) > 0$ for $a < A$.

To verify that this continuation solution also solves the stopping problem, set $D(y,x,a) = J(y,x,a) - N(a)$. For fixed $a$, the function $v \mapsto \int (kJ(z,a) - v)^+ dF(z)$ is decreasing and one-Lipschitz. Subtracting the two equations therefore gives, in remaining lifetime,

$$\begin{aligned} \dot{D} &= e^y - z_r + \psi(a)(x - x_r)N \\ &- \big[\rho + \lambda(a) - \psi(a)x + \phi p\beta\theta(y,x,a)\big]D, \qquad 0 \le \theta \le 1. \end{aligned} \tag{25}$$

Variation of constants, $D(A) = 0$, and the southwest-reference restriction imply $J \geq N$. The inequality is strict for $a < A$ whenever $y > y_r$ or $x > x_r$. Indeed, the forcing in (25) is positive on a nondegenerate remaining-time interval in either case. At the reference type it is identically zero, so $J(y_r, x_r, a) = N(a)$, including when the reference is evaluated as a hypothetical type outside the support. Consequently $J \geq N \geq kN$. For $k < 1$ the stopping obstacle is strictly slack at every type before retirement. For $k = 1$ equality can occur only at the reference type, a $\Gamma$-null set. Every accepted job therefore remains in the continuation region until an exogenous separation, an accepted outside offer, or retirement. The endogenous stopping flux vanishes in the worker occupation equations. This conclusion avoids any appeal to continuity of first-hitting times.

Set

$$g_+ \;=\; \overline{\psi}\max\{\overline{x},0\}, \qquad \overline{H} = e^{g_+A}, \qquad \overline{J} = Ae^{\overline{y}}e^{g_+A}, \qquad \overline{Q} = (1-\beta)\overline{H}\,\overline{J}.$$

These bounds are independent of $p$ and $F$. To prove the value bound, let $W(t) = \max\{N(A-t), \sup_{\mathcal{E}} J(\cdot, A-t)\}$. At a match attaining the maximum, the outside-offer gain is zero, since $k \leq 1$, and $\lambda(kN - J) \leq 0$. At a maximizing nonemployment coordinate its outside-offer gain is also zero. The upper right derivative thus satisfies $D^+W(t) \leq e^{\overline{y}} + g_+W(t)$ and $W(0) = 0$. Scalar comparison gives $0 \leq N, J \leq \overline{J}$. Actual skills satisfy $e^h \leq \overline{H}$, since all employer separations weakly reduce $h$. Because the total search weight is $S = U + \phi(1-U)$,

$$s_- \;:=\; \min\{1,\phi\} \leq S \leq \max\{1,\phi\} =: s_+.$$

Using search weights in the contact-value formula gives $0 \leq \mathcal{T}_c(Q) \leq \overline{Q}$.

The same ODE comparison gives uniform continuity and sensitivity estimates needed below. Fix any finite contact-rate bound $P$. There are constants $C_P, d_P > 0$, depending only on fixed primitives and $P$, such that solutions with $p, p' \leq P$ satisfy

$$\begin{aligned} \|J(\cdot,a) - J'(\cdot,a)\|_\infty + |N(a) - N'(a)| &\leq C_P(A-a)\big(|p-p'| + \|F-F'\|_{\mathrm{TV}}\big), \qquad (26)\\ J(y',x,a) - J(y,x,a) &\geq d_P(A-a)(y'-y), \qquad y' > y. \qquad (27) \end{aligned}$$

The total-variation norm here is the total mass of the absolute signed measure. For the first inequality subtract the coupled ODEs. Their state Lipschitz constants are uniform for $p \le P$, and changes in the right-hand side due to $p$ and $F$ are bounded by a constant times $|p - p'| + \|F - F'\|_{\mathrm{TV}}$. Gronwall's inequality, with common zero initial values in remaining lifetime, proves the bound. For the second inequality, subtract the equations at $(y', x)$ and $(y, x)$. Their difference solves

$$\dot{E} \;=\; e^{y'} - e^{y} - \big[\rho + \lambda(a) - \psi(a)x + \phi p \beta \vartheta(a)\big]E, \qquad 0 \le \vartheta \le 1.$$

If $R_P \ge 0$ is an upper bound for the coefficient in brackets, variation of constants gives $E \ge e^{\underline{y}}(A - a)e^{-R_P A}(y' - y)$. One may evaluate these passive-type equations on the bounding rectangle, with offers still drawn from $\mathcal{E}$; consequently the inequality also holds when a vertical section of $\mathcal{E}$ is disconnected. Subtracting the equations at two arbitrary firm types similarly gives a common Lipschitz bound for $J$ in $(y, x)$, uniform over $p \le P$ and all $F$.

**An invariant domain.** Assume there is $\Delta > 0$ with

$$\Gamma(\mathcal{E}_\Delta) \;>\; 0, \qquad \mathcal{E}_\Delta = \{(y, x) \in \mathcal{E} : y \ge y_r + \Delta\}.$$

Under the hiring-reference normalization, $\Delta$ need only be positive; it need not exceed $\epsilon$. Set $d_\Delta = z_r(e^{\Delta} - 1) > 0$. For any $p \le P$, equation (25) implies

$$J(y, x, a) - N(a) \;\ge\; d_\Delta (A - a)e^{-R_P A} \quad \text{on } \mathcal{E}_\Delta.$$

Thus this surplus is at least $\sigma_P = d_\Delta A e^{-R_P A}/2$ for ages $a \in [0, A/2]$. Workers who have received no offer by age $a$ remain nonemployed with $h = 0$; their mass over that age interval is at least $m_P = \frac{1}{2}e^{-PA/2}$. Define

$$L_P \;=\; \tfrac{1}{2}\Gamma(\mathcal{E}_\Delta)\frac{(1-\beta)\sigma_P m_P}{s_+} > 0, \qquad \mathcal{X}_P = \{Q \in C(\mathcal{E}) : 0 \le Q \le \overline{Q},\ \int Q d\Gamma \ge L_P\}.$$

The inequalities $\sigma_P \le \overline{J}/2$, $m_P \le 1/2$, and $\overline{H}, s_+ \ge 1$ imply $L_P < \overline{Q}$. Hence $\mathcal{X}_P$ is nonempty, closed, bounded, and convex. Retaining only these never-contacted workers in the contact-value formula shows that $\int \mathcal{T}_c(Q) d\Gamma \ge 2L_P$ whenever the candidate's contact rate is at most $P$.

