Labor Market Fluidity and Human Capital Accumulation

Niklas Engbom*
New York University
November, 2020

Abstract

In countries where job-to-job mobility is more common, wages grow more over the life-cycle. A life-cycle theory of on-the-job training and job shopping predicts that a more fluid labor market allows workers to faster reallocate to jobs where they can better use their skills, incentivizing accumulation of skills. Quantitatively, lower labor market fluidity reduces labor productivity by nine percent across the OECD relative to the US. A set of testable predictions for training are derived and confronted with comparable micro data on training across countries, and are found to offer support for the theory.

* nengbom@stern.nyu.edu. First draft: August, 2015. This is the first chapter of my PhD dissertation at Princeton University. It previously circulated under the title "Worker Flows and Wage Growth over the Life-Cycle: A Cross-Country Analysis." I am grateful for the generous support and advice of Richard Rogerson. I thank Mark Aguiar, Jorge Alvarez, Adrien Bilal, Victoria Gregory, Veronica Guerrieri, Greg Jarosch, Greg Kaplan, Burhan Kuruscu, Guido Menzio, Claudio Michelacci, Ben Moll, Chris Moser, Diego Restuccia, Todd Schoellman, Gianluca Violante, three anonymous referees, and seminar participants at CEMFI, Einaudi, Georgetown, ITAM, Melbourne, MIT, Princeton, Toronto and the SED. 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 differences in labor market flows across countries. This research finds that such flows vary markedly between countries, and that policies and institutions that impede such flows may lead to misallocation of factors of production. The literature, however, has tended to focus on the effect of such policies on firms’ job creation and destruction decisions, with less attention paid to their impact on worker flows and the behavior of workers. Yet workers’ responses to such large differences in the functioning of the labor market may have a first-order effect on aggregate economic outcomes.

In this paper, I quantify the impact of differences in the fluidity of the labor market on workers’ wage growth, where I measure labor market fluidity as how frequently workers make job-to-job (JJ) transitions. In particular, I argue that greater labor market fluidity encourages individuals to accumulate skills on-the-job. The reason is that it improves their ability to locate a job where their skills are more useful. Consequently, labor market fluidity facilitates life-cycle wage growth, both by allowing workers to move toward jobs that reward their skills better and by inducing them to grow their skills more.

I offer two main contributions to the literature. First, I establish a series of novel facts on cross-country differences in life-cycle careers. To that end, I build an internationally comparable worker-level panel data set covering close to a million observations for over 20 years across 23 OECD countries. The data suggest wide dispersion across countries in labor market flows and life-cycle wage growth. For instance, labor market fluidity differs by a factor of 2.5 across countries, while workers experience substantially greater life-cycle wage growth in some countries, mirroring findings in Lagakos et al. (2018).

My main empirical finding is that wages grow more over the life-cycle in more fluid labor markets. The panel dimension of my data allows me to rule out that this is driven by differential selection patterns over the life-cycle across countries by controlling for individual-fixed effects. Moreover, by standardizing education and occupation classifications across countries, I conclude that differences in workforce composition along such dimensions across countries do not drive the patterns. Yet it is not entirely surprising that wages grow more over the life-cycle in countries where JJ mobility is more common, as such transitions are typically associated with wage gains. Indeed, I document that wages, in a residual sense, rise by 5–6 percent more in years when a worker makes a JJ move, and that the magnitude of these gains is uncorrelated with labor market fluidity. Nevertheless, in an accounting sense, accumulating such wage gains over the life-cycle, they cumulatively account for only about a quarter of the steeper life-cycle wage growth in more fluid labor markets. That is, most of the steeper life-cycle wage growth in more fluid labor markets arises within jobs, as opposed to between jobs.

---

1See, e.g. Bentolila and Bertola (1990), Hopenhayn and Rogerson (1993), and Ljungqvist and Sargent (1998, 2008).
The second contribution of this paper is to develop a tractable life-cycle theory of optimal on-the-job accumulation of skills in a frictional labor market to interpret these patterns. The theory combines learning on-the-job as in Ben-Porath (1967) with an equilibrium search model in the spirit of Diamond (1982)—Mortensen and Pissarides (1994), and offers rich, testable predictions for how human capital accumulation on-the-job and labor market search interact. The marginal product of a worker’s human capital differs across firms, but frictions in the labor market prevent workers from immediately reallocating to their most productive use. I make one key modification relative to an earlier literature on training in frictional labor markets (Acemoglu, 1997; Acemoglu and Pischke, 1998, 1999b): I allow workers to move directly from one employer to another without an intervening spell of unemployment.2 A young worker enters the labor market with few skills and in a job that does not utilize her skills particularly well. Through on-the-job training, she gradually builds her skills. Other firms try to poach her, such that over time she reallocates toward jobs that use her skills efficiently—she climbs the job ladder.

I examine in the model the impact on workers’ wage growth of wedges to firms’ cost of creating jobs, motivated by a large literature that argues that policies such as, for instance, employment protection legislation, business regulations, and labor taxes serve to raise the cost to firms of hiring (Fonseca et al., 2001; Pries and Rogerson, 2005). Such wedges reduce labor market fluidity. As workers have a harder time finding a job that uses their skills efficiently, the expected value of human capital declines. Workers respond by accumulating less human capital, such that the stock of human capital falls.

To quantify the impact of labor market fluidity on worker careers, I estimate the model targeting the US as a high fluidity country. The model matches life-cycle dynamics well. JJ mobility declines over the life-cycle, as workers gradually settle into jobs that use their skills efficiently. Wages grow rapidly early in careers, as young workers have significant scope to climb the job ladder and face high returns to training. My estimates imply that human capital is the most important source of life-cycle wage and productivity growth, contributing roughly 25 log points between age 25 and 50. Growth in match productivity is a close second. Wages also grow as the arrival of outside offers allows workers to raise the share of rents they expropriate, and due to a declining share of time devoted to training over the life-cycle.

Subsequently, I isolate the impact of labor market fluidity by introducing in the estimated model wedges to firms’ cost of creating jobs, calibrated such that the model matches the cross-country variation in labor market fluidity. Differences in labor market fluidity account for 50 percent of the steeper life-cycle wage growth in more fluid labor markets across my sample of OECD countries. Match productivity grows less over the life-cycle in less fluid labor markets as workers climb the job ladder slower.

---

2 Acemoglu and Pischke (1998) briefly discuss the effect of allowing for poaching—what they refer to as raids—noting that “whether raids are possible or not, may have important consequences for training.” They do not pursue this further, though.
accounting for 40 percent of the lower life-cycle wage growth. Of the eight log point lower accumulation of human capital in the least fluid labor market relative to the US, 41 percent is due to the fact that workers spend more time in unemployment, where they do not accumulate skills. 35 percent is accounted for by workers increasingly being stuck at the bottom of the job ladder, where they have less incentive to accumulate skills since those are less useful. The remainder is due to the fact that conditional on a worker’s current labor market state, she is discouraged from accumulating skills by the higher expected incidence of unemployment and slower rate of climbing the job ladder. Across the OECD, aggregate labor productivity is nine log points lower relative to the US. In summary, my findings highlight that policies and regulations that reduce labor market fluidity in turn have large negative consequences on both workers’ life-cycle wage growth and aggregate economic outcomes.

In the last part of the paper, I return to the micro data to document three patterns consistent with the theory. To that end, I standardize data on vocational training for nine countries based on an identical survey fielded by the European Union’s statistical agency, Eurostat. First, workers spend more time on vocational training in more fluid labor markets. This pattern is not accounted for by cross-country differences in worker observable characteristics. Second, conditional on worker-fixed effects and time-varying covariates, within countries workers train more when employed at larger, higher paying employers. This pattern is consistent with the predictions of the theory that, *ceteris paribus*, workers train more in more productive, higher paying firms, as such firms afford them greater use of their skills. Third, workers in less fluid labor markets train disproportionately less at small, low paying employers. According to the theory, this is because in less fluid labor markets workers currently in low productive firms have a harder time moving to jobs where they can better use their skills. Expecting this, they train less.

The premise of this paper is to take differences in labor market fluidity as given to assess their consequences for worker behavior. I hence purposefully avoid taking a strong stand on the exact policies that result in differences in labor market fluidity, which has been the focus of a large literature. My reading of that literature is that several polices likely combine to discourage firms from creating jobs in some countries. Consistent with this view, I end by showing that employment protection legislation, regulations that raise the cost on firms of doing business, and employer labor taxes are all negatively correlated with labor market fluidity. That being said, I offer no new insights on this front. Rather, the novelty of this paper is to quantify the dynamic consequences of such policies on worker outcomes, in the spirit of Restuccia and Rogerson (2017, p.170)’s argument that "whereas much of the literature has focused on static misallocation, we think the dynamic effects of misallocation deserve much more attention.”

This paper contributes to three strands of the literature. First, a vast literature documents cross-country differences in labor market outcomes. Whereas much work has focused on hours worked or
unemployment (Nickell, 1997; Bick et al., 2018; Feng et al., 2020), less work assesses differences in wage growth and labor market flows, particularly JJ flows. Hobijn and Sahin (2009) document flows in and out of unemployment across 27 OECD countries, while Jolivet et al. (2006) find JJ mobility patterns consistent with this paper across 11 countries for 1994–1997. Blanchard and Portugal (2001), Schönberg (2007), and Dustmann and Pereira (2008) provide two-country comparisons. Lagakos et al. (2018) show that richer countries have steeper life-cycle wage growth across 18 countries at different stages of development, while Donovan et al. (2020) document differences in labor market flows across rich and poor countries.\footnote{Consistent with this paper, the latter find that richer countries have higher JJ flows among a set of developed countries—for instance, their Table 2 “provides the regression estimates from a sample that includes only EU countries, Switzerland, the U.K., and the United States. For this sample, we also find a positive relationship between labor market flows and development” (Donovan et al., 2020, p.14). In contrast, developing countries are characterized by higher labor market flows.} Poschke (2019) studies unemployment and self-employment rates across poor and rich countries. Guner et al. (2018) account for steeper earnings growth for managers relative to non-managers in rich countries.

Second, recent work combines theories of human capital and search in studies of worker careers (Yamaguchi, 2010; Bagger et al., 2014; Karahan et al., 2020), building on seminal work by Becker (1962), Ben-Porath (1967), Diamond (1982), Mortensen and Pissarides (1994), and Burdett and Mortensen (1998). This literature typically models human capital as exogenous via learning-by-doing (see Bowlus and Liu, 2013, for an exception). I instead follow Ben-Porath (1967) to model training as a choice. This approach provides novel insights into how search and human capital accumulation interact. Herkenhoff et al. (2018) study learning through interactions with co-workers. Gregory (2019) argues that workers grow their human capital more at some firms, which is consistent with the argument in this paper (but here it arises through optimizing training decisions rather than exogenous differences in firms’ learning environment). Engbom (2020) proposes a theory of search and entrepreneurship to study the impact of demographic change within countries, whereas the current paper assesses how search and training interacts in a model of endogenous human capital accumulation with a focus on cross-country patterns.

Third, a voluminous literature assesses the sources of cross-country income differences. Seminal work by Klenow and Rodríguez-Clare (1997), Prescott (1998) and Hall and Jones (1999) find that differences in total factor productivity (TFP) play a key role behind income differences. Erosa et al. (2010) and Manuelli and Seshadri (2014) challenge this conclusion, arguing that human capital accounts for a significant share of cross-country differences. The latter is closest to this paper in that they also allow for on-the-job accumulation of skills, but abstract from labor market frictions.

Section 2 documents new facts on life-cycle careers across countries. Section 3 develops the theory and Section 4 estimates it. Section 5 quantifies the impact of labor market fluidity. Section 6 provides supportive evidence on training across countries and Section 7 concludes.
2 Motivating correlations

I start by establishing some new facts on life-cycle labor market outcomes across countries. To that end, I build an internationally comparable worker-level panel data set covering 23 OECD countries. Specifically, I rely on data from the following sources and time periods: the US Panel Study of Income Dynamics (PSID) 1994–2015; the German Socio-Economic Panel (GSOEP) 1991–2011; the British Household Panel Survey (BHPS) 1991–2008; the European Community Household Panel (ECHP) 1994–2001; and the European Union Statistics on Income and Living Conditions (EUSILC) 2003–2014. While a cross-country comparison inevitably is subject to issues of comparability, an important advantage of these data sets is that they are modeled on the PSID, facilitating the international comparison. In the interest of space, Appendix A.1 provides a more detailed discussion of the data sources.

2.1 Sample selection

I focus on men, as female labor force participation likely varies across countries for reasons that the theory in the next section abstracts from. To limit the impact of issues associated with labor force entry and retirement, I primarily focus on ages 25–54, but present additional samples as robustness. As I discuss in Appendix A.3, male labor force participation rates are consistently high across countries between ages 25–54. Moreover, there is no statistically significant correlation between labor force participation rates and labor market fluidity, either at the aggregate level, at age 25 or at age 54. Additionally, I drop observations with missing year of birth or employment status, as well as individuals whose reported year of birth deviates by more than five years across panel years. This excludes very few observations.

I focus on employees and flows in and out of unemployment, as the theory also abstracts from self-employment and non-participation. Flows in and out of non-participation are small, however, and do not vary systematically with labor market fluidity among prime aged men. I include all wage employees, regardless of full-time status, but similar results hold among those working 30+ hours a week. In my main analysis, I do not condition on being in the private sector, partly because the EUSILC does not make available sector to researchers. Appendix A.4 argues that the public-private distinction is unlikely to be a main force driving the cross-country patterns documented here (among prime aged men).

My analysis focuses primarily on 13 developed Western European countries and the US for which I have 15 or more years of data. I report robustness results including an additional 10 OECD countries for which I have fewer years of data. The core annual sample includes over six hundred thousand

---

4For this reason, I prefer to use the PSID in my main analysis of worker flows. Appendix A.2 shows that the resulting flows are broadly consistent with those in the US Survey of Income and Program Participation (SIPP).

5The EUSILC also contains a few years of data from five non-OECD countries: Bulgaria, Cyprus, Malta, Romania and Serbia.
observations, with another two hundred thousand observations for the other OECD countries.

2.2 Variable definitions

I construct two samples. The first is an annual sample used in my wage analysis. The wage is total gross labor income in the prior calendar year divided by annual hours worked, constructed as the product of weeks worked times usual weekly hours. I top-code weekly hours at 98 hours to be consistent with the PSID. I include in labor income also income from self-employment to be consistent with the BHPS, which does not distinguish sources of labor income. I do, however, focus on those who are wage employed at the time of the survey—henceforth employees. I convert nominal values to real 2004 local currency using the national CPI, then to real US dollars using the PPP-adjusted exchange rate in 2004.

The second sample is a monthly data set, which I use to estimate labor market flows. The surveys ask for labor market status in each month during the prior calendar year. By linking subsequent years, I obtain labor force status in each month during the 12 months prior to the survey month. In particular, a worker is said to make an EU transition if she is employed in the current month but unemployed in the subsequent month. She makes an UE transition if she is unemployed in the current month but employed in the subsequent month. The available data sets do not allow for the construction of a satisfactory monthly measure of JJ mobility, because the surveys in general ask for information on only (up to) two employment spells in the prior year. As a consequence, at most one JJ move can be observed during the past 12 months, even though the worker might have made multiple transitions. For young, highly mobile workers in particular, this restriction is not innocuous. Hence, I instead construct a consistent measure of JJ mobility across countries as the fraction of employees who started working for their current employer at some point in the past 11 months while having had employment as the main employment status in every of the past 12 months. Subject to one caveat, this measure accounts for intervening months of non-employment between job switches—it is not equivalent to the fraction of employed workers who were at a different employer 12 months earlier. The one caveat is that it does not rule out short intervening

---

6 In the PSID, each subcomponent of total income is top-coded at separate thresholds that vary across years. I use a Pareto imputation to top-coded subcomponents in each year before I sum each component to get total income (Heathcote et al., 2010). The BHPS records income and hours from September to September instead of by calendar year.

7 Robustness exercises suggest that differences in self-employment income among wage employees are third-order with respect to the patterns documented here (available for all countries but the UK).

8 The PSID does not allow a distinction between wage and self-employment in the monthly calendar of events, and hence to be consistent all monthly flows include the self-employed as employed. In the other data sets as well as in the US SIPP, though, flows from employment to (and from) self-employment are an order of magnitude smaller than those to (and from) unemployment, so I believe that this issue is second-order.
spells of unemployment, as I only observe main employment status in a month. Because flows in and out of unemployment are so low, however, I doubt time aggregation majorly biases my results.9

I standardize year of birth to the modal value across panel years, education into two groups—less than college or college or more—based on an individual’s highest reported degree across panel years, and occupation into 10 internationally comparable, aggregate occupation groups based on ISCO-88.

Figure 1 plots the share of employees who made a JJ move at some point in the previous year—my measure of labor market fluidity—over the life-cycle across the core countries. JJ mobility has a common shape: it is high early in careers, and subsequently declines substantially as individuals age. There are significant differences in the level of labor market fluidity across countries, with some countries displaying higher fluidity at all ages. These high fluidity countries include the Anglo-Saxon countries (UK and US), as well as Denmark and the Netherlands. On the other end of the spectrum, Belgium, Austria and Greece have JJ mobility rates that are less than half those in the US at all ages.

I next construct life-cycle wage profiles by regressing separately by country the log hourly real wage of individual $i$ in year $t$, $w_{it}$, on age effects, $A_{it}$, year effects, $Y_t$, and individual fixed effects, $I_i$,

$$w_{it} = A_{it} + Y_t + I_i + \varepsilon_{it} \quad (1)$$

The inclusion of individual fixed effects in regression (1) addresses important concerns about differences in sample attrition across countries biasing the cross-country comparison of life-cycle wage growth. In my benchmark, I compute wage growth by age. But I also report results below controlling for education—isomorphic to wage growth by potential experience—with similar results.

Whenever an individual gets one year older, time also increases by one year, and vice versa. That is, age, time and individual fixed effects are collinear. Hence, a restriction is needed to identify regression (1). I follow Heckman et al. (1998) and Lagakos et al. (2018) in imposing that wages depreciate at some rate $d$ after some age $\bar{A}$. This restriction is sufficient to separate individual, time and age effects. Effectively, fluctuations in wages among individuals older than $\bar{A}$ identify the year effects. The age effects can then be recovered from within-individual fluctuations in wages among those aged less than $\bar{A}$.

Figure 1 plots wage growth between ages 25–50 across countries, assuming that wages do not grow after age 50. Across all countries, life-cycle wage profiles share a common shape, with wages growing

---

9As part of the estimation in Section 4, I simulate a monthly approximation to the underlying continuous time model. I have alternatively simulated a weekly model and aggregated that up to the monthly level. It has a second-order impact on my measure of JJ mobility, precisely because flows in and out of unemployment are so low. The fact that flows in and out of unemployment are estimated to be much lower in panel data relative to primarily cross-sectional data such as the US Current Population Survey has been long recognized in the literature, driven by employment status classification error (Abowd and Zellner, 1985; Poterba and Summers, 1986). See Appendix A.2 for more details and a comparison with the US SIPP.
rapidly early in careers. In contrast to the common shape, there is a remarkable degree of heterogeneity across countries in the overall wage growth over the life-cycle, as emphasized by Lagakos et al. (2018).\(^\text{10}\)

Appendix A.5 plots the EU and UE rates over the life-cycle across countries. The EU rate shares a common shape across countries, with high rates of job loss early in careers and subsequent declines. The UE rate is roughly flat between ages 25–54 in most countries, although its shape is somewhat more heterogeneous across countries than the EU and JJ rates. Labor market fluidity is negatively correlated with the aggregate EU rate and positively correlated with the aggregate UE rate.

### 2.3 Wages grow more over the life-cycle in more fluid labor markets

The left panel of Figure 2 plots life-cycle wage growth against labor market fluidity across the core set of 13 countries. Wages grow substantially more in more fluid labor markets. For instance, wages grow by 75 log points in the US between age 25–54, but only by 30 log points in Italy. The right panel shows that similar results hold when including the additional 10 countries with fewer years of data, with a correlation coefficient between life-cycle wage growth and labor market fluidity of around 0.8.

