The Aggregate and Distributional Effects of Spatial Frictions*

Sebastian Heise†
Federal Reserve Bank of New York

Tommaso Porzio‡
Columbia University and CEPR

March 25, 2021

Abstract

We develop a general equilibrium model of frictional labor reallocation across firms and regions, and use it to quantify the aggregate and distributional effects of spatial frictions that hinder worker mobility across regions in Germany. The model leverages matched employer-employee data to unpack spatial frictions into different types while isolating them from labor market frictions that operate also within region. The estimated model shows sizable spatial frictions between East and West Germany, especially due to the limited ability of workers to obtain job offers from more distant regions. Despite the large real wage gap between East and West of Germany, removing the spatial frictions leads, in equilibrium, to only a small increase in aggregate productivity and it mostly affects the within-region allocation of labor to firms rather than the between-region allocation. However, spatial frictions have large distributional consequences, as their removal drastically reduces the gap in lifetime earnings between East and West Germans.

JEL: J6, O1, R1

*We thank Michael Peters for a very insightful discussion of the paper. We also thank Andy Atkeson, Gharad Bryan, Paco Buera, Julieta Caunedo, Lorenzo Caliendo, Kevin Donovan, Niklas Engbom, Ben Faber, Pablo Fajgelbaum, Tarek Hassan, Kyle Herkenhoff, Fatih Karahan, Pete Klenow, David Lagakos, Mushfiq Mobarak, Ben Moll, Simon Mongey, Melanie Morten, Todd Schoellman, and Jonathan Vogel for very useful comments that improved the paper. We have also benefited from the reactions of several seminar and conference audience, including participants at the NBER SI EFMPL, NBER Growth, Berkeley, Columbia, LSE, UBC, UCLA, UPenn, University of Toronto. Rachel Williams provided excellent research assistance. The views and opinions expressed in this work do not necessarily represent the views of the Federal Reserve Bank of New York. This study uses the weakly anonymous Establishment History Panel (Years 1975 - 2014) and the Linked-Employer-Employee Data (LIAB) Longitudinal Model 1993-2014 (LIAB LM 9314). Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and remote data access. The study also uses data made available by the German Socio-Economic Panel Study at the German Institute for Economic Research (DIW), Berlin. Neither the original collectors of the data nor the archive bear any responsibility for the analyses or interpretations presented here.

†33 Liberty Street, New York, NY 10045, email: sebastian.heise@ny.frb.org.

‡3022 Broadway, New York, NY 10027, email: tommaso.porzio@columbia.edu.

1 Introduction

In many countries, large differences in wages and labor productivity across regions have persisted for decades (Gollin, Lagakos, and Waugh (2014), Redding and Rossi-Hansberg (2017)). Spatial frictions (e.g., moving costs that prevent workers from moving to high-wage regions) may be an important factor, as the regional differences cannot be explained by compensation for regional amenities or spatial sorting alone.[1] Spatial frictions could imply a misallocation of workers across regions and entail large aggregate productivity costs (Bryan and Morten (2019); Hsieh, Hurst, Jones, and Klenow (2019)).

Large differences in wages and labor productivity are also observed across firms, even within the same region.[2] These wage differences plausibly reflect labor market frictions (e.g., Burdett and Mortensen (1998)), which could affect the gains from removing spatial barriers. For example, moving workers to high-wage regions is not sufficient to generate aggregate gains if workers do not reach the high-paying firms in those regions. Moreover, workers' willingness to migrate influences how easily low productivity firms can expand relative to high productivity firms, and hence spatial frictions affect the within-region distribution of labor. To provide a proper quantitative assessment of the costs of spatial frictions, we must therefore take into account the allocation of labor both between and within regions, hence we should study *space* taking into account the role of *firms*.

In this paper, we build a general equilibrium framework of frictional labor reallocation across firms and regions, which we use to quantify the aggregate costs of spatial frictions in Germany. Our main result is that, despite the large and persistent real wage gap of 26% between the East and West of Germany, removing all spatial frictions only modestly raises aggregate output, and these aggregate gains are purely driven by the reallocation of labor within rather than between the two German regions.[3] This result is due to several counteracting forces. While removing spatial frictions allows workers to reach the most productive firms more easily regardless of their location – raising aggregate productivity – it also makes it easier for low productivity firms to attract unemployed workers from anywhere in Germany. This second mechanism allows low productivity firms to expand and offsets most of the gains. Moreover, we find large equilibrium effects: without spatial frictions, East-born workers move towards the more productive West, congesting the local labor market, which leads more West-born workers to move East. Overall, we do not find any net reallocation of workers towards the West in equilibrium.

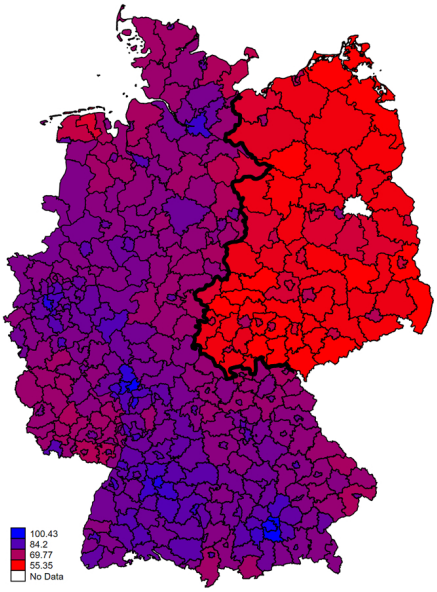
The estimated model yields two additional results. First, despite the limited aggregate effects of spatial barriers, they have large distributional consequences. Without spatial barriers, the gap

[1] See Combes, Duranton, and Gobillon (2008); Young (2013); Hicks, Kleemans, Li, and Miguel (2017); Lagakos, Marshall, Mobarak, Vernot, and Waugh (2020); Lagakos (2020)

[2] See, among others, Sorkin (2018); Song, Price, Guvenen, Bloom, and Von Wachter (2019).

[3] This result is in contrast to previous work which focused solely on the reallocation between regions and found large effects (e.g. Bryan and Morten (2019)).

Figure 1: Average Real Daily Wage, 2009-2014



Source: BHP, Bundesagentur für Arbeit. The figure shows real wages in each county, expressed in 2007 euros valued in Bonn, the former capital of West Germany, and using county-specific prices. Former East-West border is drawn in black for clarification. We exclude Berlin since we cannot assign it unambiguously to "East" or "West".

in lifetime earnings between East and West Germans is drastically reduced since workers' job decisions and opportunities no longer depend on their location. Second, our framework allows us to distinguish spatial search frictions, which prevent workers from accessing job opportunities across regions, from moving costs and home bias, which affect workers' willigness to move, and to study separately the aggregate effect of each spatial friction. We show that spatial search frictions have the largest independent effect on aggregate output while eliminating moving costs actually slightly reduces GDP since it worsens the within-region allocation of workers to firms.

In the first part of the paper, we use micro data from the German Federal Employment Agency to document three sets of facts. These facts show that Germany is an ideal setting to study regional wage gaps and motivate the ingredients of our model. First, we use the Establishment History Panel (BHP), a 50% sample of all establishments in Germany, to study wage gaps between and within the East and West of Germany. We show that there exists a large real average wage gap between the two regions (see Figure 1), and this wage gap has been constant over time and is not driven by observables such as industry or education. At the same time, we find that the within-region wage heterogeneity across establishments is larger than the average wage gap between regions, and that there is significant overlap between the East and the West German wage distributions. Second, we use the Linked Employer-Employee Data (LIAB) to show that all workers obtain large real wage gains when moving from East to West Germany. Conditional on observables, these gains are higher for East Germans than for West Germans, suggesting that workers need to be compensated to leave their home. Importantly, we find that workers also obtain substantial wage gains when moving jobs within East or West

Germany. Since any move across space is also a move across firms, the wage gains of cross-region movers thus combine the returns from switching across firms and the returns from migrating. Third, based on the LIAB data we document that there is substantial worker mobility between East and West Germany, but that, even conditional on current region, workers are much more likely to move towards their home region.

Next, we develop a general equilibrium framework to estimate the aggregate costs of spatial frictions. We build on a job-posting model à la Burdett-Mortensen (e.g., Burdett and Mortensen (1998)), in which workers move between heterogeneous firms subject to labor market frictions, but we extend this previous work to consider worker mobility across regions, subject to spatial barriers. The model is based on the micro empirical evidence: we allow for different types of spatial frictions to match the features of cross-region workers flows and wage gains, and we incorporate labor market frictions to capture the within-region wage dispersion. Our theory allows for an arbitrary number of regions characterized by an exogenous productivity distribution of firms, and an arbitrary number of worker types characterized by differences in skills, moving costs, location preferences, and search efficiency towards each region. Firms choose the wage to post and decide how many job vacancies to open. Workers decide how many job applications to submit to each region and receive random job offers, moving into and out of unemployment and across firms both within and between regions. Workers and firms meet according to a matching function that is concave in applications and vacancies, as in Diamond-Mortensen-Pissarides models (e.g., Pissarides (2000)), generating endogenous labor market tightness. We derive a tractable solution to the model represented by a system of differential equations.

Our model provides a framework to structurally identify the different components of spatial frictions and to isolate them from general labor market frictions. While all model parameters and frictions are jointly identified, we provide a heuristic identification argument. Within-region data on the joint distribution of wages and firm size, the average wage gains of job movers, and the frequency of job changes discipline the unobservable endogenous distribution of job offers in each region. Given a set of within-region distributions, the spatial frictions are identified by comparing the wage gains and job flows across regions to their within-region analogues. The key identifying assumption of the model is that the job offer distributions do not depend on the location or identity of the worker, and thus all workers draw offers from the same distributions irrespective of their current location. Under this assumption, higher observed wage gains for movers into a region compared to movers within that region reflect the presence of moving costs, as cross-region job switchers need to be compensated to move. Similarly, higher observed wage gains for workers moving out of their home region relative to other worker types making the same move identify home preferences. The relative frequency of job switches, instead, disciplines the search efficiency. Relatively lower worker flows across regions, compared to between firms within region, indicate that workers are less successful in applying for jobs in other regions.

We estimate the model with four sub-regions of Germany – which we refer to as *locations* – corresponding to the Northwest, Southwest, Northeast, and Southeast, and incorporate four worker types reflecting the four possible home locations. The model matches the data well, despite being relatively parsimonious with 21 parameters being used to match 305 micro and aggregate moments.

The model estimates imply substantial spatial barriers. The most important of these barriers is a lower search efficiency across locations: for a given search effort, workers generate only 1/20th as many job applications when searching for jobs across locations as within. Search efficiency is also biased towards each individual's home: a given search effort across locations directed towards the home location results in four times as many applications as the same effort directed to jobs outside of it. In contrast to the large spatial search frictions, we estimate a direct cost of moving between any two locations of only between 3.1% and 5.3% of life-time income, dependent on the distance, which is considerably smaller than most previous estimates.[4] These relatively small moving costs reflect our model's ability to distinguish actual moving costs from other spatial frictions and to control for the frictional labor market. In terms of home preferences, we find that workers need to be paid 7.4% of their yearly income to move away from their home location and maintain the same utility.

Our model also allows us to quantify the contribution of unobserved individual characteristics to the East-West wage gap. The wages of East German workers are approximately 10% lower than those of West German workers even while working at the same firm, which, structurally interpreted through our model, implies a 10% difference in unobserved ability. Due to the sorting of workers towards their home location, these ability differences explain more than one third of the East-West regional wage gap.

We use the model to study the aggregate and distributional impact of spatial frictions in Germany. We find that eliminating all spatial barriers would decrease the wage gap between East and West Germany by half, mostly due to the reduction in workers' sorting towards their home location. However, aggregate GDP per capita would increase by a mere 2%. The reason for this result is that the estimated labor misallocation is mainly across firms within locations, rather than between East and West Germany, and the effect of spatial frictions on the within-location allocation of labor is mediated by several counteracting forces. While removing spatial frictions allows workers to find more jobs and to climb a country-wide job ladder – thus increasing labor productivity – eliminating spatial frictions also makes it easier for low productivity firms to hire unemployed workers from far away locations, which depresses productivity. In addition, eliminating spatial barriers raises the share of East German workers in the West by 40 percentage points, but the additional competition for jobs encourages more West German workers to migrate to the East, resulting in no net labor reallocation towards the

[4]See, for example, Bryan and Morten (2019) and Kennan and Walker (2011).

West.

In contrast to the small aggregate gains, the distributional consequences of spatial frictions are large. Workers' home location has persistent effects on their earnings throughout their lifetime: the average wage per efficiency unit of East Germans is about 11% below the one of West Germans. Removing spatial frictions almost eliminates this gap since all workers have equal access to jobs in East and West Germany and climb a similar job ladder.

Overall, our results highlight the importance of studying the allocation of labor within and across regions in a unified general equilibrium framework, hence to study *space* and *firms* jointly.

Literature. Our paper contributes to several strands of literature.

First, we contribute to the literature quantifying the size of spatial barriers and their aggregate effects (Caliendo, Opromolla, Parro, and Sforza (2017) and Bryan and Morten (2019)).[5] This literature has used observed worker flows and average wage differentials across space to estimate the size of the moving costs. Since worker flows in response to average wage gaps are relatively modest, even after accounting for compensating disamenities, the papers infer large moving costs, which suggest substantial aggregate gains from reallocating workers. Our framework allows us to benchmark worker mobility across space to mobility across firms in a frictional labor market. We find that, despite sizeable spatial barriers, the aggregate gains from removing them are modest because most of the labor misallocation is within regions, and removing spatial barriers does not substantially improve the within-region allocation of workers to firms. Overall, we argue that firms, and firm level-data, should have a prominent role in the analysis of spatial wage gaps.

Second, a recent literature has used panel data to study the observational returns from migration and to quantify the contribution of workers' sorting to regional wage gaps (see Hicks, Kleemans, Li, and Miguel (2017), Alvarez (2018), and Lagakos, Marshall, Mobarak, Vernot, and Waugh (2020)).[6] We show that the interpretation of panel data used in this literature can be misleading. In our setting, the wages of East-born workers increase steeply when moving West, which the cited literature would interpret as evidence of a large causal effect of working in the West, hence of large returns from reducing spatial barriers. This conclusion does not take into account, however, that labor markets are frictional, and that all job movers are selected − they must have received a good enough job offer to move. Moreover, removing spatial barriers can lead to equilibrium effects. Our work controls for movers' selection by benchmarking the wage gains of movers between regions to those within regions, and computes the aggregate gains in equilibrium. We conclude that removing spatial frictions provides smaller gains than implied by an a-theoretical interpretation of the data.

[5]See also Artuç, Chaudhuri, and McLaren (2010), Kennan and Walker (2011), Caliendo, Dvorkin, and Parro (2019).

[6]Other relevant papers on sorting, using different methods, are Young (2013), Lagakos and Waugh (2013).

Third, our work is related to job ladder models à la Burdett and Mortensen (1998) with labor mobility across sectors or space. Schmutz and Sidibé (2018) build a partial equilibrium model where identical workers receive job offers both from their current location and from other locations. Consistent with our work, they estimate relatively small moving costs and sizable search frictions across space. However, due to the partial equilibrium assumption their paper cannot study the aggregate effects of removing these spatial barriers, and due to the assumption of homogeneous labor the paper cannot study the distributional effects of spatial frictions. Bradley, Postel-Vinay, and Turon (2017) analyze wage posting and employment in a Burdett-Mortensen setup in the presence of an exogenous public sector, and Meghir, Narita, and Robin (2015) develop a general equilibrium model with two sectors to study the allocation of labor between the formal and informal sectors in Brazil. In both papers, workers receive identical job offers from both sectors *independently* of their current employment status. As a result, there is one unified labor market, and the wage function is continuous as in the standard Burdett-Mortensen model. In our model, workers' probability of receiving and accepting offers depends on their identity and their current location due to the presence of spatial frictions, which could make the wage functions discontinuous in principle. We resolve this problem by introducing extreme value shocks, building on earlier insights to obtain tractable solutions for discrete choice problems from the trade literature (e.g., Eaton and Kortum (2002)).[7]

Last, our work is related to the literature on East German convergence (or the lack thereof) after the reunification (e.g., Burda and Hunt (2001), Burda (2006)). This literature has studied the possible drivers behind the East-West wage gap and the nature of migration between the two regions (Krueger and Pischke (1995), Hunt (2001, 2006), Fuchs-Schündeln, Krueger, and Sommer (2010)). Uhlig (2006, 2008) shows that the persistent East-West wage gap is consistent with network externalities, which could discourage firms from moving to the East. In contrast to this work, we take the distribution of firms in each region as exogenously given and do not explicitly model the source of the productivity differences.[8] Instead, we focus on spatial barriers to worker mobility and estimate the aggregate effects of removing them.

Our paper proceeds as follows. In Section 2 we describe our data, and Section 3 documents stylized facts on the German labor market. Section 4 introduces the model and Section 5 discusses how to unpack spatial frictions. We estimate the model in Section 6 and we use it to quantify the aggregate and distributional effects of spatial frictions in Section 7. Section 8 concludes.

[7]Two other related papers are Hoffmann and Shi (2016) and Bilal (2019). The first studies a two-sector Burdett-Mortensen model with no mobility frictions; the second studies unemployment differences across space.

[8]For recent related work that models the endogenous productivity differences across regions, see Fajgelbaum and Gaubert (2020); Bilal (2019); Schmutz and Sidibé (2021).

2 Data

We use two main datasets provided by the German Federal Employment Agency (BA) via the Institute for Employment Research (IAB): i) the Establishment History Panel (BHP) and ii) the longitudinal version of the Linked Employer-Employee Dataset (LIAB).

The BHP is a panel containing a 50% random sample of all establishments in Germany with at least one employee liable to social security on June 30th of a given year. The data are based on mandatory social security filings and exclude government employees and the self-employed. Each establishment in the BHP is defined as a company's unit operating in a distinct county and industry.[9] For simplicity, we will refer to these units as "firms" from now on. For each such firm in each year, the dataset contains information on location, average wages, the number of employees, and employee characteristics (education, age, gender). The data are recorded since 1975 for West Germany and since 1992 for East Germany, and they cover about 1.3 million firms per year in the recent period.

The LIAB data contain records for more than 1.9 million individuals drawn from the Integrated Employment Biographies (IEB) of the IAB, which cover all individuals that were employed subject to social security or received social security benefits since 1993. These data are linked to information about the approximately 400,000 firms at which these individuals work from the BHP. For each individual in the sample, the data provide the entire employment history for the period 1993-2014, including unemployment periods as long as the individual received unemployment benefits. Each observation is an employment or unemployment spell, with exact beginning and end dates within a given year.[10] A new spell is recorded each time an individual's employment status changes, for example due to a change in job, wage, or employment status. For individuals that do not change employment status, one spell is recorded for the entire year. Variables include the worker's firm's location at the county level, the worker's daily wage, education, and year of birth.

An important variable for our analysis is each worker's county of residence, reported in the LIAB since 1999, which we will use to analyze workers' mobility across space. In contrast to the other variables, which are newly reported at each spell, the location of residence is recorded at the end of each year for employed workers and at the beginning of an unemployment spell for unemployed workers and then added to all observations of that year. Since the social security reporting regulations do not prescribe which residence to report for workers with multiple residences, some workers can report very large distances between residence and work location even though they live in a second home closer to work. To deal with the potential measurement

[9] Since several plants of the same company may operate in the same county and industry, the establishments in the BHP do not always correspond to economic units such as a plant (Hethey-Maier and Schmieder (2013)).

[10] We use the term unemployment spell to refer to the period in which an individual is receiving unemployment benefits. After the expiration of the benefits, individuals are not in our dataset until they are employed again.

error, we will define several alternative measures of migration below.

We use three additional datasets. First, we obtain information on cost of living differences across German counties from the Federal Institute for Building, Urban Affairs and Spatial Development (BBSR (2009)), which we will use to construct real wages. The BBSR conducted a study assessing regional price variation in 2007 across 393 German micro regions covering all of Germany that correspond to counties or slightly larger unions of counties.[11] Second, we supplement our main analysis with data from the German Socio-Economic Panel (SOEP), an annual survey of around 30,000 individuals in Germany since 1984, to examine additional demographic characteristics and to corroborate some of our main findings. Finally, we use information on firms' profit shares from the ORBIS database by Bureau van Dijk to compute additional targets for the model's estimation.

Sample Construction. Since we are interested in the persistent divide between East and West Germany, our analysis focuses on the years 2009 to 2014, the latest years available in the IAB data. We refer to this period as our baseline sample. For some empirical specifications that require a longer sample, we use the years 2004 to 2014. We construct real wages for each county using the BBSR's price index, which we deflate forward and backward in time using state-specific GDP deflators from the statistics offices of the German states. We use time-consistent industry codes at the 3-digit WZ93 level provided by the IAB based on the concordance by Eberle, Jacobebbinghaus, Ludsteck, and Witter (2011). Since wages are only reported to the IAB up to the upper limit for statutory pension insurance contributions, the BHP contains an imputed average wage variable which estimates the censored wages based on Card, Heining, and Kline (2013). For the LIAB, no such variable is provided and we replicate the imputation steps ourselves. We use the corrected wages for all our analyses. We use full-time workers only, and exclude Berlin, which cannot be unambiguously assigned to East or West since it was divided between the two. We provide additional details on the datasets and on data construction in Appendix A.

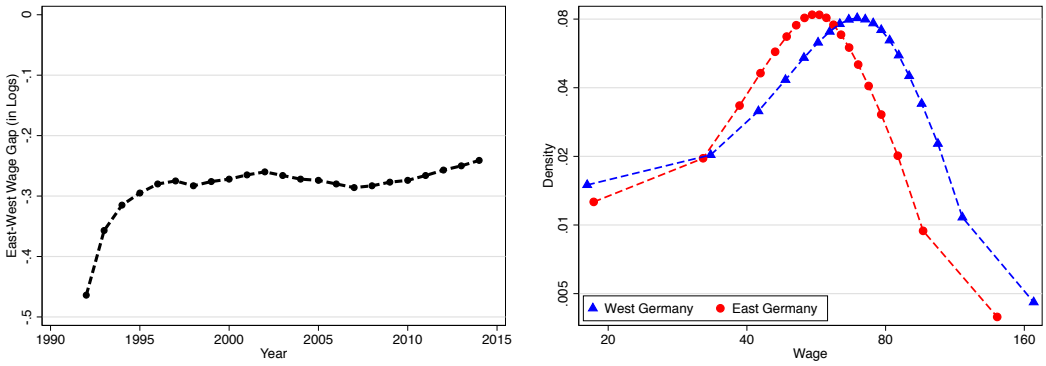
3 Motivating Facts

We next document three sets of facts on the German labor market. These facts highlight that Germany is a good setting to study regional wage gaps and serve as motivation for the main

[11] The data cover about two thirds of the consumption basket, including housing rents, food, durables, holidays, and utilities. We provide further information on the data in Appendix A and provide a map of county-level price levels. East Germany has a 7% lower population-weighted average price level.