For a fixed $c > 0$, choose

$$P \;=\; \overline{p}(c) := \left[\frac{\eta}{(1+\eta)(\rho+\iota)c}\overline{Q}^{(1+\eta)/\eta}\right]^{\kappa}, \qquad \kappa = \frac{\eta\alpha}{(1-\alpha)(1+\eta)}.$$

Then every $Q \in \mathcal{X}_P$ generates $0 < p_Q \le P$. The firm map is continuous in the uniform norm, with $F_{Q_n} \to F_Q$ in total variation when $Q_n \to Q$ uniformly. To verify the denominator bound for every $\eta > 0$, put $r = 1/\eta$. Jensen's inequality gives $\int Q^r d\Gamma \geq L_P^r$ when $r \geq 1$; when $0 < r < 1$, $\int Q^r d\Gamma \geq \overline{Q}^{r-1} L_P$. Uniform continuity of the power function on $[0, \overline{Q}]$ proves the claim. These same bounds and the bounded primitive density imply a uniform upper bound on every offer density $f_Q$.

**Continuity of worker occupations.** There is no endogenous stopping under the restrictions above. The unweighted and efficiency-weighted forward systems therefore have only bounded jump rates, the bounded drift term $\psi(a)x\widehat{g}$, and inflows with densities proportional to $f_Q$. They determine unique finite occupation measures, with type densities $g_Q$ and $\widehat{g}_Q$ for employed workers. Human capital need not have a density. Scalar comparison in the inflow equations gives, uniformly over the domain,

$$g_Q(y,x,a) + \widehat{g}_Q(y,x,a) \quad \leq \quad C_P f_Q(y,x).$$

For example, the weighted employment inflow is at most $(1+\phi)P\overline{H} f_Q/A$ and its positive drift is at most $g_+ \widehat{g}_Q$; dropping nonpositive outflows and integrating from zero proves the weighted bound. The unweighted bound follows identically.

For $a < A$, (27) and the offer-density bound imply the band estimate

$$F_Q\{(y,x) : |kJ_Q(y,x,a) - v| \leq \zeta\} \quad \leq \quad C_P \frac{\zeta}{A-a}, \qquad \zeta \geq 0.$$

Indeed, each vertical section has diameter at most $2\zeta/[kd_P(A-a)]$; integrate over the bounded range of $x$. The same estimate holds without $k$. Combining this estimate with (26) bounds the offer mass whose acceptance or poaching decision changes by a constant times $|p_Q - p_{Q'}| + \|F_Q - F_{Q'}\|_{\mathrm{TV}}$. The corresponding bound holds for the set of incumbent types a given destination can poach. At retirement the decision sets are empty for every candidate. The density-domination bound converts offer-mass bounds into incumbent-mass bounds.

Subtracting the two forward systems and taking absolute values now gives a Gronwall inequality in the sum of the scalar nonemployment differences and the $L^1(\mathcal{E})$ employment-density differences, for both weighted and unweighted systems. It follows that their age-integrated difference and $|S_Q - S_{Q'}|$ tend to zero when $|p_Q - p_{Q'}| + \|F_Q - F_{Q'}\|_{\mathrm{TV}} \to 0$. This establishes continuity of worker occupations on the invariant domain.

*Proof of Proposition 2.* The preceding bounds show that $\mathcal{T}_c$ maps $\mathcal{X}_P$ into itself. Write its two surplus integrands as $(J_Q - N_Q)^+$ and $(kJ_Q(y,x,a) - J_Q(\widetilde{y}, \widetilde{x}, a))^+$. Uniform convergence of the values, $L^1$ convergence of the weighted occupation densities, and $S_Q \geq s_- > 0$ imply continuity of $\mathcal{T}_c$ in the uniform norm. The common type-Lipschitz bound on $J_Q$ gives a common type-Lipschitz bound on $\mathcal{T}_c(Q)$; the occupation weights are uniformly bounded and the positive-part map is one-Lipschitz. Arzelà–Ascoli therefore makes $\mathcal{T}_c(\mathcal{X}_P)$ relatively compact in $C(\mathcal{E})$. Schauder's theorem yields a fixed point. At that fixed point, (6) recovers the positive leisure flow and the original nonemployment equation. The vacancy and matching conditions recover the firm-side scale, and the maintained spanning condition implements the wage promises. Thus a stationary equilibrium exists under the stated southwest-reference, activity, and admissibility restrictions. This proof applies to every $\eta > 0$. $\square$

### B.5 Uniqueness when contacts are infrequent

*Proof of Proposition 3.* Fix all primitives other than $c$ and maintain the positivity and spanning conditions. Suppose $\eta \leq 1$. Choose any fixed finite $P_0 > 0$ and use the common domain $\mathcal{X}_{P_0}$ for all $c$ with $\overline{p}(c) \leq P_0$. The activity and upper bounds show that every admissible equilibrium belongs to this common domain. All constants below are uniform over these $c$.

Because $r = 1/\eta \geq 1$, the power function $u^r$ is Lipschitz on $[0, \overline{Q}]$ and its integral is bounded away from zero on $\mathcal{X}_{P_0}$. The same holds for the integral of $u^{1+r}$. The normalized density formula and the mean-value theorem applied to the contact-rate formula therefore give

$$\|F_Q - F_{Q'}\|_{\mathrm{TV}} \quad \leq \quad C\|Q - Q'\|_\infty, \qquad |p_Q - p_{Q'}| \leq C\overline{p}(c)\|Q - Q'\|_\infty. \tag{28}$$

The factor $\overline{p}(c)$ in the second inequality follows by writing $p_Q = \overline{p}(c)[\int Q^{1+r} d\Gamma / \overline{Q}^{1+r}]^\kappa$; the bracket ranges in a fixed compact subset of $(0, 1]$.