One possible factor behind these patterns is differences in workforce composition. To assess this, I consider an augmented version of regression (1) that instead pools all countries and years,

\[
\log w_{it} = \alpha \text{fluidity}_c \times a_{it} + X_{it} \beta + I_i + Y_t + A_{it} + \varepsilon_{it}
\]

As above, \(w_{it}\) is the log real hourly wage, \(I_i\) are individual fixed effects, \(Y_t\) are year effects, and \(A_{it}\) are restricted age effects. The coefficient \(\alpha\) captures the covariance between life-cycle wage growth and labor market fluidity, while \(X_{it}\) controls for separate age slopes for two education groups or 10 occupation groups (minus one due to collinearity). I renormalize the provided survey weights such that each country in the aggregate receives a total weight of one,\(^\text{11}\) and cluster standard errors by country.

Table 1 presents regression results from several specifications, including allowing for wages to fall late in life and extending the sample to include all workers aged 22–59. Appendix A.6 shows that similar results hold for the full sample of 23 countries. Confirming the pattern in Figure 2, there is a strong, statistically significant positive correlation between labor market fluidity and life-cycle wage growth.

\(^{10}\)Where our studies overlap, my wage profiles align well with Lagakos et al. (2018). The main exception is Germany, which they find has the highest wage growth (steeper than the US). I confirm this finding in a specification that includes workers aged 22–59, which may be closer to their specifications that include all workers with 0–40 years of experience. My patterns remain robust to such alternative specifications. Nevertheless, I prefer to focus on ages 25–54 due to higher non-participation rates prior to 25 and after age 54 (see Appendix A.3), which the theory abstracts from.

\(^{11}\)For countries with observations in the ECHP and EUSILC, I adjust the weights such that each country-survey receives a relative weight equal to the number of years of data for that country in that survey, and the total weight for the country is one.
Note: Male employees aged 25–54. JJ mobility: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-age-level. Log hourly real wage profile based on (1) with worker fixed effects, time effects and age effects restricted to not grow past age 50. Source: BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.
in the baseline specification. While college graduates, for instance, experience greater life-cycle wage growth, differences in educational or occupational composition across these developed countries are much too small to change this conclusion. Allowing for a different depreciation rate late in careers makes virtually no difference to the point estimate. Moreover, extending the sample to start at age 22 and/or include those up to age 59 does not change the conclusion. In additional results I reach the same conclusion including workers up to age 64, including instead setting $\bar{A} = 55$.

**Figure 2. Life-cycle wage growth and labor market fluidity**

(A) Core countries

(B) All countries

*Note:* Male employees aged 25–54. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-level. Life-cycle wage growth: Log hourly real wage profile based on regression (1) with worker fixed effects, time effects and age effects restricted to zero growth after age 50. *Source:* BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.

It is, perhaps, not surprising that wages grow more in countries where workers on average make more JJ transitions, as such transitions typically involve a wage gain (Topel and Ward, 1992). Indeed, the left panel of Figure 3 shows that across all countries, the average worker tends to experience greater growth in wages in years when she made a JJ move. It projects the difference in median annual wage growth between those who made a JJ move in the year and those who did not on labor market fluidity, computed within age-year bins and subsequently aggregated to the country-level giving equal weight to each year-age. I use the median to limit the impact of a few outliers.  

$\text{12}^2$Because the PSID turned biannual in 1997, the estimate of wage gains for the US relies only on years 1994–1997. As I show in Section 4, however, the wage gain upon a JJ move is similar in the SIPP for more recent years and I reach very similar conclusions if I approximate the annual wage gain after 1997 in the PSID with the bi-annual wage gain.

---

12Because the PSID turned biannual in 1997, the estimate of wage gains for the US relies only on years 1994–1997. As I show in Section 4, however, the wage gain upon a JJ move is similar in the SIPP for more recent years and I reach very similar conclusions if I approximate the annual wage gain after 1997 in the PSID with the bi-annual wage gain.
percent greater residual wage growth in a year relative to a stayer with the same age in the same year. There is no systematic relationship between the wage gain upon a JJ move and labor market fluidity.

**Table 1. Life-cycle wage growth and labor market fluidity**

<table>
<thead>
<tr>
<th></th>
<th>Panel A. Ages 25–54</th>
<th>Panel B. 1% depreciation</th>
<th>Panel C. Ages 22–59</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline Educ Occup</td>
<td>Baseline Educ Occup</td>
<td>Baseline Educ Occup</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.244*** 0.217** 0.224***</td>
<td>0.244*** 0.217** 0.224***</td>
<td>0.261*** 0.230*** 0.238***</td>
</tr>
<tr>
<td></td>
<td>(0.074) (0.076) (0.069)</td>
<td>(0.074) (0.076) (0.069)</td>
<td>(0.069) (0.069) (0.065)</td>
</tr>
<tr>
<td>N</td>
<td>336,349 334,415 323,888</td>
<td>336,349 334,415 323,888</td>
<td>393,910 391,042 379,094</td>
</tr>
</tbody>
</table>

**Note:** Male employees. Panel A. Wages restricted to not grow after age 50. Panel B. Ages 25–54 with wages restricted to depreciate 1% annually after age 50. Panel C. Wages restricted to not grow after age 50. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. $\alpha$: Fluidity-age interaction in regression (2) with worker fixed effects, time effects and restricted age effects. Standard error below are clustered at the country-level. ** statistically significant at 5%; *** statistically significant at 1%. ** Source:** BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.

**Figure 3. The role of between-job wage growth**

(A) Wage gain in year of JJ move  
(B) Between vs within job

**Note:** Male employees aged 25–54. Panel A. Median annual wage growth of workers who made a JJ move in the past year relative to those who did not, computed within country-year-age groups and then collapsed to the country-level giving equal weight to each year-age. Panel B. Between-job: Product of the median wage gain from a JJ move and the average total number of JJ moves between age 25–54. Within-job: Difference between total wage growth between age 25–54 and the between-job component. Total wage growth based on (1) with worker fixed effects, time effects and age effects restricted to be zero past age 50. All panels. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. ** Source:** BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.

The right panel uses the gain upon a JJ move to estimate the contribution of job shopping toward
life-cycle wage growth. I compute between-job life-cycle wage growth by multiplying the wage gain from a JJ move with the average number of JJ moves a worker makes between age 25–54.\textsuperscript{13} I compute within-job life-cycle wage growth as the difference between total wage growth and between-job wage growth. I stress that no structural interpretation should be assigned to the components—it is simply an accounting decomposition. Less than a quarter of the steeper life-cycle wage growth in more fluid labor markets is accounted for by between-job wage growth. Because the average wage gain is not correlated with fluidity, the steeper between-job wage growth in more fluid labor markets is entirely driven by the higher frequency of moves. Hence, while the between-job component is non-trivial, the majority of the steeper life-cycle wage growth in more fluid labor markets takes place within-jobs.

2.4 Discussion

Before I go to the model, I establish some additional empirical patterns that help guide the analysis.

Entry wages. Table 2 regresses wages of labor market entrants on labor market fluidity, with or without controls. Because this exercise assesses level differences, I cannot control for worker fixed effects. Entry wages are lower in more fluid labor markets, although not statistically significantly so. Quantitatively, entry wages are 2–6 percent higher in the least fluid labor market relative to the US. Controlling for real GDP per hour in 2004 PPP-adjusted US dollars makes the pattern more pronounced.\textsuperscript{14}

What workers grow their wage more? Assessing what workers grow their wages more in more fluid labor markets may shed further light on potential driving forces behind the patterns. The left panel of Figure 4 re-estimates life-cycle wage growth based on (1) separately by education group, and projects it on labor market fluidity, also computed separately by education group. College educated workers grow their wage more over the life-cycle in all countries. Moreover, the gradient with labor market fluidity is steeper among college graduates (although Italy is a peculiar outlier).

The right panel considers an augmented version of the pooled regression (2) that includes a linear in an occupation’s wage rank, as well as its interaction with age, fluidity and age times fluidity. I rank occupations within each country-year, and then assign the occupation its (unweighted) average across country-years.\textsuperscript{15} Higher wage occupations experience steeper life-cycle wage growth. Moreover, they

\textsuperscript{13}While there is a life-cycle profile to the wage gains upon a JJ move, this computation relies only on life-cycle averages. That is, it makes no difference to the total to use the age-specific wage gains (although it does matter for the timing of the gains).

\textsuperscript{14}I am not, however, convinced that doing so makes sense. To the extent that the labor share does not covary systematically with fluidity (which it does not), controlling for labor productivity is isomorphic to controlling for the average wage. Hence with such a control, the steepness of life-cycle wage profiles and entry wages are the flip side of the same coin.

\textsuperscript{15}The resulting ranking makes intuitive sense, with engineers and doctors ranked the highest and laborers ranked the lowest.
experience disproportionately steeper wage growth in more fluid labor markets relative to lower wage occupations, although the pattern is not statistically significant (p-value of 0.25). I conclude from these two exercises that the cross-country correlation between life-cycle wage growth and labor market fluidity is not driven by low skilled, low wage workers. At face value, this speaks against explanations based on minimum wages or unions, which typically impact lower skilled workers the most.

**Table 2. Entry wages and labor market fluidity**

<table>
<thead>
<tr>
<th></th>
<th>Year</th>
<th>Edu</th>
<th>Occup</th>
<th>GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>-0.891</td>
<td>-0.358</td>
<td>-0.637</td>
<td>-3.105*</td>
</tr>
<tr>
<td></td>
<td>(2.795)</td>
<td>(2.921)</td>
<td>(2.722)</td>
<td>(1.721)</td>
</tr>
<tr>
<td>N</td>
<td>21,266</td>
<td>20,342</td>
<td>20,200</td>
<td>21,266</td>
</tr>
</tbody>
</table>

Note: Male employees 21–24. Log hourly real wage in 2004 PPP-adjusted USD. Year: Year controls. Edu: Year and education controls. Occup: Year and occupation controls. GDP: Year and GDP per hour controls (in 2004 PPP-adjusted USD). Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-level. Standard errors are clustered at the country-level. * significant at 10%. Source: BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.

**Inequality.** Patterns for inequality may provide additional guidance regarding the forces at work. For instance, Guvenen et al. (2013) emphasize that progressive taxation reduces life-cycle wage growth and inequality, which may be correlated with labor market fluidity. The top left panel of Figure 5 relates labor market fluidity to the standard deviation of log residual wages. I focus here on residual inequality, because that is what the theory in the next section is about. Appendix A.7 contains additional results.

The top right panel graphs life-cycle growth in inequality against fluidity, where the former is the difference in standard deviation of residual log wages at age 50–54 relative to at age 25–29. Neither the level of inequality nor its change over the life-cycle is systematically related to labor market fluidity. While inequality in general rises over the life-cycle across these countries, it is primarily accounted for by increasing dispersion across education and occupation groups. Within groups, there is not much of an increase in inequality, even in the US. As I believe that this finding is somewhat novel to the literature, I provide a further discussion and robustness in Appendix A.7.

**Voluntary versus involuntary.** The data sets also include information on the reason for separation, which I standardize into those who quit the job versus the firm initiated the separation, which I somewhat loosely refer to as an involuntary separation. The middle left panel shows that only about 60 percent of JJ movers report that they quit their job, with the remainder saying that the firm initiated the separation. If anything, the share of quits among JJ movers is higher in more fluid labor markets. Hence, the higher JJ rate is not driven by more involuntary movers. The distinction between an employer initiated separation and a quit, however, is murky. Indeed, the model in the next section predicts
that separations are bilaterally optimal, with no theoretical distinction between a quit and the employer letting the worker go. For this reason, I prefer to use the overall measure of JJ mobility as my benchmark.

**Figure 4. Wage Growth and Labor Market Fluidity by Education and Occupation**

<table>
<thead>
<tr>
<th>Interaction with occupation’s wage rank</th>
<th>Fluidity × age</th>
<th>-0.222*** (0.070)</th>
</tr>
</thead>
<tbody>
<tr>
<td>occupation wage rank</td>
<td>-0.021***     (0.006)</td>
<td></td>
</tr>
<tr>
<td>occupation wage rank × fluidity</td>
<td>-0.122        (0.163)</td>
<td></td>
</tr>
<tr>
<td>occupation wage rank × age</td>
<td>0.002***      (0.001)</td>
<td></td>
</tr>
<tr>
<td>occupation wage rank × fluidity × age</td>
<td>0.017         (0.014)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>359,394</td>
<td></td>
</tr>
</tbody>
</table>

Note: Male employees aged 25–54. Left panel. Life-cycle wage growth based on regression (1) separately by country and education groups with age effects restricted to not grow past age 50. Right panel. Pooled regression based on (2) with a linear in an occupation’s wage rank (10 occupations), as well as its linear interaction with age, fluidity and fluidity times age. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-level (in the left panel separately by education groups). Source: BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.

The boundary of the firm. My measure of fluidity captures only mobility between firms. Hence, less mobility across firms may be compensated for by higher mobility within firms. To the extent that large firms offer greater scope for within-firm mobility, this scenario may be particularly plausible if firms are smaller in more fluid labor markets. The middle right panel of Figure 5, however, shows that a larger share of workers work at large firms in more fluid labor markets. At face value, this suggests that the higher mobility rates in some countries is not driven by greater fragmentation of firms in such countries.

Workforce quality. The bottom panel of Figure 5 shows that labor market fluidity is not correlated with PISA test scores. Together with the finding that my results remain essentially unchanged when controlling for education or occupation-specific age slopes and the fact that my core sample includes only highly developed OECD countries, this observation leads me to discount the hypothesis that the steeper wage growth in more fluid labor markets is primarily accounted for by workforce composition.

---

16My measure is the share of employment at firms with 50 employees or more. This is the only firm size measure consistently available across countries and years, and is not available for all countries (Germany is missing).
Figure 5. Additional outcomes

(A) Level of inequality

(B) Growth in inequality

(C) Share of voluntary JJ

(D) Employment share of 50+ firms

(E) PISA test score

Note: Men aged 25–54. Panel A. St.d. of residual log hourly wages, controlling for year-education-age and year-occupation effects separately by country. Panel B. Growth between ages 25–29 and 50–54 in st.d. of residual log hourly wages, controlling for year-education-age and year-occupation effects separately by country. Panel C. Share of JJ transitions in which the worker quit as opposed to the firm initiated the separation (restricted to employees only). Panel D. Share of employees working at firms with 50 or more workers. Firm size is only available in the cross-sectional EUSILC, missing from the BHPS (I use the ECHP/EUSILC), coded into non-comparable bins in the GSOEP, and only available in 1999 and 2003–2015 in the PSID. Panel E. PISA math test score. Panels A–D. Constructed by first collapsing the outcome variable to the country-year-age level using the provided survey weights, then to the country-level. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-year-age level using the provided survey weights, then to the country-level. Source: BHPS, ECHP, EUSILC, GSOEP, OECD, and PSID 1991–2015.
Taking stock. Using detailed panel micro data, the above analysis establishes that wages grow more over the life-cycle in more fluid labor markets. This pattern is not accounted for by observable workforce characteristics, and the majority of it takes place within-jobs. While novel and intriguing, however, these correlations do not settle the main question this paper aims to answer: the impact of labor market fluidity on life-cycle wage growth. Moreover, I believe that there is little hope of finding an instrument for fluidity that offers both a sufficiently strong first stage and satisfies an exclusion restriction. For this reason, I instead proceed to develop a model to interpret the cross-country patterns.

3 Model

This section develops an equilibrium model of life-cycle careers that introduces endogenous skill accumulation following Ben-Porath (1967) into an otherwise standard search model (Bagger et al., 2014).

3.1 Environment

Time is continuous and infinite, there are no aggregate shocks and I focus on the long-run steady-state. The economy consists of a unit mass of finitely-lived workers and some positive mass of firms. All agents have linear preferences over a single output good discounted at rate $\rho$. Workers may allocate a unit of time indivisibly toward working or not working. As unemployed, a worker enjoys flow value of leisure $b(a, h)$, which may depend on her age $a$ and current human capital $h$.

Demographics. Workers are ex ante heterogeneous in initial skills $h_0 \sim \Lambda$ and enter the labor market as unemployed. They exit the labor market at age $A$ and are replaced with an equal mass of young workers. Upon retirement, workers enjoy a continuation value $\phi$ that is independent of their labor market history. Under linear utility, I may, without loss of generality, normalize the continuation value to zero, $\phi = 0$. I neither model the distribution of initial skills $\Lambda$, the retirement age $A$, nor the continuation value $\phi$, and I later take these to be identical across countries. The lack of a systematic relationship between labor market fluidity, on the one hand, and the overall labor force participation rate as well as that at age 25 and 54, on the other hand (see Appendix A.3), leads me to focus my analysis elsewhere.

Production technology. The single good of the economy is produced by one worker-one firm matches, and serves as the numeraire. A large number of potential firms may pay flow cost $c$ in return for the opportunity to meet with a worker. All costs are in terms of the final good. If a firm contacts a worker,
the two draw an idiosyncratic productivity \( z \) from exogenous offer distribution \( \Gamma \). Output is,

\[
y = zh
\]

As in Acemoglu and Pischke (1998) and Bagger et al. (2014), human capital and technology are complements. That is, a worker may use her skills more effectively in some jobs. As discussed further by Acemoglu (1997), a large empirical literature going back to Griliches (1969) has found evidence of complementarities between physical and human capital. I hence also abstract from purely firm-specific human capital, consistent with the evidence in Kambourov and Manovskii (2009) and Lazear (2009). Section 6 and Appendix D.4 document cross-sectional patterns consistent with these two assumptions.

**Human capital accumulation.** Employed workers may accumulate skills on-the-job. The human capital accumulation technology is independent of worker age \( a \). Building on Ben-Porath (1967), if a worker sets aside some fraction of her work time \( i \) toward building her skills, her human capital grows by

\[
\dot{h} = \frac{\mu}{\eta} (izh)^{\eta}, \quad \mu > 0, \ \eta \in (0, 1)
\]

The opportunity cost of training is foregone production, \( izh \). Human capital does not depreciate.\(^{17}\)

In Ben-Porath (1967), there is no productivity dispersion, \( z \), such that human capital grows by \( \mu(ih)^{\eta}/\eta \) at cost \( ih \). The current specification provides an extension to the case of employer heterogeneity. Note in particular that a given amount of time, \( i \), is assumed to grow human capital by more in a more productive firm. Absent this assumption, training falls in productivity, which is inconsistent with the positive correlation between training, on the one hand, and firm size and firm pay, on the other, that I document in Section 6 (conditional on worker fixed effects and time varying controls). In contrast, under the current assumption, the model matches these empirical patterns well. The view that workers may learn more in some jobs goes back to Rosen (1972) and more recently Jovanovic and Nyarko (1996).

**Labor market.** The labor market is characterized by informational frictions that prevent the immediate reallocation of workers to the jobs that use their skills the most efficiently. Both the unemployed and employed search for jobs at random in a common labor market, in the latter case with exogenous relative efficiency \( \phi \). Employed workers separate to unemployment at rate \( \delta(z) \). The dependence on productivity \( z \) captures in reduced-form the view that less productive jobs are more likely to separate in response to

\(^{17}\)An earlier version of this paper allowed for depreciation. As results were insensitive to (reasonable) variation in the extent of depreciation and the rate of depreciation was hard to identify from the available data, I opt to abstract from it.
idiosyncratic shocks. It allows the model to match the decline in EU mobility with age.

If firms create vacancies $V$ and workers search with efficiency $S$, total meetings are,

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

The job finding rate is hence $p = \chi (V/S)^\alpha$, while the worker finding rate of firms is $q = \chi (V/S)^{\alpha-1}$.

I adopt the bargaining protocol of Dey and Flinn (2005) and Cahuc et al. (2006), which has become a benchmark in the literature for its tractability and empirical relevance. A worker without a job who meets a firm with productivity $z$ gets a share $\beta$ of the difference between the value of the match and that of unemployment, henceforth the surplus. If a worker at a firm with productivity $z$ meets a potential new employer with productivity $z'$, the two firms first run a second price auction for the worker. This is won by the bidder with the higher valuation of the worker’s services, and it leaves the worker with the value of working for the least productive firm as her outside option. The worker and winning firm bargain over the differential surplus such that the worker receives a slice $\beta$ of the differential surplus. Following Barlevy (2002) and Bagger et al. (2014), wages are paid as a piece-rate $r$ of net output, $w = r(1 - i)zh$.