Figure 2: Wages Between and Within Regions

(a) East-West Real Wage Gap (b) Within-Region Real Wage Distributions



Source: BHP and authors' calculations. Notes: The left figure plots the coefficients on the East Germany dummy obtained from running specification (1) separately for each year without controls, where firms are weighted by size. The right figure plots the density of wages across firms separately for East and West Germany for the period 2009-2014. Wages are residualized by regressing the log real wage on 3-digit industry dummies and time dummies, for East and West Germany separately. We generate the cleaned wage as the residuals from this regression plus the mean of the log wage in the given region and transform these log wages back into levels. We then find the twentiles of the residualized wage distribution, compute the average wage within each twentile, and transform it into a density. While all firms are weighted equally, only a very small share of overall employment is at the lowest wage firms.

ingredients of our model.

3.1 Wage Gaps Between and Within Regions

We first show that a sizable and persistent real wage gap remains between East and West Germany, despite the absence of a physical or legal border or language difference, since the reunification in 1990.[12] We run, in the BHP, firm-level regressions of the form[13]

$$\log(\bar{w}_{jt}) = \gamma \mathbb{I}_{j,East} + \beta X_{jt} + \delta_t + \epsilon_{jt}, \tag{1}$$

where $\bar{w}_{jt}$ is the average real wage paid by firm j in year t, $\mathbb{I}_{j,East}$ is a dummy for whether firm j is located in the East, X_{jt} is a vector of controls, and δ_t are time fixed effects. We weight by firm size, measured by full-time workers, since we are interested in the average wage gap in Germany.

In Figure 2a we plot coefficients γ_t from running regression (1) separately for each year in the data and without any additional controls. The real wage gap between East and West has been closing very slowly since the mid-1990s, and remains at around 25%.[14]

We next pool the data for our core sample period (2009-2014) and investigate, in Table 1, the role of different controls in explaining the wage gap. Worker gender and education (column

[12]Supplemental Appendix K provides a brief discussion of the reunification process. This Supplemental Appendix is not meant for publication and includes additional material. It is available on the authors' websites.

[13]Recall that we refer to establishment units as "firms".

[14]In Supplemental Appendix L we use aggregate data on GDP to perform a growth accounting exercise to show that most of the sizable GDP gap between East and West Germany today is due to TFP differences.

Table 1: Effect of Region on Real Wage

Dep var.: $\log(\bar{w}_{jt})$	(1)	(2)	(3)	(4)
$\mathbb{I}_{j,East}$	$-\mathbf{.2609}^{***}$	$-\mathbf{.2695}^{***}$	$-\mathbf{.2467}^{***}$	$-\mathbf{.2052}^{***}$
	(.0074)	(.0058)	(.0031)	(.0027)
Year FE	Y	Y	Y	Y
Gender & Education	–	Y	Y	Y
Age & Firm Size	–	–	Y	Y
Industry FE	–	–	–	Y
Observations	4,797,798	4,741,107	4,725,435	4,725,210

Source: BHP and authors' calculations. Notes: The table presents the estimates on the East Germany dummy from specification (1) for the period 2009-2014, where firms are weighted by size. $*$, $**$, and $***$ indicate significance at the 90th, 95th, and 99th percent level, respectively. Standard errors are clustered at the firm-level.

2) and worker age and firm size (column 3) do not contribute significantly to the wage gap. Controlling for 3-digit industries narrows the gap slightly (column 4), but overall about 80% of the real wage gap remains unexplained.

While, on average, firms pay a higher wage in the West, the distribution of wages paid in the East and West of Germany show substantial overlap. In fact, many firms in the East pay higher wages than some West firms and, even looking within the same 3-digit industry, the gap between the firms at the top and at the bottom of the within-region distribution is larger than the average wage gap between regions. This result can be seen in Figure 2b, which plots PDFs of firms' average real wage from the BHP, separately for both East and West Germany. We pool all observations from our core sample period and residualize log real wages by regressing them on year dummies and 3-digit industry dummies to remove across industry variation. We then find the twentiles of the residualized wage distribution, compute the average real wage (in levels) for each twentile, and plot the associated density.

In Supplemental Appendix M[15], we provide some additional robustness checks and show that the between-region wage gap is similar for all industries and across counties of different education or gender composition. Moreover, we show that the wage gap is not driven by a few outlier counties, that there are no clearly delineated regional differences in tax rates, and that the wage gap is accompanied by a large gap in unemployment. We also provide additional details on the within-region joint distribution of wages and firm size, and show that there is substantial wage heterogeneity across firms even within the same county.

3.2 Wage Gains of Movers

Second, we investigate the wage gains that individuals obtain when moving between East and West Germany and between firms. The large East-West wage gap would suggest that workers

[15] Available on the authors' websites and not meant for publication.

can obtain substantial gains from moving to the West. We confirm this hypothesis. However, we show that focusing on migrants' wage gains alone provides a misleading picture. A move across space is also a move across firms, and job switchers within-region also obtain large wage gains. Therefore, to isolate the returns from migration, we need to compare moves across regions to those within region.

Empirical Specification. We define job-to-job movers as workers that change jobs between two firms without an intermittent unemployment spell. For moves between East and West Germany, we distinguish between migration and commuting. The distinction is useful because we expect that commuters to a new job are paid a smaller wage premium than workers that also have to move their residence. We classify job-to-job movers between East and West Germany as migrants if they report a different county of residence in the year of the move from the previous year.[16] All other moves between East and West are defined as commuting. As discussed above, the residence variable is subject to measurement error. Our migration measure only includes workers that actively change their recorded residence in the year of the move. We provide several summary statistics on our migration measure in Appendix B.

Our analysis distinguishes individuals based on their "home region", either East or West Germany. The migration literature has shown that individuals display *home bias*: the birthplace is a key determinant of job flows across space even among adults (see Kennan and Walker (2011)). Since our social security data do not contain information on birth location, we classify individuals as East (West) German if at the first time they appear in our entire dataset since 1993, either employed or unemployed, they are in the East (West). Appendix A provides additional details on the construction of the home region. Our measure is imperfect, since some individuals migrated between the reunification and 1993. In Appendix C, we use survey data from the SOEP, which include individuals' actual birth location, to provide several validation exercises of our measure of home region. Overall, our results suggest that our measure properly classifies individuals into the region in which they were born in more than 90% of the cases. For this reason, we will interpret workers' home region also as their birth region going forward, and refer to individuals whose home is East as East-born.[17]

Let d_{it}^{x} be a dummy for a job switch of type $s \in \mathbb{S}$, where $\mathbb{S}$ is the set of the six possible types of moves: i) from East to West via migration or ii) commuting; iii) from West to East via migration or iv) commuting; v) within-East, and vi) within-West. To visualize an individual's wage dynamics around the time of a job-to-job move, we run a standard system of local projections,

[16] We compare residence location across years since, as discussed above, the variable is only updated at the end of each year.

[17] None of our results hinge on the home region being the birth region, though it does alter the interpretation. An alternative interpretation would be that an individual's location when they first enter the labor market shapes their attachment and biases.

consisting of one regression for each time period $\tau \in \{t-3, ..., t-1, t+1, ..., t+5\}$ around t:[18]

$$\Delta \log(w_{i\tau}) = \sum_{s \in \mathbb{S}} \beta_{s,\tau}^{West} d_{it}^{s}(1 - \mathbb{I}_i^{East}) + \sum_{s \in \mathbb{S}} \beta_{s,\tau}^{East} d_{it}^{s} \mathbb{I}_i^{East} + B_\tau X_{it} + \epsilon_{it}, \tag{2}$$

where $w_{i\tau}$ is an individual's weighted average wage across all employment spells in year τ, where we use each spell's length as its weight. The variable $\Delta \log(w_{i\tau})$ is the log change of this average wage between year τ and the previous year except for $t+1$, where it is the difference with respect to $t-1$. We drop wages from the year of the move to avoid contaminating our results by other types of payments in the year of the move.[19] The variable $\mathbb{I}_i^{East}$ is a dummy for whether an individual's home region is East Germany. Finally, the controls X_{it} include dummies for the current work region, home region, and their interaction, distance dummies since moves further away could lead to higher wage gains, the total number of past job-to-job switches, age controls, and year fixed effects. Since the left hand side variable is wage growth, any difference across individuals in the wage level would be netted out. Therefore, we do not include individual fixed effects in our main specification. The coefficients $\beta_{s,\tau}^{West}$ and $\beta_{s,\tau}^{East}$ capture the real wage gains from making a job-to-job transition relative to the wage growth obtained by staying at the same firm, which is the omitted category.

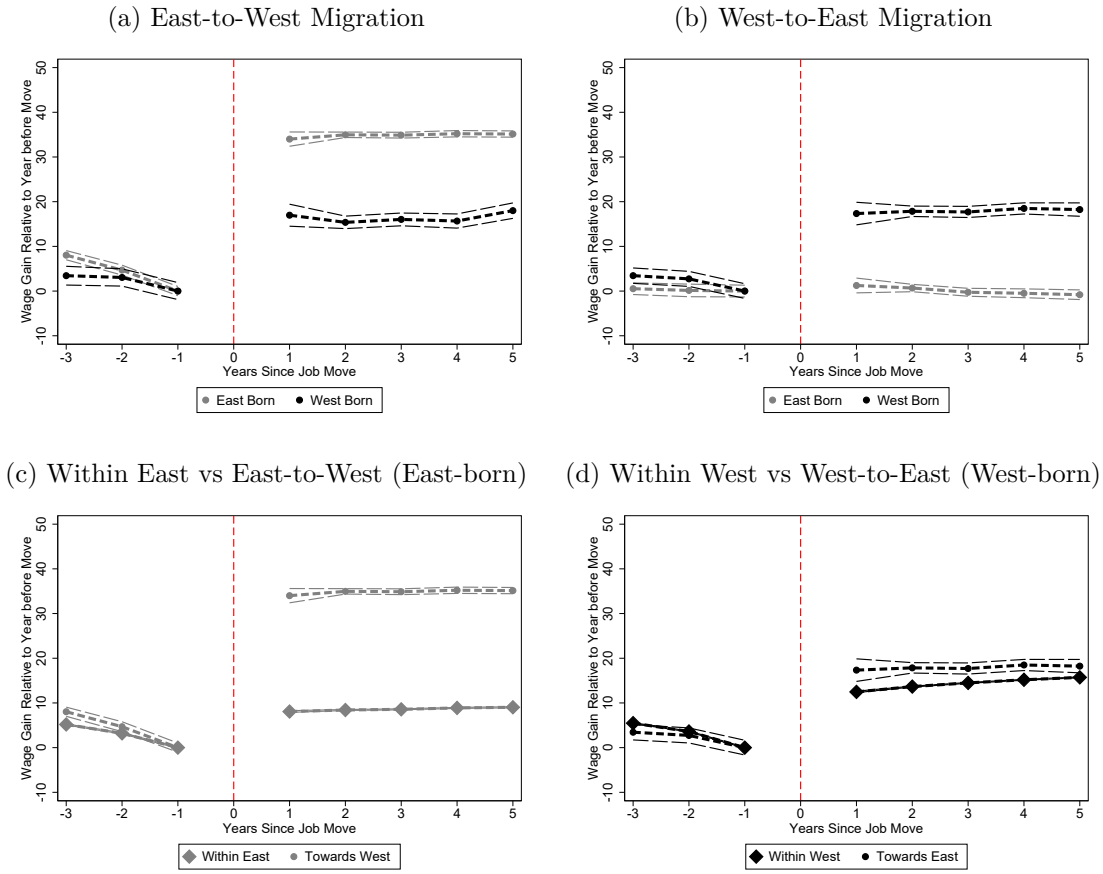
Results. Figure 3a plots the estimated wage gains for East-to-West migration – i.e. the predicted wage from the relevant coefficients $\beta_{s,\tau}^{West}$ and $\beta_{s,\tau}^{East}$, translated into levels, and normalized around the wage level prior to the year of the migration. Figure 3b presents the wage gains for West-to-East migration. The figures highlight that workers moving out of their home region see their wage increase steeply. East-born movers to the West receive on average almost a 35% real wage increase relative to their average within-firm wage growth, which is almost double the wage gain obtained by West-born workers making the same move. Moves to the East, instead, are associated with sizable wage gains for West-born workers and almost no effect for East-born ones. Average wage gains for moves to the East tend to be smaller, consistent with the lower average wage level in the East.

The large wage gains from moving West could imply the presence of substantial moving costs, and their asymmetry suggests that workers need to be compensated to leave their home region. However, any move across regions is also a move across firms. Figures 3c-3d plot the estimated wage gains for within-region job-to-job switches from regression (2) against the wage gains from a migration away from home. We find that workers experience fairly large gains even moving jobs within-region, suggesting that they are climbing a job ladder in the presence

[18] We pool together all the data for time periods t from 2004 to 2014 thus creating an unbalanced panel. In general, working with an unbalanced panel could be problematic. In our application, we are less concerned because: i) we do not observe post-trends; and ii) we are mostly interested in the wage growth on impact.

[19] The results are similar if we include year t, see Supplemental Appendix N.

Figure 3: Wage Gains for Job-to-Job Moves



Source: LIAB and authors' calculations. Notes: The figure is constructed by taking the point estimates for different sets of coefficients $\beta^{West}_{s,\tau}$ and $\beta^{East}_{s,\tau}$ from the regressions (2) for $\tau \in \{t-3, ..., t-1, t+1, t+5\}$. We then sum up the coefficients starting at $\tau = -3$ to obtain for each period τ the sum $\sum_{u=-3}^{\tau} \beta^{i}_{s,u}$, where $i \in \{\text{West, East}\}$, and subtract from this sum the term $\sum_{u=-3}^{-1} \beta^{i}_{s,u}$ to normalize the coefficients with respect to period $\tau = -1$. The dotted lines represent the 95% confidence intervals. The top left panel shows the normalized coefficients for $\beta^{West}_{EW,\tau}$ and $\beta^{East}_{EW,\tau}$. The top right panel shows the normalized coefficients for $\beta^{West}_{WE,\tau}$ and $\beta^{East}_{WE,\tau}$. The bottom left panel shows $\beta^{East}_{EE,\tau}$ and $\beta^{East}_{EW,\tau}$, and the bottom right panel shows $\beta^{West}_{WW,\tau}$ and $\beta^{West}_{WE,\tau}$.

of labor market frictions. As a result, to properly infer the cost of moving between regions, we need to benchmark the cross-regional wage changes with the within-region gains, taking into account that workers moving between regions are selected: they are the ones that received job offers sufficiently appealing to make them migrate. Our model will allow us to do so structurally.

In Supplemental Appendix N, we list the full estimates from specification (2), and show that our results are robust to alternative definitions of job-to-job switches and migration. We also present the results for commuting moves, which exhibit wage gains that are, as expected, smaller but show the same qualitative features as migration moves. We also show that the results are robust to the inclusion of individual fixed effects that account for differences across individuals in the steepness of their experience profiles (as in Guvenen, Karahan, Ozkan, and Song (2015)). Finally, we perform the regression for different demographic subgroups of workers. Across all results we find asymmetric wage gains, compatible with home bias.

3.3 Workers' Flows

Last, we study the flows of workers across regions. We show that, although the persistent wage gap would suggest otherwise, the labor markets of East and West Germany are, in fact, quite integrated: workers frequently move across the former border; however, worker flows are biased towards their home region. While many workers leave for a few years, most eventually return.

Summary Statistics on Cross-Regional Mobility. Table 2 presents mobility statistics for workers with at least one employment spell in our core sample period 2009-2014. Row 1 shows that in our core period 2.9% of employment spells by West-born workers and 17.5% of spells by East-born workers are not in their home region. Of the workers in our sample, 4.6% of West-born and 23.9% of East-born have at some point had a full-time job in the other region (row 2). However, a sizable fraction of job changes are via commuting: only 1.8% of West-born and 10.2% of East-born have resided in the other region.[20] Row 3 indicates that between one third and one half of the workers taking a job in the other region have since returned to a job at home, after spending on average only 2-3 years away (row 4). The average non-returner is employed in the other region, until her employment history ends, for longer than the average returner, as expected due to a simple selection argument (row 5). The final three rows present characteristics of workers that never left their home region ("Stayers"), took a job in the other region ("Movers"), or took a job and have returned ("Returners"). Overall, while we find some selection on observables, the data show that both workers with and without college education migrate between regions. In fact, less educated workers comprise about three quarters of all cross-region switchers.[21] We present additional statistics on movers in Appendix B, and show that the share of workers away from their home region has been relatively stable in the recent period.

Empirical Gravity Specification on Worker Flows. We next estimate a gravity equation for workers' flows between counties to separate the role of geographical barriers and identity (home bias) barriers between East and West Germany.

Let $n^h_{o,d,t}$ be the number of workers with home region h that were in a job in county o in year $t-1$ and that have made any job switch to a new job in county d in year t, including both via migration and commuting for cross-region moves.[22] We compute the share of these job-to-job switchers from county o moving to county d (where d can be equal to o) across all years in our

[20] As discussed, however, the residence variable may be misreported. We provide further statistics on workers' residence in Appendix B.

[21] We observe a higher share of males than in the general population since our sample consists only of full-time workers.

[22] We include all moves to maximize the number of county-pairs for which we observe positive flows and minimize the risk of biases due to granularity in our data, see Dingel and Tintelnot (2020). In Supplementary Appendix O we show that our results are robust to alternative definitions, such as using only migrants.

Table 2: Summary Statistics on Mobility

		(1) Home: West			(2) Home: East		
(1)	FT job spells in foreign region	2.9%			17.5%		
(2)	Crossed border (job / residence)	4.6% / 1.8%			23.9% / 10.2%		
(3)	Returned movers	46.3%			36.1%		
(4)	Mean years away (returners)	2.90			2.41		
(5)	Mean years away (non-returners)	9.41			7.47		
		Stayers	Movers	Returners	Stayers	Movers	Returners
(6)	Age at first move	–	34.4	34.0	–	32.0	31.8
(7)	Share college	0.22	0.30	0.29	0.20	0.19	0.18
(8)	Share male	0.70	0.78	0.82	0.56	0.74	0.80

Source: LIAB. Notes: The table shows statistics for workers with at least one full-time employment spell in 2009-2014. For these workers, row 1 shows the share of full-time employment spells at jobs not in their home region. Row 2 shows the share of these workers that have ever had a full-time job / resided in their non-home region over the entire sample since 1993. Row 3 shows the share of workers that returned to a job in their home region after their first job in the non-home region, and row 4 presents the average number of years away. Row 5 shows the time passed between the last year the worker is in the data and the year of the first job away from home for workers that never again take a job in their home region. Rows 6-8 present the average age at the time of the first move away from the home region, college share, and male share among workers that have never taken a job outside of their home region ("Stayers"), workers that have moved ("Movers"), and workers that have moved and returned ("Returners").

core period as

$$s^h_{o,d} = \frac{\sum_t n^h_{o,d,t}}{\sum_t \sum_{d \in \mathbb{D}} n^h_{o,d,t}}$$

where $\mathbb{D}$ is the set of all the 402 counties in both East and West Germany.[23] We use these shares to fit the gravity equation

$$\log s^h_{o,d} = \delta^h_o + \gamma^h_d + \sum_{x \in \mathbb{X}} \phi_x D_{x,o,d} + \rho \mathbb{I}_{R(o) \neq R(d)} + \epsilon^h_{o,d}, \tag{3}$$

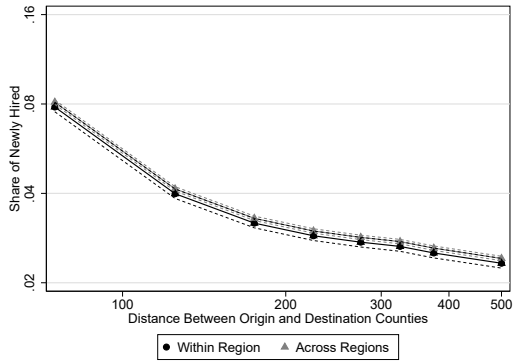
where δ^h_o and γ^h_d are county of origin and destination fixed effects, respectively, which differ by workers' home region, $D_{x,o,d}$ are dummies for buckets of distance traveled between origin and destination, and $\mathbb{I}_{R(o) \neq R(d)}$ is a dummy that is equal to one if the job switch is between East and West Germany. The set of buckets $\mathbb{X}$ contains seven 50km intervals from 50km-99km onward to 350km-399km and an eighth group for counties that are further than 399 km apart.

The regression investigates three channels that could affect worker flows. First, the dummies $D_{x,o,d}$ capture the role of distance. Second, the term involving $\mathbb{I}_{R(o) \neq R(d)}$ reflects the role of geographical barriers affecting mobility between East and West Germany. If all workers, regardless of their home region, are less likely to make a job switch that involves moving between East and West Germany, then the coefficient ρ on the dummy should be negative. Finally, the

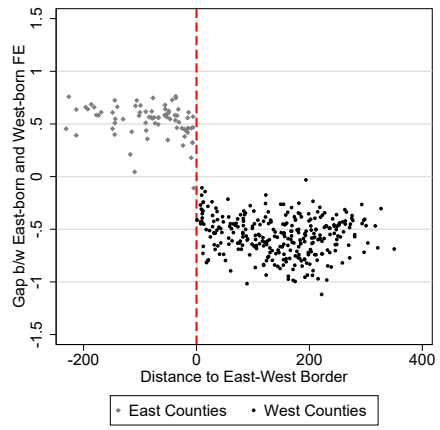
[23] We observe at least one job-to-job flow in some year for 75,937 out of the $160,801$ possible origin-destination pairs. When we include also job switches with an intermittent unemployment spell – in Supplemental Appendix O – we have 95,275.

Figure 4: Results from the Gravity Equation: Geography versus Identity

(a) Geography



(b) Identity: Destination FE



Source: LIAB. The figures plot results from specification (3). The left panel shows the point estimates for the coefficients for distance, $\hat{\phi}_x$, in black and the distance coefficients for a cross-border move, $\hat{\phi}_x + \hat{\rho}$, in gray, where each coefficient is plotted at the mid-point of the relevant distance interval and the 400+ category is plotted at 500km. All coefficients are transformed into levels by taking their exponent and then normalized into interpretable shares by dividing by their sum plus exp(0) for the omitted category of short-distance moves. Dotted lines represent the 95% confidence interval. The right panel plots the difference between the destination fixed effects for East- and West-born, $\gamma_d^{East} - \gamma_d^{West}$, as a function of the distance of each county d to the East-West former border. We normalize the fixed effect coefficients for each worker type by their mean and plot counties in the East with a negative distance.

home-region specific fixed effects δ_o^h and γ_d^h capture the fact that some counties may be more attractive to workers of home region h, for example due to preferences, comparative advantage, or possibly due to a social network that allows them to find job opportunities. For example, if γ_d^h is high for a destination then a high share of job-to-job movers of type h move into that county regardless of the location of their previous job and regardless of whether these workers have to cross the East-West border. We refer to this channel as home bias.