Set $\Delta = \|Q - Q'\|_\infty$. In the worker equations, a difference in offer distributions is multiplied by the contact rate. Subtracting those equations, using (28), and applying Gronwall's inequality therefore strengthens the earlier continuity bound to

$$\|J_Q(\cdot, a) - J_{Q'}(\cdot, a)\|_\infty + |N_Q(a) - N_{Q'}(a)| \quad \leq \quad C(A - a)\overline{p}(c)\Delta.$$

The band estimate consequently bounds changes in acceptance and poaching regions by $C\overline{p}(c)\Delta$. In the forward equations, changes in the contact-rate/density product are bounded by

$$\|p_Q f_Q - p_{Q'} f_{Q'}\|_1 \quad \leq \quad |p_Q - p_{Q'}| + p_{Q'}\|f_Q - f_{Q'}\|_1 \leq C\overline{p}(c)\Delta.$$

Changes in decision indicators are also multiplied by contact rates and are controlled by the band and density-domination bounds. Their contribution is at most $C\overline{p}(c)^2\Delta$, and hence at most $CP_0\overline{p}(c)\Delta$. The resulting forward-system inequality is $Z(a) \leq C \int_0^a Z(s) ds + C\overline{p}(c)\Delta$, where $Z$ includes both weighted and unweighted differences. Gronwall gives

$$|S_Q - S_{Q'}| + \int_0^A \left(|\widehat{n}_Q - \widehat{n}_{Q'}| + \|\widehat{g}_Q - \widehat{g}_{Q'}\|_1\right) da \quad \leq \quad C\overline{p}(c)\Delta.$$

Finally subtract the two contact-value formulas. The surplus functions are bounded and Lipschitz in the values, and $S \geq s_- > 0$. The value and occupation bounds prove

$$\|\mathcal{T}_c(Q) - \mathcal{T}_c(Q')\|_\infty \leq C_T \overline{p}(c) \|Q - Q'\|_\infty.$$

Since $\overline{p}(c)$ tends to zero as $c$ tends to infinity, there exists $c_* > 0$ such that $\overline{p}(c) \leq P_0$ and $C_T \overline{p}(c) < 1$ for $c > c_*$. For these cost composites, $\mathcal{T}_c$ is a contraction on the common invariant domain. Existence was established above, and every admissible equilibrium lies in that domain, so the equilibrium is globally unique. The threshold holds with the other primitives fixed; this argument does not establish monotone comparative statics of the uniqueness region with respect to every parameter entering the constants. It also does not cover $\eta > 1$.

□

## B.6 Policy equivalence

*Proof of Lemma 3.* Write $\theta \equiv 1 - \tau > 0$. A worker and firm receive $w\theta e^{y+h}$ and $(1-w)\theta e^{y+h}$, respectively. Divide every monetary flow, value, and bargaining promise by $\theta$. The normalized match flow is $e^{y+h}$, and homogeneity of the positive-part operator preserves every surplus comparison and every participation constraint.

The leisure flow scales in the same way under the hiring-reference normalization. In monetary units, this normalization sets $J^\tau(y_r, x_r, a) = N^\tau(a)$ and implies

$$\begin{aligned} B^\tau(a) &= \theta e^{y_r} + \left[\psi(a)x_r - d(a)(1-k)\right] N^\tau(a) \\ &\quad + \phi p \beta \int_{\mathcal{E}} \left(kJ^\tau - N^\tau\right)^+ dF - p\beta \int_{\mathcal{E}} \left(J^\tau - N^\tau\right)^+ dF. \end{aligned}$$

Dividing by $\theta$ yields exactly the normalized residual leisure flow derived above. Thus the normalized worker equations coincide with those of the untaxed problem, conditional on the same $(p, F)$. This result uses the recalibration of leisure required by the reference-job condition; it is not a statement about a tax change with the monetary leisure schedule held fixed.

If $Q$ is the normalized value of a contact, its monetary value is $\theta Q$. For any vacancy choice,

$$\theta^{-1}\left(qv\theta Q - \frac{c_v}{1+\eta} v^{1+\eta}\right) = qvQ - \frac{\widetilde{c}_v}{1+\eta} v^{1+\eta}, \qquad \widetilde{c}_v \equiv \frac{c_v}{\theta}.$$

Dividing the entry cost by the same factor gives $\widetilde{c}_e = c_e/\theta$. Hence the normalized first-order condition and free-entry equation have the untaxed form with costs $(\widetilde{c}_v, \widetilde{c}_e)$.

To derive the composite explicitly, let $s \equiv 1/\eta$ and $r \equiv 1 + s$. The first-order condition gives

$v = (qQ/\widetilde{c}_v)^s$, and free entry gives

$$\widetilde{c}_e \;=\; \frac{\eta}{(1+\eta)(\rho+\iota)}\widetilde{c}_v^{-s} q^r \int_{\mathcal{E}} Q^r d\Gamma.$$

Substitute $q = \chi^{1/\alpha} p^{-(1-\alpha)/\alpha}$ and rearrange:

$$\begin{aligned}
p^{1/\kappa} &= \frac{\eta}{(1+\eta)(\rho+\iota)} \frac{1}{c} \int_{\mathcal{E}} Q^r d\Gamma, \\
c &\equiv \widetilde{c}_e \widetilde{c}_v^{1/\eta} \chi^{-(1+\eta)/(\alpha\eta)} = (1-\tau)^{-(1+\eta)/\eta} c_e c_v^{1/\eta} \chi^{-(1+\eta)/(\alpha\eta)}.
\end{aligned}$$

This is (12), with $\kappa = \eta\alpha/[(1-\alpha)(1+\eta)] > 0$. $\square$

**Entry accounting and firm size.** The factor $(\rho+\iota)^{-1}$ in the entry value discounts the time until future recruitment opportunities. The contact value $Q$ already discounts the profit stream after a hire, including the possibility that the firm subsequently exits. Conditional on its survival to date $t$, a type-$(y,x)$ firm faces the same stationary recruiting opportunities as a newly created firm. Writing $\widetilde{C}_v(v) = \widetilde{c}_v v^{1+\eta}/(1+\eta)$, its normalized value at birth is therefore

$$\begin{aligned}
V^e(y,x) &= \int_0^\infty e^{-(\rho+\iota)t} \left[ qv(y,x)Q(y,x) - \widetilde{C}_v(v(y,x)) \right] dt \\
&= \frac{qv(y,x)Q(y,x) - \widetilde{C}_v(v(y,x))}{\rho+\iota}.
\end{aligned}$$

The two appearances of the exit hazard concern survival before and after a hire, respectively; they do not charge for the same interval twice. Under the Poisson exit process, the residual lifetime conditional on survival has the same distribution at every recruitment date. If $\iota > 0$, stationary gross entry equals $\iota M$.