I assume that the worker makes the decision of how much to train, but that the amount of training can be contracted on between the worker and firm. That is, the firm can condition pay on the amount of training done by the worker. Without this assumption, the bargaining set may be non-convex, potentially rendering the bargaining protocol above invalid (Shimer, 2006). As I discuss further below, this assumption implies that it does not matter who pays for the training—the match will agree on the bilaterally optimal level and share the cost. The latter is consistent with lower entry wages in more fluid labor markets. More directly, Appendix B.2 shows that workers—conditional on worker fixed effects and time-varying covariates—earn lower hourly wages in years when they spend more hours on training.

### 3.2 Value functions

Let $U(a, h)$ be the value of unemployment to a worker of age $a$ with human capital $h$, $W(a, z, h, r)$ the value of employment of an age $a$ worker with human capital $h$ in a firm with productivity $z$ when paid piece rate $r$, and $J(a, z, h)$ the value of a match between a worker of age $a$ with human capital $h$ and a firm with productivity $z$. As I discuss in further detail below, the allocation is independent of how value is split between the worker and incumbent firm, and can hence be determined without alluding to the value of the worker. This appealing feature of the bargaining protocol serves to speed up estimation.

---

18Subject to the constraint that the worker cannot be made worse off by receiving a new offer.
Unemployment. The value of unemployment $U(a, h)$ solves the differential equation

$$
\rho U(a, h) = b(a, h) + p\beta \int_0^\infty \max \left\{ J(a, z, h) - U(a, h), 0 \right\} d\Gamma(z) + \frac{\partial U(a, h)}{\partial a}
$$

subject to the terminal condition $U(A, h) = 0$. The unemployed worker enjoys flow value of leisure $b(a, h)$. She meets potential employers at rate $p$, who are sampled from the offer distribution $\Gamma$. She accepts the job if it provides a positive surplus and she gets a slice $\beta$ of the surplus.

The value of a match. The value of a match $J(a, z, h)$ solves

$$
\rho J(a, z, h) = \max_{i \in [0,1]} \left\{ (1 - i)zh + \frac{\mu}{\eta} (izh)^\eta \frac{\partial J(a, z, h)}{\partial h} \right\} + \frac{\partial J(a, z, h)}{\partial a} + \phi p\beta \int_z^\infty J(a, z', h) - J(a, z, h) d\Gamma(z') + \delta(z) \left( U(a, h) - J(a, z, h) \right)
$$

subject to the boundary conditions $J(A, z, h) = 0$ and $J(a, z, h) \geq U(a, h)$. The match optimally invests in training and produces net output $(1 - i)zh$. The worker finds a new potential job at rate $\phi p$, drawn from the offer distribution $\Gamma$. If the new job is better than the current, she switches employer and gets a slice $\beta$ of the differential surplus. The match breaks up exogenously at rate $\delta(z)$, in which case the worker becomes unemployed and the firm gets continuation value zero.

The value of a worker and wages. Given an training rule $i(a, z, h)$ and reservation threshold $z(a, h)$ that solve the problem of the match (3), the value of a worker, $W(a, z, h, r)$, is for $z \geq z(a, h)$,

$$
\rho W(a, z, h, r) = r(1 - i(a, z, h))zh + \frac{\mu}{\eta} (i(z, h)zh)^\eta \frac{\partial W(a, z, h, r)}{\partial h} + \frac{\partial W(a, z, h, r)}{\partial a} + \phi p \int_0^z \max \left\{ J(a, z', h) + \beta \left( J(a, z, h) - J(a, z', h) \right) - W(a, z, h, r), 0 \right\} d\Gamma(z') + \phi p \int_z^\infty \left( J(a, z, h) + \beta \left( J(a, z', h) - J(a, z, h) \right) - W(a, z, h, r) \right) d\Gamma(z') + \delta(z) \left( U(a, h) - W(a, z, h, r) \right)
$$

subject to the boundary conditions $W(A, z, h, r) = 0$, $W(a, z, h) \geq U(a, h)$ and $W(a, z, h) \leq J(a, z, h)$. The worker receives a share $r$ of net output and grows her human capital by $\mu (i(z, h)zh)^\eta / \eta$. At rate $\phi p$, she receives outside job offers from offer distribution $\Gamma$. If the new productivity is lower than the current, she remains with her current firm, but potentially with an updated piece rate. If the new match is better
than the current, she switches jobs. Finally, she is subject to separation shocks.

Four wage policies determine wages. First, the wage \( w^u(a,z,h) \) characterizes a worker’s wage out of unemployment. Second, the wage \( w^e(a,z,z',h) \) gives the worker’s wage when simultaneously in contact with two employers \( z \) and \( z' < z \). Third, the wage \( w^r(a,z,h) \) ensures that a worker always prefers working in a viable match relative to unemployment. Fourth, the wage \( w^f(a,z,h) \) ensures that the firm always prefers to employ a worker in a viable match. These are defined by the conditions

\[
W(a,z,h, w^u(a,z,h)) = U(a,h) + \beta \left( J(a,z,h) - U(a,h) \right)
\]

\[
W(a,z,h, w^e(a,z,z',h)) = J(a,z',h) + \beta \left( J(a,z,h) - J(a,z',h) \right)
\]

\[
W(a,z,h, w^r(a,z,h)) = U(a,h)
\]

\[
W(a,z,h, w^f(a,z,h)) = J(a,z,h)
\]

**Equilibrium.** Appendix B.1 characterizes further how the evolution of the distribution \( G(a,z,h) \) of workers over age, productivity and human capital evolves as well as the distribution of unemployed workers over age and human capital, \( u(a,h) \). Given these objects, free entry requires

\[
c = (1 - \beta)q \frac{\mu}{S} \int_0^\infty \int \max \left\{ J(a,z,h), 0 \right\} du(a,h) d\Gamma(z)
\]

\[
+ \ (1 - \beta)q \frac{\phi(1 - u)}{S} \int_0^\infty \int J(a,z,h) - J(a,z',h) dG(a,z',h) d\Gamma(z)
\]

where \( u \) is the aggregate unemployment rate. In return for flow cost of a vacancy \( c \), the firm contacts a potential hire at rate \( q \). The first term is the return from meeting an unemployed potential hire and the second term the payoff from contacting an employed potential hire. In both cases, the new potential match draws a productivity from \( \Gamma(z) \) and the firm gets a slice \( 1 - \beta \) of any match that is formed.

### 3.3 Qualitative insights

Before turning to a quantitative analysis, I provide some qualitative insights in a simplified version of the model. To that end, suppose instead that time is discrete and has two periods, there is no discounting, and productivity can take a low, \( z = z_1 \), or a high value, \( z = z_2 > z_1 \), with equal probability. Moreover, I abstract from unemployment to focus on how training and poaching interact. Specifically, I assume that young workers enter in low productive matches, workers do not separate exogenously, employed workers search with \( \phi = 1 \) intensity, and recruiting firms compete for young workers going into the second period. Figure 6 illustrates the timing of the simplified two-period model.
Denote by $W(r, i)$ the value to a young worker of being in a low productive firm under some amount of training $i$ and piece rate $r$. Without loss of generality, I normalize initial human capital, $h_0 = 1$, and abstract from it as an argument to simplify the notation. The value $W(r, i)$ solves,

$$W(r, i) = (1 - i)z_1 r + \left((1 - p)z_1 r + \frac{p}{2}z_1 + \frac{p}{2}(z_1 + \beta(z_2 - z_1))\right) \left(1 + \frac{\mu}{\eta}(z_1 i)^\eta\right)$$

(4)

The worker is paid piece rate $r$ on net output $(1 - i)z_1$ (the level of training need not be optimal at this point). The worker’s human capital rises to $1 + \mu(z_1 i)^\eta / \eta$ in the second period. The worker receives no new job offer with probability $1 - p$, in which case she continues to be paid piece rate $r$. With probability $p/2$, the worker receives a job offer from another low productive firm. I impose the tie-breaking rule that an indifferent worker switches employer (this only impacts the measured JJ mobility, though). The worker extracts the full surplus of the current match, i.e. she gets an updated piece rate equal to the productivity of the match, $z_1$. Finally, with probability $p/2$, the worker receives a job offer from a high-productive firm, in which case she switches employer and gets a piece rate that reflects the full value of the least productive match, $z_1$, plus a share $\beta$ of the differential value, $z_2 - z_1$.

Denote by $F(r, i)$ the value to an incumbent firm of employing a young worker under some piece rate $r$ and training policy $i$ (again, not necessarily the optimal one). It satisfies,

$$F(r, i) = (1 - i)z_1 (1 - r) + (1 - p)z_1 (1 - r) \left(1 + \frac{\mu}{\eta}(z_1 i)^\eta\right)$$

(5)

The incumbent firm makes profits $1 - r$ per net output, $(1 - i)z_1$. With probability $1 - p$, the worker receives no outside offer and the incumbent firm makes profits $1 - r$ per net output in the second period. If the worker receives an outside offer, the incumbent firm makes no profit in the second period.\(^\text{19}\)

**Proposition 1.** For any level of training $i$, the joint value of a match between a young worker and an incumbent firm

\(^{19}\)Even if the worker had remained, the incumbent firm would have had to pay the worker her full marginal product.
firm, \( J(i) = W(r, i) + F(r, i) \), is independent of how it is split between the worker and firm, \( r \),

\[
J(i) = (1 - i)z_1 + \left( z_1 + \frac{p}{2} \beta (z_2 - z_1) \right) \left( 1 + \frac{\mu}{\eta} (z_1 i)^{\eta} \right)
\]  

(6)

**Proof.** All proofs are in Appendix B.3.

Differentiating the joint value (6) with respect to the job finding rate, \( p \), holding training fixed,

\[
\frac{\partial J(i)}{\partial p} \bigg|_{i = 1} = \frac{1}{2} \beta (z_2 - z_1)(1 + \mu (z_1 i)^{\eta} / \eta) > 0.
\]  

Hence, *ceteris paribus*, the joint value increases in the rate at which the worker switches employer (assuming \( \beta > 0 \)). Although the incumbent employer ex post loses value when the worker leaves, the worker is compensated by the poaching firm with the full value of the incumbent match plus a share \( \beta \) of the differential surplus. From the joint perspective of the incumbent match, the gain to the worker more than offsets the loss to the incumbent firm. Ex ante, an incumbent firm benefits from the opportunity of the worker leaving by having to pay the worker less.

**Proposition 2.** Optimal investment chosen by the worker maximizes the bilateral surplus,

\[
i = \frac{1}{z_1} \left( \mu \left( z_1 + p \beta z_2 - z_1 \right) \right)^{1/\eta}
\]  

(7)

The optimal training policy (7) increases in the job finding rate, \( p \). Although as in Acemoglu and Pischke (1998) a higher probability that the worker meets a new employer lowers the value of human capital to the incumbent firm, it increases the value of human capital to the worker. When workers’ bargaining power, \( \beta \), is positive, the latter effect outweighs the former. The reason is that a higher arrival rate of outside offers allows the worker to use her skills at an employer that values them higher. Because the incumbent match gets (partly) compensated for this, it raises the value of human capital to the incumbent match. As a result, the match invests more in response to a higher arrival rate of outside offers. Allowing for JJ mobility is critical to this argument as it gives the worker the chance to re-bargain using the value of the current match as benchmark, and not the value out of unemployment.

This conclusion clearly depends on the stipulated bargaining protocol, which ensures that a JJ mover obtains a share of the additional value of human capital in a new match. Nonetheless, I believe that the insight is more general. For instance, in a partial equilibrium version of the model in the spirit of McCall (1970) in which workers sample exogenously given piece rates per human capital, if workers sample outside offers more frequently and hence expect to grow their piece rate faster, it would encourage human capital accumulation. The current framework may be viewed as a general equilibrium version of that model, in which the piece rate is determined in equilibrium via bargaining.
That being said, there exists alternative models of equilibrium wage setting—most prominently those of wage posting—in which cases could arise where the incumbent firm loses more value than the worker gains when she moves to a new employer. In such cases, I hypothesize that training may fall in the poaching rate (over parts of the domain for productivity, the gain to the worker may still outweigh the loss to the firm, though). Such a model, however, would likely be significantly more difficult to solve, as the firm would have to internalize how its wage offer affects workers’ incentives to train. Hence, this remains only a hypothesis. In any case, such cases are arguably the least appealing feature of wage posting models, as the worker and incumbent firm would have a very strong incentive to renegotiate the contract since they could both benefit from doing so. But for some (unmodeled) reason, such renegotiation is ruled out. For this reason, I prefer my framework for the particular question at hand. Still, the model allows for the possibility that the incumbent match does not benefit from poaching (i.e. \( \beta = 0 \)).

**Equilibrium.** If there is poaching, the cost of doing so, \( c \), must equal its expected return,

\[
c = \frac{1}{2}q(1 - \beta)(z_2 - z_1) \left( 1 + \frac{\mu}{\eta}(z_1i)^\eta \right)
\]

(8)

A poaching firm contacts a worker at rate \( q \) and with probability 0.5 the potential match draws a high productivity. Only in this case is the poaching firm successful and gets slice \( 1 - \beta \) of the surplus. If the worker invested \( i \) in training in the first period, human capital in the second period is \( 1 + \mu(z_1i)^\eta / \eta \).

**Definition 1 (Stationary search equilibrium).** A stationary search equilibrium with positive vacancy creation consists of a value function, \( J \); a training policy, \( i \); and a mass of vacancies, \( v \), such that

1. The value function and training policy maximizes (6) given a mass of vacancies;
2. The mass of vacancies is consistent with free entry (8);
3. And the economy is time invariant.

The equilibrium is characterized by two curves. The first is a training curve, which can be derived from the first-order condition (7) by substituting for the job finding rate using the matching function,

\[
i(v) = \frac{1}{z_1} \left( \mu \left( z_1 + v^\alpha \beta \frac{z_2 - z_1}{2} \right) \right)^{\frac{1}{\eta}}
\]

(9)

The second curve is a job creation curve, derived by substituting the worker finding rate in the free
entry condition (8) (I normalize the scalar in the matching technology, $\chi = 1$, to minimize notation),

$$v(i) = \left(1 - \beta \left(\frac{z_2 - z_1}{2c^\alpha} \left(1 + \frac{\mu}{\eta (z_1 i)^\eta}\right)^{\frac{1}{\alpha}}\right)\right)^{\frac{1}{1-\alpha}}$$

Figure 7 graphs the training and job creation curves (9)–(10) in $v$-$i$ space. At zero vacancies, $v = 0$, the training curve (9) is positive. It subsequently rises in vacancies. A higher job finding rate raises the return to human capital since workers expect to use it more efficiently, encouraging training. At zero training, $i = 0$, the job creation curve (10) is positive. Optimal job creation subsequently increases in the training rate of workers. If workers invest more, matches produce more output, and poaching firms get a share of this. The two curves may never cross or may cross multiple times.

**Figure 7. Comparative-static impact of a higher cost of hiring**

(A) Low cost

(B) High cost

Note: The comparative-static equilibrium impact of a higher cost of creating jobs, $c$. Training: Equation (9). Job creation: Equation (10).

**Proposition 3.** If $(1 - \eta) / \eta > \alpha / (1 - \alpha)$, the economy admits a unique stationary equilibrium.

The key parameters governing whether the economy may display multiple stationary equilibria are the curvatures of the matching technology, $\alpha$, and the human capital accumulation technology, $\eta$. If it is easier to scale up training (i.e. $\eta$ is higher), a given increase in job creation leads to a stronger optimal increase in training. If the worker finding rate declines less with an increase in vacancies (i.e. $\alpha$ is higher), a given increase in training leads to a stronger optimal increase in job creation.

**Proposition 4.** Suppose $(1 - \eta) / \eta > \alpha / (1 - \alpha)$ and that workers’ bargaining power is positive, $\beta > 0$. A higher cost of creating jobs, $c$, is associated with lower average match productivity and human capital.

In response to a higher cost of vacancy creation, poaching firms create fewer vacancies for any given level of training of workers. That is, the job creation curve in Figure 7 shifts to the left. The lower job
finding rate, in turn, reduces the expected value of human capital to workers, lowering training.

The planning problem. Before I go to the data, I briefly turn to the question of the efficiency of the decentralized equilibrium. The planning problem is to maximize life-time output minus costs,

$$\max_{v,i} \left\{ (1 - i)z_1 + \left( z_1 + v^\alpha \frac{1}{2} (z_2 - z_1) \right) \left( 1 + \frac{\mu}{\eta} (z_1 i) \right)^\eta - cv \right\}$$

with first order conditions,

$$i_{sp}(v) = \frac{1}{z_1} \left( \mu \left( z_1 + v^\alpha \frac{z_2 - z_1}{2} \right) \right)^{1/\eta}, \quad v_{sp}(i) = \left( \frac{\alpha z_2 - z_1}{2c} \left( 1 + \frac{\mu}{\eta} (z_1 i) \right)^\eta \right)^{1/\alpha}$$

Lemma 1. Suppose \((1 - \eta)/\eta > \alpha/(1 - \alpha)\) and that training in the decentralized economy equaled the constrained first best, \(i(v) = i_{sp}(v)\). Then the decentralized equilibrium attains the constrained first best number of vacancies iff the elasticity of matches w.r.t. vacancies equals firms’ bargaining power, \(\alpha = 1 - \beta\).

Abstracting from the training decision, the model reproduces a well-known property of this class of models going back to Hosios (1990). Poaching firms face two externalities in their vacancy creation decision. On the one hand, they do not internalize the fact that part of the gain from job creation accrues to workers. On the other hand, they do not internalize the fact that when they create jobs, they lower the worker finding rate of other firms. It turns out that when firms’ bargaining power equals the elasticity of the matching function with respect to vacancies, these two forces exactly offset.

Lemma 2. Suppose \((1 - \eta)/\eta > \alpha/(1 - \alpha)\) and that job creation in the decentralized economy equaled the constrained first best, \(v(i) = v_{sp}(i)\). Then the decentralized equilibrium attains the constrained first best amount of training iff workers’ bargaining power is one, \(\beta = 1\).

Due to a positive externality of training on future employers, for a given number of vacancies the decentralized economy in general features less training in human capital relative to the social optimum, as in Stevens (1994) and Acemoglu (1997). Only if incumbent matches enjoy the full marginal benefit from additional training upon a JJ move would they undertake the socially optimal level of training.

Proposition 5. Suppose \((1 - \eta)/\eta > \alpha/(1 - \alpha)\). There exists no bargaining power of workers \(\beta \in [0, 1]\) such that the decentralized search equilibrium coincides with the constrained optimal allocation.

Only if workers’ bargaining power \(\beta\) is one would training in the decentralized and planned economy coincide, given vacancies. For vacancy creation in the decentralized equilibrium to coincide with the constrained first best under such a high \(\beta\), the elasticity of matches with respect to vacancies \(\alpha\) would
have to be zero, violating the assumption that $\alpha \in (0, 1)$. Moreover, by leaving nothing for the poaching firm, such a high bargaining power of workers is inconsistent with positive job creation in equilibrium.

### 4 Estimation

This section estimates the model targeting the US as a high-fluidity country. The next section introduces wedges to firms’ cost of creating jobs in the spirit of Hsieh and Klenow (2009) to assess the impact of cross-country differences in labor market fluidity on workers’ optimal behavior in the estimated model.

#### 4.1 Methodology

I externally calibrate the discount rate, $\rho$, to a four percent annual real interest rate. The available data do not allow identification of the curvature of the matching function, so I set $\alpha = 0.5$ based on Petrongolo and Pissarides (2001). I estimate nine parameters using SMM (Gourieroux et al., 1993) to minimize the sum of weighted squared percentage deviations between a set of moments in the model and the data,

$$p^* = \arg \min_{p \in \mathcal{P}} \sum_i w_i \left( \frac{m^\text{model}_i(p) - m^\text{data}_i}{m^\text{data}_i} \right)^2$$

As discussed further below, some moments particularly inform some parameters. I set the weights $w_i$ such that each set of moments particularly informing one parameter receives the same total weight (that is, if $n$ moments particularly inform parameter $p$, each of these moments receives a weight $1/n$). The two exceptions are total life-cycle wage growth and aggregate labor market fluidity, which I assign five times this weight given their key role in the analysis.