Results. We show the full list of estimated coefficients of regression (3) in Supplemental Appendix O and present here the key take-aways. In Figure 4a, the black line plots the distance coefficients ϕ_x, which we re-normalize into interpretable shares of switchers. As expected, workers are less likely to move to counties that are further away. The gray line plots the same results for cross border flows (the coefficients $\hat{\phi}_x + \hat{\rho}$), taking the origin and destination effects as constant. The lines are almost on top of each other. Thus, conditional on distance and fixed effects, we do not find a substantial role for geographical barriers at the former East-West border.

Figure 4b shows that there is strong home bias. For each county, we compute the difference between the destination fixed effect for East- and West-born workers, $\gamma_d^{East} - \gamma_d^{West}$. We then plot these differences against each county's distance to the East-West border, defined so that

East counties have negative distance.[24] The figure shows that East individuals have significantly higher destination fixed effects for the East, indicating that they are relatively more likely to move to counties in the East than West workers regardless of their current county. Conversely, East-born workers are less likely to move to counties in the West.

In Supplemental Appendix O, we plot the origin fixed effects and find that workers are also less likely to move out of counties in their home region. We also show that the results are similar for different sub-groups of the population, for different definitions of cross-border mobility, if we include controls for the distance of the origin county to the former border, and if we include transitions between jobs that are separated by an unemployment spell.

4 Model

We develop a model to leverage our matched employer-employee data to quantify and structurally decompose the spatial barriers that impede worker mobility, taking into account labor market frictions. Our general equilibrium framework embeds the on-the-job search model of Burdett and Mortensen (1998) into a multi-region economy inhabited by heterogeneous firms and workers that are subject to different types of spatial frictions.

The design of our model is motivated by four empirical facts. First, since the East-West wage gap has been constant and the number of workers away from their home region has been relatively balanced in recent years (see Appendix B), we perform our analysis in steady state. Second, since job movers obtain significant wage increases even within-region, a model with heterogeneous firms and labor market frictions is needed. Third, the presence of frequent and repeated moves across East and West leads us to design a framework in which multiple regions are partially integrated and individuals draw (infrequently) jobs from different regions. Finally, the salient asymmetries, in both wage gains and job flows, call for a model with home biases.

4.1 Environment

Let time be continuous and all agents discount future income at rate r. There are $\mathbb{J} = \{1, ..., J\}$ sites, which we refer to as *locations*, in an economy inhabited by a continuum of workers of types $i \in \mathbb{I}\{1, .., I\}$ with mass $\bar{D}^i$, where $\sum_{i \in \mathbb{I}} \bar{D}^i = 1$.[25] Throughout the text, we will use superscripts for worker types and subscripts for locations. Workers of type i have a preference parameter τ_j^i for being at location j, and consume both a tradable and a local good, such as housing. Their utility is $\mathcal{U}_j^i = \tau_j^i c^\eta h^{1-\eta}$, where c and h are the amounts of tradable good and

[24] As known in gravity equations, the level of the fixed effects is not identified. We normalize the fixed effects for both East-born and West-born workers relative to their average value. This normalization is without loss of generality since we are interested only in the relative fixed effects across counties.

[25] We introduce the term "locations" to differentiate it from the two regions in the empirical section. We will estimate the model below with four locations.

local good, respectively. A worker of type i produces θ_j^i units of output per time unit in location j. Hence if this worker is employed at wage rate w per efficiency unit, she earns an income of $w\theta_j^i$. Worker i's indirect utility from receiving wage rate w in location j is then $\mathcal{V}_j^i = w\theta_j^i\tau_j^i/P_j$, where $P_j = (P_c)^\eta (P_{h,j})^{1-\eta}$ is the location's price level, P_c is the price of the tradable good, and $P_{h,j}$ the price level of the local good in location j.[26] We normalize $P_c = 1$.

Workers and firms operate in a frictional and local labor market. We define by e_j^i and u_j^i the mass of employed and unemployed workers of type i in location j, respectively. Workers of type i currently in location j must spend search effort s_x to send $a_{jx}^i(s_x) = z_{jx}^i s_x$ job applications towards location x. Here, z_{jx}^i is the worker's relative search efficiency, which depends on the worker's current and destination locations (j, x) to capture that it may be easier to find job opportunities locally. Search efficiency also depends on the worker's type i, reflecting that it may be easier for workers to find open positions in their home location, for example due to reliance on social networks or referrals (as in, e.g., Galenianos (2013)). Search effort is subject to a cost, to be paid in each location x in which the worker files applications, given by $\psi(s_x) = \frac{s_x^{1+\epsilon}}{1+\epsilon}$ for employed workers. Unemployed workers face a cost $\psi_u(s_x) = \nu^{-\epsilon}\frac{s_x^{1+\epsilon}}{1+\epsilon}$, where $\nu \geq 1$ modulates a potential difference in search intensity between employed and unemployed workers along the lines of Moscarini and Postel-Vinay (2016).

On the firm side, there is a continuum of firms exogenously assigned to locations $j \in \mathbb{J}$, where M_j is the mass of firms in location j and $\sum_{j\in\mathbb{J}} M_j = 1$. Within each location, firms are distributed over labor productivity p according to density function $\frac{\gamma_j(p)}{M_j}$ with support in a location-specific closed set $[\underline{p}_j, \bar{p}_j] \subseteq \mathbb{R}^+$.[27] Each firm p in location j decides how many vacancies $v_j(p)$ to post, subject to a vacancy cost $\xi_j(v)$, and what wage rate $w_j(p)$ to offer, determining the endogenous distributions of wage offers $\{F_j\}_{j\in\mathbb{J}}$. Firms cannot discriminate between worker types, hence they must offer identical wages per efficiency unit to all their workers.

Matches in location j are created as a function of the total mass of applications filed by workers, $\bar{a}_j$, and vacancies posted by firms, $\bar{v}_j$, according to a matching function $M(\bar{a}_j, \bar{v}_j) = \bar{a}_j^\chi \bar{v}_j^{1-\chi}$. We define market tightness in location j as $\vartheta_j \equiv \frac{\bar{v}_j}{\bar{a}_j}$. Thus, the rate at which a vacancy is filled is $\vartheta_j^{-\chi}$, and the rate at which an application is accepted and becomes a job is $\vartheta_j^{1-\chi}$. Offers are randomly drawn from the endogenous wage offer distributions $\{F_j\}_{j\in\mathbb{J}}$.

Upon receiving an offer from location x, workers draw idiosyncratic preference shocks for locations x and j and decide whether to accept or decline the offer. Movers between j and x incur a utility cost κ_{jx}^i that captures any monetary and non-monetary one-time cost associated with the move across locations, similar to Caliendo, Dvorkin, and Parro (2019). Workers can always separate from a match and engage in home production with a backyard technology that has productivity per efficiency unit given by R_j. Workers separate into unemployment at

[26] We omit the constant in the indirect utility.

[27] Thus, $\gamma_j(p)$ will integrate to the mass of firms in location j, M_j. This definition will simplify notation below.

location-type-specific rate δ_j^i and receive an unemployment benefit rate equal to b_j^i per efficiency unit when unemployed.

We denote by l_j^i the measure of workers of type i employed per vacancy of a firm, and thus $\sum_{i\in\mathbb{I}} \theta_j^i l_j^i$ is the measure of efficiency units of labor used by one vacancy. Vacancies can produce any combination of the two goods according to the production functions $c = pn_c$ and $h = (pn_h)^{1-\alpha}k^\alpha$, where $0 < \alpha(1-\eta) < 1$, and n_c and n_h are the efficiency units of labor per vacancy used in the production of the two goods, which satisfy $n_c + n_h = \sum_{i\in\mathbb{I}} \theta_j^i l_j^i$. The term k is a factor that is in fixed supply, such as land, with aggregate supply in location j of K_j and equilibrium price ρ_j. Firms decide how to allocate labor across the production of the two goods, taking prices in the output market as given.

In our model, firms compete for all worker types in one unified labor market. That seems an adequate description of the German labor market since we will define worker types based on their home region below, and firms cannot explicitly hire only West Germans, for example. Previous work with heterogeneous types (e.g. Moser and Engbom (2018)) assumes that the labor market is segmented by type. In our framework, each firm posts a single wage rate $w_j(p)$, which determines the composition of worker types it attracts.

We next describe the equilibrium in the goods market, which pins down local price levels. We then turn to the workers' and firms' optimization problems and the labor market equilibrium.

Goods Market. Consider a firm that has hired $n_j(w) \equiv \sum_{i\in\mathbb{I}} \theta_j^i l_j^i(w)$ efficiency units of labor per vacancy by posting wage w. The firm's remaining problem is

$$\hat{\pi}_j(w) = \max_{n_h, n_c, k} pn_c + P_{h,j}\left(pn_h\right)^{1-\alpha} k^\alpha - \rho_j k \tag{4}$$

subject to $n_c + n_h = n_j(w)$. Standard optimization and market clearing conditions imply that in equilibrium the relative price between any two locations j and x satisfies

$$\frac{P_j}{P_x} = \left(\frac{P_j Y_j}{P_x Y_x}\right)^{\alpha(1-\eta)} \left(\frac{K_j}{K_x}\right)^{-\alpha(1-\eta)}, \tag{5}$$

where $P_j Y_j$ is the nominal output of location j. If more labor moves to location j, increasing output Y_j relative to Y_x, then the relative local price index P_j/P_x rises, due to the presence of the fixed factor. As a result, there is local congestion as typical in spatial models (e.g. Allen and Arkolakis (2014)). Substituting in the optimal choices and equilibrium price, we can simplify $\hat{\pi}(w)$ to

$$\hat{\pi}_j(w) = pn_j(w) = p\sum_{i\in\mathbb{I}} \theta_j^i l_j^i(w). \tag{6}$$

The firm's profits thus boil down to a linear expression in the total number of workers, as in the standard Burdett-Mortensen framework. We provide details in Appendix D.1.

Workers. Workers choose search effort for each location x, file applications, and randomly and infrequently receive offers from firms. Workers accept an offer if it provides higher expected value than the current one. As is known, this class of models yields a recursive representation (e.g., Burdett and Mortensen (1998)). We next derive the expected value of a job offer and the value functions for employed and unemployed workers, respectively.

Given an offer from a firm in location x paying wage w', the acceptance decision of an employed worker of type i earning wage w in location j solves

$$\max\left\{W_j^i\left(w\right)+\varepsilon_j; W_x^i\left(w'\right)-\kappa_{jx}^i+\varepsilon_x\right\},$$

where $W_j^i(w)$ is the value of employment at wage w in location j, $W_x^i(w')$ is the value of employment in location x at wage w', and $\kappa_{jx}^i=0$ if $j=x$. The terms ε_j and ε_x are idiosyncratic shocks drawn from a type-I extreme value distribution with zero mean and standard deviation σ, as in, for example, Caliendo, Dvorkin, and Parro (2019), which capture shocks to workers' preferences for being in a specific location. These shocks simplify the model characterization and computation. We assume that workers operating the backyard technology are subject to the same shocks, which fixes a lower bound for wages in each location.

Given the properties of the type-I extreme value distribution, the probability that an employed worker of type i accepts an offer is given by

$$\mu_{jx}^{E,i}\left(w,w'\right)\equiv\frac{\exp\left(W_x^i\left(w'\right)-\kappa_{jx}^i\right)^{\frac{1}{\sigma}}}{\exp\left(W_j^i\left(w\right)\right)^{\frac{1}{\sigma}}+\exp\left(W_x^i\left(w'\right)-\kappa_{jx}^i\right)^{\frac{1}{\sigma}}}$$

and the expected value of an offer is

$$V_{jx}^{E,i}\left(w,w'\right)\equiv\sigma\log\left(\exp\left(W_j^i\left(w\right)\right)^{\frac{1}{\sigma}}+\exp\left(W_x^i\left(w'\right)-\kappa_{jx}^i\right)^{\frac{1}{\sigma}}\right).$$

Similarly, an unemployed worker of type i in location j receiving an offer w' from x solves

$$\max\left\{U_j^i+\varepsilon_j; W_x^i\left(w'\right)-\kappa_{jx}^i+\varepsilon_x\right\}.$$

The probability of an unemployed worker accepting this offer is $\mu_{jx}^{U,i}\left(b_j^i,w'\right)$, defined analogously to the acceptance probability of employed workers. The expected value of an offer is

$$V_{jx}^{U,i}\left(b_j^i,w'\right)\equiv\sigma\log\left(\exp\left(U_j^i\right)^{\frac{1}{\sigma}}+\exp\left(W_x^i\left(w'\right)-\kappa_{jx}^i\right)^{\frac{1}{\sigma}}\right).$$

The discounted expected value of employment $W_j^i(w)$ of a worker i earning wage w in location j consists of the flow value of employment, $w\theta_j^i\tau_j^i/P_j$, a continuation value for drawing new job offers from location x at rate $a_{jx}^i\left(s_x\right)\vartheta_x^{1-\chi}$, which is a function of the optimal search effort s_x,

and a continuation value for separating into unemployment at rate δ_j^i

$$\begin{aligned} rW_j^i(w) &= \frac{w\theta_j^i\tau_j^i}{P_j} + \max_{\{s_x\}_{x\in\mathbb{J}}} \sum_{x\in\mathbb{J}} \left(a_{jx}^i(s_x)\,\vartheta_x^{1-\chi}\left[\int V_{jx}^{E,i}(w,w')\,dF_x(w') - W_j^i(w)\right] - \psi(s_x)\right) \\ &+ \delta_j^i\left[U_j^i - W_j^i(w)\right]. \end{aligned} \tag{7}$$

Defining the expected value gain from location x as $\bar{V}_{jx}^{E,i}(w) \equiv \int V_{jx}^{E,i}(w,w')\,dF_x(w') - W_j^i(w)$, replacing the functional forms for $a_{jx}^i(\cdot)$ and $\psi(\cdot)$, and solving out the optimal search effort of a worker at wage w searching in location x, we get

$$rW_j^i(w) = \frac{w\theta_j^i\tau_j^i}{P_j} + \frac{\epsilon}{1+\epsilon}\sum_{x\in\mathbb{J}}\left[z_{jx}^i\vartheta_x^{1-\chi}\bar{V}_{jx}^{E,i}(w)\right]^{\frac{1+\epsilon}{\epsilon}} + \delta_j^i\left[U_j^i - W_j^i(w)\right]. \tag{8}$$

Similar steps, with $\bar{V}_{jx}^{U,i}(b) \equiv \int V_{jx}^{U,i}\left(b_j^i, w'\right)dF_x(w') - U_j^i$, yield the unemployment value:

$$rU_j^i = \frac{b_j^i\theta_j^i\tau_j^i}{P_j} + \nu\frac{\epsilon}{1+\epsilon}\sum_{x\in\mathbb{J}}\left[z_{jx}^i\vartheta^{1-\chi}\bar{V}_{jx}^{U,i}(b)\right]^{\frac{1+\epsilon}{\epsilon}}. \tag{9}$$

We denote by $s_{jx}^{E,i}(w)$ and $s_{jx}^{U,i}(b)$ the optimal search efforts of an employed worker with wage w and an unemployed worker with benefit b, respectively, that are currently in location j and searching in location x. We define by $a_{jx}^{E,i}(w)$ and $a_{jx}^{U,i}(b)$ the associated mass of applications. The total mass of applications filed for jobs in location j by workers of type i is then

$$\bar{a}_j^i \equiv \sum_{x\in\mathbb{J}}\left[\int a_{xj}^{E,i}(w)\,dE_x^i(w) + a_{xj}^{U,i}(b)\,u_x^i\right],$$

where $E_j^i(w)$ is the mass of employed workers of type i at firms in location j receiving at most w, with $E_j^i(w(\bar{p}_j)) = e_j^i$. The total number of applications by location is $\bar{a}_j \equiv \sum_{i\in\mathbb{I}}\bar{a}_j^i$.

Firms. Since the firms' production functions are linear, the firm-level problem of posting vacancies and choosing wages can be solved separately. Employers choose the wage rate that maximizes their steady state profits for each vacancy

$$\pi_j(p) = \max_w (p-w)\sum_{i\in\mathbb{I}}\theta_j^i l_j^i(w), \tag{10}$$

where $p\sum_{i\in\mathbb{I}}\theta_j^i l_j^i(w)$ are the net revenues from the goods market from (6). As in the standard Burdett-Mortensen setup, a higher wage rate allows firms to hire and retain more workers, but cuts down the profit margin, $p-w$. The complementarity between firm size and productivity implies that more productive firms offer a higher wage.

Firms choose the number of vacancies by solving

$$\varrho_j(p) = \max_v \pi_j(p)\, \vartheta_j^{-\chi} v - \xi_j(v), \tag{11}$$

where $\pi_j(p)$ are the maximized profits per vacancy from (10). The overall size of a firm p in location j is given by $l_j(w_j(p))v_j(p)$, where $w_j(p)$ is the profit-maximizing wage.

Firms' vacancy posting policy gives the total mass of offers posted in each location,

$$\bar{v}_j = \int_{\underline{p}_j}^{\bar{p}_j} v_j(p)\, \gamma_j(p)\, dp, \tag{12}$$

and the wage policy gives the endogenous distribution of offers

$$F_j(w) = \frac{1}{\bar{v}_j} \int_{\underline{p}_j}^{\hat{p}_j(w)} v_j(p)\, \gamma_j(p)\, dp, \tag{13}$$

where $\hat{p}_j(w) \equiv w_j^{-1}(w)$ is the productivity of the firm paying wage w. This inverse of the wage function exists since the wage function within a given location is strictly increasing as in the standard framework.

Labor Market Clearing. To close the model, we need to describe how the distribution of workers to firms is determined. We obtain the steady state value of $l_j^i(w)$ from its law of motion

$$\dot{l}_j^i(w) = \vartheta_j^{-\chi} \frac{\bar{a}_j^i}{\bar{a}_j} \mathcal{P}_j^i(w) - q_j^i(w)\, l_j^i(w) \qquad \text{if} \quad w \geq R_j, \tag{14}$$

and $\dot{l}_j^i(w) = 0$ if $w < R_j$. The first term is the hiring rate, which consists of the product of three endogenous terms: i) $\vartheta_j^{-\chi}$, the arrival rate of workers for vacancies posted in location j, which is a decreasing function of the local market tightness ϑ_j; ii) $\frac{\bar{a}_j^i}{\bar{a}_j}$, the share of applications going towards location j that is filed by workers of type i; and iii) $\mathcal{P}_j^i(w) \in [0,1]$, the probability that an offer w posted in location j is accepted by workers of type i. Since there is random matching within location, the acceptance probability is a weighted average of the acceptance probabilities of workers of type i that are submitting applications to location j,

$$\mathcal{P}_j^i(w) \equiv \frac{1}{\bar{a}_j^i} \sum_{x \in \mathbb{J}} \left[\int a_{xj}^{E,i}(w')\, \mu_{xj}^{E,i}(w', w)\, dE_x^i(w') + a_{xj}^{U,i}(b)\, \mu_{xj}^{U,i}(b, w)\, u_x^i \right]. \tag{15}$$

The second term in (14) is the separation rate, where

$$q_j^i(w) \equiv \delta_j^i + \sum_{x\in\mathbb{J}} \vartheta_x^{1-\chi} a_{jx}^{E,i}(w) \int \mu_{jx}^{E,i}(w,w')\, dF_x(w'), \tag{16}$$

which consists of the exogenous separation rate into unemployment plus the rate at which workers receive and accept offers from other firms – i.e. poaching within and across locations.

In steady state, the mass of workers per vacancy solves $\dot{l}_j^i(w) = 0$, and thus

$$l_j^i(w) = \frac{\mathcal{P}_j^i(w)\, \vartheta_j^{-\chi} \frac{\bar{a}_{\underline{j}}^i}{\bar{a}_j}}{q_j^i(w)} \qquad \text{if } w \geq R_j \tag{17}$$

and zero otherwise.

The mass of employed workers i in location j at firms paying at most w satisfies

$$E_j^i(w) \quad = \quad \int_{\underline{p}_j}^{\hat{p}_j(w)} l_j^i(w_j(z))\, v_j(z)\, \gamma_j(z)\, dz, \tag{18}$$

where $l_j^i(w)$ is given by (17). The mass of unemployed workers is defined via the flow equation

$$\dot{u}_j^i \quad = \quad \delta_j^i e_j^i - \varphi_j^i u_j^i,$$

where φ_j^i is the rate at which workers leave unemployment, given by

$$\varphi_j^i = \sum_{x\in\mathbb{J}} \vartheta_x^{1-\chi} a_{jx}^{U,i}(b) \int \mu_{jx}^{U,i}(b,w')\, dF_x(w').$$

In steady state, the mass of unemployed workers is then

$$u_j^i \equiv \frac{\delta_j^i}{\varphi_j^i + \delta_j^i} \bar{D}_j^i, \tag{19}$$

where $\bar{D}_j^i \equiv e_j^i + u_j^i$.

4.2 Stationary Equilibrium

As discussed, we focus on the steady state equilibrium of the economy, which we now define.

Definition 1: Stationary Labor Market Equilibrium. *A stationary equilibrium in the labor market consists of a set of wage and vacancy posting policies* $\{w_j(p), v_j(p)\}_{j\in\mathbb{J}}$*, search efforts* $\left\{s_{jx}^{E,i}(w), s_{jx}^{U,i}(b)\right\}_{j\in\mathbb{J},x\in\mathbb{J},i\in\mathbb{I}}$*, wage offer distributions* $\{F_j(w)\}_{j\in\mathbb{J}}$*, acceptance probabili-*

ties $\left\{\mu_{jx}^{E,i}(w,w'),\mu_{jx}^{U,i}(b,w')\right\}_{j\in\mathbb{J},x\in\mathbb{J},i\in\mathbb{I}}$, *labor per vacancy for each worker type* $\left\{l_j^i(w)\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$, *unemployment* $\left\{u_j^i\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$, *and market tightness* $\{\vartheta_j\}_{j\in\mathbb{J}}$ *such that*

1. *workers file applications and accept offers to maximize their expected present discounted values taking as given tightness* $\{\vartheta_j\}_{j\in\mathbb{J}}$ *and the wage offer distributions,* $\{F_j(w)\}_{j\in\mathbb{J}}$*;*

2. *firms set wages to maximize per vacancy profits, and choose vacancies to maximize overall firm profits, taking as given the function mapping wage to firm size,* $\left\{l_j^i(w)\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$*;*

3. *the arrival rates of offers and wage offer distributions are consistent with aggregate applications, vacancy posting, and wage policies, according to equations (10), (12) and (13);*

4. *firm sizes and worker distributions satisfy the stationary equations (17), (18), and (19).*

Our model does not admit an analytical solution. However, the following proposition shows that the wage policies follow a system of differential equations, which facilitates significantly the computation of the model.