For completeness, set

$$I_s \;\equiv\; \int_{\mathcal{E}} Q^{1/\eta} d\Gamma, \qquad I_r \equiv \int_{\mathcal{E}} Q^{(1+\eta)/\eta} d\Gamma, \qquad K_v \equiv (q/\widetilde{c}_v)^{1/\eta}.$$

Vacancy aggregation and matching imply $V = MK_v I_s$ and $qV = pS$. Free entry implies

$$qK_v I_r \;=\; \widetilde{c}_e \frac{(1+\eta)(\rho+\iota)}{\eta}.$$

Eliminating $qK_v$ gives the average employment of all incumbent firms,

$$\frac{1-U}{M} \;=\; \widetilde{c}_e \frac{1-U}{S} \frac{(1+\eta)(\rho+\iota)}{\eta p} \frac{I_s}{I_r}.$$

This average includes firms with no employees. Under the model's convention of computing size moments from type-level mean employment, let $\pi_+$ be the $\Gamma$-mass of types with positive employment. Average size conditional on those active types is $(1-U)/(M\pi_+)$, so the displayed expression must be divided by $\pi_+$ when compared with that conditional measure. The worker allocation fixes $\pi_+$, while $M$ scales inversely with the effective entry cost at a fixed cost composite.

*Proof of Proposition 1.* Hold fixed the nonpolicy primitives, including employed-search efficiency, the firm-type distribution, and the reference-job normalization. Consider two configurations of $(\tau, c_v, c_e, \chi)$ with the same $c$. For any candidate normalized contact-value function $Q$, equations (13) and (14) generate the same contact rate $p$ and offer distribution $F$. The normalized worker problem, its policy regions, and its forward law are then identical. They therefore return the same contact-value function through (24). The two configurations have the same normalized fixed-point operator and hence the same equilibrium correspondence for $(Q, p, F)$ and the associated worker allocation. No uniqueness assumption is required for this correspondence result. Firm mass, vacancy levels, firm size, and conversion from normalized to monetary values depend on the separate policy primitives and need not coincide. □

### B.7 Contact-rate comparative statics

Hold all primitives other than the cost composite fixed. Call a positive equilibrium locally regular if the equilibrium conditions excluding free entry admit a continuously differentiable local branch $p \mapsto Q_p$, with

$$I(p) \quad \equiv \quad \int_{\mathcal{E}} Q_p(y,x)^{(1+\eta)/\eta} d\Gamma(y,x) > 0.$$

The reduced free-entry map on that branch is

$$G(p;c) \quad \equiv \quad \left[\frac{\eta}{(1+\eta)(\rho+\iota)} \frac{I(p)}{c}\right]^{\kappa}.$$

Call the equilibrium linearly stable under the adjustment rule $\dot{p} = G(p;c) - p$ if $G_p(p;c) < 1$. This specifies the stability concept used in the proposition; it makes no claim about stability under other adjustment rules.

*Proof of Proposition 4.* At an equilibrium, $p = G(p;c)$. Once $p$ is fixed, the remaining normalized equilibrium conditions do not contain $c$. In particular, the reference leisure schedule is recomputed from $(p, Q_p)$ under the same normalization on both sides of the comparison. Hence

$$\begin{aligned} G_c(p;c) &= -\frac{\kappa}{c} G(p;c) = -\frac{\kappa p}{c}, \\ G_p(p;c) &= \kappa p \frac{I'(p)}{I(p)}. \end{aligned}$$

Local linear stability implies $1 - G_p > 0$, so the implicit function theorem gives

$$\frac{dp}{dc} = \frac{G_c(p;c)}{1 - G_p(p;c)} = -\frac{\kappa p/c}{1 - \kappa p I'(p)/I(p)} < 0.$$

This is a local result along the specified equilibrium branch. It does not require the contact-value response $I'(p)$ to have a particular sign. □

### B.8 Career indifference curves

*Proof of Lemma 2.* Write $k = e^{-\epsilon}$ and $d(a) = \delta(a) + \iota$. Hold the equilibrium objects $(p, F, N)$ and the destination-value functions fixed when varying the type of a current match. The scalar worker problem is an optimal-stopping problem with obstacle $kN(a)$. An offer can be represented, for valuation purposes, as an optional switch to the value $kJ(\tilde{y}, \tilde{x}, a)$ at rate $\beta\phi p$. This representation reproduces the positive-part term in the joint-value equation.

Applying the same stopping and switching rule to two current types shows that value is weakly increasing in either attribute: higher $y$ raises current output, and higher $x$ raises the human capital multiplying all subsequent positive payoffs while the current match survives. At a continuation point there is a nondegenerate initial interval over which remaining in the match has positive probability. Increasing output or learning quality strictly increases the payoff on that event. Hence value is strictly increasing from a continuing lower type. Finite-horizon comparison also gives continuity in the firm attributes. At an interior continuation point, continuity and strict monotonicity in $y$ permit a local continuous inverse for any nearby $x$. Strict monotonicity in $x$ makes the resulting iso-value graph strictly decreasing.

The value of a hypothetical current type can be evaluated on an open neighborhood of the support, keeping destination offers drawn from $F$. Derivatives at a support boundary refer to this extension. For differentiability, fix a type with $y > y_r$ and $x \geq x_r$. The maintained hiring normalization gives $J(y_r, x_r, a) = N(a)$. Type comparison, or direct comparison of the reference and match recursions, gives

$$J(y, x, a) > N(a) \geq kN(a), \qquad a < A.$$

The strict productivity advantage makes this continuation property stable to small changes in both attributes over the entire remaining horizon. To see this near retirement, put $t = A - a$. Uniform terminal estimates give

$$N(A - t) = e^{y_r} t + O(t^2), \qquad J(y, x, A - t) = e^{y} t + O(t^2).$$

Since $e^y > e^{y_r} \geq ke^{y_r}$, the quitting gap is uniformly positive to first order in $t$ throughout a sufficiently small type neighborhood. On any interval bounded away from retirement, continuity and the strictly positive quitting gap give the same conclusion by compactness. Consequently there is a neighborhood of the specified type on which the worker problem is an ordinary terminal-value equation without stopping.