I solve the model and compute moments in continuous time, which allows me to correctly time-aggregate to any desired frequency.\(^{20}\) It is difficult, however, to derive a law of motion for labor market fluidity. Hence, I compute it based on a simulated discrete-time, monthly approximation for 13 months for 800,000 individuals uniformly distributed between ages 25–54, where each age group is initialized from its age-conditional distribution over employment state, human capital and productivity. This corresponds to my empirical measure, which assigns equal weight to each age 25–54.

I assume that firm productivity, $z$, is Pareto distributed with tail index $1/\zeta$. I assume that initial human capital, $h_0$, is drawn from a Pareto distribution with tail index $1/\sigma$. I parameterize the separation

---

\(^{20}\)I solve the model on a discretized grid for productivity, human capital, piece rates and age with 25, 10, 8 and 6 grid points, respectively. The age grid points are set to reflect real life ages of 24–29, 30–34, 35–39, 40–44, 45–49, 50–64. I approximate between these age bins to obtain more disaggregated age outcomes.
rate, \( \delta(z) = \delta_0 e^{-\delta_1 (\ln z - \ln z_0)/(\ln z - \ln z_0)} \), where \( z_0 \) is the maximum (minimum) productivity on the discretized grid for productivity. The flow value of leisure \( b(a, h) \) is recovered such that workers of each age and human capital level are indifferent between unemployment and working at the second lowest grid point for productivity. The nine parameters to estimate internally are,

\[
p = \left\{ \mu, \eta, \sigma, \zeta, p, \phi, \delta_0, \delta_1, \beta \right\}
\]

While the estimation is joint, it is nevertheless useful to provide a heuristic discussion of what moments particularly inform what parameter. The scalar in the human capital accumulation technology, \( \mu \), and its curvature, \( \eta \), are jointly informed by the life-cycle wage profile. If \( \mu \) is higher, wages in general grow more over the life-cycle. If \( \eta \) is higher, the marginal return to investment falls less with investment such that investment is more front loaded and the wage profile more concave.

Dispersion in initial human capital, \( \sigma \), as well as the shape of the offer distribution, \( \zeta \), are informed by the life-cycle profile of the standard deviation of wages. If \( \sigma \) is larger, inequality is greater. If \( \zeta \) is larger—the tail of the productivity distribution is fatter—inequality grows more with age. The job finding rate \( p \) targets the aggregate UE rate. While \( p \) is an endogenous outcome, the cost of a vacancy, \( c \), is a free parameter. I set it ex post to rationalize the job finding rate. The relative search efficiency of employed workers, \( \phi \), is set to target the aggregate JJ mobility rate. If \( \phi \) is higher, JJ is greater. The intercept, \( \delta_0 \), and slope, \( \delta_1 \), of the separation rate are jointly informed by the life-cycle EU rate. If \( \delta_0 \) is higher, the EU rate is generally higher, while if \( \delta_1 \) is higher, the separation rate falls more with productivity and hence also with age as workers move up the job ladder with age. Finally, workers’ bargaining power, \( \beta \), is informed by wage gains upon a JJ move, because if it is larger, wage gains from moving are more front loaded. To compute the empirical counterpart, I additionally rely on data from the US SIPP.

### 4.2 Estimates and model fit

Table 3 summarizes the parameter estimates, expressed at a monthly frequency. While it is not known whether the condition for uniqueness in Proposition 3 extends to the richer model, the low estimated curvature of the human capital accumulation technology, \( \eta \), suggests that the equilibrium may be unique.

---

21I use the SIPP for this particular moment because the PSID became biannual in 1997, leaving me with only a few years of annual wage growth observations. An additional advantage of the SIPP is that I can compute the measure of wage gains upon a JJ move at a monthly frequency, which avoids the need to simulate this moment in estimation (it is difficult to derive analytically a KFE for the annual wage growth of JJ movers). The annual measure in the PSID for 1994–1997 lines up well with the monthly measure in the SIPP for 1996–2013, however. I describe the SIPP in more detail in Appendix A.1.

22I normalize the flow value of leisure such that workers are indifferent between unemployment and employment at the second grid point on the discretized grid for human capital to avoid any potential numerical issues associated with the boundary of the grid. Hence, the estimated job finding rate \( p \) is higher than the actual UE rate, since some job offers are not accepted.
Moreover, I have not uncovered any evidence of multiplicity. I estimate a relatively high concavity of the Ben-Porath (1967) technology because job shopping serves to increase the concavity of the wage profile, requiring a less elastic human capital margin. The employed search by roughly 40 percent of the intensity of the unemployed. Workers’ bargaining power $\beta$ is 0.32, with an implied labor share of 82 percent (for comparison, Bagger et al., 2014, estimate $\beta \approx 0.3$ with an implied labor share of 81–85 percent).

### Table 3. Parameter estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Targeted moment</th>
<th>Model</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Externally set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$ Discount rate</td>
<td>0.003</td>
<td>4% annual real interest rate</td>
<td>0.761</td>
<td>0.760</td>
</tr>
<tr>
<td>$\alpha$ Elasticity of matches w.r.t. vacancies</td>
<td>0.500</td>
<td>Petrongolo and Pissarides (2001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B. Internally estimated</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu$ Drift of human capital</td>
<td>0.001</td>
<td>Life-cycle wage growth</td>
<td>0.761</td>
<td>0.760</td>
</tr>
<tr>
<td>$\eta$ Curvature of human capital production function</td>
<td>0.497</td>
<td>Life-cycle wage profile</td>
<td>See Figure 8</td>
<td></td>
</tr>
<tr>
<td>$\sigma$ Initial human capital dispersion</td>
<td>0.491</td>
<td>Life-cycle inequality profile</td>
<td>See Figure 8</td>
<td></td>
</tr>
<tr>
<td>$\zeta$ Shape of productivity distribution</td>
<td>0.154</td>
<td>Life-cycle inequality profile</td>
<td>See Figure 8</td>
<td></td>
</tr>
<tr>
<td>$p$ Job finding rate</td>
<td>0.472</td>
<td>Aggregate UE rate (monthly)</td>
<td>0.226</td>
<td>0.230</td>
</tr>
<tr>
<td>$\phi$ Relative search efficiency of employed</td>
<td>0.394</td>
<td>Aggregate JJ rate (annual)</td>
<td>0.100</td>
<td>0.100</td>
</tr>
<tr>
<td>$\delta_0$ Separation rate, intercept</td>
<td>0.016</td>
<td>Life-cycle EU rate</td>
<td>See Figure 8</td>
<td></td>
</tr>
<tr>
<td>$\delta_1$ Separation rate, slope in $z$</td>
<td>3.131</td>
<td>Life-cycle EU rate</td>
<td>See Figure 8</td>
<td></td>
</tr>
<tr>
<td>$\beta$ Worker bargaining power</td>
<td>0.321</td>
<td>Wage gain upon a JJ move</td>
<td>0.079</td>
<td>0.082</td>
</tr>
</tbody>
</table>


Figure 8 illustrates the model fit to life-cycle dynamics. Wages grow rapidly early in careers, as workers have much scope to move up the job ladder and bargain up their wage, and face high returns to training. As noted above and discussed further in Appendix A, the standard deviation of residual wages is essentially flat in the data. The model matches this reasonably well. It understates somewhat the decline in the EU rate. The EU rate falls as workers move away from low productive matches with a high likelihood of breaking up. To a first order, the UE rate is flat over the life-cycle in both the model and data. Finally, the model matches well the fall in JJ mobility, as workers gradually find a good job.

### 4.3 The workings of the model

Before I turn to the cross-country analysis, I pause to illustrate the workings of the model in the cross-section. The left panel of Figure 9 plots the optimally chosen growth rate of human capital, $h(a, z, h)/h$. Assuming that the solution for optimal investment is away from the corner, $i \in (0, 1)$, taking first-order

---

23The slight "jumpiness" of the model wage profile is due to the approximation to six underlying age groups. The model does not fully match the curvature of the wage profile in the data, potentially raising concern about the empirical restriction to zero wage growth after age 50. While in my benchmark results I do not impose this restriction in the model, I show in Appendix C.1 that the cross-country patterns for life-cycle wage growth are virtually identical if I impose it.
Figure 8. Model fit

(A) Mean Log Wage

(B) Std. of Log Wages

(C) Monthly EU Rate

(D) Monthly UE Rate

(E) Annual JJ Rate

Note: Men aged 25–54. Panel A. Average log real hourly wage. Data: Based on regression (1) controlling for worker fixed effects, year effects and age effects restricted to not grow after age 50. Panel B. Standard deviation of residual log hourly wages. Data: Residual wages controlling for year-education-age and year-occupation effects and dropping the bottom 1% of wages. Panel C. Share of employed workers who are unemployed in the subsequent month. Data: Employment includes self-employment due to data limitations. Panel D. Share of unemployed workers who are employed in the subsequent month. Data: Employment includes self-employment due to data limitations. Panel E. Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Data: Employment in the past 12 months includes self-employment due to data limitations. All panels. Empirical moments constructed by first collapsing to the age-year level using the provided survey weights, then to the age level. Source: Model and PSID 1994–2015.
conditions of the problem of the match (3), and rearranging, it is given by

\[
\frac{\dot{h}(a, z, h)}{h} = \frac{\mu}{\eta} \left( \frac{\partial J(a, z, h)}{\partial h} \right)^{\frac{\eta}{\mu}} \times \frac{1}{h}
\]

The value function is close to affine in human capital, so growth in human capital is close to independent of \( h \). The growth rate hence falls in human capital. The marginal value of human capital rises in productivity, \( \partial^2 J(a, z, h) / \partial h \partial z > 0 \), due to the complementarity in production. Hence, the growth rate of human capital increases in productivity. That is, workers who find a high productive job respond by growing their human capital more, propagating the impact of luck in the labor market.\(^{24}\)

**FIGURE 9. WORKINGS OF THE MODEL**

(A) OPTIMAL GROWTH RATE OF HUMAN CAPITAL

(B) DECOMPOSING LIFE-CYCLE WAGES

*Note: Panel A. Optimally chosen growth rate of human capital, \( \dot{h}/h \), of first age group as a function of productivity and human capital. Panel B. Decomposition of average life-cycle wage growth based on the accounting identity (12). Source: Model.*

The right panel of Figure 9 decomposes life-cycle wage growth based on the identity,

\[
\ln w = \ln z + \ln h + \ln(1 - i) + \ln r
\]

Growth in human capital, \( h \), is the most important source of life-cycle wage growth, with growth in productivity a close second. Human capital accumulation, however, does not come for free. Indeed, a

\(^{24}\)An earlier version of this paper found that because of this interaction between search and training, the model can generate a significant increase in human capital dispersion over the life-cycle if estimated to fit the increasing dispersion in raw inequality over the life-cycle in the US, despite not featuring heterogeneity in learning ability. As discussed further in Appendix A.7, however, most of the increase in inequality over the life-cycle takes place between groups. Since the model is about residual inequality, the current version of the paper instead targets life-cycle profiles of residual inequality.
fall in the amount of time spent on training is also an important source of growth in life-cycle wages, with young workers earning lower wages partly to offset the cost of training. Finally, the piece rate grows as workers gradually bargain up their wage through counteroffers.

5 Labor market fluidity and human capital accumulation

I now turn to the effect of differences in labor market fluidity on worker careers. In the spirit of Hsieh and Klenow (2009), I introduce wedges to firms’ optimal job creation condition (8) such that the model matches cross-country differences in labor market fluidity (i.e. I adjust $c$).\textsuperscript{25} This approach purposefully avoids taking a strong stand on the specific policies that lead to differences in labor market fluidity, as this has already been extensively studied in the literature (see, e.g., Pries and Rogerson, 2005, and references therein). In Section 7, I return to the question of what drives differences in fluidity to document that, consistent with the literature, fluidity correlates negatively with measures of employment protection legislation, the cost of doing business and labor taxes. In this section, however, I focus on the key novelty of this paper, which is to take such differences as given to quantify their impact on worker behavior.

5.1 The effect of labor market fluidity

I start by contrasting the predictions of the model with the data, focusing on the core sample (similar results hold for the full sample in Appendix C.2). The top left panel of Figure 10 shows that life-cycle wage growth is greater in more fluid labor markets, while the top right panel illustrates that unemployment is lower (a similar pattern holds for wage employment rates in Appendix C.6).\textsuperscript{26} As I show in Appendix C.5, this is the result of both a higher UE and a lower EU rate in more fluid labor markets.

The bottom panel highlights that labor productivity is higher in more fluid labor markets, constructed as real PPP-adjusted GDP per hour in the data. An important caveat is that the model is estimated for men, while productivity is only available at the economy level. I construct two outcomes in the model. The first is total output net of training costs and the cost of job creation including the wedge, $\Delta$, divided by employment: $\left(1-u\right)\int\left(1-i\right)zhdG(a,z,h) - (c + \Delta)V / \left(1-u\right)$. I refer to this specification as regulation since it reflects the view that the wedge is a waste of resources, for instance due to red tape. The second is total output net of training costs and the cost of job creation excluding the wedge:

\textsuperscript{25}I allow the flow value of leisure to adjust such that each human capital group is indifferent between unemployment and employment at the least productive firm on the grid for productivity. Not doing this, however, only marginally affect results.

\textsuperscript{26}The model understates the unemployment rate in the US, because the targeted EU and UE rates are inconsistent with the implied unemployment rate from a flow-balance equation. I prefer to get the flows right, as opposed to the stocks, but I have re-estimated the model targeting instead the EU and unemployment rates (hence missing the UE rate), with similar results.
\((1 - u) \int (1 - i) zhdG(a, z, h) - cV) / (1 - u)\). This specification, which I refer to as tax, reflects the view that the wedge is tax revenue that is ultimately used productively. Across the core sample of OECD countries, labor productivity is 8.8 percent lower than the US under the tax view and 16.0 percent lower under the regulation view (7.9 percent and 14.0 percent across the full sample, respectively). The real world is likely somewhere in between these two extreme views, highlighted in purple.

**Figure 10. The Impact of Labor Market Fluidity, Model versus Data**

(A) Life-cycle wage growth

(B) Unemployment rate

(C) Labor productivity

*Note:* Men aged 25–54. Panel A. Wage growth between age 25–50. Data: Based on (1) with worker fixed effects, year effects and age effects restricted to not grow after age 50. Panel B. Data: Collapsed to the country-year-age level using the provided survey weights, then to the country-level. Model: Collapsed to the country-age level, then to the country-level. Panel C. Data: Log real output per hour in 2014 in 2004 PPP-adjusted US dollars. Regulation (model): Output per worker net of training and job creation costs including the wedge, \(\Delta,(1 - u) \int (1 - i(a, z, h)) zhdG(a, z, h) - (c + \Delta)^V) / (1 - u)\). Tax (model): Output per worker net of training and job creation costs excluding the wedge per worker, \((1 - u) \int (1 - i(a, z, h)) zhdG(a, z, h) - cV) / (1 - u)\). Both model series are in logs and normalized to the data for the US. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Data: Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the survey weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age-level using the survey weights, then to the country-level. Source: Model, BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.

27 The levels are not comparable, so I normalize the model moment to the data moment in the US. Neither of these specifications is isomorphic to welfare, since none accounts for the forsaken value of leisure. Adjusting the tax view to also reflect forgone leisure, welfare is 11 percent lower across the core countries relative to the US (10 percent across the full sample).
Table 4 summarizes these predictions for life-cycle wage growth (see Appendix C.4 for the same tables for other outcomes). It shows for each country the difference in life-cycle wage growth to the US in the model and the data. The mechanism emphasized here accounts for 50 percent of the systematic relationship between life-cycle wage growth and labor market fluidity in the data (60 percent in the full sample of 23 countries). Moreover, while the correlation between life-cycle wage growth and fluidity is high, there is also a significant amount of orthogonal variation. Hence, while non-trivial, the forces emphasized here are far from the sole driver of cross-country dispersion in life-cycle wage growth—see, for instance, Guvenen et al. (2013) for a complementary mechanism.

Table 4. Life-cycle wage growth and labor market fluidity, model versus data

<table>
<thead>
<tr>
<th></th>
<th>AT</th>
<th>BE</th>
<th>DK</th>
<th>FI</th>
<th>FR</th>
<th>DE</th>
<th>EL</th>
<th>IE</th>
<th>IT</th>
<th>NL</th>
<th>ES</th>
<th>UK</th>
<th>US</th>
<th>Core</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>0.07</td>
<td>0.12</td>
<td>0.48</td>
<td>0.34</td>
<td>0.24</td>
<td>0.54</td>
<td>0.09</td>
<td>0.53</td>
<td>0.21</td>
<td>0.60</td>
<td>0.54</td>
<td>0.64</td>
<td>0.44</td>
<td>0.70</td>
<td>0.76</td>
</tr>
<tr>
<td>∆US_d</td>
<td>-0.69</td>
<td>-0.64</td>
<td>-0.22</td>
<td>-0.55</td>
<td>-0.22</td>
<td>-0.67</td>
<td>-0.21</td>
<td>-0.46</td>
<td>-0.22</td>
<td>-0.32</td>
<td>-0.06</td>
<td>0.00</td>
<td>-0.39</td>
<td>-0.29</td>
<td></td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>0.51</td>
<td>0.49</td>
<td>0.70</td>
<td>0.57</td>
<td>0.48</td>
<td>0.53</td>
<td>0.45</td>
<td>0.53</td>
<td>0.48</td>
<td>0.64</td>
<td>0.57</td>
<td>0.77</td>
<td>0.76</td>
<td>0.56</td>
<td>0.58</td>
</tr>
<tr>
<td>∆US_d</td>
<td>-0.25</td>
<td>-0.27</td>
<td>-0.06</td>
<td>-0.19</td>
<td>-0.28</td>
<td>-0.23</td>
<td>-0.23</td>
<td>-0.28</td>
<td>-0.12</td>
<td>-0.19</td>
<td>0.01</td>
<td>0.00</td>
<td>-0.20</td>
<td>-0.17</td>
<td></td>
</tr>
<tr>
<td>∆US_m/∆US_d</td>
<td>0.37</td>
<td>0.41</td>
<td>0.22</td>
<td>0.45</td>
<td>0.51</td>
<td>1.02</td>
<td>0.45</td>
<td>1.10</td>
<td>0.61</td>
<td>0.63</td>
<td>0.60</td>
<td>-0.26</td>
<td>- -</td>
<td>0.50</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Note: Men aged 25–54. ∆US: Difference in life-cycle wage growth relative to the US. ∆US_m/∆US_d: Difference in life-cycle wage growth relative to the US in the model relative to the data. Core: Average difference relative to the US across the 12 core countries. All: Average difference relative to the US across the 22 countries in the full sample. Data moments are based on regression (1) with worker fixed effects, year effects and age effects restricted to not grow after age 50. Source: Model, BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.

Appendix C.5 shows that the mechanism matches well the lack of a pronounced relationship between labor market fluidity and various inequality outcomes. The one aspect of the data it fails to fully capture is the lack of correlation between the wage gain upon a JJ move and labor market fluidity. The model predicts that this should decline with fluidity, as workers in more fluid labor markets are higher up the job ladder and hence have less scope to further climb it. I discuss in Appendix C.5 an extension to the model that allows also for so called godfather shocks (Jolivet et al., 2006), which resolves this tension without changing any of my main conclusions. Appendix C.3 conducts a sensitivity analysis.

5.2 Understanding the effects

I now turn to understanding the channels through which labor market fluidity impacts worker careers via a series of decomposition and counterfactual experiments in the model.