Proposition 1. *The* J *location-specific equilibrium wage functions* $\{w_j(p)\}_{j\in\mathbb{J}}$ *solve a system of differential equations*

$$w_j(p) \;=\; w_j\left(\underline{p}_j\right)+\int_{\underline{p}_j}^{p}\frac{\partial w_j(z)}{\partial z}\gamma_j(z)\,dz$$

where, defining $\tilde{x}(p)\equiv x(w(p))$ *for any* x*,*

$$\frac{\partial w_j(p)}{\partial p}=\frac{(p-w_j(p))\left(\sum_{i\in\mathbb{I}}\theta_j^i\frac{\frac{\partial\tilde{\mathcal{P}}_j^i(p)}{\partial p}\tilde{q}_j^i(p)-\tilde{\mathcal{P}}_j^i(p)\frac{\partial\tilde{q}_j^i(p)}{\partial p}}{\tilde{q}_j^i(p)^2}\vartheta_j^{-\chi}\frac{\bar{a}_j^i}{\bar{a}_j}\right)}{\left(\sum_{i\in\mathbb{I}}\theta_j^i\frac{\tilde{\mathcal{P}}_j^i(p)}{\tilde{q}_j^i(p)}\vartheta_j^{-\chi}\frac{\bar{a}_j^i}{\bar{a}_j}\right)}$$

and

$$\tilde{q}_j^i(p)\equiv\delta_j^i+\sum_{x\in\mathbb{J}}\vartheta_x^{1-\chi}\tilde{a}_{jx}^{E,i}(p)\int\tilde{\mu}_{jx}^{E,i}(z,z')\,d\tilde{F}_x(z')$$

$$\tilde{\mathcal{P}}_j^i(p)\equiv\frac{1}{\bar{a}_j^i}\sum_{x\in\mathbb{J}}\left[\int\tilde{a}_{xj}^{E,i}(z')\,\tilde{\mu}_{xj}^{E,i}(z',z)\,d\tilde{E}_x^i(z')+a_{xj}^{U,i}(b)\,\tilde{\mu}_{xj}^{U,i}(b,p)\,u_x^i\right]$$

together with J *boundary conditions for* $w_j\left(\underline{p}_j\right)$ *satisfying*

$$w_j\left(\underline{p}_j\right)=\max\left\{R_j,\arg\max_{\hat{w}}\left(\underline{p}_j-\hat{w}\right)\sum_{i\in\mathbb{I}}\theta_j^i l_j^i(\hat{w})\right\}.$$

Proof. See Appendix D.2. □

Comparison to Burdett-Mortensen. Our framework is a rich generalization of the benchmark Burdett-Mortensen model (see Mortensen (2005)). In the benchmark case – as is well known – the equilibrium wage policy is as follows: the lowest productivity firm sets the minimum wage that allows it to hire workers from unemployment, i.e. $w\left(\underline{p}\right) = b$, and the wage policy is an increasing and continuous function of productivity. Workers separate either exogenously or upon receiving a job offer from any firm with a higher productivity. Lemma 1 shows that our model collapses to the standard framework under the appropriate assumptions.

Lemma 1. *If $a^i_{jx}(s_x) = 1$ and $\kappa^i_{jx} = 0$ for all i, j, and x, $\theta^i_j = 1$, $\tau^i_j = \tau_j$, $\delta^i_j = \delta$, $b^i_j \tau_j P_j^{-1} = \hat{b}$, and $R_j \tau_j P_j^{-1} = \hat{R}$ for all i and j, $\nu = 1$, $\chi = 0$, and $\sigma \to 0$, then the ODEs for the wage functions simplify to*

$$\frac{\partial \hat{w}\left(p\right)}{\partial p} = \frac{-2\left(p - \hat{w}\left(p\right)\right)\frac{\partial \tilde{q}(p)}{\partial p}}{\tilde{q}\left(p\right)}$$

where

$$\tilde{q}(p) = \delta + \bar{v}[1 - \tilde{F}(p)]$$

$$\tilde{\mathcal{P}}\left(p\right) = \tilde{E}(p) + u$$

and

$$\hat{w}(\underline{p}) = \hat{R},$$

where $\hat{w} \equiv w\tau_j P_j^{-1}$ is the real wage in terms of utility, hence accounting for local amenities and prices.

Proof. See Appendix D.3. □

Our setting generalizes the insights of the benchmark model, subject to some refinements. First, since workers receive wage offers from firms in any location, their decision to quit to another firm no longer depends only on the wage offered but instead on the overall value of the job, reflected in the acceptance probability $\tilde{\mathcal{P}}^i_j\left(p\right)$. Second, firms take into account that by changing the posted wage rate they can affect the composition of workers they attract. Within a given type i-location j pair, the firm size depends only on the ranking of firms' wage offers, just as in the benchmark model. However, across locations and worker types, the level of the wage is also relevant. While in principle this feature of the model can lead to discontinuities in the wage policy, in practice the presence of the preference shocks preserves the continuity of the wage function. Third, due to the presence of the stochastic shocks, the lowest productivity firms might be willing to offer a higher or lower wage than the value of unemployment within

their location. Our solution bounds the minimum wage by R_j since the backyard technology is subject to the same shocks as regular production.

The framework closest to ours is Meghir, Narita, and Robin (2015), which builds a two-sector Burdett-Mortensen framework. Crucially, however, their model does not consider switching costs between sectors, and assumes one unified labor market rather than multiple semi-integrated local labor markets as in our framework. As a result, their model cannot be easily adapted to study our research question.

5 Unpacking Spatial and Labor Market Frictions

Our model contains four frictions that could hinder the mobility of workers across space: i) moving costs κ^i_{jx}, as in spatial models with frictional labor mobility (e.g., Bryan and Morten (2019), Caliendo, Dvorkin, and Parro (2019)); ii) comparative advantage towards the home location, governed by θ^i_j; iii) location preference τ^i_j, as in migration models with home bias (e.g., Kennan and Walker (2011); Caliendo, Opromolla, Parro, and Sforza (2017)); and iv) differences in search efficiency across locations, governed by z^i_{jx}. In addition, the model contains three sources of general labor market frictions that prevent the most productive firms from hiring all the workers within in a given location: i) vacancy posting costs, $\xi_j(v)$; ii) costs faced by workers to file applications, $\psi(s)$; and iii) preference shocks ε that limit the allocative power of wages.

While all model parameters and frictions are jointly identified, we next provide a heuristic argument to clarify how our data can be used to identify the different spatial frictions, taking into account the presence of labor market frictions.

5.1 Overall Identification Strategy

To isolate the spatial frictions from the labor market frictions, we rely on the insight that the labor market frictions directly impact the within-location moments as in the standard Burdett-Mortensen model, while spatial frictions mostly affect cross-location moments. Specifically, labor market frictions impact the within-location job flows of workers and the within-location wage gains of job movers, and therefore the joint distributions of firm sizes and wages. As a result, they directly influence the market tightness, ϑ_j, the probability that within-location offers are accepted, $\mu^{E,i}_{jj}(\cdot,\cdot)$, the distributions of wage offers, $F_j(\cdot)$, the mass of employed workers, $E_j(\cdot)$, and the mass of within-location applications, $a^{E,i}_{jj}(\cdot)$. In our empirical implementation in Section 6, we target a rich set of within-location moments to discipline these endogenous objects, and hence the labor market frictions, following a large literature on estimating Burdett-Mortensen models (see, e.g., Bontemps, Robin, and Van den Berg (2000)). Conditional on the labor market frictions, the cross-location moments help to isolate the spatial frictions.

5.2 Identifying the Spatial Frictions

We next show how different cross-location moments discipline the size of the spatial frictions.

Comparative Advantage: θ. The model, due to wage posting, yields a log additive wage equation decomposing the wage into an individual-location effect and a firm effect

$$\log w_j^i(p) \quad = \quad \log \theta_j^i + \log w_j(p).$$

This equation is similar to the specification by Abowd, Kramarz, and Margolis (1999), which relates wages to an individual and a firm fixed effect. The main difference is that in our specification the individual fixed effect is location-specific. As we describe in Appendix E, we can include additional dummies for workers not in their home location in the AKM regression to identify the comparative advantage term.

Moving Costs and Location Preferences: τ and κ. The key empirical moments that help to pin down the moving costs and location preferences are the wage gains of cross-location job-to-job movers. The average wage gain conditional on a move for an individual of type i, employed in location j, and taking a job in location x is[28]

$$\underbrace{\mathbb{E}\left[\log(w_x^i\theta_x^i) - \log(w_j^i\theta_j^i)\right]}_{\text{Average Observed Wage Gain}} = \underbrace{\log\left(\theta_x^i\right) - \log\left(\theta_j^i\right)}_{\text{Comparative Advantage}} + \tag{20}$$

$$\int\left(\int \underbrace{(\log w' - \log w)}_{\text{Wage Gain}} \underbrace{\frac{\mu_{jx}^{E,i}(w,w')}{\bar{\mu}_{jx}^{E,i}(w)}}_{\text{Rel. Prob. Accept}} \underbrace{\mathrm{d}F_x(w')}_{\text{Offers CDF}}\right) \underbrace{\frac{a_{jx}^{E,i}(w)}{\bar{a}_{jx}^{E,i}}\mathrm{d}E_j^i(w)}_{\text{Weighted Employment CDF}},$$

where $\bar{a}_{jx}^{E,i} \equiv \int a_{jx}^{E,i}(w)\,\mathrm{d}E_j^i(w)$ and $\bar{\mu}_{jx}^{i,E}(w) \equiv \int \mu_{jx}^{i,E}(w,w')\,\mathrm{d}F_x(w')$. Thus, the average wage gain depends on: i) whether the individual has comparative advantage towards location x relative to j; ii) which offers she is willing to accept, as given by $\mu_{jx}^i(w,w')$; iii) the offer distribution in the destination location x, $F_x(w')$; and iv) the employment distribution in the origin location, weighted by the relative share of applications.

Given offer distributions $F_x(\cdot)$, employment distributions $E_j^i(w)$, and the share of applications coming from each firm $\frac{a_{jx}^{E,i}(w)}{\bar{a}_{jx}^{E,i}}$, which are all mostly shaped by labor market frictions, as well as an estimate of skills θ, the equation directly relates the moving costs κ and local preferences τ to the relative wage gains of cross-location movers. Consider the limiting case when $\sigma \to 0$ – i.e. there are no preference shocks. In that case, workers accept an offer if and only if

[28] The flow utility of an individual i employed at a firm that pays wage w per efficiency unit in location j is given by $\frac{1}{P_j}\tau_j^i\theta_j^i w$. However, the observed nominal wage is simply $\theta_j^i w$, since τ_j^i does not enter into the wage.

$W_x^i(w') - \kappa_{jx}^i \geq W_j^i(w)$. Since the value functions are increasing, the cutoff wage level $\hat{w}_{jx}^i(w)$ at which an individual of type i employed in location j would accept an offer from location x is an increasing function of w. An increase in κ_{jx}^i, or a decrease in τ_x^i, would raise this cutoff wage for any level of w, since workers need to be compensated for the moving cost or the lower utility received from income earned in location x. As the worker accepts only relatively better offers, the expected wage gain of a move increases in κ_{jx}^i and decreases in τ_x^i.

The identification argument relies on a key assumption of our model: individuals draw random offers from the same offer distribution $F_x(\cdot)$, irrespective of their current location. In other words, firms cannot discriminate and post different wages for workers that are in different locations, nor can workers direct their search effort to specific firms within the location. As a result, given $F_x(\cdot)$, which is mostly disciplined by within-location moments, the cross-location moments allow us to pin down spatial frictions.

Without further restrictions, we cannot separate the moving costs from the location preference parameters. In our empirical implementation, we will disentangle the two by assuming that moving costs are identical for all worker types. Under this assumption, we can identify the location preferences from the differences in wage gains for individuals of different types that make the same migration move.

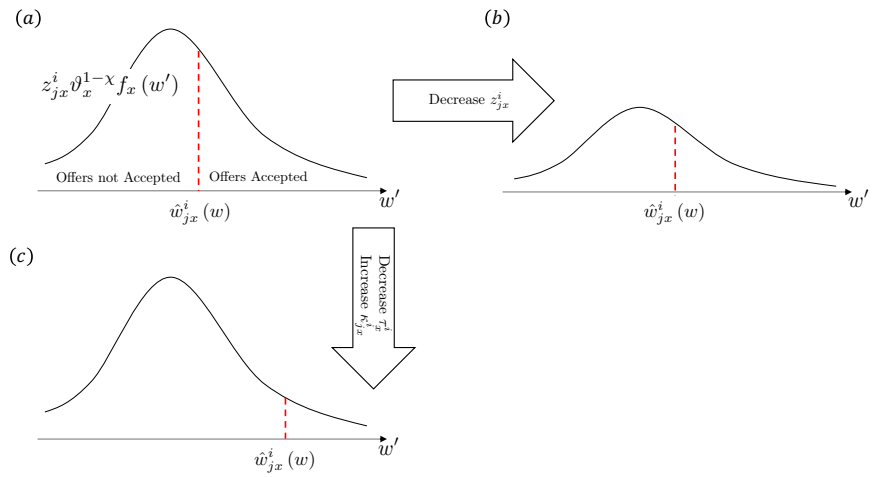
Search Efficiency: z. Given an estimate of the labor market frictions, as well as estimates of skills, moving costs, and preferences (θ, κ, τ), we can recover the relative search efficiencies from the relative job-to-job flows within and between locations. The rate at which workers of type i currently employed in location j move towards a job in location x is given by

$$\underbrace{\psi_{jx}^i}_{\text{Quit Rate}} = \left[\underbrace{\vartheta_x^{1-\chi}}_{\text{Tightness}} \underbrace{\bar{a}_{jx}^{E,i}}_{\text{Applications}}\right] \times \left[\int \left(\int \underbrace{\mu_{jx}^{E,i}(w,w')}_{\text{Prob. Accept}} \underbrace{\mathrm{d}F_x(w')}_{\text{Offer CDF}}\right) \underbrace{\frac{a_{jx}^{E,i}(w)}{\bar{a}_{jx}^{E,i}} \mathrm{d}E_j^i(w)}_{\text{Weighted Employment CDF}}\right] \tag{21}$$

The quit rate is the product of the rate at which offers arrive (first bracket) and the average probability that an offer is accepted (second bracket). Since $\bar{a}_{jx}^{E,i} = z_{jx}^i \bar{s}_x^{E,i}$, where $\bar{s}_x^{E,i} \equiv \int s_{jx}^{E,i}(w)\,\mathrm{d}E_j^i(w)$, a lower search efficiency z_{jx}^i leads to lower job-job flows from location j to x, given the acceptance probability $\mu_{jx}^{E,i}(w,w')$, which is not directly affected by z_{jx}^i itself.

The argument relies on the same assumption as before: irrespective of their current location, workers file job applications to the same labor market within each location. Thus, they draw offers from the same distributions and face the same job market tightness ϑ_x. Therefore, we can compare job flows within and across locations to infer the implied search efficiencies.

Figure 5: Identifying Spatial Frictions



Notes: Each panel shows the mass of job offers with a given wage w that is generated by a unit of search effort directed towards location x from location j, $\vartheta_x^{1-\chi} z^i_{jx} f_x(w)$. Moving from panel (a) to panel (b) illustrates the effect of an increase in spatial search frictions (i.e. a decrease in z^i_{jx}) on the the distribution of accepted offers; moving from panel (a) to panel (b) illustrates the effect of either an increase in moving costs (κ^i_{jx}) or a decrease in preferences for the destination location $\left(\tau^i_x\right)$.

Summary of the Identification. Figure 5 illustrates how the search efficiency, moving costs, and location preferences can be separetely identified. Each panel shows the mass of job offers with a given wage w that is generated by a unit of search effort directed towards location x from location j, $\vartheta_x^{1-\chi} z^i_{jx} f_x(w)$. The accepted offers, assuming again that $\sigma \to 0$, are at the right of $\hat{w}^i_{jx}(w)$, and hence the mass of job flows per unit of search effort is the integral under the wage offer density to the right of $\hat{w}^i_{jx}(w)$. Going from panel (a) to (b), a decrease in the search efficiency z^i_{jx} reduces the mass of offers received, and hence the worker flows. For comparison, panel (c) shows the effect of a decline in the worker's preference for location x, τ^i_x, which shifts the acceptance location to the right (a similar argument applies for the moving cost). This shift changes the average wage gain. Since τ and κ also affect worker flows across locations, we need both flows and wage gains to separate the effect of the search efficiency from location preferences and moving costs.

5.3 Discussion of Model Assumptions

Our identification argument is based on two assumptions that are at the core of the Burdett and Mortensen (1998) framework: wage posting and random search.

The wage posting protocol implies that firms cannot discriminate based on workers' type or current location. This assumption is supported by recent evidence that shows that the outside option has a very limited effect on workers' wages (Jäger, Schoefer, Young, and Zweimüller

(2020)) and that, conditional on the current firm, a worker's previous firm has almost no effect on current wages (Kline, Saggio, and Sølvsten (2019)). Nonetheless, we note that under a different wage setting method what we infer as a lower skill level of a given type i could represent some type of discrimination from firms, rather than a lower level of human capital. Similarly, larger wage gains for movers between locations could be driven by firms offering wage premia to compensate workers that have to migrate to accept a job offer. In our framework, these premia would be identified as moving costs as long as they are common across workers.

Random search within location implies that, for any given application, workers are equally likely to draw offers from each firm in the distribution. Since we do not observe offers received, this is an unverifiable assumption. It affects the interpretation of the search efficiencies z^i_{jx}. For example, lower observed flows from location j to location x could be driven not by a low search efficiency, but, for example, by workers i employed in location j being more likely to sample from the left tail of the distribution in location x. While our assumption is strong, it does not affect the overall interpretation of z^i_{jx}: whether workers receive fewer or worse offers from a particular location, they still have a hard time accessing job opportunities, hence a low search efficiency. A related assumption of our model is that only workers can direct their search effort towards locations, while firms cannot post vacancies targeted to a specific labor market. This is an identifying assumption driven by the fact that, given our data, we cannot distinguish between firms' or workers' behavior in generating matches.

6 Estimation

We use simulated method of moments to estimate our model for the German labor market. First, we discuss how we parametrize the model and present the parameters that are directly calibrated outside of the model. Next, we discuss the targeted moments and describe our estimation algorithm. Finally, we present the model fit and the estimation results.

6.1 Parametrization and Calibrated Parameters

Solving the model requires the computation of several endogenous and interrelated distributions. To keep the problem tractable, we limit the number of locations to four, two in the West and two in the East. Analogously, we choose four worker types, which are distinguished by their home location. This is the minimum number of locations and types that allows us to distinguish the role of the former East-West border from more general local identity and migration frictions. Going forward, we will continue to refer to East and West Germany as "regions", as in the empirical section. The four locations we use – Northwest ($j = NW$), Southwest ($j = SW$), Northeast ($j = NE$), and Southeast ($j = SE$) – combine federal states so that, within the

East and within the West, each has approximately the same number of workers. Appendix F provides further details.

Even with four locations and types, the model entails a very large number of parameters. We thus directly calibrate all the ones that have an empirical counterpart and, facing the usual trade-off between model flexibility and parsimony, we choose functional forms and structural restrictions, explained below, to reduce the total number of estimated parameters to 21.

Functional Forms. We set a unit interval of time to be one month.[29] Firms' log productivity is drawn from a log-normal distribution with equal variance in all locations, Σ, and mean A_j. We normalize $A_{NW} = 1$ and refer to A_j as the relative aggregate productivity in location j.

We parametrize the vacancy cost function as $\xi_j(v) = \frac{\xi_{0,j}^{-\xi_1}}{1+\xi_1} v^{1+\xi_1} \bar{\pi}_j(p)$, where $\xi_{0,j}$ and ξ_1 are parameters to be estimated, and $\bar{\pi}_j(p)$ is the average firm profit in location j. This parametrization implies that the equilibrium mass of vacancies posted by a firm with productivity p is $v_j(p) = \xi_{0,j} \left(\frac{\pi_j(p)}{\bar{\pi}_j(p)}\right)^{\frac{1}{\xi_1}}$. We assume that the curvature ξ_1 is constant across locations but allow $\xi_{0,j}$ to be specific to the overall region – i.e. we estimate $\xi_{0,W}$ and $\xi_{0,E}$.

We fix the unemployment benefits b_j^i so that $U_j^i = W_j^i\left(w_j\left(\underline{p}_j\right)\right)$. Under this assumption our model simplifies to the standard Burdett and Mortensen (1998) formulation, with $w_j\left(\underline{p}_j\right) = R_j$, once we remove preference shocks and spatial frictions.

Finally, we set the backyard technology to $R_j = \iota \underline{p}_j$, where $\iota \leq 1$ determines how profitable it is to set up a firm since R_j provides a lower bound on workers' wages.

Spatial Frictions. We interpret the moving cost as an opportunity cost of foregone wages (Sjaastad (1962)), and assume that the moving cost of a given worker type is symmetric and proportional to her average value, $\kappa_{jx}^i = \hat{\kappa}_{jx} \bar{W}^i$, where $\bar{W}^i = \frac{1}{e^i} \sum_{j \in \mathbb{J}} \int W_j^i(w)\, dE_j^i(w)$ and $e^i \equiv \sum_{j \in \mathbb{J}} e_j^i$. Otherwise, if κ_{jx}^i were a constant for all i, then the moving cost would be more binding for East-born workers since these have on average lower wages at any firm, as we show below.

We assume that $\hat{\kappa}_{jx}$ is equal to zero within each location and that it is a symmetric function of distance between locations j and x, identical for all workers,

$$\hat{\kappa}_{jx} = \begin{cases} 0 & \text{if } j = x \\ \kappa_0 e^{\kappa_1 dist_{jx}} & \text{if } j \neq x \end{cases}.$$

The symmetry across worker types is important for identification because it loads all asymmetries on the preference parameter τ_j^i.

[29] For example, we measure empirically the average probability that a worker moves into unemployment during a month, call it $Prob_u$, and then – since the model is in continuous time – we can recover the Poisson rate δ at which unemployment shocks arrive such that $Prob_u = 1 - e^{-\delta}$.