Offer-value ties at a continuing current value are $F$-null. A tied destination satisfies $J(\widetilde{y}, \widetilde{x}, a) = J(y, x, a)/k > N(a)$ and is therefore itself continuing. For each $\widetilde{x}$, strict monotonicity in $\widetilde{y}$ permits at most one such productivity. Since $F$ has a density, Fubini's theorem gives zero probability to the tie set. The positive-part integral is therefore continuously differentiable in the current value along this neighborhood. Its derivative is bounded, so finite-horizon parameter differentiation, with dominated convergence at the terminal endpoint, gives continuously differentiable values and the sensitivity equations

$$\begin{pmatrix} (J_y)_a \\ (J_x)_a \end{pmatrix} = L(a) \begin{pmatrix} J_y \\ J_x \end{pmatrix} - \begin{pmatrix} e^y \\ \psi(a) J(y, x, a) \end{pmatrix},$$

where

$$L(a) = \rho + d(a) - \psi(a)x + \beta\phi p\, F\{(\widetilde{y}, \widetilde{x}) : kJ(\widetilde{y}, \widetilde{x}, a) > J(y, x, a)\}.$$

The terminal sensitivities are zero. Define $D(a, t) = \exp\left\{ - \int_a^t L(s)\, ds \right\} > 0$. Backward integration yields

$$J_y(y, x, a) = e^y \int_a^A D(a, t)\, dt, \qquad J_x(y, x, a) = \int_a^A D(a, t)\psi(t) J(y, x, t)\, dt.$$

Both derivatives are strictly positive. The implicit-function theorem therefore gives the slope $-J_x/J_y$.

The continuation equation implies

$$\begin{aligned} \frac{J_a}{J} &= \rho + d(a) - \psi(a)x - \frac{e^y + d(a)kN(a) + \beta\phi p \int (kJ(\widetilde{y}, \widetilde{x}, a) - J(y, x, a))^+ dF}{J} \\ &< \rho + d(a) - \psi(a)x. \end{aligned}$$

Under (7) and $x \geq x_r$, $(\psi J)' < 0$. The derivative ratio is consequently a strictly weighted average of values smaller than its current endpoint:

$$R(a) := \frac{J_x}{J_y} = \frac{\int_a^A D(a, t)\psi(t) J(y, x, t)\, dt}{e^y \int_a^A D(a, t)\, dt} < \frac{\psi(a) J(y, x, a)}{e^y}.$$

Using the sensitivity equations,

$$R'(a) = \frac{e^y R(a) - \psi(a) J(y, x, a)}{J_y(y, x, a)} < 0.$$

Integration proves the stated age comparison. □

### B.9 Learning-preference comparative statics

*Proof of Lemma 4.* Use an index $j \in \{L, H\}$ for the contact rate, set $k_\epsilon = e^{-\epsilon}$, and write $d(a) = \delta(a) + \iota$. The hiring-reference equation implies

$$\begin{aligned}(\psi N_j^\epsilon)' &= \left[\frac{\psi'}{\psi} + \rho + d(a)(1-k_\epsilon) - \psi x_r\right] \psi N_j^\epsilon \\ &- \psi \left[e^{y_r} + \beta\phi p_j \int (k_\epsilon J_j^\epsilon - N_j^\epsilon)^+ dF\right] < 0.\end{aligned}$$

Indeed, the coefficient in square brackets is at most $-d(a)k_\epsilon$ by (7), and the final bracket is strictly positive. Thus $\psi N_j^\epsilon$ strictly decreases with age.

First let $\epsilon = 0$. Acceptance and stopping have the same boundary $J_j^0 = N_j^0$. At that boundary, subtracting the reference recursion from the match recursion gives the instantaneous continuation advantage

$$B_j(y, x, t) = e^y - e^{y_r} + \psi(t)(x - x_r)N_j^0(t).$$

For $x \geq x_r$, this expression is nonincreasing in $t$. If it is positive at age $a$, continuing for a sufficiently short interval and then taking the outside value yields a strictly positive gain over stopping. If it is nonpositive at age $a$, it remains nonpositive thereafter. The function $N_j^0(t)$ is then a supersolution of the match problem on $[a, A]$ and equals its stopping obstacle. Scalar comparison implies $J_j^0 = N_j^0$ at age $a$. Therefore

$$J_j^0(y, x, a) > N_j^0(a) \iff e^y > e^{y_r} - \psi(a)(x - x_r)N_j^0(a).$$

At an interior cutoff,

$$z_j^{A,0}(x, a) = e^{y_r} - \psi(a)(x - x_r)N_j^0(a).$$

The zero-loss system also has a useful auxiliary control representation. Treat the reference state as producing $e^{y_r+h}$ and accumulating human capital at rate $\psi x_r$. A nonreference state produces $e^{y+h}$, accumulates at rate $\psi x$, and returns to the reference state at rate $d(a)$; it may also return voluntarily to the reference state at any time. Optional offers arrive at rate $\beta\phi p_j$ in every state. Its optimal value equations are exactly the reduced zero-loss equations for $(N_j^0, J_j^0)$; this is a valuation representation, not a claim about actual transition rates.

Couple the high-rate offer process with the low-rate process plus independent additional offers. Ignoring all additional offers reproduces the low-rate policy and payoffs, so $N_H^0 \geq N_L^0$. The maintained condition $F\{y > y_r\} > 0$ makes the inequality strict. There are a productivity gap $\Delta > 0$ and a positive-probability set of offers with $e^y \geq e^{y_r} + \Delta$. Uniform terminal expansions imply that these offers strictly improve on the reference state over a terminal age interval. Starting from any $a < A$, the event of no common offer before that interval and an additional offer from this set has positive probability. Accepting such an offer yields a strict gain. Consequently

$$N_H^0(a) \quad > \quad N_L^0(a) > 0, \qquad a < A.$$

The cutoff formula gives

$$z_H^{A,0}(x_1, a) - z_H^{A,0}(x_2, a) \quad = \quad \psi(a)(x_2 - x_1)N_H^0(a) > \psi(a)(x_2 - x_1)N_L^0(a) > 0.$$

It remains to justify continuity of acceptance cutoffs at zero portability loss. For fixed $(p_j, F)$, finite-horizon comparison for the coupled reference equation and match variational inequality gives uniform convergence of $(J_j^\epsilon, N_j^\epsilon)$ to $(J_j^0, N_j^0)$ as $\epsilon \downarrow 0$. The coefficients, offer payoff, and obstacle $k_\epsilon N_j^\epsilon$ vary continuously with $\epsilon$ and are uniformly Lipschitz in the bounded continuation values.