The importance of human capital. The left panel of Figure 11 decomposes life-cycle wage growth into the components in accounting identity (12). Specifically, I take the difference in log wages and each of its components between ages 50 and 25. Match productivity grows by 13 log points more in the US relative
to the least fluid labor market, while human capital grows by eight log points more. Because the cost of training is shared with workers through a lower wage and young workers train more in more fluid labor markets, the investment component contributes four log points to steeper wage growth in the US relative to the least fluid labor market market. Finally, the piece rate grows by six log points more, as workers receive more outside offers allowing them to bargain up the wage.

Unemployment, job ladder and incentive effects. Let \( e(a, h) \) be the number of employed workers of age \( a \) with human capital \( h \), \( \tilde{g}(z|a, h) \) the conditional distribution of employed workers of age \( a \) with human capital \( h \) over productivity \( z \), and \( h(a, z, h) \) the optimally chosen growth of human capital of workers of age \( a \) with human capital \( h \) employed in a firm with productivity \( z \). Since \( 1/A \) is the number of workers of age \( a \), the average growth of human capital of workers of age \( a \), \( d\bar{h}(a) \), is

\[
\frac{d\bar{h}(a)}{\text{growth of human capital}} = \frac{1}{A} \int \int \tilde{g}(z|a, h) h(a, z, h) dz \, \tilde{g}(z|a, h) dh
\]

incentive effect  
job ladder effect  
unemployment effect

Hence, labor market fluidity impacts human capital accumulation through three channels.

First, workers spend more time in unemployment in less fluid labor markets, since the job finding rate is lower (i.e. \( e(a, h) \) changes). Although skills do not depreciate when unused, workers cannot accumulate skills in unemployment. Consequently, \( ceteris paribus \), a higher incidence of unemployment reduces skill accumulation. I refer to this as the unemployment effect. I isolate its role by holding all decision rules as well as the job finding rate from employment fixed at their estimated US values, but adjust the job finding rate from unemployment as in the benchmark cross-country experiment.\(^{28}\)

Second, workers move up the job ladder faster in more fluid labor markets (i.e. \( \tilde{g}(z|a, h) \) changes). Workers higher up the job ladder grow their human capital more, because they expect to have greater use for it (recall Figure 9). Hence, workers accumulate more skills in more fluid labor markets, since they move faster to matches were they train a lot. I refer to this as the job ladder effect. Starting from the unemployment effect counterfactual, I compute it as the incremental effect of also letting the job finding rate from employment adjust as in the benchmark.\(^{29}\) Decision rules are still kept fixed at their US values.

Finally, a worker grows her human capital more in more fluid labor markets conditional on her current state, as she expects to use her skills more efficiently in the future (\( h(a, z, h) \) changes). This is the force highlighted in the qualitative analysis in Section 3. I refer to this as the incentive effect and compute

\(^{28}\)I achieve this by adjusting the job finding rate \( p \) as in the cross-country experiment, while at the same time raising relative search efficiency from employment, \( \phi \), such that the job finding rate from employment, \( \phi p \), remains fixed at its US value.

\(^{29}\)Achieved by instead not letting relative search efficiency, \( \phi \), adjust to hold \( \phi p \) fixed as the job finding rate \( p \) changes.
it as the residual between the total effect and the sum of the unemployment effect and job ladder effect.

The right panel of Figure 11 plots the results from these counterfactual exercises for life-cycle human capital accumulation. Going from the US to the least fluid labor market, life-cycle human capital accumulation falls by eight percent. Initially, the job ladder effect is the most important channel, accounting for half of the fall in human capital accumulation, with the unemployment effect accounting for less than 15 percent and the incentive effect for the remaining 36 percent. The relative importance of the unemployment effect, however, rises as labor market fluidity falls further relative to the US, rising to 41 percent in the least fluid labor market. Correspondingly, the relative importance of the job ladder effect falls to 35 percent and the incentive effect to 24 percent for the least fluid labor market.

**Figure 11. Decomposing the effect, model**

![Diagram](image)

Note: Panel A. Decomposition of the effect of labor market fluidity on life-cycle wage growth based on (12). Panel B. Unemployment effect: Counterfactual impact of labor market fluidity holding decision rules and all flows fixed at their US values apart from the job finding rate from unemployment, which evolves as in the benchmark cross-country experiment. Job ladder effect: Incremental counterfactual impact of labor market fluidity of also letting the JJ mobility rate evolve as in the benchmark cross-country experiment, but still under decision rules fixed at their US values. Incentive effect: Incremental effect of also letting decision rules adjust as in the benchmark cross-country experiment, computed as the difference between the total effect and the combined unemployment and job ladder effects. Source: Model.

One interpretation of this key finding is that the most important channel through which labor market fluidity impacts human capital accumulation is backward-looking, as opposed to forward-looking. Greater labor market fluidity lets workers faster locate jobs where they have greater use for their skills and where they consequently optimally accumulate more skills. This includes leaving unemployment faster. In contrast, faster expected future growth in match productivity plays a non-trivial, yet quantitatively less
important role. Nevertheless, I stress that it remains critical to model training as endogenous in order to reach this conclusion, since otherwise training would not differ over the job ladder.

**The role of labor market fluidity.** The incentive effect in Figure 11 is in turn a response to both a slower rate of leaving unemployment and a slower rate of climbing the job ladder in less fluid labor markets. To decompose the importance of each, I resolve workers’ problem holding the job finding rate from unemployment, \( p \), fixed at the estimated value for the US, instead adjusting the relative search efficiency from employment, \( \phi \), so as to match observed cross-country variation in labor market fluidity.

![Figure 12. The role of a higher rate of climbing the job ladder, model](image)

Note: JJ: Counterfactual effect of only letting the job finding rate from employment adjust as in the benchmark cross-country experiment, holding fixed the job finding rate from unemployment, \( p \), at its US value, computed by varying relative search efficiency of the employed, \( \phi \), to match the cross-country variation in labor market fluidity. UE: Counterfactual effect of only letting the job finding rate from unemployment adjust as in the benchmark cross-country experiment, holding fixed the job finding rate from employment, computed as the difference between the total effect and the JJ effect. Source: Model.

The left panel of Figure 12 shows that a slower rate of climbing the job ladder accounts for 92 percent of the slower life-cycle growth in wages in less fluid labor markets. This finding is, per se, not surprising given that less growth in match productivity plays an important role behind weaker life-cycle wage

---

30 For this reason, the impact of labor market fluidity on human capital accumulation is not very sensitive to workers’ bargaining power, as I discuss further in Appendix C.3. Workers’ bargaining power primarily impacts training through the incentive effect, and this effect only accounts for a quarter to a third of the overall impact of fluidity on human capital accumulation.

31 Because workers still move away from low productive jobs with a higher chance of breaking up at a slower rate when the search intensity from employment, \( \phi \), is lower, the incidence of unemployment remains higher in low-\( \phi \) countries. Nevertheless, over 90 percent of unemployment differences disappear once differences in the job finding rate from unemployment are shut down, such that remaining differences in the incidence of unemployment play a second-order role in driving results.
growth in less fluid labor markets (Figure 11). When only the UE rate varies, life-cycle growth in match productivity remains equally large as in the US by construction.\textsuperscript{32} The right panel shows that 40 percent of the lower human capital accumulation in less fluid labor markets is accounted for by the lower UE rate. The remaining 60 percent is accounted for by a slower rate of climbing the job ladder, which reduces human capital accumulation both by slowing down workers’ reallocation up the job ladder to jobs where they train a lot and by discouraging training conditional on place in the job ladder.

6 Direct evidence on training

This section confronts the predictions of the model with cross-country micro data on training. To that end, I exploit the ECHP, which records days and hours spent on on-the-job training, who paid for the training, etc. A key advantage of the ECHP is that it was conducted by the statistical agency of the EU using a common questionnaire and processing routines, facilitating the cross-country comparison. This limits concerns about comparability in cross-country studies of training (Acemoglu and Pischke, 1999a).

6.1 Aggregate patterns

I regress various measures of training on labor market fluidity in Table 5. Panel A considers days on training in the past 12 months, expressed as a fraction of total work days (5*52).\textsuperscript{33} The first column shows the point estimate from a projection on aggregate labor market fluidity with year effects. The second column adds controls for worker observable characteristics, including age, education, one digit occupation and one digit sector. The third column adds also a control for whether the worker made a JJ move in the past year. Panels B–C repeat the analysis for hours on training in the past 12 months and whether the worker did any training since January in the year prior to the survey, respectively.

Workers train more in more fluid labor markets. While older workers train less and college graduates train more, controlling for compositional differences does not change the main takeaway of a positive correlation between aggregate fluidity and incidence of training. Moreover, there is no statistical correlation between training and JJ mobility in the past year. Hence, the positive correlation between labor

\textsuperscript{32}With only the UE rate varying, wage growth initially \textit{rises} as labor market fluidity falls, because workers are less willing to remain unemployed. Consequently, they start at a lower piece rate. Greater growth in the piece rate more than offsets lower growth due to human capital and investment for small changes in fluidity when the JJ rate is held fixed.

\textsuperscript{33}The reported time in training is since January in the year prior to the survey. If the training started more than 12 months prior to the survey date, I use reported start dates of training to adjust training to a per-past-12-months basis (i.e. if the worker reports 20 training days since January last year, her survey was conducted in April and she reports starting training in February last year, my measure of training days is 12/14 * 20). I top-code the training measures at 13 weeks of full time training per year. This concerns 1.5 percent of observations. Results without top-coding in Appendix D.3 are more pronounced.
market fluidity and training does not appear to be due to a somewhat mechanical effect of new hires being trained to perform the required tasks of the new job.

As illustrated by the figure, the model significantly overstates the amount of training in the data. This may reflect issues associated with fully capturing on-the-job training in survey data, as respondents may not perceive all training as such (although the model overstates the level, Appendix D.1 shows that it matches well the relative decline in training with age). Assuming that such underreporting is equally common across countries, the mechanism accounts for a large share of the higher training in more labor market fluidity. Yet caution is warranted given the discrepancy in levels between the model and data.

### Table 5. Training and labor market fluidity, data

<table>
<thead>
<tr>
<th></th>
<th>Panel A. Days (fraction of year)</th>
<th>Panel B. Hours (fraction of year)</th>
<th>Panel C. Whether trained</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw Controls Direct</td>
<td>Raw Controls Direct</td>
<td>Raw Controls Direct</td>
</tr>
<tr>
<td>Fluidity</td>
<td>0.448*** (0.091)</td>
<td>0.305*** (0.031)</td>
<td>9.915*** (1.196)</td>
</tr>
<tr>
<td></td>
<td>0.386*** (0.083)</td>
<td>0.312*** (0.038)</td>
<td>9.438*** (1.366)</td>
</tr>
<tr>
<td></td>
<td>0.388*** (0.066)</td>
<td>0.360*** (0.050)</td>
<td>9.545*** (1.280)</td>
</tr>
<tr>
<td>JJ</td>
<td>0.000 (0.002)</td>
<td>0.001 (0.001)</td>
<td>-0.011 (0.012)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.000*** (0.000)</td>
<td>-0.000*** (0.000)</td>
<td>-0.005*** (0.001)</td>
</tr>
<tr>
<td></td>
<td>-0.011 (0.012)</td>
<td>-0.004*** (0.001)</td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>0.006*** (0.002)</td>
<td>0.003*** (0.001)</td>
<td>0.108*** (0.027)</td>
</tr>
<tr>
<td></td>
<td>0.007*** (0.002)</td>
<td>0.004*** (0.001)</td>
<td>0.106*** (0.029)</td>
</tr>
<tr>
<td>N</td>
<td>135,563 115,897 63,634</td>
<td>135,318 115,676 63,494</td>
<td>135,989 116,261 63,852</td>
</tr>
</tbody>
</table>

Note: Men aged 25–54. Projection of training outcome on labor market fluidity. Raw: Controls for year. Controls: Controls for age, education, occupation, sector and year. Direct: Controls for age, education, occupation, sector, year and JJ mobility in previous year. Panel A: Data: Days on vocational training in the past 12 months, expressed as a fraction of potential work days (5*52). Model: Average time on training, \(f(a,z,h)(c|h,a)da/da = \int f(a,z,h)(c|h,a)da/da\). Panel B: Hours on vocational training in the past 12 months, expressed as a fraction of potential work hours (40*52). Hours and days on training are top-coded at 13 weeks of full time training per year. Panel C: Whether worker undertook any vocational training since January last year. All training measures: Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Standard errors are clustered at the country-level. *** statistically significant at 1%. Source: ECHP 1995–2001.

Appendix D.2 contains additional outcomes. On average over 70 percent of workers report that their employer paid for their training (see also the evidence in Appendix B.2 that training is associated with lower wages). If anything, the share rises with labor market fluidity, but the pattern is not statistically significant.
significant. Appendix D.4 finds no statistically significant correlation between vocational training in the previous year and current JJ mobility. That is, workers who train more are not more likely to remain with their current employer, in contrast to what one may have hypothesized if skills were firm-specific.

6.2 Dissecting the patterns

The theory also makes rich predictions for how training varies in the cross-section across countries. As highlighted by Figure 9, workers grow their human capital less when employed in less productive firms, as they have less use for it. In more fluid labor markets, however, workers currently in low productive firms expect to be able to move away from such firms faster. Consequently, while workers in general train less in less productive firms, in a relative sense they train more in low productive firms in more fluid labor markets. In contrast, absent separation shocks, training at the top of the job ladder is unaffected by labor market fluidity. Separation shocks moderate this prediction, but the intuition remains: training should be particularly low in less productive firms in less fluid labor markets. By differencing out any country-specific factors, this may provide a more robust test of the predictions of the model.

Testing this prediction is somewhat complicated by the fact that the available data do not contain productivity. They do, however, report firm size, grouped into size bins 1–4, 5–19, 20–49, 50–99, 100–499, and 500+ employees. Pooling across countries, these size categories contain close to equal shares of employment. Under the assumption of a monotone relationship between productivity and size, the six size bins allow me to group firms into (roughly) employment-weighted productivity sextiles. I similarly bin firms based on productivity into employment-weighted sextiles in the model. As an alternative ranking of firms, I compute firm pay by regressing log wages on indicators for the sextiles—controlling for age, education, occupation and country-year fixed effects (age and country fixed effects in the model),

\[
\begin{align*}
  w_{it} &= \sum_{s=1}^{6} \alpha_s \text{sextile}_s^{it} + X_{it} \beta + \epsilon_{it} \\
\end{align*}
\]

where \( \text{sextile}_s^{it} \) takes value one if worker \( i \) in year \( t \) works for an employer in sextile \( s \). My measure of firm pay, \( f_{\text{wage}}_{it} \), is the estimated coefficients, \( \{\alpha^s\} \). Consistent with typical findings in the literature, larger firms pay better. Firm pay has a standard deviation of 0.114 in the data and 0.123 in the model.

I subsequently regress hours on training in the past 12 months on either firm productivity sextile or firm pay, \( x_{it} \), as well as its linear interaction with fluidity, controlling for worker-fixed effects and age,

\[
\begin{align*}
  \text{training}_{it} &= \alpha_0 x_{it} + \alpha_1 x_{it} \times \text{fluidity}_c + I_i + X_{it} \beta + \epsilon_{it} \\
\end{align*}
\]
The model data are based on a simulated monthly approximation aggregated to yearly data such that employer outcomes are recorded in "May" every year, while annual training is the sum of $i(a, z, h)$ over the 12 months prior to the "survey," and the wage is the average wage over the past "calendar year."

Table 6 provides results, offering two main takeaways. First, controlling for worker-fixed effects, workers train more when employed at high-paying, productive (large) firms, consistent with evidence in other data sets (Arellano-Bover, 2020). The reason is that workers have greater use for their skills in more productive firms, and as a consequence train more. In contrast, as I discuss further in Appendix D.6, if the production function had not featured a complementarity between productivity and human capital, one may expect training to decline in productivity (and hence firm pay). Moreover, as highlighted by Figure 5, a larger share of workers are employed at high-paying, productive (large) employers in more fluid labor markets, where ceteris paribus they train more. In fact, as shown in Appendix D.7, the model matches this pattern fairly well. These findings relate to the job ladder effect discussed in Section 5.

Second, workers in low fluidity countries particularly train less when employed at low paying, unproductive (small) firms. The reason is that a worker currently in a low productive firm is particularly stuck with her current employer in a low fluidity country, where she has less use for her human capital. She responds by training less. This relates to the incentive effect discussed in Section 5.

Finally, recall from Figure 10 that the mechanism also matches well the higher unemployment rate in less fluid labor markets—the unemployment effect in Section 5. I conclude that these aggregate and disaggregate patterns for training lend empirical support for the key predictions of the theory.

<table>
<thead>
<tr>
<th>Panel A. Data</th>
<th>Panel B. Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod. Wage</td>
<td>Prod. Wage</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>0.002** 0.035**</td>
</tr>
<tr>
<td>(0.001) (0.013)</td>
<td>(0.000) (0.003)</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-0.028* -0.515**</td>
</tr>
<tr>
<td>(0.012) (0.211)</td>
<td>(0.002) (0.034)</td>
</tr>
<tr>
<td>N</td>
<td>77,386 77,386</td>
</tr>
</tbody>
</table>

Note: Men aged 25–54. Training on firm productivity sextile / firm wage, as well as its interaction with fluidity. Panel A. Hours on training, top-coded at 13 full-time weeks per year. Firm productivity sextile and firm wage imputed based on size. Panel B. $i(a, z, h)$. Firm wage imputed based on productivity. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Standard errors are clustered at the country-level. * statistically significant at 10%; ** statistically significant at 5%; *** statistically significant at 1%. Source: ECHP 1995–2001.

7 Factors behind fluidity and concluding remarks

I end with a brief discussion of potential factors behind fluidity. Figure 13 correlates labor market fluidity with a few factors that the previous literature has emphasized as important drivers of firms’ incentives to create jobs. Labor market fluidity is negatively correlated with employment protection, costs of doing
business, and labor taxes on firms. I conclude from this assessment that, consistent with a large existing literature, policies that raise the cost on firms of hiring are correlated with lower labor market fluidity.

**FIGURE 13. DETERMINANTS OF LABOR MARKET FLUIDITY**

(A) EMPLOYMENT PROTECTION LEGISLATION  
(B) DOING BUSINESS INDEX  
(C) LABOR TAX RATE  
(D) COST OF STARTING BUSINESS

Note: Male employees aged 25–54. Panel A. Employment protection legislation index. Panel B. Doing business score. Panel C. Average tax rate on labor on firms. Panel D. Data: Cost of starting a firm relative to GNI per capita. Model: Cost of vacancy, \( c \), relative to total output. Model moments are normalized to match the empirical value for the US by first dividing all country outcomes in the model by the model outcome for the US and then multiplying all country outcomes by the data outcome for the US. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Source: BHPS, ECHP, EUSILC, GSOEP, OECD, PSID and World Bank 1991–2015.

A question remains of whether also quantitatively the required wedges in the model are consistent
with empirical measures of differences in the cost of creating jobs. While several of the available measures are difficult to map into the model, the World Bank estimates the cost of starting a firm relative to GNI per capita. I construct a model counterpart of this as the cost of creating a job, $c$, relative to GDP per capita. As the levels are non-comparable, I normalize the model measure to the empirical measure for the US. Although this measure only captures the cost of starting a new firm, it is well-known that young firms contribute a disproportionate share of total job creation. The required wedges in the model are quantitatively consistent with observed dispersion in the cost of creating jobs across countries.

I argue that lower labor market fluidity, in turn, reduces life-cycle wage growth of workers and the stock of human capital. The reason is that makes it harder for workers to find jobs that fully utilize their skills, discouraging human capital accumulation. Differences in labor market fluidity account for half of lower life-cycle wage growth in less fluid labor markets across OECD countries and result in nine percent lower labor productivity relative to the US. More broadly, my findings highlight that the dynamic consequences of misallocation can be large. They deserve more attention in future work.

References


44


A Data — FOR ONLINE PUBLICATION

This section describes the data sources in some more detail and provides additional empirical results.

A.1 Data sources

Table 7 summarizes the annual data set and Table 8 the monthly data set. In total, the annual sample contains roughly nine hundred thousand individual-years and the monthly sample roughly 11 million individual-months. I describe the data sources in some more detail below.