We specify worker preferences τ_j^i to be the product of three terms:

$$\tau_j^i \quad = \quad \underbrace{\tau_j}_{\text{Amenities}} \underbrace{\left(1 - \tau_l \mathbb{I}_{(i \neq j) \cap (r(i) = r(j))}\right)}_{\text{Home Location Bias}} \underbrace{\left(1 - \tau_r \mathbb{I}_{r(i) \neq r(j)}\right)}_{\text{Home Region Bias}},$$

where τ_j captures general amenities of location j, τ_l captures a worker's utility cost to live outside of her home location but inside her home region, and τ_r is the cost to live outside the home region, where $r(i)$ maps locations to regions. This specification allows individuals to value both their home location and their overall home region, i.e., East or West Germany.

We specify the search efficiency z_{jx}^i to be a function of both geography and identity:

$$z_{jx}^i = \begin{cases} (1 - z_{l,1} \mathbb{I}_{i \neq j}) & \text{if } j = x \\ \left(z_0 e^{-z_1 dist_{jx}}\right) (1 + z_{l,2} \mathbb{I}_{i=x}) \left(1 + z_r \mathbb{I}_{(r(i)=r(x)) \cap (i \neq x)}\right) & \text{if } j \neq x \end{cases}.$$

In the first expression, which governs within-location moves, the parameter $z_{l,1}$ captures that workers might be less effective in filing applications when they are away from their home location. In the second expression, which governs across-location moves, the parameters z_0 and z_1 allow workers' search efficiency to decay with distance. The parameters $z_{l,2}$ and z_r allow workers' search efficiency to be relatively higher towards their home location, if $z_{l,2} \geq 0$, and region, if $z_r \geq 0$.

To reduce the number of parameters to be estimated we make two further assumptions. First, we restrict $A_{NE} = A_{SE}$ since average wages and GDP per capita are similar in the Northeast and the Southeast, see Appendix F. Second, matching this assumption, we assume that local amenities are the same, $\tau_{NE} = \tau_{SE} = \tau_E$. In our estimation below, we show that despite these restrictions, we match well the location-specific moments of the Northeast and Southeast.

Calibrated Parameters. We calibrate eight sets of parameters, which we discuss briefly here. We provide more details on how they are computed in Appendix G.

We compute the mass of firms in each location, M_j, from the BHP data. We obtain the share of workers born in each location, $\bar{D}^i$, based on the population shares in 1991 from the the Growth Accounting of the States (Volkswirtschaftliche Gesamtrechnung der Länder, VGRdL) since the LIAB data start only in 1993 and are not nationally representative. More than 80% of firms are in West Germany, and 76% of the workers are born there. Within the East and the West, the North and the South have roughly equal sizes.

We assume that the separation rates δ_j^i depend only on the work location j and set them equal to the monthly probabilities, computed in the LIAB data, that workers separate into unemployment or permanent non-employment (i.e. either retired or dropping out of the labor

force).

To set the price levels P_j, we take the prices from the BBSR (2009) for each state and compute a population-weighted average across all the states within each of the locations.

Interpreting the fixed factor in the model as land, we set $\alpha(1-\eta)$ equal to 5%, which is the estimate of the aggregate share of land in GDP for the United States, see Valentinyi and Herrendorf (2008). We are not aware of estimates for Germany. It is worthwhile to note that $\alpha(1-\eta)$ does not affect the estimation of the model since we feed in the local price levels directly. It is only relevant for the general equilibrium counterfactuals.

We assume that the matching function has constant returns to scale - as standard in the literature, see Petrongolo and Pissarides (2001) - and puts equal weight on applications and vacancies, $\chi = 0.5$. The value of χ only affects the parameters of the vacancy costs and does not influence the other parameters in the estimation procedure, as it is not separately identified from $\xi_{0,j}$ and ξ_1.

Since individuals in our model are infinitely lived, the interest rate r accounts for both discounting and rates of retirement or death. We pick a monthly interest rate equal to 0.5%.

Finally, we estimate workers' skills θ_j^i using the augmented-AKM regression described in Appendix E. Our estimation indicates that workers do not have regional comparative advantages, and therefore we set $\theta_j^i = \theta^i$ for all j. We recover $\{\theta^i\}_{i\in\mathbb{I}}$ as the average individual fixed effects of workers with home location i, and we find that conditional on age, gender, and schooling, West-born workers earn, within the same firm, nearly 10% higher wages. The differences between locations within the East and within the West are small.

A recent literature has shown several concerns related to the estimation of second moments in AKM regressions.[30] For our application, these concerns do not apply since we focus on first moments, which are unbiased (Andrews, Gill, Schank, and Upward (2008)).

6.2 Targeted Moments and Identification

We are left with 21 parameters that we jointly estimate through simulated method of moments. We target the 305 moments shown in Table 3. The table indicates the appendices where we list the values of the moments and provide details on how each moment is constructed. The table also lists the parameters that are primarily pinned down by each set of moments, as we describe below through an heuristic identification argument.

Identification of Parameters and Choice of Moments. The spatial frictions are disciplined by the wage gains of job-to-job movers and the worker flows across firms within and across locations, as discussed. Therefore, we target the 64 wage gains and 64 rates of job flows, by type i, location of current firm j, and location of destination firm x (rows 1 and 2 of Table 3).

[30] See Andrews, Gill, Schank, and Upward (2008, 2012); Bonhomme, Lamadon, and Manresa (2019).

Since the model is in steady state, the size of the spatial frictions together with firms' vacancy costs determine labor demand and supply in each location. Therefore, we also target the distribution of employed and unemployed workers across locations and the firm component of wages in each location and for each type relative to $(i, j) = (NW, NW)$ (rows 3, 4 and 5). Overall, these moments help us to pin down the preferences $\left\{\tau_j^i\right\}$, search efficiencies $\left\{z_{jx}^i\right\}$, moving costs $\left\{\kappa_{jx}^i\right\}$, and vacancy costs $\{\xi_j\}$.

The productivity shifters $\{A_j\}$ are mainly related to the relative average wage paid by firms in each location, since a higher productivity leads firms, everything else equal, to offer higher wages. A higher productivity is also reflected in a higher relative GDP per worker, which we target as well (rows 6 and 7). The local unemployment rates (row 8) allow us to identify the relative search intensity from unemployment ν, given the separation rates that we calibrated directly.

As described, our model needs to be consistent with the joint distributions of firm wage and size, $G_j^i(w)$, in each location. Therefore, we target the share of employment in each decile of the firm size distribution (row 9) and the relationship between firm wage and size (row 10). These moments are relevant to discipline firms' vacancy costs $\{\xi_j\}$ since lower posting costs imply that more labor is concentrated at the most productive firms. The moments also help determine the variance of the firm productivity Σ since the variance of wages increases in Σ.

The variance of taste shocks σ governs how directed workers' moves are. As $\sigma \to \infty$, the idiosyncratic preference shocks dominate workers' acceptance decisions and workers become equally likely to accept offers that give a wage increase or decrease. The cost of search effort ϵ modulates the relationship between workers' search intensity and the value of employment at their current firm. When $\epsilon \to \infty$, workers search at equal intensity irrespective of their current job's value, while for any $\epsilon < \infty$ workers in low paying jobs search more intensively. To separately identify σ and ϵ, we target the relationship between workers' wage and their wage gains upon a job-to-job move (row 11), and the relationship between workers' separation rates (including job-to-job moves) and their wage (row 12). The former increases in σ, while the latter increases in ϵ.[31] We also target the standard deviation of the job-to-job wage gains by type i, location of the current firm j, and location of the destination firm x (row 13). This standard deviation is increasing in σ because a high σ makes workers more likely to accept offers with a negative wage change.

Finally, the ratio of firms' profits to labor costs (row 14) helps us to pin down the productivity of the backyard technology ι. Since workers have the possibility to leave employment and get $\iota \underline{p}_j$, a larger ι implies that workers need to be compensated more and firms' profits are lower.

[31] Both relationships are negative, hence when they increase, they become less steep.

Computing the Moments. While all moments have a clear correspondence between the model and the data, there are two conceptional issues that arise in their empirical computation.

First, we need to decide which controls to add in the empirical regressions that construct the moments. In our model, differences across firms in size and wages are purely driven by differences in productivity and by labor market frictions. We thus want to control for other sources of empirical variation that may be affecting firm size and wage but that are not in our model. For this reason, when calculating the moments for the joint distributions of firm wage and size, we control for industry dummies and for a set of demographic controls that capture the composition of the labor force along observable characteristics (but not birth place, of course, since it is a key variable in the model). Similarly, when we calculate the wage gains for individuals that make a job-to-job move, we control for age, gender, and education to avoid for example capturing the fact that young individuals might be more likely to move across firms. Further details on all the controls are in Appendix G.

Second, we need to take a stand on how to define a cross-location move for the estimation. While we do not introduce separate residence and work location choices in the model to keep it tractable, a sizable share of individuals in our data report to be working in a location different from their residence.[32] As a result, we face a trade-off. Defining cross-location movers as only those workers that change the location of their job and update their residence, similar to the cross-region definition in Section 3, could overestimate the search frictions and moving costs since some of the received offers lead workers to commute rather than migrate. On the other side, including all job-to-job moves regardless of residence could underestimate the frictions since commuters most likely do not pay the same fixed costs of relocating as migrants. Moreover, some movers report to be already living in the location of their new job, and hence are in fact reducing their commuting cost as a result of the move.

To strike a balance between these concerns, our baseline definition of a cross-location move includes all movers that change their work location and update their residence, similar to Section 3, plus all cross-location moves that take the worker further away from her current residence as long as the worker's residence remains within 200km of her job.[33] To be consistent with this definition we target the distribution of labor in rows 3, 4, and 5 of Table 3 using workers' residence location, as it more closely reflects the way in which we define a cross-location move.

In Supplemental Appendix P, we re-estimate our model with two alternative definitions of cross-location moves: first, the broadest possible definition by defining all job changes across locations as cross-location moves, regardless of residence. In this alternative, we target the distribution of labor using individuals' work location rather than their residence. Second, we re-

[32] About 7% of workers work in a location different from their residence.

[33] As mentioned in Section 2, the living location is self-reported and subject to misreporting. We therefore exclude individuals that report to be living far away from their job as it is likely that these observations are misreported.

estimate the model with the narrow definition of only workers who update their residence when they move. Appendix F provides statistics on the number of movers and their distance. The estimation results are broadly consistent across the three alternatives, with the main difference being that, as expected, the estimated search frictions and moving costs in our benchmark estimation lie between the two alternatives.

Estimation. Proposition 1 allows us to solve the model in just a few seconds despite its high dimensionality. Appendix H provides further details on the solution algorithm. Our estimation algorithm is otherwise standard: we solve for a vector of parameters ϕ satisfying

$$\phi^* = \arg\min_{\phi \in \mathbb{F}} \sum_x \left[\omega_x \left(T \left(m_x \left(\phi \right), \hat{m}_x \right) \right)^2 \right]$$

where $T(\cdot)$ is the log difference between the model, $\{m_x(\phi)\}$, and data, $\{\hat{m}_x\}$, moments, unless the moments are in logs. In this latter case, $T(\cdot)$ is the difference expressed as a percentage of the empirical moment. We also pick a weighting vector, $\{\omega_x\}$, so that each row of moments in Table 3 receives the same weight.[34] Otherwise, rows that are by origin-destination-type would receive higher weight than moments that are by region only. For example, we have 64 moments for standard deviation of wage gains, but we only have three for the GDP differences across regions. Our procedure assigns to the GDP differences $\frac{64}{3}$ times more weight.

6.3 Model Fit

The model provides a good fit to the data along several dimensions: most importantly, the model closely matches the key moments that help to identify the spatial frictions – the wage gains associated with the different types of job-to-job moves, and the frequency of job flows within and across locations. The left panel of Figure 6 plots the wage gains of job-to-job movers in the data against those in the model (from row 1 of Table 3).[35] Each dot is for one of the 64 different types of moves by origin-destination-home location, which we color code by direction and type of worker. The dots are close to the 45 degree line, indicating a good fit. As in the data, the model generates larger wage gains for moves towards the West (blue symbols), for within-region moves away from the home location (gray stars) and for moves away from the home region, in particular to the West (blue stars). The right panel presents a similar plot for the monthly share of movers in all employed workers (from row 2). As in the data, in our model individuals are more likely to move within-location (gray circles) and to move back to their home location and region (diamonds) than away from home (stars).

[34] Details are in Appendix H. Crucially, Figure A5 shows that all parameters seem to be properly estimated, at least based on the likelihood being locally single-peaked.

[35] For brevity, we present the model fit in figures in the main draft. In Supplemental Appendix Q, we list all the targeted and estimated moments explicitly.

Table 3: Targeted Moments

	Moments	N	Source	Model Fit	Key Parameters
(1)	Wage gains of job-job moves, by (i, j, x)	64	Sect G.2.1	Fig 6	$\{\tau_j^i\}$; $\{\kappa_{jx}^i\}$
(2)	Frequency of job flows, by (i, j, x)	64	Sect G.2.2	Fig 6	$\{z_{jx}^i\}$; $\{\xi_j\}$
(3)	Employment shares, by (i, j)	16	Sect G.2.3	Fig A6	$\{\tau_j^i\}$; $\{z_{jx}^i\}$; $\{\xi_j\}$
(4)	Unemployment shares, by (i, j)	16	Sect G.2.4	Fig A6	$\{\tau_j^i\}$; $\{z_{jx}^i\}$; $\{\xi_j\}$
(5)	Firm component of wages, by (i, j)	15	Sect G.2.5	Fig A6	$\{\tau_j^i\}$; $\{Z_j\}$
(6)	Average firm component of wages, by j	3	Sect G.2.6	Fig A6	$\{Z_j\}$
(7)	Relative GDP per worker, by j	3	Sect G.2.7	Fig A6	$\{Z_j\}$
(8)	Unemployment rates, by j	4	Sect G.2.8	Fig A6	ν
(9)	Deciles of firm-size distributions, by j	40	Sect G.2.9	Fig A7	σ, ϵ, $\{\xi_j\}$
(10)	Slope of wage vs firm size relationship, by j	4	Sect G.2.10	Table A29 and Fig A8	Σ, $\{\xi_j\}$
(11)	Slope of J2J wage gain vs firm wage, by j	4	Sect G.2.11	Table A29 and Fig A8	σ, ϵ, Σ
(12)	Slope of separation rate vs firm wage, by j	4	Sect G.2.12	Table A29 and Fig A8	σ, ϵ
(13)	Std of job-job wage gains, by (i, j, x)	64	Sect G.2.13	Table A29 and Fig A9	σ, Σ
(14)	Profit to labor cost ratio, by j	4	Sect G.2.14	Table A29	ι

Notes: The table reports the moments used in the estimation. The column titled "N" lists the number of moments in the group. Column "Source" links to the appendix section where the moment is computed, and column "Model fit" lists the table or figure that compares the empirical moment to the model-computed moment. The last column lists the key parameters that are pinned down by each set of moments.

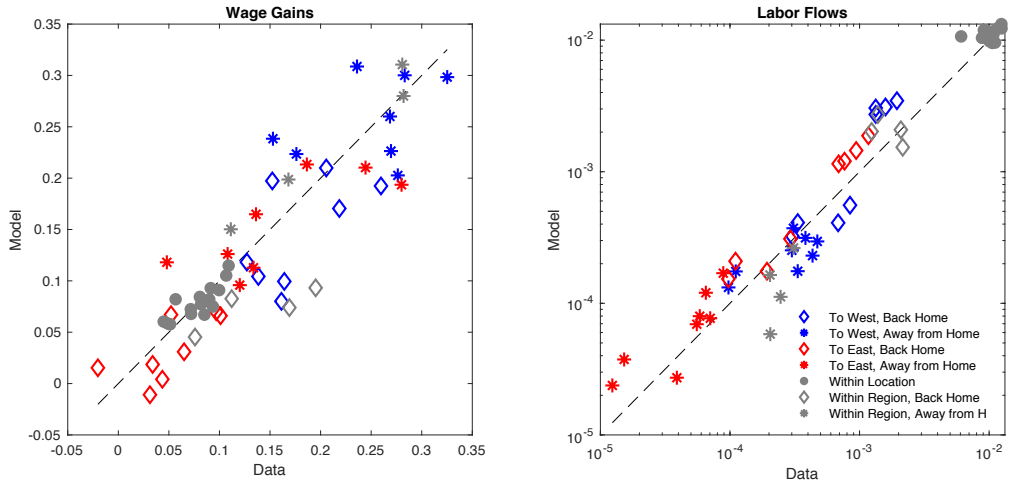
We discuss the fit of all other moments in Appendix I, and we summarize here the main takeaways.

The model matches well the steady state distributions of workers and the average GDP, wages, and unemployment rates by location, consistent with the hypothesis that the German labor market is in a steady state. Higher productivity firms offer higher wages to increase their size, leading workers to climb a job ladder and to separate with higher probability from low productivity firms. This core mechanism allows the model to do a reasonable job in matching the empirical joint distribution of firm wages, sizes, and separation rates, as well as the standard deviations of the wage gains of job movers and firms' profit shares. The model somewhat overestimates the relationship between firm wage and firm size, and generates a smaller standard deviation of wage gains of movers than the data. These results are possibly expected: in the model, wage dispersion across firms is purely generated by labor market frictions, while in the data there may be other sources of wage dispersion that our empirical controls are not capturing.[36]

Overall, the model displays a good fit. Several structural restrictions imposed by the model on the joint distributions of firm wages, employment, wage gains, and labor flows are satisfied in the data, building confidence in our estimated spatial frictions, which we discuss next.

[36]In Figure A9, we show that adding individual fixed effects in wage growth brings the empirical estimates for the standard deviations of wage growths very close to the model's ones. In Figure A8 we show the non-parametric relationships for the moments in rows 10, 11, and 12 of Table 3.

Figure 6: Wage Gains and Frequency of Job Flows



Notes: The left panel shows the average wage gains of different types of job-to-job moves in the data (x-axis) against the average wage gains in the model (y-axis). The right panel shows the frequency of each type of job-to-job move in the data (x-axis) against the frequency in the model (y-axis). Different types of moves are identified by a mix of colors and symbols, listed in the right panel. In total, there are 64 possible types of moves by origin location, destination location, and home location. The data moments are listed in Appendix G.2.1 and G.2.2. Gray symbols are moves within-region, blue symbols are moves to the West, and red symbols are moves to the East. Diamonds symbolize cross-location moves within-region back to the home location (in gray) or cross-region moves back to the home region (blue or red). Stars symbolize cross-location moves within-region away from the home location (in gray) or cross-region moves away from the home region (blue or red). Gray circles are moves within-location.

6.4 Estimation Results

Table 4 presents the spatial frictions estimated by the model. Row (1) reports the estimated one-time moving costs as a fraction of the present discounted value of income, $\hat{\kappa}_{jx}$. Since these costs vary with distance, we present a range of costs for moves between the closest two locations and moves between the farthest two locations. Our estimates indicate moving costs in the range of $3-5\%$ of the PDV of income, implying that an individual earning a yearly salary of 36,000€ for a work life of 45 years faces a moving cost of between 17,453 € and 29,704 €.[37]

Rows (2) and (3) show strong preferences for living in the home location and region. A worker employed not in her home location but still in her home region would need to be paid, in real terms, about 7.4% more than in her home location to obtain the same utility. Moving away from both home location and region requires a yearly compensation almost 10% higher.

Both the estimated migration and preference costs are an order of magnitude smaller than previous estimates in the literature (see Kennan and Walker (2011); Bryan and Morten (2019)) for two main reasons: first, our estimation identifies these costs by comparing the wage gains of cross-location movers to those of within-location job-to-job movers. Since any cross-location move is also a move between firms, we should expect migrants to experience a wage increase even in the absence of moving costs, simply due to general frictions in the labor market. Only the difference between across- and within-location wage gains reflects moving costs. Second,

[37] We discount at the model interest rate of 0.5% per month.

our framework allows us to distinguish between moving costs and search frictions. A lack of movement away from the home location can either be due to migration and preference costs or due to a lack of job opportunities resulting from search frictions.

Rows (4) and (5) report the estimated search efficiencies, relative to the within-home location level, which is normalized to 100%. Individuals that are in a location away from home and search within that location are slightly less effective than at home, filing only about 90% as many applications per unit of search effort as at home (row 4). More importantly, however, all individuals have a much lower search efficiency for cross-location searches (row 5). Since the search efficiency depends on distance, we again provide a range for searches between the two closest locations and between the two farthest locations. Row (5.i) shows that one unit of search effort used to search across locations in the non-home region translates into filing only about 1/20th as many applications as in the home location. Cross-location searches directed towards the home region, but not to the home location, are only slightly more effective (5.ii). Row (5.iii) shows that the search efficiency has a sizable home bias: one unit of search effort by workers currently away from their home location that is directed towards the home location generates 24.11% to 17.22% as many applications as searches within the home location. Hence, workers searching across locations are about four times as efficient in searching in their home location than in their non-home region, possibly due to stronger social connections (Bailey, Farrell, Kuchler, and Stroebel (2020), Burchardi and Hassan (2013)). These results show that search frictions play an important role in hampering labor market integration.

To illustrate more formally the importance of search frictions, we decompose the variance of workers' log job-to-job flows from equation (21), $\mathrm{Var}\left[\log \psi^i_{jx}\right]$, into variation due to differences in workers' search efficiency, $\mathrm{Var}\left[\log z^i_{jx}\right]$, and variation due to the remaining endogenous components.[38] These endogenous components explain only approximately 21% of the variance of log flows for all the pairs (j, i, x). The remaining empirical variance is due to the large estimated differences in the search efficiency.

We present all 21 estimated parameters in Appendix H. While the main parameters of interest are those related to the spatial frictions, it is worthwhile to mention that unemployed workers have approximately six times the search intensity of employed workers and that the estimated application cost is quite convex, making it very costly to improve labor market outcomes simply by searching more intensively.

We present the estimated parameters for the alternative definitions of cross-location moves

[38]From equation (21), we can write $\log \psi^i_{jx} = \log z^i_{jx} + \log \tilde{\psi}^i_{jx}$ where $\log z^i_{jx}$ is the search efficiency and $\log \tilde{\psi}^i_{jx} \equiv \vartheta_x^{1-\chi} \bar{s}^{E,i}_{jx} \left(\int \left(\int \mu^{E,i}_{jx}(w, w') \, \mathrm{d}F_x(w') \right) \frac{a^{E,i}_{jx}(w)}{\bar{a}^{E,i}_{jx}} \mathrm{d}E^i_j(w) \right)$ is the endogenous component given by the matching rate, the search intensity and the acceptance probability. We can then decompose the variance of job flows as $\mathrm{Var}\left[\log \psi^i_{jx}\right] = \mathrm{Var}\left[\log z^i_{jx}\right] + \mathrm{Var}\left[\log \tilde{\psi}^i_{jx}\right] + 2\mathrm{Cov}\left[\log z^i_{jx}, \log \tilde{\psi}^i_{jx}\right]$. We find that $\mathrm{Var}\left[\log \tilde{\psi}^i_{jx}\right]$ explains 21% of the variance, $\mathrm{Var}\left[\log z^i_{jx}\right]$ 38%, and the remaining 41% is due to the covariance term.