Fix one of the four interior zero-loss cutoffs $z_0$ and feasible nearby outputs $z_- < z_0 < z_+$. Above it, $J_j^0(\log z_+, x, a) - N_j^0(a) > 0$, so uniform convergence preserves acceptance for small $\epsilon$. Below it, uniform value convergence alone would be insufficient because the zero-loss surplus is zero on the stopping region. Instead, the strict margin and monotonicity of $\psi N_j^\epsilon$ imply, for sufficiently small $\epsilon$,

$$z_- - e^{y_r} + \psi(t)(x - x_r)N_j^\epsilon(t) \quad < \quad 0 \qquad \text{for all } t \in [a, A].$$

Substitute $N_j^\epsilon$ as a candidate match value. At $J = N_j^\epsilon$, the offer integrals and separation terms in the match and reference equations cancel, leaving precisely this expression as the continuation advantage. Thus $N_j^\epsilon$ is a supersolution of the match equation; it also dominates the quitting obstacle $k_\epsilon N_j^\epsilon$ and has the same terminal value. Scalar variational-inequality comparison gives

$$J_j^\epsilon(\log z_-, x, a) \quad \leq \quad N_j^\epsilon(a).$$

Hence $z_-$ is not accepted. Continuity and monotonicity in output now place the cutoff between $z_-$ and $z_+$. Taking arbitrarily close brackets, using the interior-section hypothesis, proves $z_j^{A,\epsilon}(x, a) \to z_j^{A,0}(x, a)$.

The two zero-loss inequalities are strict for the given age and pair of learning environments, so they persist for all sufficiently small nonnegative $\epsilon$. Continuity gives equality $J_j^\epsilon = N_j^\epsilon$ at each interior acceptance cutoff. When $\epsilon > 0$, this equality lies strictly above the quitting obstacle; when $\epsilon = 0$, it is the stopping boundary itself. $\square$

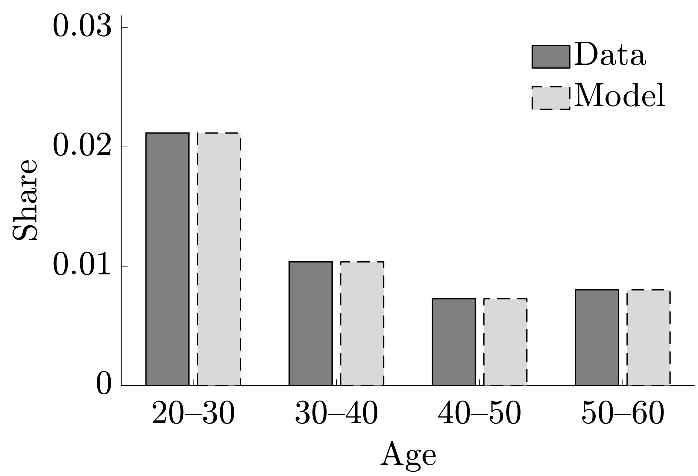
# C Estimation Appendix

This appendix reports the separation rates used in the calibration, validates the training measure used in the estimation, reports estimates for the restricted model, and presents additional cross-country predictions.

## C.1 Separation Rates

The exogenous separation rates $\{\delta(a)\}$ are set from the empirical EN rates by age, net of firm exit. Figure 19 reports the empirical EN profile.

Figure 19: EN Rates




*Notes:* The exogenous separation rates $\{\delta(a)\}$ are set from the empirical EN rates by age, net of firm exit. *Source:* ECHP 1993–2001; EU-SILC 2003–2024; PSID 1993–2023; author's calculations.

## C.2 Validation of the Training Measure

The dispersion in learning environments is identified from the excess wage growth that workers experience at training firms, which I classify based on whether the employer provides on-the-job training in the ECHP. This appendix validates the measure using within-worker variation across employers.

The ECHP asks each employed respondent whether their current employer provides them with on-the-job training. The measure is therefore a worker-reported, employer-specific indicator of formal, job-related training; it is distinct from general education, which the ECHP records separately, and does not capture informal learning by doing. I code a worker as employed at a *training firm* in a given year if they report receiving such training. Two features make the measure well-suited to identifying $\sigma_x$. First, because it refers to the worker's *current* employer, it varies within a worker across the firms they hold over the panel; this within-worker, across-firm variation—isolated by the individual fixed effects in (29)—is what identifies the wage premium at training firms, and hence the dispersion in learning environments $\sigma_x$. Second, the training question is unique to the ECHP, with no counterpart in EU-SILC or the PSID, so the training-firm classification is available only over the 1993–2001 period.

I regress wage growth on an indicator for whether the worker is at a firm that offers on-the-job training—and, for comparison, at a large firm with 50 or more employees—separately for young workers (ages 20–39) and older workers (ages 40–59). The training measure is available only in the 1993–2001 ECHP. I control for individual and country-year fixed effects:

$$\Delta \log w_{it} \quad = \quad \beta \times \text{large/training}_{it} + \xi_i + \xi_{cy} + \varepsilon_{it}. \tag{29}$$

Because the specification includes individual fixed effects, the estimates compare wage growth for the same worker across different firm environments. Table 11 shows that young workers experience faster wage growth while at a training firm: for workers aged 20–39, wage growth is 1.6 log points higher at training firms (and 2.4 log points higher at large firms). Older workers experience little differential wage growth across training and non-training firms.

Table 11: Wage Growth at Large and Training Firms

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Large | | Training | |
| | 20-39 | 40-59 | 20-39 | 40-59 |
| Estimate | 0.024*** | 0.001 | 0.016*** | -0.001 |
| | (0.005) | (0.004) | (0.005) | (0.005) |
| Observations | 239,523 | 356,209 | 91,265 | 80,320 |
| Years | 24 | 24 | 7 | 7 |
| Countries | 14 | 14 | 12 | 12 |

*Notes:* OLS estimates of (29). Standard errors clustered by country in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
*Source:* ECHP 1993–2001; EU-SILC 2003–2024; PSID 1993–2023.

## C.3 Calibration of the Restricted Model

The restricted model makes skill accumulation policy invariant by setting $\sigma_x = 0$, $\varrho = 0$, and $\epsilon = 0$. Five parameters $(\phi, \mu, \psi, \sigma_y, \beta)$ are re-estimated by minimum distance; the vacancy-cost curvature $\eta$ is imposed at its full-model estimate, and the job-finding rate $p$ is preset. Target families that no longer have a counterpart are omitted: the training-firm wage-growth premium, the training share by age, and the wage loss of job losers. The wage gain of job switchers and post-re-entry wage growth remain informative about $\sigma_y$ and $\beta$.

Table 12 reports the mapping between parameters and moments for the restricted model. Table 13 reports the preset parameters and Table 14 the re-estimated, imposed, and implied values; the corresponding model fit appears in Figure 20. Table 15 reports the elasticities of the targeted moments with respect to the re-estimated parameters in the restricted model.