PSID. The PSID has been collected annually since 1968 (biannually since 1997) based on an initially representative sample of households and their offspring. Over time, additional households have been added, but I focus on the original core sample and their descendants (the so called Survey Research Center sample). Initially, no weights were provided for this sample, since it was representative of the US population. Subsequent attrition and non-response, however, necessitate the use of survey weights, which I employ throughout my analysis. As a large share of the questions in the PSID center around the "head" of the household, I restrict attention to heads of households. Starting in 1988, the PSID asks respondents for a monthly calendar of labor market events during the prior calendar year (during the prior two calendar years starting with the 2003 survey). It also asks for information on up to two employment spells (in some years more), including start and end dates, earnings, hours, occupation, etc.

GSOEP. The GSOEP was modeled on the PSID and has followed the same individuals annually since 1984. Additional samples have been added over time, but I restrict attention to the original, representative samples for West and East Germany. I start the analysis with German reunification in 1991, which also corresponds well with the sample period available from the other data sources. I end the analysis in 2011 because it was the last year available when I applied for the data several years ago. The GSOEP asks a rich set of questions about demographics, income and hours worked, as well as labor force status in each month during the past calendar year and the start date of the current employment spell. I use survey weights throughout my analysis to adjust for nonrandom attrition and non-response.

BHPS. The BHPS began in 1991 and was discontinued in 2008. It is similar to the PSID. The sample has expanded over time, but I focus on the original core sample. As the PSID/GSOEP, the BHPS contains demographic characteristics on the respondent, as well as annual information on gross income and hours

---

34I am awaiting an extension of my project to add the last few years from the GSOEP as well as the EUSILC.
worked. The BHPS also contains information on start and end dates on all labor market spells since the last survey, which I use to construct a monthly calendar of labor market events as in the PSID/GSOEP. I use survey weights throughout my analysis to adjust for nonrandom attrition and non-response.

### Table 7. Overview of annual data set

<table>
<thead>
<tr>
<th>Source</th>
<th>Years</th>
<th>T</th>
<th>NT</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Core Western European countries plus the US</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>ECHP/EUSILC</td>
<td>1995–2014</td>
<td>18</td>
<td>39,700</td>
</tr>
<tr>
<td>Belgium</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>19</td>
<td>36,070</td>
</tr>
<tr>
<td>Denmark</td>
<td>ECHP/EUSILC</td>
<td>1994–2013</td>
<td>18</td>
<td>25,883</td>
</tr>
<tr>
<td>Finland</td>
<td>ECHP/EUSILC</td>
<td>1996–2014</td>
<td>16</td>
<td>48,432</td>
</tr>
<tr>
<td>France</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>16</td>
<td>58,981</td>
</tr>
<tr>
<td>Germany</td>
<td>GSOEP</td>
<td>1991–2011</td>
<td>21</td>
<td>60,346</td>
</tr>
<tr>
<td>Ireland</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>17</td>
<td>25,018</td>
</tr>
<tr>
<td>Italy</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>15</td>
<td>85,926</td>
</tr>
<tr>
<td>Netherlands</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>18</td>
<td>61,312</td>
</tr>
<tr>
<td>Spain</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>19</td>
<td>92,355</td>
</tr>
<tr>
<td>UK</td>
<td>BHPS</td>
<td>1991–2008</td>
<td>18</td>
<td>32,910</td>
</tr>
<tr>
<td>US</td>
<td>PSID</td>
<td>1994–2015</td>
<td>13</td>
<td>19,047</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>223</strong></td>
<td><strong>622,321</strong></td>
<td><strong>191,594</strong></td>
</tr>
</tbody>
</table>

| **Panel B. Other OECD countries** | | | | |
| Czech Republic | EUSILC | 2005–2014 | 9 | 27,761 | 10,805 |
| Estonia | EUSILC | 2004–2014 | 10 | 19,682 | 8,113 |
| Hungary | EUSILC | 2005–2014 | 10 | 37,406 | 16,059 |
| Iceland | EUSILC | 2004–2014 | 10 | 10,407 | 4,522 |
| Latvia | EUSILC | 2007–2014 | 7 | 12,684 | 6,207 |
| Lithuania | EUSILC | 2005–2014 | 9 | 13,495 | 5,414 |
| Norway | EUSILC | 2004–2014 | 10 | 22,994 | 8,241 |
| Poland | EUSILC | 2005–2014 | 9 | 46,972 | 19,850 |
| Slovak Republic | EUSILC | 2005–2013 | 8 | 15,404 | 6,166 |
| **Total** | | **71** | **221,112** | **96,101** |

*Note: Men aged 25–54. T: Number of years; NT: Number of individual-years; N: Number of individuals.*

**ECHP.** The ECHP was run 1994–2001 across the original set of European Union countries. Because of confidentiality restrictions, however, data from Germany and Portugal are not released to researchers. The data from Sweden are only available for a few years and miss several key variables, forcing me to exclude Sweden. Luxembourg switched to collecting data via a separate, national survey after three years in the ECHP, dropping several key variables in the process. As the Luxembourg sample for the first three years in the ECHP is small, I drop also Luxembourg from my analysis (similar results hold including the few available years, though). While data from the UK are available in the ECHP, several years are missing so I opt to use the larger and consistently collected BHPS instead. The ECHP follows the same individuals annually for up to eight years. The survey is similar to the PSID, including a similar
set of variables. In particular, it asks for a monthly calendar of events in the prior calendar year and the start date of the current employment spell. I weigh all results using the provided survey weights.

### Table 8. Overview of monthly data set

<table>
<thead>
<tr>
<th>Source</th>
<th>Years</th>
<th>T</th>
<th>NT</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Core Western European countries plus the US</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>ECHP/EUSILC</td>
<td>1995–2014</td>
<td>214</td>
<td>544,535</td>
</tr>
<tr>
<td>Belgium</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>225</td>
<td>482,697</td>
</tr>
<tr>
<td>Denmark</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>213</td>
<td>243,673</td>
</tr>
<tr>
<td>Finland</td>
<td>ECHP/EUSILC</td>
<td>1996–2014</td>
<td>189</td>
<td>430,234</td>
</tr>
<tr>
<td>France</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>201</td>
<td>833,033</td>
</tr>
<tr>
<td>Germany</td>
<td>GSOEP</td>
<td>1991–2011</td>
<td>239</td>
<td>756,609</td>
</tr>
<tr>
<td>Greece</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>177</td>
<td>491,048</td>
</tr>
<tr>
<td>Ireland</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>201</td>
<td>342,759</td>
</tr>
<tr>
<td>Italy</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>177</td>
<td>1,162,774</td>
</tr>
<tr>
<td>Spain</td>
<td>ECHP/EUSILC</td>
<td>1994–2014</td>
<td>225</td>
<td>1,247,297</td>
</tr>
<tr>
<td>UK</td>
<td>BHPS</td>
<td>1991–2014</td>
<td>228</td>
<td>466,563</td>
</tr>
<tr>
<td>US</td>
<td>PSID</td>
<td>1994–2015</td>
<td>225</td>
<td>374,570</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>2,727</td>
<td>7,926,703</td>
<td>194,919</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Years</th>
<th>T</th>
<th>NT</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel B. Other OECD countries</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Czech Republic</td>
<td>EUSILC</td>
<td>2005–2014</td>
<td>106</td>
<td>387,008</td>
</tr>
<tr>
<td>Estonia</td>
<td>EUSILC</td>
<td>2004–2014</td>
<td>118</td>
<td>267,939</td>
</tr>
<tr>
<td>Hungary</td>
<td>EUSILC</td>
<td>2005–2014</td>
<td>118</td>
<td>523,418</td>
</tr>
<tr>
<td>Iceland</td>
<td>EUSILC</td>
<td>2004–2014</td>
<td>82</td>
<td>95,374</td>
</tr>
<tr>
<td>Latvia</td>
<td>EUSILC</td>
<td>2007–2014</td>
<td>82</td>
<td>172,982</td>
</tr>
<tr>
<td>Lithuania</td>
<td>EUSILC</td>
<td>2005–2014</td>
<td>106</td>
<td>191,482</td>
</tr>
<tr>
<td>Norway</td>
<td>EUSILC</td>
<td>2004–2014</td>
<td>118</td>
<td>225,178</td>
</tr>
<tr>
<td>Poland</td>
<td>EUSILC</td>
<td>2005–2014</td>
<td>106</td>
<td>651,946</td>
</tr>
<tr>
<td>Slovak Republic</td>
<td>EUSILC</td>
<td>2005–2013</td>
<td>94</td>
<td>204,716</td>
</tr>
<tr>
<td>Slovenia</td>
<td>EUSILC</td>
<td>2005–2014</td>
<td>106</td>
<td>646,751</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>1,036</td>
<td>3,366,794</td>
<td>122,228</td>
</tr>
</tbody>
</table>

Note: Men aged 25–54. T: Number of months; NT: Number of individual-months; N: Number of individuals.

**EUSILC.** The EUSILC is the successor to the ECHP starting in 2003. It gradually expanded the set of countries covered to eventually include all EU members plus a set of affiliated countries. The survey is annual and uses a rotating panel design, which differs somewhat in length across countries. Most countries follow the same individuals for up to four years, but France follows individuals for eight years and Norway for six. As the other surveys, it contains the standard demographic and income variables, including gross annual labor income and hours worked during the previous calendar year. It also contains a monthly calendar of labor market events for the prior calendar year. It differs slightly from the other surveys in that it does not contain the start date of the current employment spell, instead recording whether the individual switched employer in the past 12 months. All results use survey weights.
**SIPP.** The SIPP has been conducted in separate panels since the mid-1980s, but a break in the survey in 1996 implies that data on job-to-job mobility in the earlier panels are not directly comparable to that in the later panels. To align with the time period covered by the other data sets used by this paper, I focus on SIPP data from 1996–2012 (i.e. the 1996, 2001, 2004 and 2008 panels). Each panel of the SIPP follows a group of individuals over time. Data are collected in "waves", with the respondent in each wave being asked to recall labor market events during the prior four months. The survey asks for information regarding up to four employment spells during the past four months (two as employee and two as self-employed), including start dates and end dates (if applicable), income, hours, occupation, sector, etc. It also contains standard demographic characteristics. I use the provided survey weights throughout my analysis to make results representative of the overall US population.

**A.2 Benchmarking the PSID with the SIPP**

The EU rate in the PSID is lower than what researchers typically find in the CPS, but more in line with what other research has found in the SIPP (Engbom, 2020). Figure 14 compares the monthly EU and UE rates in the PSID against the SIPP. Broadly, the series compare reasonably well. In particular, both the PSID and SIPP show significantly lower mobility than, for instance, the CPS. This is consistent with well-known issues with classification error leading to substantially overstated gross worker flows in the CPS (Abowd and Zellner, 1985; Poterba and Summers, 1986).

**Figure 14. Monthly EU and UE rates, PSID versus SIPP**

(A) EU

(B) UE

*Note:* Men 25–54. Panel A. Share of employed workers who are unemployed in the subsequent month. Panel B. Share of unemployed workers who are employed in the subsequent month. All panels. Employment includes self-employment due to data limitations. Definitions follow ILO standard. Constructed by first collapsing the data to the country-year-age level using the provided survey weights, then to the country-level. *Source:* PSID and SIPP 1994–2015.
Figure 15. Life-cycle labor market states

A.3 Life-cycle labor market states

Figure 15 plots the distribution of men across four labor market states—wage employment, self-employment, unemployment and non-participation—over the life-cycle, offering three main takeaways. First, workers enter the labor market at a declining pace up to age 30. There is no pronounced covariation between age of entry and labor market fluidity, although the US (and somewhat less the UK) appears to be an outlier. The way the PSID is collected, however, may bias participation rates at young ages if
those who participate in the labor market are also more likely to have formed their own households (the SIPP, in contrast, shows an increase in male labor force participation rates between age 20–25). By not focusing on the head of household like the PSID, the other surveys may be less prone to this.

Second, participation rates remain high and roughly constant until age 50, after which they start to decline. The decline becomes pronounced after around age 55. Third, the share of wage employed falls gradually after age 30, while the share of self-employed rises. There is also some evidence of a higher self-employment rate in less fluid labor markets. As I show in Appendix C.6, the model is consistent with this pattern if self-employment is interpreted as out of necessity, i.e. akin to unemployment. While the share of self-employment remains modest across all countries, it would be interesting to develop a deeper theory of this than necessity entrepreneurs in future research (see, for instance, Engbom, 2020, for a joint model of labor market search and entrepreneurship).

Figure 16 plots labor force participation rates for men 25–54, men aged 25 and men aged 54 against labor market fluidity. While there is dispersion across countries in participation, it is only weakly correlated with labor market fluidity. Figure 17 plots the share of men who are in the labor force in year $t$ but not in the labor force in year $t+2$. I focus on two-year later outcomes to be able to use also the biannual PSID years after 1997. The share remains low up to around age 50, when it starts to gradually rise.

### A.4 The public sector

The left panel of Figure 18 shows that the share of prime aged male employees that work in the public sector is about 25 percent and that it declines with labor market fluidity, although the relationship is driven by the US. The right panel plots the estimated life-cycle wage profile of private and public employees based on a regression of log hourly real wages, $w_{it}$, of individual $i$ in year $t$ on separate age effects for the private, $A_{it}$, and public sector, $PA_{it}$, country-year effects, $Y_{ct}$, and worker fixed effects, $I_i$,

$$w_{it} = A_{it} + PA_{it} + Y_{c(i)t} + I_i + \epsilon_{it}$$  \hspace{1cm} (14)

I focus on the sample of 25–54 year olds and restrict wages to not grow after age 50. While in principle this restriction is only required for one of the two sectors, to treat both identically I impose it in both. Inclusion of worker fixed effects in (14) controls for differences in worker composition across sectors, i.e. identification of the level wage difference is based on workers switching between sectors. Note also that the required data for this exercise are only available from the ECHP, GSOEP and PSID. Young workers earn less when employed in the public sector, but the public sector is associated with steeper within-worker residual life-cycle growth in wages. Hence, the fact that the share of public sector employees
Figure 17. Labor force exit rates

Note: Male employees aged 25–63. Share of labor force participants in year $t$ who are not in the labor force at year $t + 2$. Employment status is based on ILO standards. Computed by first collapsing the data to the age-year-country level using the provided survey weights, then to the country-age-level. Source: BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.
declines with labor market fluidity tends to, *ceteris paribus, flatten* the relationship between life-cycle wage growth and labor market fluidity. The differences in life-cycle wage growth, however, are modest, so it is unlikely that this would have a major impact on the patterns documented in this paper.

**A.5 Flows in and out of unemployment**

The top panel of Figure 19 plots the monthly EU rate over the life-cycle across countries. The EU rate shares a common shape across countries, with high rates of job loss early in careers and subsequent declines. Aggregate cross-country differences are less pronounced than for JJ mobility, with Spain as an exception. Moreover, there is no evidence that high-fluidity countries also have higher EU mobility rates. In fact, the correlation between the aggregate EU rate and labor market fluidity is negative.

**Figure 18. Public sector**

![Graph](image)

*Note:* Male employees aged 25–54. Panel A. Share of employees that work in public sector, computed by first collapsing the data to the age-year-country level using the provided survey weights, then to the country-age-level. Panel B. Wage profile for ages 25–54 with wages restricted to not grow past age 50 based on regression (14). Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-level. *Source:* ECHP, GSOEP and PSID 1991–2015.

The bottom panel plots the monthly UE rate over the life-cycle across countries. The profiles are noisier than the other mobility rates, because the sample of unemployed is much smaller than the sample of employed. The life-cycle pattern of UE mobility is somewhat more heterogeneous across countries. Most countries display modest declines in the UE rate over the life-cycle, but a few countries have more pronounced declines, while others see increases. There are also significant differences also in the UE rate
across countries, and the aggregate UE rate is positively correlated with labor market fluidity.

A.6 The role of composition, full sample

Table 9 replicates the regression results for life-cycle wage growth and labor market fluidity controlling for compositional differences for the full sample of 23 countries. Results are very similar to the core sample.

<table>
<thead>
<tr>
<th></th>
<th>Panel A. Ages 25–54</th>
<th>Panel B. 1% depreciation</th>
<th>Panel C. Ages 22–59</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Educ</td>
<td>Occup</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.242***</td>
<td>0.217**</td>
<td>0.219**</td>
</tr>
<tr>
<td>(0.081)</td>
<td>(0.083)</td>
<td>(0.078)</td>
<td></td>
</tr>
<tr>
<td>( N )</td>
<td>474,919</td>
<td>472,836</td>
<td>449,644</td>
</tr>
</tbody>
</table>

Note: Male employees. Panel A. Wages restricted to not grow after age 50. Panel B. Ages 25–54 with wages restricted to depreciate 1% annually after age 50. Panel C. Wages restricted to not grow after age 50. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. \( \alpha \): Fluidity-age interaction in regression (2) with worker fixed effects, time effects and restricted age effects. Standard error below are clustered at the country-level. ** statistically significant at 5%; *** statistically significant at 1%. Source: BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.

A.7 Raw and residual life-cycle inequality

Figure 20 correlates various measures of inequality and life-cycle growth in inequality with labor market fluidity. There is possibly some evidence that inequality is higher in more fluid labor markets, while growth in inequality is lower. As it turns out, these patterns are consistent with the predictions of the model. The patterns, however, are not pronounced (which is also consistent with the theory).

Figure 21 plots the standard deviation of log hourly wages over the life-cycle for various measures. I drop the bottom one percent of wages to reduce the impact of a few outliers, but this primarily impacts the level of inequality and not the life-cycle patterns. Inequality rises over the life-cycle. This, however, is largely accounted for by increasing dispersion across education and occupation groups. I believe that this finding is novel to the literature. At least, it has received little attention as far as I am aware.
Figure 19. EU (top) and UE rate (bottom)

Note: Men aged 25–54. EU: Share of employed who are unemployed in the subsequent month. UE: Share of unemployed who are employed in the subsequent month. Employment includes self-employed due to data limitations. Constructed by first collapsing to the country-age-year level using the provided survey weights, then to the country-age level. Source: BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.
Figure 20. Inequality and labor market fluidity, alternative measures

(A) Inequality, raw

(B) Inequality, w > p1

(C) Growth in inequality, raw

(D) Growth in inequality, w > p1

Note: Male employees aged 25–54. Panel A. Cross-sectional standard deviation of residual log hourly real wages. Panel B. Cross-sectional standard deviation of residual log hourly real wages dropping the bottom 1% of residual wages in each country. Panel C. Growth in standard deviation of residual log hourly real wages between age 25–29 and age 50–54. Panel D. Growth in standard deviation of residual log hourly real wages between age 25–29 and age 50–54 after first dropping the bottom 1% of residual wages in each country. All panels: Raw: Only country-year controls. Education: Country-year-education-age controls. Education+occupation: Country-year-education-age and country-occupation-year controls. All moments: Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-level. Source: BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.
Note: Male employees aged 25–54. Solid black: Standard deviation of log hourly wages controlling for country-year effects. Dashed blue: Standard deviation of residual log hourly wages controlling for country-year-education-age effects. Dotted purple: Standard deviation of residual log hourly wages controlling for country-year-education-age effects and country-year-occupation effects (10 occupations). All moments first drop the bottom 1% of wages in each country and are computed by first collapsing the data to the age-year-country level using the provided survey weights, then to the country-age-level. Source: BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.
B Model — FOR ONLINE PUBLICATION

This appendix contains additional model details, including proofs of propositions.