Table 4: Estimated Spatial Frictions

Moving Costs $\{\kappa\}$			
(1)	Moving cost as share of PDV of income: $\kappa_0 e^{\kappa_1 dist_{jx}}$ (b/w closest to b/w furthest locations)		3.12% to 5.31%
Preferences $\{\tau\}$			
(2)	Cost of not living in the home location but in the home region, as share of income: τ_l		7.41%
(3)	Cost of not living in the home region, as share of income: τ_r		9.88%
Relative Search Efficiency $\{z\}$			
(4)	w/i location, away from home location: $1 - z_{l,1}$		90.52%
(5)	b/w locations (closest to furthest locations)	5.i) not to home region: $z_0 e^{-z_1 dist_{jx}}$	6.10% to 4.95%
		5.ii) to home region: $\left(z_0 e^{-z_1 dist_{jx}}\right)(1 + z_r)$	7.32% to 5.23%
		5.iii) to home location: $\left(z_0 e^{-z_1 dist_{jx}}\right)\left(1 + z_{l,2}\right)$	24.11% to 17.22%

Notes: The table shows the estimated values of the spatial frictions. All parameters used to compute them, according to the formula included in each row, are in Table A28. Row 1 provides a range of the estimated moving costs, ranging from costs for moves between the closest two locations to moves between the furthest two locations. Rows 2-3 present the values of the estimated preference parameters. Search efficiencies in rows 4 and 5 are expressed as a percentage of the efficiency within the home location, z^j_{jj}, which is normalized to 1. Rows 5i-5iii show the efficiencies for searching across locations outside of the home region, in the home region but not the home location, and in the home location, respectively. The efficiencies are again reported as a range for searching between the two closest locations to searching between the two furthest locations.

in Supplementary Appendix P. When we define cross-location moves including all job switches, the model estimates moving costs roughly one quarter as large and a search efficiency towards the home location that is roughly twice as high as in the baseline to match the higher worker mobility. On the other hand, with our narrower definition using only workers that update their residence, estimated moving costs double and the search efficiency towards home approximately halves.

In Supplementary Appendix R, we further explore one potential source of home preferences using the SOEP. We show that workers' likelihood of moving back home increases sharply after the birth of a child, possibly highlighting the importance of family ties.

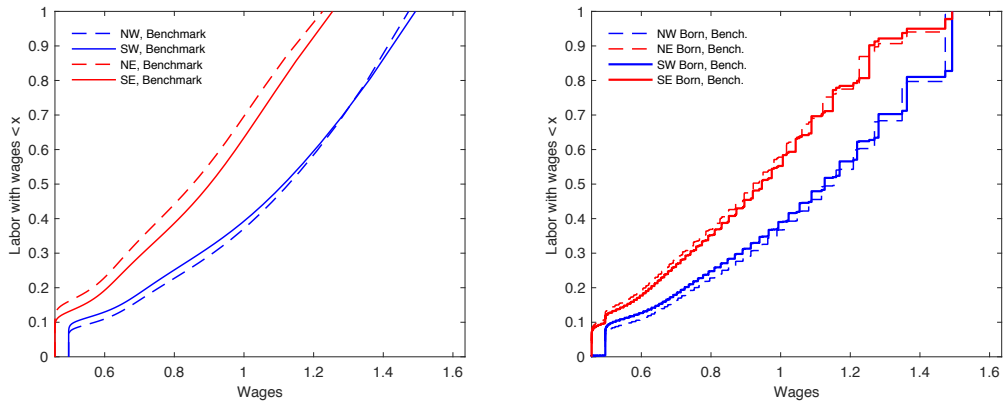
7 Spatial Frictions' Aggregate and Distributional Effects

We use the estimated model to study the aggregate costs of spatial frictions and their role in generating earnings inequality between East- and West-born individuals. Beyond the specific German context, the discussion clarifies the mechanism through which different spatial frictions impact the economy.

Our analysis is motivated by Figure 7, which shows the model-generated CDFs of real wages per efficiency unit, $w_j(p) P_j^{-1}$, by location of the firm (Fig. 7a) and by home location of the

Figure 7: Wage Distributions by Work and Birth Regions

(a) Work Location (b) Home Location



Notes: The figure shows the CDFs of real wages per efficiency unit by firms' location (left panel) and by workers' birth location (right panel). The two locations in the East are in red and the two in the West are in blue. The level of wages is normalized based on $A_{NW} = 1$, which implies a support of firm productivity in the North-West of $[0.5, 2]$.

worker (Fig. 7b).[39] These CDFs are consistent with the data as they reflect several targeted moments. Figure 7a reveals that firms in the East pay, on average, lower wages than in the West. The figure also shows that there is large wage dispersion within each region. Spatial frictions could generate misallocation of labor across locations, as workers might be trapped in the lower productivity East, and they could keep workers at lower productivity firms within each location by hampering their ability to climb a more integrated Germany-wide job ladder. The spatial frictions could also play a key role in generating a persistent effect of birth location on wages, as shown in Figure 7b.

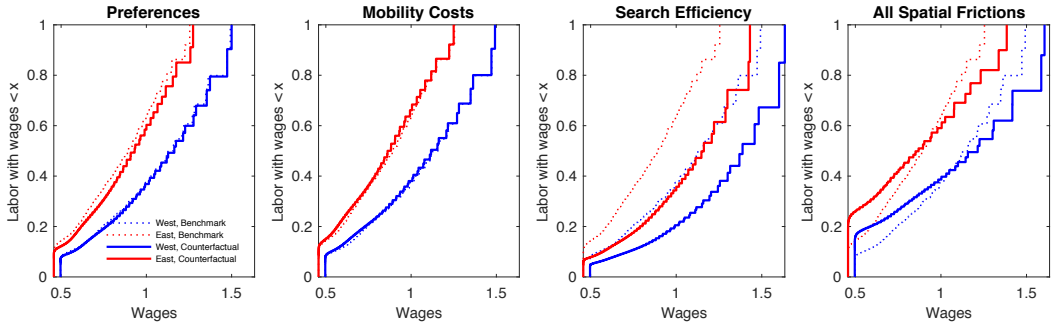
7.1 The Aggregate Effects of Removing Spatial Frictions

We recompute the equilibrium, varying the different spatial frictions while keeping all other primitive parameters unchanged, to study the role of spatial frictions in the allocation of labor within and across locations. To simplify the exposition, our discussion focuses on the comparison between East and West Germany. This choice is justified by Figure 7, which shows that the within-region differences are minor. While we report the results at the regional level, however, the removal of frictions also affects the allocation of workers within regions across locations, which contributes to the aggregate results.

Figure 8 and Table 5 summarize the results. Each panel in the figure shows the CDFs of real wages per efficiency unit in East and West Germany in the benchmark (dotted) and in the counterfactual (solid line). In the table, the first five columns report the values of various key variables for the aggregate German economy, with the baseline in column 1 normalized to one.

[39] Recall that individual of type i working in location j for firm p earns the real wage per efficiency unit multiplied by θ_j^i, i.e., $w_j(p) P_j^{-1} \theta_j^i$.

Figure 8: Effects of Spatial Frictions on Wage Distributions by Region



Notes: The figure reports the CDFs of real wages per efficiency unit by firms' region, East in red and West in blue. In each panel, we compare the benchmark economy (dotted lines) with the counterfactual economy (solid line). We consider four counterfactuals, left to right: i) no home preferences; ii) no moving costs; iii) equal search efficiency towards each region; iv) no spatial frictions.

The last five columns compute the percentage gap between West and East Germany for each of these variables. Appendix J shows all the aggregate statistics separately for East and West Germany.

Home Preferences. First, we consider an economy with no *home preferences,* $\tau_l = \tau_r = 1$. Column 2 of Table 5 and panel 1 of Figure 8 show that removing home preferences has only a small effect on the real wages per efficiency unit paid by firms (row 1). The main effect is that the share of East-born workers in the West increases while the share of West-born workers declines, since workers are no longer attached to their home regions (row 2i and 2ii).

The increased mobility of workers puts pressure on East German firms, which lose their comparative advantage in hiring East German workers, shifting the wage distribution slightly to the right (red lines in Figure 8). This shift reduces the gap in real wages per efficiency unit between the East and the West from 15% to 12%. While competitive pressure also increases in the West, the effect on the wage distribution is smaller since East firms are less able to hire West German workers due to the firms' lower average productivity.

The West-East difference in efficiency units per capita, θ_j^i, falls, since more East workers are in the West, reducing the gap in average wages paid – i.e., the average of $w_j(p) P_j^{-1} \theta_j^i$ (rows 3 and 4). Aggregate GDP per capita rises (row 5), but just marginally, implying that home bias does not entail large aggregate productivity costs. The effect on workers' average value is larger but still modest (row 6).[40]

Moving Costs. Next, we compute an economy with no *moving costs* – $\kappa_0 = 0$. Removing moving costs also has very small aggregate effects (Column 3 of Table 5). However, there is an important difference relative to home preferences: when home preferences are eliminated, East German workers stay in the West and climb the West German job ladder, putting pressure

[40]The average value is computed using the estimated value functions – U_j^i and $W_j^i(w)$ – and the distribution of labor across firms, regions, and employment status.

Table 5: Aggregate Effects of Spatial Frictions

		Germany Aggregate					Difference West - East (%)				
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Base	τ	κ	z	*All*	*Base*	τ	κ	z	*All*
(1)	Wage per efficiency unit	1	1.013	0.990	1.140	0.988	15.10%	12.30%	16.00%	11.30%	13.20%
(2)	% of Labor in West	80.40%	80.40%	79.67%	77.18%	72.33%	/	/	/	/	/
	2.i West-born	97.26%	92.36%	93.54%	80.94%	73.80%	/	/	/	/	/
	2.ii East-born	24.28%	41.08%	33.77%	65.03%	67.60%	/	/	/	/	/
(3)	Efficiency units pc	1	1.000	1.000	1.000	0.999	8.10%	5.63%	6.39%	1.54%	0.52%
(4)	Average wage paid	1	1.012	0.990	1.142	0.988	24.40%	18.40%	23.10%	11.60%	10.80%
(5)	GDP per capita	1	1.006	0.997	1.073	1.020	20.47%	16.12%	18.68%	9.65%	8.13%
(6)	Average value	1	1.018	1.001	1.129	1.259	13.67%	8.99%	12.20%	3.34%	1.16%

Notes: The table shows the aggregate effects of spatial frictions by comparing the benchmark (Base) with four hypothetical economies: i) No home preference (τ); ii) No moving costs (κ); iii) No differences in search efficiency across regions (z); iv) No spatial frictions (All). Columns 1-5 present statistics for the aggregate German economy. Statistics in rows 1, 3, 4, 5, and 6 are normalized relative to the benchmark. Columns 6-10 present, for each statistic, the difference between West and East Germany, computed for the same scenarios. Row 1 presents the average wage per efficiency unit, $w_j(p) P_j^{-1}$, averaged across all employed workers in j. Row 3 shows efficiency units per capita, the average of θ_j^i across all workers in j. Row 4 displays the average wage paid, $w_j(p) P_j^{-1} \theta_j^i$. Row 5 presents the average output per capita, $p \sum_{i \in \mathbb{I}} \theta_j^i l_j^i(w)$. Row 6 shows the average value, obtained by averaging across U_j^i and $W_j^i(w)$ using the distribution of labor across firms, regions, and employment status.

on East German firms to compete. In contrast, when moving costs alone are eliminated, East German workers are continually attracted back to the East because of their preferences. As a result, they do not climb much of the West German job ladder and remain at relatively lower productivity firms. Overall, East firms reduce their wages in response to the larger labor supply, which decreases aggregate wages per efficiency unit and increases the West-East wage gap (row 1, and panel 2 of Figure 8). Since more workers are employed at low productivity firms, aggregate GDP actually slightly declines (row 5).

Search Efficiency. In column 4 of Table 5 and panel 3 of Figure 8, we show an economy in which individuals' *search efficiency* is identical towards each region – $z_0 = 1$, $z_1 = 0$, $z_{l,1} = z_{l,2} = z_r = 0$. Eliminating differences in search efficiency has large effects on the distribution of labor both across and within regions. As workers draw more job opportunities from the whole German labor market, firms compete more fiercely for workers, which leads all firms to increase their wages. The increased opportunities allow workers to climb the job ladder more quickly, which concentrates labor at the more productive firms. Average wages per efficiency unit therefore rise by 14% and GDP per capita increases by 7%.

All Spatial Frictions. Finally, Column 5 in Table 5 shows our main result: *eliminating all spatial frictions* – $\tau_l = \tau_r = 1$, $\kappa_0 = 0$, $z_0 = 1$, $z_1 = 0$, $z_{l,1} = z_{l,2} = z_r = 0$ – causes the average wage per efficiency unit to decrease slightly and GDP per capita to rise by a mere 2%. These gains are significantly smaller than the gains from eliminating differences in search efficiency alone.

The aggregate effects are significantly weaker than the large gains found in previous work (e.g. Bryan and Morten (2019)). Three factors drive this unexpected result in our framework. First, while a large average wage gap exists between regions, the bulk of the labor misallocation is within regions across firms, reflected in the large within-region wage dispersion, which is not directly impacted by spatial frictions. Second, while eliminating all spatial frictions allows workers to draw more job opportunities from high productivity firms, it also gives firms in each region access to the pool of unemployed workers of the entire German market. As a result, low productivity firms can now more easily attract workers. Taken together, these two forces lead to a rotation of the CDF of wages which concentrates more labor both at the top and at the bottom of the wage distribution (panel 4 of Figure 8). Through general equilibrium effects, most firms end up paying a lower wage, which further decreases average wages (row 1). This mechanism is not strong enough when we shut down only one of the spatial frictions at a time, since the remaining frictions are strong enough to keep the regional labor markets separated. Third, the increase of the share of East German workers in the West creates extra competition for jobs, which, in equilibrium, favors the migration of about one quarter of West German workers to the lower productivity East, reducing output (rows 2i and 2ii).

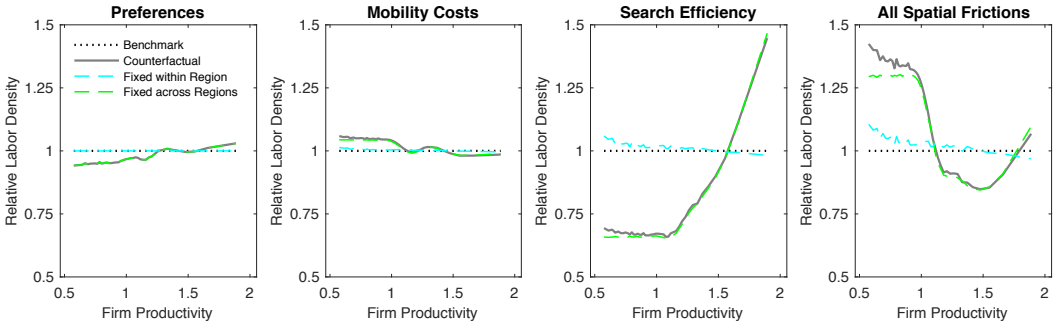
It is relevant to notice that, while the effect on GDP is larger if we only shut down search efficiency, the effect on the average value (row 6), which is the relevant welfare measure in the model, is larger, as expected, when we shut down all the frictions.

Shutting down all spatial frictions has a large effect on the West-East gap. The gap in wages paid falls from almost 25% to 11% (row 4), and the gap in GDP per capita declines from 20% to 8%. This result, however, is mostly due to the decline in spatial sorting, hence in the West-East gap in efficiency units (row 3), rather than to an increase in firm wages per efficiency units or aggregate productivity in the East.

Within Versus Across-Region Allocation of Labor. We next show that the main allocative effect of shutting down spatial frictions is to change the within-region allocation of labor to firms, rather than the share of labor in each region.

To illustrate this point, in Figure 9 we compute the density of labor at each level of firm productivity, aggregated across all regions, and take its ratio with the benchmark density in each of the four scenarios (gray solid lines). If the densities are identical, the gray solid lines should lie on top of the horizontal dotted line at 1. As previously discussed, removing preference frictions

Figure 9: Unpacking Productivity Gains from Within- and Between-Region Reallocation



Notes: The figure shows, for the four counterfactual economies, how the distribution of labor to firms changes relative to the benchmark. We compute the densities of labor at each level of firm productivity, in the whole German economy, and take their ratios (in gray solid line) relative to the benchmark. If the densities are identical, the gray solid line should lie on top of the horizontal dotted line at 1. The dashed blue line shows hypothetical densities computed keeping the same within-region allocation of labor to firms as in the benchmark and only varying the share of labor in each region. The dashed green line is the opposite: it keeps constant the share of labor in each region and only varies the within-region allocations of workers.

allows East-born workers to more easily access the West German job ladder, slightly increasing the mass of workers at more productive firms (panel 1). The economy without differences in search efficiency (panel 3) has significantly more mass at the highest productivity firms, while the economy without any spatial frictions (panel 4) has a larger mass of workers not only at the top but also at the bottom.

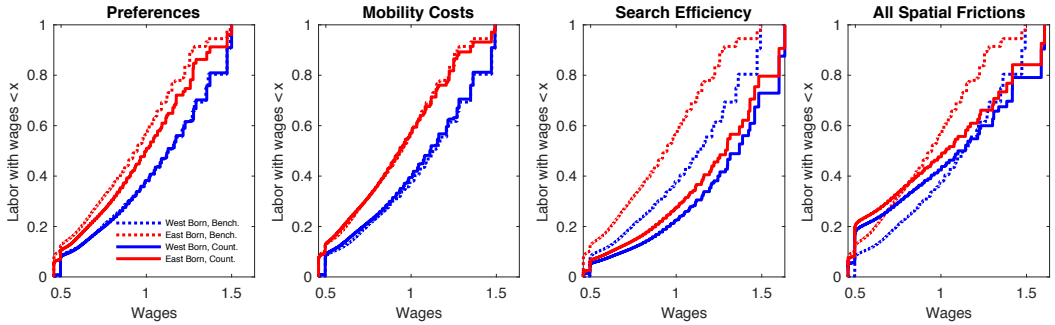
The differences in allocation could either be the result of more workers employed at more productive firms within each region, or due to a change in the shares of workers located in East and West Germany, holding the within-region allocation constant. For example, since the average productivity in the West is higher than in the East, moving more workers West would increase the mass of workers at high productivity firms.

We analyze the contribution of each of these channels separately, holding constant the other. The result is stark: keeping the within-region distribution of labor fixed and only varying the share of labor in each region barely changes the overall allocation relative to the benchmark (blue dashed line). Consequently, the effect of reducing spatial frictions on the overall distribution of labor mostly operates through within-region changes (green dashed line). Eliminating spatial frictions allows workers to climb a country-wide job ladder, significantly altering the within-region allocation of labor.

7.2 The Persistent Effects of Birth Location

Next, we show that spatial frictions have significantly larger effects on the wage inequality between East- and West-born workers than on the gap between East and West Germany. The results are presented in Figure 10 and Table 6.

Figure 10: Effects of Spatial Frictions on Wage Distributions by Birth Region



Notes: The figure reports the CDFs of real wages per efficiency units by workers' birth region, East in red and West in blue. In each panel, we compare the benchmark economy (dotted lines) with one counterfactual economy (solid line). We consider four counterfactuals, left to right: i) no home preferences; ii) no moving costs; iii) equal search efficiency towards each region; and iv) no spatial frictions.

Large Effects on Birth-Place Inequality. Column 2, row 1 of Table 6 illustrates that eliminating home preferences alone reduces the gap in the average wage per efficiency unit between East and West German workers by 40%. Removing mobility costs, instead, has a relatively small impact on wage inequality since the effect is similar for East and West-born individuals. Eliminating differences in search efficiency has the biggest independent effect, shrinking the gap in the average wage per efficiency unit by more than two thirds (column 4 of Table 6). In this scenario, all workers have identical job opportunities, irrespective of where they are born, and the only remaining difference is their willingness to accept these job offers.

When all spatial frictions are shut down, only a 2% gap in average wages per efficiency unit remains between East- and West-born workers, much smaller than the wage gap between regions. All worker types now have equal access to jobs in both regions and accept similar offers. A small gap remains due to the different efficiency units of East- and West-born workers, which affect their job search behavior due to the complementarity between firms' wages and workers' efficiency units.[41] Summing up, while spatial frictions have a relatively minor role on regional wage gaps, they are a key driver of the persistent effect of workers' birth region on their lifetime earnings.

East- and West-born Workers Climb Parallel, but Distinct, Ladders. To further investigate the mechanism through which spatial frictions generate the persistent effect of workers' home region, we use the model to simulate job histories for 100,000 workers. We assume that each worker starts unemployed in her home region and simulate her employment history for 30 years. We perform the exercise using the policy functions of the benchmark and then repeat the analysis for each one of the four alternative economies. The dotted lines in each panel of Figure 11 display the resulting paths for workers' wages per efficiency unit in the benchmark economy, and the solid lines show the counterfactual paths. To illustrate the general equilibrium effects,

[41] West-born workers search slightly more intensively and have a slightly better allocation to firms.

Table 6: Distributional Effects of Spatial Frictions

		Base	τ	κ	z	*All*
(1)	Gap in avg. wage per efficiency unit	11.02%	6.65%	9.84%	3.84%	2.11%
(2)	Gap in avg. wage paid	22.00%	17.20%	20.70%	14.11%	12.20%
(3)	% of W-born in West	97.26%	92.36%	93.54%	80.94%	73.80%
(4)	% of E-born in West	24.28%	41.08%	33.77%	65.03%	67.60%

Notes: The table shows the effects of spatial frictions on the earnings inequality between East and West-born by comparing the benchmark (Base) with four hypothetical economies: i) No home preference (τ); ii) No moving costs (κ); iii) No differences in search efficiency across regions (z); iv) No spatial frictions (All). Row 1 presents the gap in the average wage per efficiency unit, $w_j(p) P_j^{-1}$. Row 2 displays the average wage paid, $w_j(p) P_j^{-1} \theta_j^i$.

we also consider partial equilibrium counterfactuals, in which we let individuals make choices according to the economy without the corresponding spatial friction but keep the equilibrium wage, vacancy posting, and regional prices of the benchmark economy (dashed lines).