Table 12: Targeted Moments, Restricted Model

| Parameter | Moment | Data Model | Weight |
|---|---|---|---|
| $\phi$ | JJ mobility rate by age | See Figure 20a | 1 |
| $\mu$ | Wages by age | See Figure 20b | 1 |
| $\psi$ | Share of life-cycle wage growth by age | See Figure 20c | 1 |
| $\sigma_y$ | Wage gain of JJ movers | See Figure 20d | 1 |
| $\beta$ | Wage growth after re-entry from nonemployment | See Figure 20e | 1 |
| $\beta$ | Wage on re-entry, large vs. small employers | See Figure 20e | 1 |

*Notes:* For each moment, I run a linear regression of the country-level value on log aggregate poaching and evaluate the regression at the average log poaching rate across countries. The final column reports the aggregate weight placed on each moment group in the estimation objective.

Table 13: Preset or Normalized Parameters, Restricted Model

| Parameter | Description | Value | Source |
|---|---|---|---|
| **Panel A. From Literature or Normalized** | | | |
| $\rho$ | Discount rate | 0.003 | 4% annual real rate |
| $\chi$ | Matching efficiency | 1.000 | Normalized |
| $\alpha$ | Matching elasticity | 0.630 | Moscarini and Postel-Vinay (2025) |
| $\sigma_x$ | Std. dev. of learning environment | 0.000 | Normalized |
| $\varrho$ | Gaussian-copula dependence of $y$ and $x$ | 0.000 | Normalized |
| $\epsilon$ | HC loss on accepting a new job | 0.000 | Normalized |
| **Panel B. Externally Calibrated** | | | |
| $\tau$ | Labor tax wedge | 0.417 | OECD Taxing Wages, 2000–2020 |
| $\iota$ | Firm exit hazard | 0.004 | 5% annual, employment-weighted |
| $M$ | Mass of firms | 0.328 | Eurostat, 2004–2020 |
| $\delta(a)$ | Separation hazard by age | | EN rate |
| | Age 20–30 | 0.017 | |
| | Age 30–40 | 0.006 | |
| | Age 40–50 | 0.003 | |
| | Age 50–60 | 0.004 | |
| $p$ | Job finding rate | 0.036 | Micro data |

Table 14: Parameter Values, Restricted Model

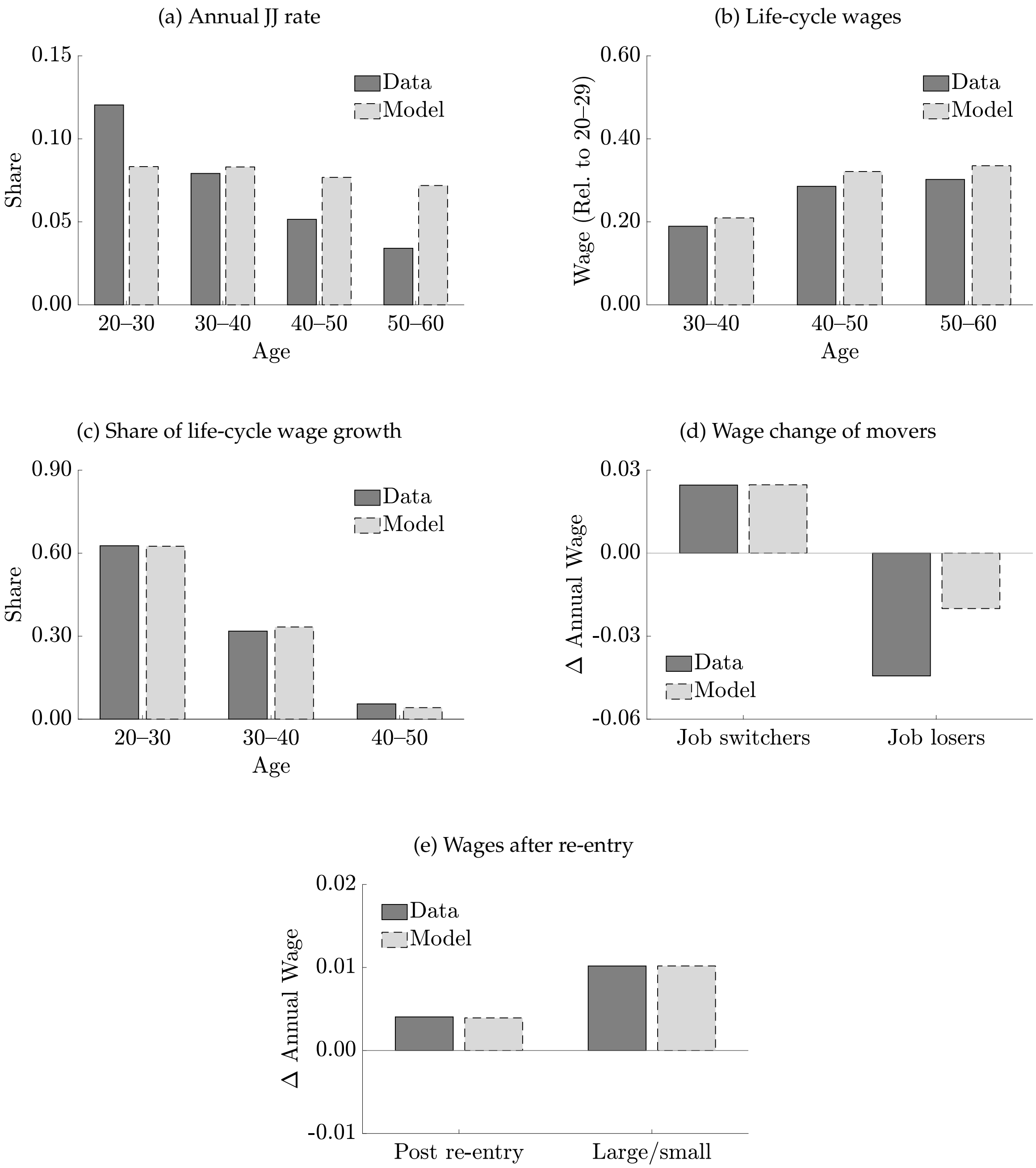
| Parameter | Description | Value | Identifying moment |
|---|---|---|---|
| **Panel A. Re-estimated Parameters** | | | |
| $\phi$ | On-the-job search efficiency | 0.507 | JJ rate (panel (a)) |
| $\mu$ | Mean learning environment ($\times$100) | 0.199 | Wages by age (panel (b)) |
| $\psi$ | Decline in learning ability | 0.889 | Wage growth shares (panel (c)) |
| $\sigma_y$ | Std. dev. of productivity | 0.039 | Job switchers (panel (d)) |
| $\beta$ | Worker bargaining power | 0.058 | Post re-entry, large/small (panel (e)) |
| **Panel B. Imposed from Full Model** | | | |
| $\eta$ | Vacancy cost curvature | 1.308 | Full-model estimate |
| **Panel C. Implied Parameters** | | | |
| $c_v$ | Vacancy cost scale | 1095 | |
| $c_e$ | Cost of starting a business | 14.22 | |
| $c$ | Cost composite | 7736 | |
| $e^{b(a)}$ | Flow value of leisure | | |
| | Age 20–30 | 1.021 | |
| | Age 30–40 | 1.019 | |
| | Age 40–50 | 1.019 | |
| | Age 50–60 | 1.022 | |