B.1 Evolution of states

The quantitative analysis sets the flow value of leisure, \( b(a, h) \), such that workers of all ages and human capital share the same match productivity reservation threshold, \( z \). That is, \( b(a, h) = J(a, z, h) = U(a, h) \), \( \forall a, h \). Given this, the distribution of workers over age, match productivity and human capital, \( g(a, z, h) \), solves for \( z \geq z_0 \) and \( a \in (0, A) \) the Kolmogorov Forward Equation (KFE)

\[
\frac{\partial g(a, z, h)}{\partial a} = -\left( \delta(z) + \phi p(1 - \Gamma(z)) \right) g(a, z, h) - \frac{\mu}{\eta} (i(a, z, h) z h) \eta \frac{\partial g(a, z, h)}{\partial h}
\]

\[+ \gamma(z) \left( p \frac{u(a, h)}{e} + \phi p \int_z^\infty g(a, z', h) dz' \right) \]

subject to the boundary conditions \( g(0, z, h) = 0, g(A, z, h) = 0 \) and \( g(a, z, h) \) for all \( a, z, \) and \( h \), where the number of unemployed of age \( a \) with human capital \( h \), \( u(a, h) \), is given by

\[
\frac{\partial u(a, h)}{\partial a} = -pu(a, h) + \int_z^\infty \delta(z) g(a, z, h) dz
\]

subject to \( u(A, h) = 0 \) for all \( h \) and \( u(0, h) = \lambda(h) / A \). and \( e = 1 - \int u(a, h) dadh \) is the total number of employed. Workers flow out due to exogenous separations at rate \( \delta(z) \) and up the job ladder at rate \( \phi p(1 - \Gamma(z)) \). With probability \( \gamma(z) \), workers who receive an offer contact a match with productivity \( z \), which they accept if they are either unemployed or employed lower down the job ladder. Finally, the density changes due to accumulation of human capital and aging.

B.2 Wages and training

This subsection provides support for the assumption that workers share the cost of training with firms through a lower wage using the ECHP training data (see Section 6 for a description of the data). Table 10 regress wages in the past calendar year on days or hours on training in the past 12 months (expressed as a fraction of total work days/hours), controlling for worker fixed effects and year effects. The "sector" columns additionally also controls for sector, exploiting individuals who switch sectors for identification. In years when a worker spends more time on training, she is paid less per hour. This is not driven by confounding movements of workers across sectors, and a correlation between the incidence of training
and wages across sectors. Quantitatively, when a worker spends one percent more of her work time on training, she is paid 0.28 percent less per hour.

**Table 10. Wages and Training**

<table>
<thead>
<tr>
<th></th>
<th>Panel A. Days</th>
<th>Panel B. Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw Sector</td>
<td>Raw Sector</td>
</tr>
<tr>
<td>β</td>
<td>-0.126***</td>
<td>-0.210***</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.066)</td>
</tr>
<tr>
<td>N</td>
<td>80,903</td>
<td>80,711</td>
</tr>
<tr>
<td></td>
<td>78,875</td>
<td>78,683</td>
</tr>
</tbody>
</table>

Note: Male employees 25-54. Log hourly real wage on days/hours on training in the past 12 months (fraction of total work days/hours). Hours and days on training are top-coded to 13 weeks of full time training per year. Raw: Year and worker-fixed effects. Sector: Year, worker-fixed effects and sector controls. Standard errors are clustered at the individual-level. *** statistically significant at 1%. Source: ECHP 1995–2001.

B.3 Proofs

**Proof of proposition 1.** Adding the value of the worker (4) and the value of the firm (5) and cancelling terms, it is immediate that the piece rate $r$ drops out.

**Proof of proposition 2.** The problem of the worker is to chose investment to maximize her value,

$$W(r, i) = \max_i \left\{ (1 - i)r(i)z_1 + \left( (1 - p)z_1 r(i) + \frac{p}{2} z_1 + \frac{p}{2} (z_1 + \beta (z_2 - z_1)) \right) \left( 1 + \frac{\mu}{\eta} (z_1 i)^\eta \right) \right\}$$  \hspace{1cm} (15)

Because the bargaining protocol ensures that the worker is delivered a share of the surplus, the piece rate a worker gets paid is implicitly a function of the training she undertakes. In particular, when a young worker is hired, the bargaining protocol ensures that she gets paid a piece rate such that

$$W(r, i) = U + \beta (J(i) - U)$$

$$(1 - i)r(i)z_1 = U + \beta (J(i) - U) - \left( (1 - p)z_1 r(i) + \frac{p}{2} z_1 + \frac{p}{2} (z_1 + \beta (z_2 - z_1)) \right) \left( 1 + \frac{\mu}{\eta} (z_1 i)^\eta \right)$$  \hspace{1cm} (16)

where $U$ is the value of unemployment. Substituting for $(1 - i)r(i)z_1$ in (15) using (16) and simplifying

$$W(r, i) = \max_i \{ U + \beta (J(i) - U) \}$$

Hence, the optimal choice of training by the worker satisfies $J'(i) = 0$, i.e. it coincides with that which maximizes the joint value.
Proof of proposition 3. Define \( i_w(v) \) and \( i_e(v) \) based on (9)–(10) to be

\[
\begin{align*}
i_w(v) &= \left( \mu \left( z_1 + v^\alpha \beta^{z_2 - z_1} \right) \right) \left( \frac{\eta}{1 - \eta} \right) \\
i_e(v) &= \frac{\eta}{\mu} \left( \frac{2cv^{1 - \alpha}}{(1 - \beta)(z_2 - z_1)} - 1 \right)
\end{align*}
\]

Then,

\[
\lim_{v \to 0} i_w(v) = \mu \frac{\eta}{1 - \eta} > 0, \quad \lim_{v \to 0} i_e(v) = -\frac{\eta}{\mu} < 0
\]

Hence, if \( \lim_{v \to \infty} \frac{i_e(v)}{i_w(v)} > 1 \), there is at least one equilibrium. Consider the limit \( \lim_{v \to \infty} \left( \frac{i_e(v)}{i_w(v)} \right) \frac{1 - \eta}{\eta} \),

\[
\lim_{v \to \infty} \left( \frac{i_e(v)}{i_w(v)} \right)^{1 - \eta} = \lim_{v \to \infty} \left( \frac{2cv^{1 - \alpha}}{(1 - \beta)(z_2 - z_1)} - 1 \right) \frac{1 - \eta}{\eta} \frac{\mu}{(z_1 + v^\alpha \beta^{z_2 - z_1})}
\]

In the limit \( v \to \infty \), the second term in (18) is strictly positive. Hence, \( \lim_{v \to \infty} \frac{i_e(v)}{i_w(v)} > 1 \) if and only if

\[
(1 - \alpha) \frac{1 - \eta}{\eta} - \alpha > 0
\]

Hence, under this condition, the limit tends to infinity and there is at least one solution.

To see that the solution is unique, take derivatives of the best-response functions,

\[
\begin{align*}
\frac{\partial i_w(v)}{\partial v} &= \frac{\eta}{1 - \eta} \frac{2v^{1 - \alpha} - 1}{\mu v^\alpha \beta^{z_2 - z_1}} \\
\frac{\partial i_e(v)}{\partial v} &= \frac{1 - \alpha}{v} \left( i_e + \frac{\eta}{\mu} \right)
\end{align*}
\]

Recall that \( i_w(0) > i_e(0) \) and that, under the assumption that \( \frac{1 - \eta}{\eta} > \frac{\alpha}{1 - \eta} \), there exists at least one \( v \) such that \( i_w(v) = i_e(v) \). Consider any such equilibrium point. If it is the case that \( \frac{\partial i_e(v)}{\partial v} > \frac{\partial i_w(v)}{\partial v} \), then it must be that the equilibrium is unique. Substituting using (19)–(20), I hence require that,

\[
\frac{1 - \alpha}{v} \left( i + \frac{\eta}{\mu} \right) > \frac{\eta}{1 - \eta} \frac{2v^{1 - \alpha} - 1}{\mu v^\alpha \beta^{z_2 - z_1}} \\
\]

\[
i + \frac{\eta}{\mu} > \frac{\eta}{1 - \eta} \frac{2v^{1 - \alpha} - 1}{\mu v^\alpha \beta^{z_2 - z_1}} \frac{\alpha}{1 - \alpha} \frac{v^\alpha \beta^{z_2 - z_1}}{2}
\]

62
Since \( i + \frac{\eta}{\mu} > i \), the above inequality is certainly true if

\[
i > \frac{\eta}{1 - \eta} i \eta \mu \frac{\alpha}{1 - \alpha} v^\alpha \beta \frac{z_2 - z_1}{2}
\]

\[
\frac{1 - \eta}{\eta} > \frac{\eta}{1 - \eta} \frac{\alpha}{1 - \alpha} \mu v^\alpha \beta \frac{z_2 - z_1}{2}
\]

Using (17) to substitute for \( i \),

\[
\left( z_1 + v^\alpha \beta \frac{z_2 - z_1}{2} \right) \mu > \frac{\eta}{1 - \eta} \frac{\alpha}{1 - \alpha} \mu v^\alpha \beta \frac{z_2 - z_1}{2}
\]

\[
z_1 + v^\alpha \beta \frac{z_2 - z_1}{2} > \frac{\eta}{1 - \eta} \frac{\alpha}{1 - \alpha} \mu v^\alpha \beta \frac{z_2 - z_1}{2}
\]

\[
z_1 > \left( \frac{\eta}{1 - \eta} \frac{\alpha}{1 - \alpha} - 1 \right) v^\alpha \beta \frac{z_2 - z_1}{2}
\]

For this to be guaranteed to hold for any \( v \), it is sufficient that

\[
\frac{\eta}{1 - \eta} \frac{\alpha}{1 - \alpha} - 1 < 0 \iff \frac{\alpha}{1 - \alpha} < \frac{1 - \eta}{\eta} \quad (21)
\]

**Proof of proposition 4.** Holding investment fixed, an increase in the cost of job creation shifts the job creation curve to the left in Figure 7, \( \frac{\partial i}{\partial c} |_i < 0 \). Holding fixed vacancy creation, an increase in the cost of job creation has no impact on the training curve, \( \frac{\partial i}{\partial c} |_v = 0 \). As established above, the unique equilibrium is characterized by the job creation curve crossing the training curve from below. Hence, in equilibrium job creation falls by more and training declines. As the job finding rate of workers falls, average match quality declines. As training falls, average human capital declines.

**Proof of lemmas 1–2 and proposition 5.** Compare the first-order condition for optimal investment and the free entry condition in the decentralized equilibrium to those in the planned economy,

\[
i(v) = \frac{1}{z_1} \left( \mu \left( z_1 + v^\alpha \beta \frac{z_2 - z_1}{2} \right) \right)^{\frac{1}{1 - \eta}} \quad (22)
\]

\[
i_{sp}(v) = \frac{1}{z_1} \left( \mu \left( z_1 + v^\alpha \frac{z_2 - z_1}{2} \right) \right)^{\frac{1}{1 - \eta}} \quad (23)
\]

\[
v(i) = \left( 1 - \beta \right) \frac{(z_2 - z_1)}{2c} \left( 1 + \frac{\mu}{\eta} (z_1 i)^\eta \right)^{\frac{1}{1 - \alpha}} \quad (24)
\]

\[
v_{sp}(i) = \left( \alpha \frac{z_2 - z_1}{2c} \left( 1 + \frac{\mu}{\eta} (z_1 i)^\eta \right) \right)^{\frac{1}{1 - \alpha}} \quad (25)
\]
Suppose that the decentralized economy created total vacancies equal to the constrained first best, \( v_{sp} = v \). Then only for \( \beta = 1 \) would investment coincide with the first-best solution. To ensure that \( v_{sp} = v \) would in turn require that \( \alpha = 0 \), which is inconsistent with the assumption that \( \alpha \in (0,1) \) (moreover, even if \( \alpha = 0 \) was allowed, it would imply that there was no vacancy creation in this equilibrium).

C  Results — FOR ONLINE PUBLICATION

This section provides additional quantitative results.

C.1 Restricting wage growth

In the benchmark results reported in the paper, I do not impose the restriction of zero wage growth after age 50. To verify that this does not bias results, I estimate a version of the empirical specification (2) on model generated data with or without imposing the restriction of zero wage growth after age 50. To be precise, I simulate a monthly approximation to the continuous time model for 18 countries for 50,000 individuals in each country for the entire life-cycle starting at age 24 and ending at age 54. Individuals enter as unemployed at age 24 with human capital drawn from the initial skill distribution \( \Lambda \).

I aggregate the data to the annual level and designated "May" as the survey month, and compute the wage during the prior "calendar" year. The wage is the sum of monthly income during the prior "calendar" year divided by the number of months worked. I construct labor market fluidity identically to the data in the simulated data, i.e. as the share of employed workers in "May" who were at a different main employer in "May" the previous year without any month of unemployment in between. I randomly select five consecutive years for each individual and drop the other years to mimic the amount of time an individual on average remains in the panel. Finally, as employment rates differ across these "countries," I design weights such that each of the 18 simulated countries receives the same weight in the aggregate.

Table 11 reports results from the benchmark model specification in the paper, which does not impose the empirical restriction of no wage growth after age 50, and an alternative specification that imposes no wage growth after age 50. Results are virtually identical.

C.2 Full sample results

Figure 25 plots life-cycle wage growth, unemployment and labor productivity against labor market fluidity in the model and the data for the full sample of 23 OECD countries. Results are similar.
Table 11. Life-cycle wage growth and labor market fluidity, model

<table>
<thead>
<tr>
<th></th>
<th>Unrestricted</th>
<th>Restricted</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.156***</td>
<td>0.156***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>N</td>
<td>3,982,295</td>
<td>3,982,295</td>
</tr>
</tbody>
</table>

Note: Simulated monthly data for 18 countries for 50,000 individuals per country, aggregated to the annual level to replicate as closely as possible the real data. See text for more details. $\alpha$: Fluidity-age interaction in regression (2) with worker fixed effects, time effects and restricted age effects. Standard error below are clustered at the country-level. Unrestricted: No restriction on wage effects. Restricted: No wage growth after age 50. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Constructed by first collapsing the simulated data to the country-age level, then to the country level. *** statistically significant at 1%. Source: Model.

Figure 22. The impact of labor market fluidity, model versus data

(A) Life-cycle wage growth

(B) Unemployment rate

(C) Labor productivity

Note: Men aged 25–54. Panel A. Wage growth between age 25–50. Data: Based on (1) with worker fixed effects, year effects and age effects restricted to not grow after age 50. Panel B. Data: Collapsed to the country-year-age level using the provided survey weights, then to the country-level. Model: Collapsed to the country-age level, then to the country-level. Panel C. Data: Log real output per hour in 2014 in 2004 PPP-adjusted US dollars. Regulation (model): Output per worker net of training and job creation costs including the wedge, $\Lambda$, $\left(1-u\right)/\left(1-i(a,z,h)zhdG(a,z,h)-(c+\Delta)V/(1-u)\right)$. Tax (model): Output per worker net of training and job creation costs excluding the wedge per worker, $\left(1-u\right)/\left(1-i(a,z,h)zhdG(a,z,h)-cV/(1-u)\right)$. Both model series are in logs and normalized to the data for the US. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Data: Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the survey weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, then the country-level. Source: Model, BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.
C.3 Sensitivity analysis

Figure 23 considers the impact of a 67 percent fall in the job finding rate, $p$, on life-cycle growth in human capital, $H$, and match productivity, $Z$, as each of the nine estimated parameters varies around its estimated value, holding the other parameters fixed at their estimated values. The predicted impact of changes in labor market fluidity on life-cycle growth in match productivity is sensitive in particular to the tail parameter of the match productivity distribution, $\zeta$. This makes sense—a fatter tail (higher value of $\zeta$) implies a greater scope for climbing the job ladder, and hence greater scope for labor market fluidity to influence productivity growth. This is also reflected in greater scope for human capital differences, since human capital accumulation responds to match productivity growth. Human capital accumulation is also sensitive to the scalar in the human capital technology, $\mu$. This also makes sense—if $\mu$ is higher, it implies greater growth in general in human capital over the life-cycle. Because human capital becomes more important with a higher $\mu$, it implies greater scope for labor market fluidity to reduce it.

Human capital is also somewhat sensitive to the curvature of the human capital technology, $\eta$, but this particular exercise confounds two forces. Because I scale the technology by $\mu/\eta$, changing $\eta$ impacts both the curvature and the scalar in the technology. I prefer to do it this way because it simplifies the first order conditions, but I have verified the following intuition by only changing the curvature of the technology, holding the scalar fixed. A higher $\eta$ (holding fixed $\mu/\eta$), makes the impact of fluidity on human capital accumulation more pronounced. This makes sense, as it effectively means investment is more elastic. The first order effect of a higher $\eta$, however, is through the change in the scale, $\mu/\eta$, and hence an increase in $\eta$ impacts the results the same way as a fall in $\mu$.

Finally, notice that workers’ bargaining power, $\beta$, impacts results in the expected way—a higher $\beta$ increases the scope for labor market fluidity to affect human capital accumulation. The impact, however, is not particularly pronounced, as it only impacts human capital accumulation through the incentive effect, whereas the unemployment effect and job ladder effect remain unaffected.

C.4 Unemployment and labor productivity

Table 12 shows the explanatory power of the mechanism for unemployment, EU and UE flows, and labor productivity. The mechanism overstates the empirical covariance between unemployment and labor market fluidity, predicting 161 percent of the higher unemployment across the OECD relative to the US (135 percent across the full sample). In an accounting sense, this is because the mechanism overstates the relationship between the EU rate and labor market fluidity in the data (138 and 126 percent across the core and full sample, respectively), whereas it gets that between the UE rate and labor market fluidity
about right (106 and 89 percent across the core and full sample, respectively). It matches 82 percent of the lower labor productivity across the OECD relative to the US (42 percent across the full sample). Figure 24 illustrates the EU and UE flows across countries in the model and data.