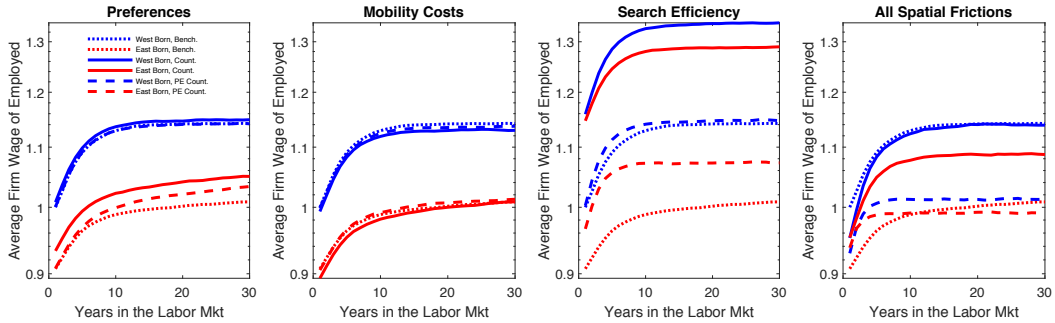
Focusing first on the benchmark economy, we notice that East- and West-born workers climb parallel job ladders starting from unemployment. This outcome is due to the fact that workers mostly remain in their home region, and the West and East regional labor markets function similarly.[42] Removing spatial frictions shifts the wage profile up, since spatial frictions similarly affect individuals at any point of their life-cycle. As noted above, the effect for moving costs is negligible due to their small size. When we remove all spatial frictions, East and West-born workers are climbing roughly the same job ladder.[43]

Large General Equilibrium Effects. We finally note that accounting for firms' general equilibrium response is important to understand the overall effects of spatial frictions on wage inequality. Comparing the dashed and the dotted lines in Figure 11, the first panel shows that in partial equilibrium, East-born gains from eliminating home preferences are only approximately half as large as in general equilibrium. While East workers' wages increase in partial equilibrium due to their willingness to accept more offers from West Germany, a large fraction of their overall wage gains is due to the equilibrium response of East firms, which increase their wages to retain workers. The equilibrium effects are even stronger when we eliminate spatial differences in search efficiency in the third panel. For West-born workers, all the gains come from the fact that firms increase their wages to retain workers since search frictions no longer shield them from competition.

[42]This result is consistent with detailed analysis in Dauth, Lee, Findeisen, and Porzio (2019). In Appendix J, we show the share of workers in each region and frequency of cross-region moves.

[43]The West-born profile is slightly steeper as they search more intensively due to their higher efficiency units, as explained earlier.

Figure 11: Simulated Employment Histories: Average Firm Wages



Notes: The figures show the average wages per efficiency unit computed from 100,000 simulated job histories, each one started with individuals being unemployed in their home region. Red is East and blue is West. Dotted lines are the benchmark economy. Solid lines represent each one of our four counterfactuals. The dashed lines are partial equilibrium counterfactuals, for which we keep constant the distribution of wage offers, prices, vacancies, and aggregate applications, but we solve for individual decision rules based on the relevant set of spatial frictions.

8 Conclusion

This paper has developed a quantitative labor market framework that encompasses frictional reallocation both across firms and across space to quantify the aggregate and distributional effects of spatial frictions. Bringing the model to matched employer-employee data from Germany, we learn three new insights that are relevant beyond our context. First, eliminating even large spatial frictions can have, as in our estimates, only modest effects on aggregate wages and productivity. Second, the aggregate effects of spatial frictions are mediated by their impact on the allocation of labor within regions across firms, which can dominate quantitatively. In fact, in our estimated economy with labor market frictions, the main effect of removing spatial frictions is to change the within-region allocation of labor, rather than generating net flows towards the high productivity region. Third, regional wage gaps and inequality of opportunities by birth region are not necessarily intertwined. Shutting down spatial frictions does not close the wage gap between East and West Germany, as labor market frictions are enough to shield low paying firms in the East from competition. However, it does substantially reduce the wage inequality between East and West-born individuals, as all workers now have equal access to jobs in all regions.

Overall, our analysis shows the importance of studying the labor allocation across firms and space in a unified framework. The model we build in this paper enables us to do so, and may prove helpful for future work on regional wage gaps and on the spatial and distributional consequences of policy interventions.

References

Abowd, J., F. Kramarz, and D. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67(2), 251–333.

Allen, T., and C. Arkolakis (2014): "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 129(3), 1085–1140.

Alvarez, J. (2018): "The agricultural wage gap: Evidence from brazilian micro-data," .

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008): "High wage workers and low wage firms: negative assortative matching or limited mobility bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(3), 673–697.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2012): "High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias," *Economics Letters*, 117(3), 824–827.

Artuç, E., S. Chaudhuri, and J. McLaren (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100(3), 1008–10045.

Bailey, M., P. Farrell, T. Kuchler, and J. Stroebel (2020): "Social connectedness in urban areas," *Journal of Urban Economics*, p. 103264.

BBSR (2009): "Regionaler Preisindex," in *Berichte*, vol. 30. BBSR, Bonn.

Bilal, A. (2019): "The Geography of Unemployment," .

Bird, R., and E. Slack (2002): "Land Taxation in Practice: Selected Case Studies. Germany.," World Bank Conference on Innovations in Local Revenue Mobilization.

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A distributional framework for matched employer employee data," *Econometrica*, 87(3), 699–739.

Bontemps, C., J.-M. Robin, and G. J. Van den Berg (2000): "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation," *International Economic Review*, 41(2), 305–358.

Bradley, J., F. Postel-Vinay, and H. Turon (2017): "Public Sector Wage Policy and Labor Market Equilibrium: A Structural Model," *Journal of the European Economic Association*, 15(6), 1214–1257.

Bryan, G., and M. Morten (2019): "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*.

Burchardi, K. B., and T. A. Hassan (2013): "The economic impact of social ties: Evidence from German reunification," *The Quarterly Journal of Economics*, 128(3), 1219–1271.

Burda, M. C. (2006): "Factor Reallocation in Eastern Germany after Reunification," *American Economic Review*, 96(2), 368–374.

Burda, M. C., and J. Hunt (2001): "From Reunification to Economic Integration: Productivity and the Labor Market in Eastern Germany," *Brookings Papers on Economic Activity*, 32(2), 1–71.

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

Caliendo, L., M. A. Dvorkin, and F. Parro (2019): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 87(3), 714–835.

Caliendo, L., L. D. Opromolla, F. Parro, and A. Sforza (2017): "Good and Factor Market Integration: A Quantitative Assessment of the EU Enlargement," NBER Working Paper No. 23695.

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 128(3), 967–1015.

——— (2015): "CHK Effects," FDZ Methodenreport 6.

Chernozhukov, V., and H. Hong (2003): "An MCMC Approach to Classical Estimation," *Journal of Econometrics*, 115(2), 293–346.

Combes, P.-P., G. Duranton, and L. Gobillon (2008): "Spatial wage disparities: Sorting matters!," *Journal of Urban Economics*, 63(2), 723–742.

Dauth, W., S. Y. T. Lee, S. Findeisen, and T. Porzio (2019): "Labor Reallocation and Convergence: Evidence from East Germany," in *2019 Meeting Papers*, no. 139. Society for Economic Dynamics.

Dingel, J. I., and F. Tintelnot (2020): "Spatial Economics for Granular Settings," Discussion paper, National Bureau of Economic Research.

Eaton, J., and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

Eberle, J., P. Jacobebbinghaus, J. Ludsteck, and J. Witter (2011): "Generation of Time-Consistent Industry Codes in the Face of Classification Changes - Simple Heuristic Based on the Establishment History Panel (BHP)," FDZ Methodenreport, 05/2011, Nuernberg.

Fajgelbaum, P. D., and C. Gaubert (2020): "Optimal spatial policies, geography, and sorting," *The Quarterly Journal of Economics*, 135(2), 959–1036.

Franzmann, G. (2007): "Bevölkerung in der ehemaligen DDR 1946-1989," *GESIS Datenarchiv, Köln histat., Studiennummer 8267*.

Fuchs-Schündeln, N., D. Krueger, and M. Sommer (2010): "Inequality Trends for Germany in the Last Two Decades: A Tale of Two Countries," *Review of Economic Dynamics*, 13(1), 103–132.

Galenianos, M. (2013): "Learning about Match Quality and the Use of Referrals," *Review of Economic Dynamics*, 16(4), 668–690.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2013): *Bayesian data analysis*. Chapman and Hall/CRC.

Gollin, D., D. Lagakos, and M. E. Waugh (2014): "The Agricultural Productivity Gap," *Quarterly Journal of Economics*, 129(2), 939–993.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2015): "What do data on millions of US workers reveal about life-cycle earnings risk?," Discussion paper, National Bureau of Economic Research.

Hethey-Maier, T., and J. F. Schmieder (2013): "Does the Use of Worker Flows Improve the Analysis of Establishment Turnover? Evidence from German Administrative Data," NBER Working Paper No. 19730.

Hicks, J. H., M. Kleemans, N. Y. Li, and E. Miguel (2017): "Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata," Discussion paper, National Bureau of Economic Research.

Hoffmann, F., and S. Shi (2016): "Burdett-Mortensen Model of On-the-Job-Search with Two Sectors," *Review of Economic Dynamics*, 19, 108–134.

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

Hunt, J. (2001): "Post-Unification Wage Growth in East Germany," *Review of Economics and Statistics*, 83(1), 190–195.

——— (2006): "Staunching Emigration from East Germany: Age and the Determinants of Migration," *Journal of the European Economic Association*, 4(5), 1014–1037.

Jäger, S., B. Schoefer, S. Young, and J. Zweimüller (2020): "Wages and the Value of Nonemployment," *The Quarterly Journal of Economics*, 135(4), 1905–1963.

Jarosch, G. (2016): "Searching for Job Security and the Consequences of Job Loss," University of Chicago.

Kennan, J., and J. R. Walker (2011): "The effect of expected income on individual migration decisions," *Econometrica*, 79(1), 211–251.

Kline, P., R. Saggio, and M. Sølvsten (2019): "It Ain't Where You're From, It's Where You're At..," Presented at the Models of Linked Employer-Employee Data Conference 2019.

Krueger, A. B., and J.-S. Pischke (1995): "A Comparative Analysis of East and West German Labor Markets: Before and After Unification," in *Differences and Changes in Wage Structures*, ed. by R. B. Freeman, and L. F. Katz, pp. 405–446. University of Chicago Press.

Lagakos, D. (2020): "Urban-rural gaps in the developing world: Does internal migration offer opportunities?," *Journal of Economic perspectives*, 34(3), 174–92.

Lagakos, D., S. Marshall, A. M. Mobarak, C. Vernot, and M. E. Waugh (2020): "Migration costs and observational returns to migration in the developing world," *Journal of Monetary Economics*.

Lagakos, D., and M. E. Waugh (2013): "Selection, agriculture, and cross-country productivity differences," *American Economic Review*, 103(2), 948–80.

Leiby, R. A. (1999): *The Unification of Germany, 1989-1990*. Greenwood Press, Westport, CT.

Lise, J., C. Meghir, and J.-M. Robin (2016): "Matching, sorting and wages," *Review of Economic Dynamics*, 19, 63–87.

Meghir, C., R. Narita, and J.-M. Robin (2015): "Wages and informality in developing countries," *American Economic Review*, 105(4), 1509–46.

Mortensen, D. T. (2005): *Wage dispersion: why are similar workers paid differently?* MIT press.

Moscarini, G., and F. Postel-Vinay (2016): "Did the Job Ladder Fail after the Great Recession?," *Journal of Labor Economics*, 34(S1), S55–S93.

Moser, C., and N. Engbom (2018): "Earnings Inequality and the Minimum Wage: Evidence from Brazil," mimeo, Columbia University.

Paetzold, J., and M. Tiefenbacher (2018): "Distributional and Revenue Effects of a Tax Shift from Labor to Property," *International Tax and Public Finance*, 25(5), 1215–1251.

Petrongolo, B., and C. A. Pissarides (2001): "Looking into the black box: A survey of the matching function," *Journal of Economic literature*, 39(2), 390–431.

Pissarides, C. A. (2000): *Equilibrium unemployment theory*. MIT press.

Redding, S. J., and E. Rossi-Hansberg (2017): "Quantitative Spatial Economics," *Annual Review of Economics*, 9, 21–58.

Schmutz, B., and M. Sidibé (2018): "Frictional Labour Mobility," *The Review of Economic Studies*, 86(4), 1779–1826.

——— (2021): "A Dynamic Theory of the Urban Network," .

Sjaastad, L. A. (1962): "The costs and returns of human migration," *Journal of political Economy*, 70(5, Part 2), 80–93.

Smolny, W. (2009): "Wage Adjustment, Competitiveness and Unemployment - East Germany after Unification," *Jahrbücher für Nationalökonomie und Statistik*, 229(2+3), 130–145.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019): "Firming up inequality," *The Quarterly journal of economics*, 134(1), 1–50.

Sorkin, I. (2018): "Ranking firms using revealed preference," *The quarterly journal of economics*, 133(3), 1331–1393.

Uhlig, H. (2006): "Regional Labor Markets, Network Externalities, and Migration: The Case of German Reunification," *American Economic Review*, 96(2), 383–387.

——— (2008): "The Slow Decline of East Germany," *Journal of Comparative Economics*, 36(4), 517–541.

Valentinyi, A., and B. Herrendorf (2008): "Measuring factor income shares at the sectoral level," *Review of Economic Dynamics*, 11(4), 820–835.

Young, A. (2013): "Inequality, the Urban-Rural Gap, and Migration," *Quarterly Journal of Economics*, 128(4), 1727–1785.

Online Appendix

A Further Details on Data and Data Construction

In this section, we provide further details on the variables we use and our data preparation. We use the Establishment History Panel (BHP) version 7514, covering the years 1975-2014. For the Linked Employer-Employee Data (LIAB), we use the longitudinal model, version 9314, covering 1993-2014. We analyze these two datasets separately, since IAB regulations do not allow us to merge them. However, as part of the LIAB data, we obtain some variables from the BHP for those establishments that are matched to a worker in the LIAB, as we describe below. This supplemental BHP data in the LIAB is a smaller subsample of the overall BHP data.

BHP Data

The BHP is a 50% sample of all establishments throughout Germany with at least one employee subject to social security as of the 30th of June of a given year. For establishments in West Germany the observation period is 1975-2014 and for establishments in East Germany it is 1992-2014. The data are reported as a panel dataset at the establishment-year level. Our version of the data includes the county location of each establishment (ao_kreis), a sensitive variable that has to be requested. As discussed in the main text, we will refer to establishments as "firms" going forward.

We create a dummy for whether a firm is in East Germany based on the firm's county, and we code the dummy as missing if the firm is in Berlin. We obtain the number of full-time employees (variable az_vz) and the number of female full-time employees (az_f_vz) for each firm-year, and we construct from this the number of male full-time employees. We use the number of employees by age group to compute each firm's number of young full-time employees (15-29 years old, az_15_19_vz + az_20_24_vz + az_25_29_vz), the number of medium-aged employees (30-49 years old, az_30_34_vz + az_35_39_vz + az_40_44_vz + az_45_49_vz), and the number of older employees (50-64 years old, az_50_54_vz + az_55_59_vz + az_60_64_vz). We obtain the number of full-time workers with low qualifications (az_gq_vz), covering individuals with a lower secondary, intermediate secondary or upper secondary school leaving certificate but no vocational qualifications. We obtain the number of full-time workers with medium qualifications (az_mq_vz), which includes workers with a lower secondary, intermediate secondary, or upper secondary school leaving certificate and a vocational qualification. Finally, we use the number of full-time workers with high qualifications (az_hq_vz), which encompasses workers who have a degree from a university of applied sciences (Fachhochschule) or a university.

We obtain the mean gross daily wage paid to full-time employees by each firm in each year. Since the social security notifications contain earnings only reported up to the upper

limit for earnings for statutory pension insurance contributions, approximately 10% of full-time employees' earnings are censored. To remedy this issue, the BHP provides a corrected mean gross daily wage for each firm (te_imp_mw), which we use for all our analyses. This variable imputes the missing wages for each worker before the mean firm wage is calculated. The imputation procedure follows Card, Heining, and Kline (2015).

We use the time-consistent 3-digit industry codes at the WZ93 level for each firm (variable w93_3_gen). These time-consistent codes were constructed by Eberle, Jacobebbinghaus, Ludsteck, and Witter (2011) and are provided to us by the IAB. The WZ93 code is based on the statistical system of economic activities in the European Community, NACE Rev.1.

We only keep our core period 2009-2014. This dataset contains 8.8 million firm-year observations. We drop firms with no full-time workers, which reduces the sample size by 3.8 million. We also drop firms located in Berlin, which removes a further 200,000 observations. We verify that all observations report county (ao_kreis) information and wage information. We then adjust the wages for cost of living differences and deflate them using county-specific price indices, described in more detail below. The final dataset contains 4,797,798 firm-year observations. Table A1 provides some summary statistics.

Table A1: Summary Statistics of the BHP Dataset

	Variable	Obs	Mean	Std. Dev.
(1)	Wage of FT workers	4,797,798	74.319	40.370
(2)	Number of FT workers	4,797,798	11.516	78.068
(3)	Share male	4,797,798	0.562	0.417
(4)	Share young	4,781,174	0.222	0.310
(5)	Share medium-aged	4,781,174	0.515	0.360
(6)	Share older	4,781,174	0.263	0.329
(7)	Share low-skilled	4,741,107	0.070	0.196
(8)	Share medium-skilled	4,741,107	0.804	0.310
(9)	Share high-skilled	4,741,107	0.125	0.264

Notes: The table presents summary statistics across all firm-year observations in our data for some key variables in 2009-2014. "Wage of FT workers" is the wage of full-time workers. Young workers are defined as those between 15-29 years old. Medium-aged workers are those between 30-49 years old. Older workers are those between 50-64 years old. Low-skilled workers are those with a lower secondary, intermediate secondary or upper secondary school leaving certificate but no vocational qualifications. Medium-skilled workers are those with a lower secondary, intermediate secondary, or upper secondary school leaving certificate and a vocational qualification. High-skilled workers are those with a degree from a university of applied sciences (Fachhochschule) or a university.

LIAB Data

The LIAB data provide matched employer-employee data that link more than 1.9 million individuals to about 400,000 firms for which these individuals work, for 1993-2014. The data contain information for the unemployment spells during which workers receive unemployment insurance benefits. Workers do not appear in the data if they are self-employed, in the public

sector, or unemployed without receiving UI benefits.

We record an individual as unemployed if her employment status (erwstat) is 1 (ALG Arbeitslosengeld, which means "Unemployment benefit"), 2 (ALHI Arbeitslosenhilfe, "Unemployment benefits"), 3 (UHG Unterhaltsgeld, "Maintenance allowance"), or 5 (PFL Beitraege zur Pflegeversicherung, "Contributions to long-term care insurance"). The remaining workers are employed. We define full-time employed workers as those that do not have a part-time flag (teilzeit), that are not in semi-retirement (Altersteilzeit), interns, working students, marginally employed, or apprentices based on their employment status (erwstat).

The LIAB data report a new employment spell each time an individual's employment status changes, for example due to a change in job, wage, or employment status. Since our data provide the exact start and end date of each spell, time aggregation is not an issue. For employed workers, one spell is recorded in every calendar year even if there is no change in employment status. For unemployed workers the spell length may exceed one year. We split such long episodes into separate records so that each spell begins and ends in the same calendar year. To deal with overlapping spells, we use the variables spell start date (begepi) and spell end date (endepi). These variables are provided by the IAB and replace partially overlapping employment spells with artificial observations with new dates so that completely parallel and completely non-overlapping periods are created. We find that about 10% of worker-start date-end date episodes are associated with multiple spells, with nearly all of these cases consisting of two spells. If we exclude part-time work (which will be our sample below), 7% of worker-start date-end date episodes are associated with multiple spells. We keep only the worker's highest-paying job in such cases, which, on average, accounts for 81% of the worker's period income (median: 86%).

We obtain an individual's daily wage or unemployment benefit (tentgelt). As in the BHP, earnings are only reported up to the upper earnings limit for statutory pension insurance contributions, and hence some wages are censored. Since no imputed earnings variable is provided by the IAB, we perform our own imputation of the censored earnings, replicating the methodology described in Card, Heining, and Kline (2015).

We obtain each worker's county of residence (wo_kreis), which is available since 1999, and for employed workers the county of their job (ao_kreis). We set each individual's home county as the earliest available county of residence (wo_kreis) or county of work (ao_kreis) recorded for the worker, from any record, including part-time or unemployed. If for a given worker the earliest county of work and county of residence are from the same spell, we set the home county to the county of residence. We generate separate dummy variables that indicate whether a worker lives, works, or has her home county in East Germany, respectively, and set these dummies to missing for Berlin. To capture the distance between counties, we merge in a matrix of distances between any county pair from Google maps, where the distance is computed from

the mid point of the counties. We also compute each county's distance to the former East-West German border.

We compute each worker's age (variable jahr - gebjahr) and construct eight age dummies (26-30 years, 31-35 years, 36-40 years, 41-45 years, 46-50 years, 51-55 years, 56-60 years, older than 60 years). Additionally, we compute a dummy for whether a worker is male (from variable frau) and a dummy for whether the worker has a college education (from ausbildung), either from a university or a university of applied sciences. The education variable is only available for employed workers. Since for employed workers this variable is less than 85% complete, we set the dummy to zero if education is missing and include in our analyses an additional dummy to capture missing cases.

We obtain firm-level information from the matched BHP data. These data include only those firms in which at least one worker in the LIAB has an employment spell. We obtain each firm's number of full-time workers (az_vz) and the firm's mean gross daily wage paid to full-time employees (te_imp_mw). As described above, the latter variable imputes the wages for workers whose earnings exceed the upper earnings limit for statutory pension insurance contributions. We also obtain the time-consistent 3-digit industry codes at the WZ93 level for each firm (variable w93_3_gen). The overall firm-year level dataset contains 2.4 million observations for the period 2009-2014. As in the BHP above, we keep only firms with at least one full-time worker, which reduces the number of observations to 2.0 million. Table A2 presents some statistics on the matched BHP data. We find that this sample contains about 40% of the firm-year observations of our BHP sample above. Firms that are matched to the LIAB pay on average about 10% higher wages and are on average about three times larger than firms in the stand-alone BHP. The skew towards larger firms is expected since larger firms are more likely to be matched to at least one worker. Due to this lack of representativeness of the matched LIAB-BHP matched sample, we rely on the BHP sample to compute the firm-level moments we use in our model estimation.