Table 15: Elasticity of Moments to Parameters, Restricted Model

| Moment | $\phi$ | $\mu$ | $\psi$ | $\sigma_y$ | $\beta$ |
|---|---|---|---|---|---|
| JJ mobility rate | **0.71** | -0.00 | 0.00 | -0.03 | -0.00 |
| Life-cycle wage growth | 0.37 | **0.12** | -0.24 | 1.65 | 0.31 |
| $(w_{30\text{–}40} - w_{20\text{–}30}) \,/\, (w_{50\text{–}60} - w_{20\text{–}30})$ | 0.36 | 0.08 | **-0.15** | 0.30 | 0.30 |
| Wage gain of job switchers | 0.10 | 0.00 | -0.00 | **1.42** | 0.08 |
| Wage growth post re-entry | 2.11 | 0.00 | -0.00 | 1.60 | **-1.11** |

*Notes:* Each entry is the elasticity of the row moment with respect to the column parameter, evaluated at the calibrated parameter vector. The elasticity is computed by re-solving the model with the column parameter perturbed by $\pm$10% around its estimated value, recomputing the moment, and dividing the centered log-difference in the moment by the log-difference in the parameter (for parameters whose baseline is at or near zero, a small additive step replaces the multiplicative one). All other parameters are held at their estimated values. Entries in bold mark the moment most informative about each parameter – the anchor moment used in the identification discussion. *Source:* author's calculations.

Figure 20: Targeted Moments, Restricted Model

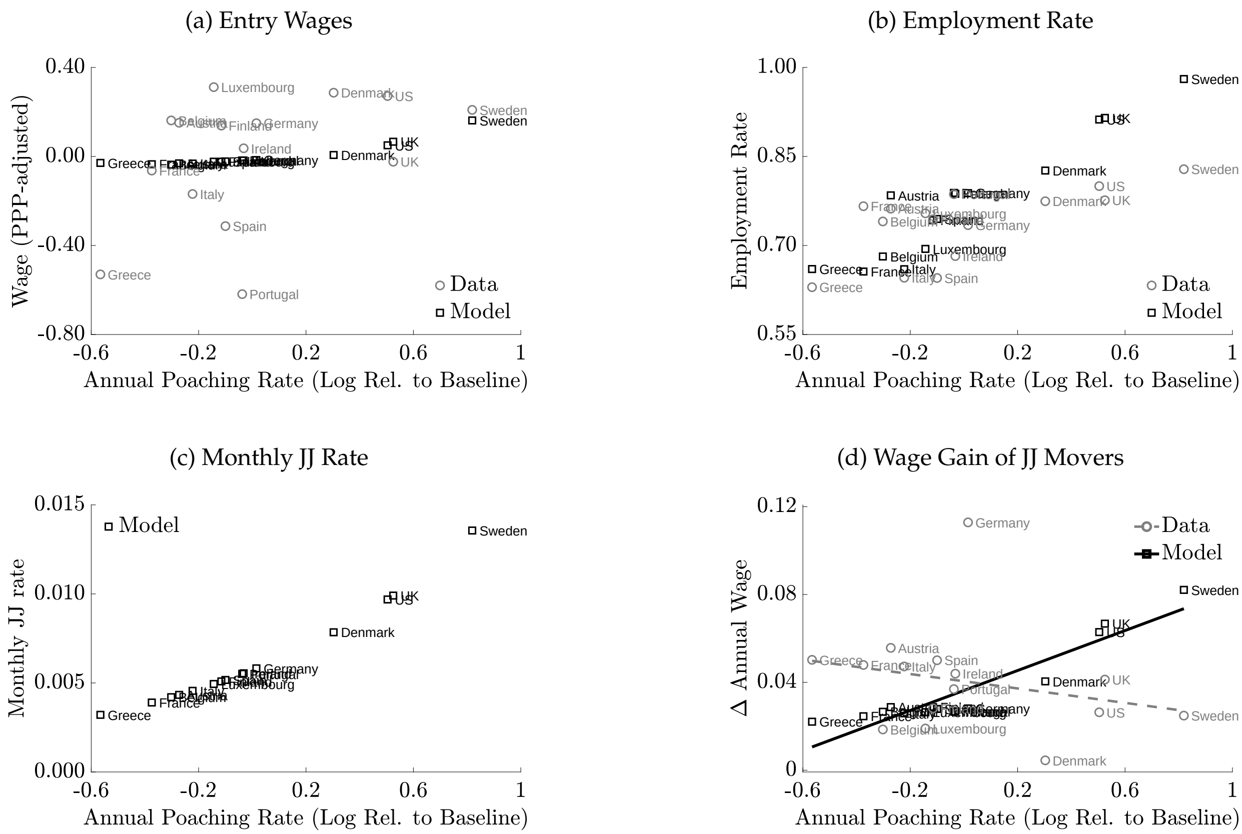



*Notes:* Panels are defined as for the full model in Figure 5, restricted to the moments that remain targeted when $\sigma_x = \varrho = \epsilon = 0$. *Source:* ECHP 1993–2001; EU-SILC 2003–2024; PSID 1993–2023; author’s calculations.

## C.4 Additional Cross-Country Predictions

Figure 21 contrasts the predictions of the model with the data along four untargeted dimensions. Entry wages are higher in high-poaching countries in both model and data (panel (a)), as are the employment rate and the monthly JJ rate (panels (b) and (c)). The model matches the absence of a systematic relationship between the wage gain of JJ movers and aggregate poaching (panel (d)).

Figure 21: Additional Cross-Country Outcomes in Model and Data




*Notes:* Squares show model outcomes and circles show data outcomes across countries. *Source:* ECHP 1993–2001; EU-SILC 2003–2024; PSID 1993–2023; EU-LFS; World Bank Doing Business; author’s calculations.