**Figure 23. Change in growth in match productivity and human capital in response to a 67% decline in the job finding rate as a function of each parameter**

(A) $\sigma$  
(B) $\mu$  
(C) $\eta$  
(D) $\phi$  
(E) $\delta_0$  
(F) $\delta_1$  
(G) $\zeta$  
(H) $\beta$  
(I) $p$

*Note:* Change in growth in match productivity $Z$ (dash-dotted purple) and human capital $H$ (dashed blue) between age 25–50 in response to an increase in the cost of job creation, $c$, such that the job finding rate falls by 67% as each of the parameters varies around its estimated value (solid black) holding all other parameters fixed. *Source:* Model.
### Table 12. The impact of labor market fluidity, model versus data

<table>
<thead>
<tr>
<th></th>
<th>AT</th>
<th>BE</th>
<th>DK</th>
<th>FI</th>
<th>FR</th>
<th>DE</th>
<th>EL</th>
<th>IE</th>
<th>IT</th>
<th>NL</th>
<th>ES</th>
<th>UK</th>
<th>US</th>
<th>Core</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Unemployment rate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>0.06</td>
<td>0.08</td>
<td>0.05</td>
<td>0.11</td>
<td>0.08</td>
<td>0.08</td>
<td>0.17</td>
<td>0.19</td>
<td>0.13</td>
<td>0.04</td>
<td>0.18</td>
<td>0.06</td>
<td>0.05</td>
<td>0.10</td>
<td>0.11</td>
</tr>
<tr>
<td>ΔUS&lt;sub&gt;d&lt;/sub&gt;</td>
<td>0.01</td>
<td>0.03</td>
<td>-0.00</td>
<td>0.06</td>
<td>0.03</td>
<td>0.03</td>
<td>0.12</td>
<td>0.14</td>
<td>0.08</td>
<td>-0.02</td>
<td>0.12</td>
<td>0.01</td>
<td>0.00</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>0.13</td>
<td>0.14</td>
<td>0.03</td>
<td>0.08</td>
<td>0.16</td>
<td>0.11</td>
<td>0.18</td>
<td>0.12</td>
<td>0.15</td>
<td>0.05</td>
<td>0.09</td>
<td>0.03</td>
<td>0.03</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td>ΔUS&lt;sub&gt;d&lt;/sub&gt;</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.06</td>
<td>0.03</td>
<td>0.09</td>
<td>0.15</td>
<td>0.09</td>
<td>0.13</td>
<td>0.03</td>
<td>0.06</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>ΔUS&lt;sub&gt;m&lt;/sub&gt;/ΔUS&lt;sub&gt;d&lt;/sub&gt;</td>
<td>10.72</td>
<td>3.90</td>
<td>-2.04</td>
<td>4.76</td>
<td>3.03</td>
<td>1.31</td>
<td>0.65</td>
<td>1.68</td>
<td>-1.65</td>
<td>0.51</td>
<td>-0.03</td>
<td>- -</td>
<td>- -</td>
<td>1.61</td>
<td>1.35</td>
</tr>
<tr>
<td><strong>Panel B. Monthly EU rate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>ΔUS&lt;sub&gt;d&lt;/sub&gt;</td>
<td>0.00</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.01</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>ΔUS&lt;sub&gt;d&lt;/sub&gt;</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ΔUS&lt;sub&gt;m&lt;/sub&gt;/ΔUS&lt;sub&gt;d&lt;/sub&gt;</td>
<td>1.12</td>
<td>-1.48</td>
<td>-1.27</td>
<td>0.51</td>
<td>3.80</td>
<td>1.87</td>
<td>0.76</td>
<td>3.24</td>
<td>-0.66</td>
<td>0.21</td>
<td>0.08</td>
<td>- -</td>
<td>1.38</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td><strong>Panel C. Monthly UE rate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>3.87</td>
<td>3.93</td>
<td>3.90</td>
<td>3.80</td>
<td>3.87</td>
<td>3.86</td>
<td>3.28</td>
<td>3.91</td>
<td>3.63</td>
<td>3.91</td>
<td>3.57</td>
<td>3.79</td>
<td>3.78</td>
<td>3.62</td>
<td>3.62</td>
</tr>
<tr>
<td>ΔUS&lt;sub&gt;d&lt;/sub&gt;</td>
<td>-0.13</td>
<td>-0.06</td>
<td>-0.10</td>
<td>-0.19</td>
<td>-0.12</td>
<td>-0.13</td>
<td>-0.72</td>
<td>-0.08</td>
<td>-0.36</td>
<td>-0.08</td>
<td>-0.42</td>
<td>-0.20</td>
<td>0.00</td>
<td>-0.21</td>
<td>-0.37</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>1.23</td>
<td>1.22</td>
<td>1.40</td>
<td>1.28</td>
<td>1.21</td>
<td>1.25</td>
<td>1.19</td>
<td>1.25</td>
<td>1.21</td>
<td>1.34</td>
<td>1.28</td>
<td>1.47</td>
<td>1.45</td>
<td>1.28</td>
<td>1.30</td>
</tr>
<tr>
<td>ΔUS&lt;sub&gt;d&lt;/sub&gt;</td>
<td>-0.22</td>
<td>-0.23</td>
<td>-0.06</td>
<td>-0.17</td>
<td>-0.24</td>
<td>-0.20</td>
<td>-0.26</td>
<td>-0.21</td>
<td>-0.24</td>
<td>-0.11</td>
<td>-0.18</td>
<td>0.02</td>
<td>0.00</td>
<td>-0.18</td>
<td>-0.15</td>
</tr>
<tr>
<td>ΔUS&lt;sub&gt;m&lt;/sub&gt;/ΔUS&lt;sub&gt;d&lt;/sub&gt;</td>
<td>1.79</td>
<td>3.87</td>
<td>0.61</td>
<td>0.90</td>
<td>1.96</td>
<td>1.60</td>
<td>0.36</td>
<td>2.59</td>
<td>0.68</td>
<td>1.38</td>
<td>0.42</td>
<td>-0.08</td>
<td>- -</td>
<td>0.82</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Note: Men aged 25–54. Labor productivity is for the entire private economy due to data availability. Panel A. Share of unemployed workers. Panel B. Share of employed workers who are unemployed in the subsequent month. Panel C. Share of unemployed workers who are employed in the subsequent month. Panels A–C. Data: Constructed by first collapsing the data to the country-year-age level using the provided survey weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, then to the country-level. Panel D. Log real output per hour. Data: 2014 real GDP per hour in 2004 PPP-adjusted US dollars. Model: Total net output, \( (1 - i(a, z, h))zhdG(a, z, h) \), divided by total employment, and in logs. \( \Delta US \): Difference in the outcome variable relative to the US. \( \Delta US_m/\Delta US_d \): Difference in the outcome variable relative to the US in the model relative to the data. Core: Average difference relative to the US across the 12 core countries. All: Average difference relative to the US across the 22 countries in the full sample. Source: Model, BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.

#### C.5 Additional outcomes

This section compares additional model predictions with the cross-country variation. The top left panel of Figure 25 plots the level of inequality against labor market fluidity, illustrating only a minor increase in inequality with labor market fluidity in both the model and the data. As shown in the top right panel, the higher inequality in more fluid labor markets is not accounted for by greater life-cycle growth in inequality, in either the model or the data. As shown in the bottom left panel, wages of labor market entrants are somewhat lower in more fluid labor markets, by a similar amount in the model and the data. Note that entry wages have been normalized to match those in the data for the US, as the levels are...
non-comparable in the model and data.

Panels A and B of Figure 26 plots the wage gain associated with JJ mobility in the model and data. Note that the US moment is higher in the model than the data because I target the monthly gain in the SIPP, which deviates (modestly) from the annual gain in the PSID. It illustrates the biggest discrepancy between the predictions of the model and the data. The model implies that the wage gain associated with JJ mobility should decline with labor market fluidity, as workers in more fluid labor markets on average are higher up the job ladder and hence have less scope to move further up the job ladder. An earlier version of this paper also allowed for so called godfather shocks—involuntary JJ mobility shocks meant to capture, for instance, the need to move to accompany a spouse, with the worker bargaining with the employer as though she is coming from unemployment Jolivet et al. (2006). To the extent that such shocks are equally common across countries, they serve to flatten the relationship between the wage gain from JJ mobility and labor market fluidity. The reason is that a larger share of JJ moves in less fluid labor markets is involuntary in nature, reducing the average gain from JJ mobility in such countries.

**Figure 24. The Impact of Labor Market Fluidity, Model Versus Data**

![Graph showing the impact of labor market fluidity, model versus data](image)

*Note:* Men aged 25–54. Panel A. Share of employed workers who are unemployed in the subsequent month. Panel B. Share of unemployed workers who are employed in the subsequent month. All panels. Data: Constructed by first collapsing the data to the country-age-year level using the provided survey-weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, and then the country-level. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Data: Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, then the country-level. Source: Model, BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.

To illustrate this point, Panel C provides the results from a rough calibration of this extended model,
with the frequency of godfather shocks set arbitrarily to 0.15 percent monthly. I recalibrate the scalar in the human capital technology, $\mu$, and the relative search efficiency from employment, $\phi$, to hit the same life-cycle wage growth and aggregate labor market fluidity, but leave other parameters unchanged. Even a small frequency of such shocks is sufficient to significantly flatten the relationship between wage gains associated with a JJ move and labor market fluidity, as a larger share of JJ moves in low fluidity countries now is due to such involuntary moves. The main prediction for life-cycle wage growth in Panel D (as well as other outcomes) remains essentially unchanged.

**Figure 25. The Impact of Labor Market Fluidity, Model versus Data**

(A) INEQUALITY  
(B) GROWTH IN INEQUALITY  
(C) ENTRY WAGES

*Note: Men aged 25–54. Panel A. St. d. of residual log hourly wages, controlling for year-education-age and year-occupation effects separately by country. Panel B. Growth in st. d. of residual log hourly wage between age 25–29 and age 50–54. Panel C. Log hourly real wages of workers aged 21–24 (data) / 24 (model). Model is normalized to equal the data for the US. All panels. Data: Constructed by first collapsing the data to the country-age-year level using the provided survey-weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, and then the country-level. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Data: Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, then the country-level. Source: Model, BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.*
Figure 26. Wage gain from JJ move, model versus data

(A) Core countries

(B) All countries

(C) Godfather shock extension

(D) Life-cycle wage growth, extended model

Note: Men aged 25–54. Panels A–C. Difference in annual wage growth between those who made a JJ move in the year and those who did not. Data: Constructed by first collapsing the data to the country-age-year level taking the median using the provided survey-weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, and then the country-level. Panel D. Wage growth between age 25–50. Data: Based on regression (1) with worker fixed effects, year effects and age effects restricted to not grow after age 50. Panels A–B. Benchmark model. Panel C–D. Extended model that incorporates godfather shocks. See text for more details. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Data: Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, then the country-level. All panels. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Data: Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, then the country-level. Source: Model, BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.
C.6 Wage employment

As noted in Appendix A.3, the share of self-employed workers is higher in less fluid labor markets. The model is consistent with this pattern if self-employment is akin to unemployment—something workers do when they cannot find a wage employment job. Figure 27 illustrates this by plotting the share of not-wage-employed men aged 25–54 against labor market fluidity in the model and data. The share in the US in the model and data differs by construction, since the model is estimated to match the unemployment rate in the US excluding the self-employed. Although the view that a large share of self-employed are necessity entrepreneurs does receive support in the literature (Poschke, 2019), it would be interesting to have a deeper theory of self-employment. That, however, is beyond this paper.

**Figure 27. Not-wage-employment (unemployed, self-employed and non-participants) and labor market fluidity**

![Graph showing not-wage-employment rate against labor market fluidity](image)

*Note: Men aged 25–54. Not wage employment rate: Share of men aged 25–54 that is not wage employed, i.e. unemployed, self-employed and non-participants. Data: Constructed by first collapsing the data to the country-year-age level using the provided survey weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, then to the country-level. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Data: Employment in the past 12 months includes self-employment due to data limitations. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country-level. Model: Constructed by first collapsing the data to the country-age level, then the country-level. Source: Model, BHPS, ECHP, EUSILC, GSOEP and PSID 1991–2015.*

D Training — FOR ONLINE PUBLICATION

This appendix contains additional details on training.
D.1 Training over the life-cycle

Figure 28 plots training over the life-cycle—days on training in the left panel and hours on training in the right panel, both expressed as a fraction of total work days/hours per year. The model moment is \( \frac{1}{A} \int i(a,z,h)g(z|a,h)dz e(a,h)dh \). As noted in the main text, the model significantly overstates the level of training in the data. As the level difference makes it difficult to assess the relative behavior of training over the life-cycle, I normalize the model moments to the empirical moment by first dividing by the model moment at age 25 and then multiplying by empirical moment at age 25. Since comparable training data is missing for the US, I use Denmark and construct the corresponding model moments by matching labor market fluidity in Denmark, holding all parameters fixed at their estimated US values. While the model overstates the amount of training in the data, it captures well the relative decline in training with age.

**Figure 28. Training over the life-cycle, model vs data**

(A) Days on training

(B) Hours on training

*Note: Men aged 25–54. Panel A. Days on vocational training in the past 12 months, expressed as a fraction of potential work hours (40*52). Panel B. Hours on vocational training in the past 12 months, expressed as a fraction of potential work hours (40*52). Both panels. Model moment is average investment, \( \frac{1}{A} \int i(a,z,h)g(z|a,h)dz e(a,h)dh \), normalized to the empirical moment by first dividing by the model moment at age 25 and then multiplying by empirical moment at age 25. Hours and days on training are top-coded at 13 weeks of full time training per year, and refer to Denmark as the highest fluidity country with training data. Model moments are for a hypothetical country with the same aggregate labor market fluidity as Denmark, holding all other parameters fixed at their estimated US values. All training measures: Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Source: ECHP 1995–2001.*

D.2 Additional training outcomes

Figure 29 plots hours on training in the past 12 months, whether the worker did any training since January in the year prior to the survey and the share of workers who report that their employer paid for
the training against labor market fluidity. If anything, the share rises with labor market fluidity, but the pattern is not statistically significant.

**Figure 29. Training and labor market fluidity, data**

(a) Hours on training (b) Whether trained

(c) Share paid for by employer

*Note: Men aged 25–54. Panel A. Hours on vocational training in the past 12 months, expressed as a fraction of potential work hours (40*52). Hours and days on training are top-coded at 13 weeks of full-time training per year. Panel B. Whether worker undertook any vocational training since January last year. Panel C. Share of workers who report that their employer paid for their vocational training since January last year. All training measures: Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Source: ECHP 1995–2001.*

**D.3 Not top-coding training**

My benchmark training results top-code days and hours on training at 13 weeks of full-time training during the past 12 months to limit the role of a few extreme observations. Table 13 reports results from
a regression of various training measures on fluidity without top-coding. The point estimates grow in magnitude, but the takeaway remains the same.

**Table 13. Training and Labor Market Fluidity without Top-Code**

<table>
<thead>
<tr>
<th>Panel A. Days (fraction of year)</th>
<th>Panel B. Hours (fraction of year)</th>
<th>Panel C. Whether trained</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw</td>
<td>Controls</td>
</tr>
<tr>
<td>Fluidity</td>
<td>1.947***</td>
<td>(0.253)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JJ</td>
<td>-0.007</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.002***</td>
<td>(0.001)</td>
</tr>
<tr>
<td>College</td>
<td>0.022***</td>
<td>(0.005)</td>
</tr>
<tr>
<td>N</td>
<td>135,563</td>
<td>115,897</td>
</tr>
</tbody>
</table>

Note: Men aged 25–54. Projection of training outcome on labor market fluidity without or with controls. Panel A. Days on vocational training in the past 12 months, expressed as a fraction of potential work days (5*52). Panel B. Hours on vocational training in the past 12 months, expressed as a fraction of potential work hours (40*52). Panel C. Whether worker undertook any vocational training since January last year. All panels. Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Standard errors are clustered at the country-level. * statistically significant at 10%; ** statistically significant at 5%; *** statistically significant at 1%. Source: ECHP 1995–2001.

**D.4 Training and Subsequent Mobility: the Generality of Human Capital**

This subsection uses the panel data on training and mobility to offer support for the notion that human capital is general. In particular, I hypothesize that if firm-specific human capital is important, then the incidence of training should be associated with a subsequent decline in the probability that the worker leaves the firm. The reason is that firm-specific human capital raises the value of the current match relative to alternative matches.

Table 14 relates JJ in the past 12 months to training in the year prior to that, i.e. 13–24 months ago (I have alternatively related mobility to training in the current year as well as two years ago with similar results). The first column shows the raw correlation with only country and year controls, the second adds age, education, occupation and sector controls, and the third controls for worker-fixed effects and year effects.35 There is no evidence that workers who trained more in the past year are less likely to make a JJ move in the current year. In fact, the raw data suggest the opposite, although this appears to be driven by selection. In any case, if skills had been primarily firm-specific, one may have hypothesized that worker mobility would decline with training.

35 The same conclusion holds with age instead of year controls.
Table 14. JJ Mobility and Lagged Training

<table>
<thead>
<tr>
<th></th>
<th>Panel A. Days (fraction of year)</th>
<th>Panel B. Hours (fraction of year)</th>
<th>Panel C. Whether trained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>Raw Controls Within</td>
<td>Raw Controls Within</td>
<td>Raw Controls Within</td>
</tr>
<tr>
<td></td>
<td>0.068* 0.044 0.014</td>
<td>0.115** 0.058 0.034</td>
<td>0.001 0.001 -0.004</td>
</tr>
<tr>
<td></td>
<td>(0.036) (0.033) (0.040)</td>
<td>(0.053) (0.047) (0.062)</td>
<td>(0.004) (0.003) (0.004)</td>
</tr>
<tr>
<td>N</td>
<td>49,114 48,289 46,247</td>
<td>49,011 48,190 46,134</td>
<td>49,285 48,460 46,434</td>
</tr>
</tbody>
</table>

Note: Men aged 25–54. Projection of JJ in the past year on various measures of vocational training 13–24 months ago, without or with controls. Raw: Country and year controls; Controls: Country, year, age, education and occupation controls; Within: Individual fixed effects and year controls. Panel A. Days on vocational training in the past 12 months, expressed as a fraction of potential work days (5*52). Panel B. Hours on vocational training in the past 12 months, expressed as a fraction of potential work hours (40*52). Panel C. Whether worker undertook any vocational training since January last year. All panels. JJ mobility: Whether worker started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment. Standard errors are clustered at the country-level. * statistically significant at 10%; ** statistically significant at 5%; *** statistically significant at 1%. Source: ECHP 1995–2001.

D.5 Firm wage and firm size/productivity

Table 15 reports results from the first stage regression (13) projecting pay on firm size bins (data) / match productivity bins (model) with controls. Larger firms pay better, consistent with previous research. As expected more productive matches on average pay better. In fact, the magnitude of the empirical and model relationship is about the same.

Table 15. Inferring Firm Pay

<table>
<thead>
<tr>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2$</td>
<td>0.101*** (0.005)</td>
</tr>
<tr>
<td></td>
<td>0.078*** (0.001)</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.188*** (0.006)</td>
</tr>
<tr>
<td></td>
<td>0.131*** (0.001)</td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>0.214*** (0.006)</td>
</tr>
<tr>
<td></td>
<td>0.182*** (0.001)</td>
</tr>
<tr>
<td>$\alpha_5$</td>
<td>0.265*** (0.006)</td>
</tr>
<tr>
<td></td>
<td>0.260*** (0.001)</td>
</tr>
<tr>
<td>$\alpha_6$</td>
<td>0.338*** (0.006)</td>
</tr>
<tr>
<td></td>
<td>0.370*** (0.001)</td>
</tr>
<tr>
<td>N</td>
<td>77,793 5,154,592</td>
</tr>
</tbody>
</table>

Note: Men aged 25–54. Projection of pay on firm size (data) / match productivity (model), each binned into 6 roughly equally large bins (in an employment-weighted sense), controlling for age, education, occupation and country-year (data) / age and country (model). Standard errors are clustered at the individual-level. *** statistically significant at 1%. Source: ECHP 1995–2001.
D.6 Training and firm size: the complementarity in production

The following intuition is clearest in the simple two period model, so I will restrict attention to that (but it carries through numerically to the general model). Recall from (7) that optimal investment is given by

\[ i = \frac{1}{z_1} \left( \mu \left( z_1 + p \beta z_2 - z_1 \right) \right)^{1/\eta} \]

Whether training rises or falls with match productivity, \( z_1 \), is ambiguous because of the opposing effects of the first and second terms.

Consider instead a model with no complementary in production, \( z + h \). Given the normalization \( h_0 = 1 \), the natural extension of the Ben-Porath (1967) investment problem would be

\[
\max_{i \in [0,1]} \left\{ (1-i)(z_1 + 1) + z_1 + \frac{p}{2} \beta (z_1 - z_1) + 1 + \mu \left( i(z_1 + 1) \right)^{\eta} \right\}
\]

with first-order condition

\[ i = \frac{1}{z_1 + 1} \mu^{\frac{1}{\eta}} \]

which is unambiguously declining in current productivity, \( z_1 \). Hence, one may hypothesize that this model would be inconsistent with the positive correlation between firm level pay and training. While a more thorough assessment is left for future work, I note that the current model matches well the empirical relationship between training and firm pay.

D.7 Share of employment at large firms

As noted in Figure 5, a larger share of workers work at large, higher paying firms in more fluid labor markets, where workers always train more.\(^{36}\) Given that the model abstracts from size, to compare this prediction with the model, I first note that the average share of workers in 50+ firms across these countries is very close to 50 percent in the data. Hence in the model, I compute the share of workers in the upper half of employment-weighted match productivity, constructed in the pooled data set of all the simulated countries. Figure 30 plots the share of workers in upper half of firm size in the data / upper half of match productivity in the model. Given that firm size is likely an imperfect proxy for match productivity, the mechanism does a decent job matching the empirical patterns.

\(^{36}\)The data on firm size is only available in the cross-sectional version of the EUSILC, which does not allow me to track individuals over time. It is also missing from the BHPS, so I use the available data from the EUSILC for the UK. The GSOEP codes firm size in non-comparable bins, and the PSID only asks for firm size in 1999 and from 2003 onwards.
Figure 30. SHARE OF WORKERS IN 50+ FIRMS, MODEL VS DATA

Note: Male wage employees aged 25–54. Data: Share of workers at 50+ employee firms. Model: Share of workers in the upper half of employment-weighted match productivity, constructed across all simulated countries. The upper half of employment-weighted match productivity corresponds closely to the average share of workers at 50+ firms across these countries (51.1%). Labor market fluidity: Share of employees who started working for their current employer at some point in the past 11 months while having been employed in all of the past 12 months. Employment in the past 12 months includes self-employment. Constructed by first collapsing the data to the country-age-year level using the provided survey weights, then to the country level. Source: ECHP, EUSILC and PSID 1994–2015.