We combine the individual-level data of the LIAB with the firm-level information from the matched BHP. For our baseline analysis, we keep only the years 2009-2014. Our dataset for this period contains 15.1 million employment or unemployment spells. We drop part-time workers, which removes 5.0 million spells, 60% of which are spells by female workers. We also remove $32,032$ spells where the worker is employed abroad, and $9,666$ spells where the residence county is missing. Finally, we also drop $657,487$ observations where the worker is employed in Berlin. We verify that all remaining observations report a work county. We adjust the wages for cost of living differences and deflate them using county-specific price indices, described in more detail below. The final dataset contains $9,485,701$ observations. Table A3 provides some summary statistics.

Table A2: Summary Statistics of the Matched BHP Dataset in the LIAB

	Variable	Obs	Mean	Std. Dev.
(1)	Wage of FT workers	2,003,150	81.510	40.921
(2)	Number of FT workers	2,003,150	38.971	207.164

Notes: The table presents statistics across firm-years in the BHP data that is matched to the LIAB for 2009-2014. We only keep firm-year observations with at least one full-time worker. "Wage of FT workers" presents the mean and standard deviation of the average wage of full-time workers across firm-years.

Table A3: Summary Statistics of the LIAB Dataset

	Variable	Obs	Mean	Std. Dev.
(1)	Wage of FT employed	7,963,537	111.890	76.967
(2)	Wage of unemployed	1,254,063	27.580	12.469
(3)	Employed dummy	9,485,701	0.849	0.358
(4)	Age	9,485,701	40.172	11.538
(5)	Male dummy	9,485,701	0.696	0.460
(6)	College dummy	5,904,697	0.205	0.403
(7)	Work county East	9,485,701	0.294	0.455
(8)	Live county East	9,485,701	0.310	0.463
(9)	Home county East	9,376,568	0.321	0.467

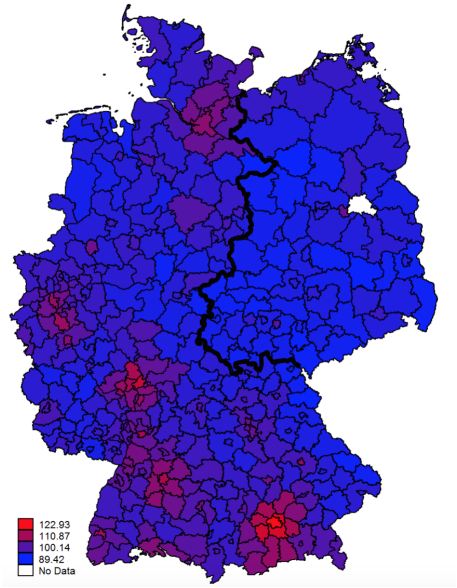
Notes: The table presents unweighted averages across all employment and unemployment spells in our core sample period for some key variables. Row 1 shows the wage of full-time employed workers. Row 2 shows the wage of unemployed workers. Row 3 presents the value of a dummy that is one for employment spells. Row 4 shows the average age, and row 5 shows the average of a dummy that is one for male workers. Row 6 shows the average of a dummy that is one for college educated workers. This variable is only available for employed individuals. Rows 7-9 present the averages for dummies that are one if the individual works, lives, and has home county in the East, respectively.

Price Deflators

We obtain data on regional prices from a study of the Federal Institute for Building, Urban Affairs and Spatial Development (BBSR (2009)). The study computed prices in 2007 for 393 micro regions covering all of Germany that correspond to cities, counties, or slightly larger unions of counties. The data cover about two thirds of the consumption basket, including housing rents, food, durables, holidays, and utilities. Of the 402 counties in the IAB data, 311 are directly represented in the BBSR data. A further 81 counties in the IAB data can be mapped to 41 regions in the BBSR data that are slightly larger than a county or combine multiple counties. The remaining 10 counties in the IAB data are represented in the BBSR data by the main town within them. Using this mapping, we obtain a price level in 2007 for each of the 402 counties in the IAB data. We then obtain for each federal state the GDP deflator

from the growth accounting of the states (Volkswirtschaftliche Gesamtrechnungen der Länder, VGRdL). We apply each state's deflator to all counties in that state to obtain a county-level price index for each year in 2009-2014. Figure A1 shows a map of the price levels in 2007.

Figure A1: Price Level, 2007



Source: BBSR. Notes: The figure plots the price level in 2007 for each county, in euros valued in Bonn, the former capital of West Germany, from the BBSR.

B Additional Statistics on Worker Mobility

In this section, we provide some additional statistics on worker mobility.

Column 1 in the top panel of Table A4 presents the number of cross-region migrants in our core sample. Migrants are defined, as in the main text, as all workers moving job-to-job between East and West Germany that change their residence in the year of the move compared to one year earlier. Job-to-job moves are defined as job switches between two firms without an intermittent unemployment spell (but possibly non-employment), as in the main text. Our sample contains about 14,000 job-to-job migrants between East and West Germany (row 1), with slightly more switches from East-to-West than from West-to-East (rows 2 and 3). Column 2 of the top panel presents the same statistics using all job-to-job switchers across regions, including those that do not change their residence. Comparing the total number of job-to-job movers in column 2 to the number of migrants in column 1, we find that about 80% of cross-region job moves are done without a reported change in residence. We will refer to such moves as "commuting". However, as discussed in the main text, social security reporting regulations do not prescribe which residence to report for individuals with multiple residences, and therefore some individuals may not list the residence closest to their job. It is therefore not possible to know with certainty whether individuals that do not report a change in residence are in fact commuting or whether their residence location is misclassified. As we discuss in Section 6.2, in our estimation we therefore consider a third, "intermediate" version of cross-region migration. This variable is defined as all migration moves plus all cross-region job switches without a change in residence where the distance between residence and work is less than 200km at both the origin and the destination, provided that the move takes the worker further away from her current residence. We impose the upper bound on the distance between work and residence to remove workers with implausibly long commutes. Moreover, we require the distance to the residence to increase to remove job changes that take the worker closer to her current residence, since such moves do not really impose a moving cost on the worker. Column 3 presents statistics for moves based on this definition. We explore the sensitivity of our structural estimates to different alternatives in a dedicated Supplemental Appendix P.

The bottom panel of Table A4 shows some selected statistics for cross-region job-to-job movers. The columns titled "Work" show moments for the distance of the cross-region job-to-job move. The first column shows that the average migration job mover changes jobs between counties that are 305km apart, with some job migrants moving jobs that are more than 500km away from each other. Once we consider all job switchers, including commuters, in column 3, the average distance between jobs drops to 278km. This still relatively large distance indicates that some workers likely have another residence closer to their job which they did not report. The intermediate definition in column 5 adds to the migrants workers that move further away from their residence but remain within 200km of their location of residence. Adding these workers

lowers the average distance between jobs for cross-region movers slightly, to 234km.

The columns titled "To Live" present analogous statistics for the distance between the worker's new job after the cross-region job switch and the worker's residence. The distances at the 5th percentile and the median highlight that most workers live close to their work location. The relatively short median and average distance for migrants in the second column suggest that workers that update their residence location when moving tend to provide their residence closest to the new job. However, even for migrants some workers in the upper tail of the distribution remain very far from their residence location. When we include all movers, the average distance to the residence increases to almost 140 km (fourth column). This result suggests that a large share of these workers have a misclassified living location, which motivates the intermediate definition, shown in the sixth column. In this case, the average, median and 95th percentile drop significantly relative to the case with all movers. In fact, by construction, this sample combines the migrants from columns 1-2 with workers that remain within 200km of their residence location, which lowers the average distance.

Table A5 presents statistics on worker mobility similar to Table 2, but considers only job-to-job migration movers as opposed to all movers that take a full-time job in their non-home region. Compared to the table in the main text, Table A5 therefore excludes job-to-job switches via commuting and moves to a new job via unemployment. Moreover, since migration can only be identified since 1999 due to the lack of residence data before then, the migration statistics are computed for this shorter period. Row 1 shows that only 0.9% of West Germans have ever migrated job-to-job to the East, and 3.9% of East Germans have migrated in the opposite direction. Row 2 presents the share of out-migrants that take up a job again in their home region at some point after their migration move. We find that about 30.1% of West Germans and 15.8% of East Germans at some point move back to a job in their home region. The number of years spent in the other region is 2-3 years for these returners (row 3). For non-returners, the average number of years passed between the migration move and their last employment spell in the data is about 5 years. To make these numbers comparable to those for all movers, Table A6 presents the table for all movers, as in the main text, using only their employment history since 1999. Comparing Table A5 and Table A6, we find that the share of workers that migrate away from their home region is significantly smaller than the share of workers that take up a job in the other region. However, conditional on migrating, migrants are considerably less likely to return home than all movers. Moreover, West German migrants that return home spend on average a longer time in the East before moving back than all West German movers. We do not find such a difference for East German migrants.

The bottom panels of Table A5 and A6 show some characteristics of stayers, movers, and movers that return home. We find that the share of college-educated migrants is significantly higher than the share of college-educated movers overall. West German migrants and movers

are significantly more likely to be college-educated than East German migrants and movers. Considering the gender of migrants, we find that the male share among migrants is comparable to the male share among non-migrants for both East and West Germans. However, East German movers overall are significantly more likely to be male than stayers.

Table A7 shows the distribution of the number of cross-border moves for workers with at least one full-time employment spell in our core sample in 2009-2014, using these workers' employment history for as many years as possible. Columns 1-2 present all cross-border moves, i.e., the number of times a worker switched full-time jobs to the other region. While the vast majority of West German workers move across regions at most three times, a small number of East German workers move up to six times. Columns 3-4 count cross-border moves since 1999 only. The distribution is similar, but shifted towards a smaller number of moves, as expected. Columns 5-6 present the number of job-to-job migration moves. These moves are significantly rarer than general moves across regions by definition, with the majority of migrants moving only once. Columns 7-8 present the distribution for moves under the intermediate definition. This distribution is similar to the distribution for migration moves, with a slight increase in the count of moves.

Table A8 looks at different cohorts of workers based on when they first took a full-time job outside of their home region, using all movers. As expected, we find that a higher share of workers returned home in the cohort that moved outside of their home region earlier. However, even in the later cohort about one third of workers that have moved away have since taken up a job in their home region. East Germans were significantly more likely to return home than West Germans in the earlier cohort, but not in the later one.

Finally, Figure A2 presents the share of workers of a given type that is employed or unemployed away from their home region in a given year (circles) and the share of workers that are living away from home (triangles). Each worker is counted at most once in a given region per year, even if she reports multiple spells in that region. The figure shows that the share of individuals working and living away from their home region has leveled off, suggesting that population shares have arrived near a steady state. Based on this figure, we perform our model analysis in steady state.

Table A4: Number of Movers Between East and West Germany

	Migration	All Cross-Region	Intermediate
Number of movers	13,853	59,603	21,199
- East-to-West	7,919	31,673	13,350
- West-to-East	5,934	27,930	7,849
Avg. moves per year	0.003	0.010	0.004

	Migration		All Cross-Region		Intermediate	
Distance	Work	To Live	Work	To Live	Work	To Live
Mean	305.054	72.498	277.848	136.381	233.558	79.956
P5	73.258	0	36.662	0	28.532	0
P50	308.840	5.661	289.260	48.387	210.635	35.203
P95	530.993	389.323	510.573	463.083	499.491	339.766

Source: LIAB. Notes: The first column of the top panel considers job-to-job migration moves (i.e., the worker changes her residence location in the same year), the second column contains all job-to-job switches between East and West, i.e., migrants plus commuters, and the third column considers migration moves plus other moves that increase the distance to the home location, as long as the distance to the residence does not exceed 200km, as described in the text. All figures are for our sample period 2009-2014. The first three rows of the top panel show the number of cross-region movers between East and West overall, East-to-West, and West-to-East, respectively. The fourth row computes for each worker the average number of moves between East and West divided by the number of years the worker is in the data, and averages across all workers. The bottom panel presents some statistics on the distance of moves. The "Work" columns show the average distance between the county of the origin job and the county of the destination job for cross-region movers, as well as some selected moments of the distribution. The "To Live" present similar statistics for the distance between the work and the residence county of the worker at the destination job for cross-region movers.

Table A5: Summary Statistics for Migrants

		(1)			(2)		
		Home: West			Home: East		
(1)	Crossed border	0.9%			3.9%		
(2)	Returned movers	30.1%			15.8%		
(3)	Mean years away (returners)	2.27			2.31		
(4)	Mean years away (non-returners)	4.67			5.16		
		Stayers	Movers	Returners	Stayers	Movers	Returners
(5)	Age at first move	–	33.5	33.2	–	30.6	29.5
(6)	Share college	0.22	0.50	0.51	0.20	0.32	0.30
(7)	Share male	0.70	0.67	0.73	0.60	0.61	0.69

Source: LIAB. Notes: The table shows statistics for workers with at least one full-time employment spell in our core sample period 2009-2014. Row 1 shows the share of these workers that has ever migrated to their non-home region, over the sample since 1999 since we do not have residence information prior to that year. Migration is defined as a job switch to the non-home region associated with a change in the county of residence in the year of the job move. Row 2 shows the share of workers that have ever taken up a job again in their home region after their first migration to the non-home region. Row 3 presents the average number of years passed between the first migration to the non-home region and the worker's job back home for returners. Row 4 shows the time passed between the last year the worker is in the data and the year of the first migration out of the home region for workers that never again take a job in their home region. Rows (5)-(7) present the average age at the migration move away from home, college share, and male share among workers that have never migrated out of their home region ("Stayers"), workers that have migrated ("Movers"), and workers that have migrated and returned to a job ("Returners").

Table A6: Summary Statistics for Job Moves since 1999

			(1)			(2)	
			Home: West			Home: East	
(1)	Crossed border		3.8%			21.9%	
(2)	Returned movers		41.9%			32.3%	
(3)	Mean years away (returners)		1.86			2.34	
(4)	Mean years away (non-returners)		5.38			6.65	
		Stayers	Movers	Returners	Stayers	Movers	Returners
(5)	Age at first move	–	35.9	35.5	–	32.3	32.2
(6)	Share college	0.22	0.34	0.32	0.19	0.19	0.19
(7)	Share male	0.70	0.75	0.80	0.57	0.73	0.78

Source: LIAB. Notes: The table shows statistics for workers with at least one full-time employment spell in our core sample period 2009-2014, and considers their employment history since 1999 only. Row 1 shows the share of these workers that have ever worked in their non-home region, over the sample since 1999. Row 2 shows the share of workers that returned to a job in their home region after their first job in the non-home region. Row 3 presents the average number of years passed between the first job in the non-home region and the worker's return to a job at home for returners. Row 4 shows the time passed between the last year the worker is in the data and the year of the first job outside of the home region for workers that never again take a job in their home region. Rows (5)-(7) present the average age at the first move away from the home region, college share, and male share among workers that have never taken a job outside of their home region ("Stayers"), workers that have moved ("Movers"), and workers that have moved away and returned to a job in the home region ("Returners").

Table A7: Distribution of Cross-Region Moves Throughout Workers' Lifetime

	Share of Workers Throughout Lifetime							
Number of cross-border moves	All Movers		All Movers 99		Migration		Intermediate	
Time period	1993-2014		1999-2014		1999-2014		1999-2014	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Home:	West	East	West	East	West	East	West	East
0	95.4%	76.1%	96.2%	78.1%	99.1%	96.1%	98.7%	93.8%
...1	2.3%	13.0%	1.9%	12.5%	0.7%	3.5%	1.1%	5.4%
...2 – 3	1.9%	8.6%	1.6%	7.6%	0.2%	0.4%	0.3%	0.8%
...4 – 6	0.4%	1.8%	0.3%	1.5%	0.0%	0.0%	0.0%	0.0%
...7+	0.1%	0.4%	0.1%	0.3%	0.0%	0.0%	0.0%	0.0%

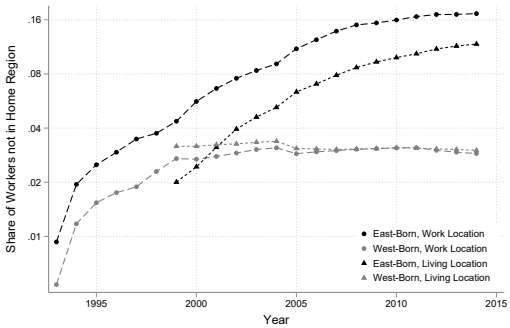
Source: LIAB. Notes: The table shows statistics for workers with at least one employment spell in our core sample period 2009-2014. For these workers, we compute the distribution of the number of cross-border moves throughout their lifetime, going back as many years as available. The first two columns present the number of times workers take up a job in the region different from the region of their last job since 1993. Columns 3-4 show the same distribution of moves but counting only moves since 1999. Columns 5-6 present the distribution of migration job-to-job moves between East and West Germany since 1999. Columns 7-8 present the number of job-to-job moves based on our intermediate definition since 1999. The intermediate definition includes migration moves plus other moves that increase the distance to the home location, as long as the distance to the residence does not exceed 200km, as described in the text.

Table A8: Mobility by Cohort

	(1)	(2)	(3)	(4)
	Movers before 1996		Movers after 2004	
	Home: West	Home: East	Home: West	Home: East
Returned movers	52.0%	71.2%	39.6%	29.6%
Mean years away (returners)	5.58	2.55	1.41	1.66
Mean years away (non-returners)	19.29	19.08	3.34	4.02

Source: LIAB. Notes: The table shows statistics for our cleaned data for 1993-2014 for workers with at least one employment spell in our core sample period 2009-2014, but distinguishes between two cohorts: workers that took the first job outside of their home region prior to 1996 (columns 1-2) and workers that first took a job outside of their home region after 2004 (columns 3-4). Row 1 presents the share of workers, among these movers, that have since moved back to a job in their home region. Row 2 presents the average number of years passed between the first job in the non-home region and the worker's return home for returners. Row 3 shows the time passed between the last year the worker is in the data and the year of the first job outside of the home region for workers that never again take a job in their home region.

Figure A2: Stock of Individuals away from Home Region



Source: LIAB. Notes: The circles plot the share of workers of a given type that is working or receiving unemployment benefits in their non-home region, for East Germans (black) and West Germans (gray). Each worker is counted once per year and region, regardless of the number of spells in that region. The triangles analogously plot the share of workers reporting their residence in their non-home region.

C Results from the Socio-Economic Panel

We use survey data from the German Socio-Economic Panel (SOEP) to examine how accurately our imputed home region in the LIAB reflects the individual's true region of birth and upbringing. The SOEP data consist of different samples drawn at different times, called "waves", and a reliable measure of birth region is available for two of them. First, the wave of individuals in the SOEP drawn in 1984 covered only West German individuals, while a wave in 1990 covered only East German individuals. For these waves the birth location is known with certainty. We will refer to individuals from these waves that are still in the labor force in 2009-2014 as the "Old SOEP Sample". Second, for individuals that entered the survey while they were still in their childhood, the data contain information on the location of individuals' preschool, primary school, and secondary school. We code the home region as the location of the individual's earliest observed non-tertiary schooling. We refer to these individuals as the "Young SOEP Sample". While the SOEP also asks some individuals about their place of residence in 1989, coverage of that variable is very low. It is only available for about 0.5% of individual-year observations in our data.

To validate our LIAB-based measure of home region, we construct an imputed home region in the same way as in the LIAB. Specifically, we keep only observations since 1993 and working age individuals under the age of 65, and drop the residence information in the SOEP before 1999 since that is not available in the LIAB. We then code an individual's imputed home region as the first residence location at which we observe the individual in employment or unemployment after 1999, or as the first job location from 1993, whichever is earlier. Table A9 compares the imputed home region to the actual home region for individuals that are employed or unemployed in 2009-2014. We find that in the "Old SOEP Sample" the imputed home region corresponds to workers' true birth region for 88% of workers born in East Germany and 99% of workers born in the West. In the "Young SOEP Sample", the imputed home region matches the region in which we observe the earliest non-tertiary schooling for an individual in 92% and 99% of cases, respectively.

As a second step, we compare the wage gap between individuals classified as East and West German under our imputation to the wage gap calculated with the true birth/schooling region. Given the limited data, we extend the period to 2004-2014, and run for employed workers the regression

$$\log(w_{it}) = \gamma \mathbb{I}_{i,East,r} + \beta X_{it} + \delta_t + \epsilon_{it},$$

where w_{it} is worker i's wage in year t and $\mathbb{I}_{i,East,r}$ is a dummy for the worker's home region, with either the true home location ($r = true$) or the imputed location ($r = imp$). The controls X_{it} contain a dummy for the worker's gender, two dummies for age (30-49 years and 50+ years), two dummies for school – i) Realschule or technical school; ii) Gymnasium or equivalent – and

two dummies for post-secondary education – indicating i) at most a vocational degree; ii) a college degree.

The first four columns of Table A10 show the results for the "Old SOEP Sample", with and without controls, and the last four columns show the results for the "New SOEP Sample". The wage gap is similar under both the true and the imputed location definitions. Thus, we find no evidence that our misclassification of some workers quantitatively alters the wage gap. Given this evidence, we also interpret workers' home region as their "birth" region.

Table A9: Imputed Home Region in the LIAB vs. Birth Region in the SOEP

	Old SOEP Sample		New SOEP Sample	
	East	West	East	West
	(1)	(2)	(3)	(4)
LIAB = SOEP	.8752	.9891	.9200	.9923
Observations	769	1,285	350	1,306

Notes: We compute in the SOEP an imputed home region in the same way as in the LIAB. Specifically, we use only SOEP data from 1993, exclude Berlin, and drop the location of residence prior to 1999. We then use the worker's location of residence at the first time he/she is observed in employment or unemployed, but not outside of the labor force, from 1999 onwards, or the worker's job location prior to 1999, to assign an imputed home region. We compare this imputed home region to the birth region based on the SOEP for individuals that are either employed or unemployed in 2009-2014. The birth region is known perfectly in the Old SOEP Sample. In the New SOEP Sample, it is equal to the region in which the individual was located at the earliest schooling for which we have data (prior to tertiary education). The figures show the proportion of observations for which the two match.

Table A10: Individual-Level Wages by Imputed Home Region versus Birth Region in the SOEP

	Old SOEP Sample				New SOEP Sample			
$\log(w_{it})$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{I}_{i,East,imp}$	$\mathbf{-.3460}^{***}$	$\mathbf{-.4042}^{***}$			$\mathbf{-.1603}^{***}$	$\mathbf{-.1632}^{***}$		
	(.0212)	(.0196)			(.0325)	(.0309)		
$\mathbb{I}_{i,East,true}$			$\mathbf{-.3377}^{***}$	$\mathbf{-.4055}^{***}$			$\mathbf{-.1326}^{***}$	$\mathbf{-.1273}^{***}$
			(.0207)	(.0192)			(.0319)	(.0303)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Age/edu/male	–	Y	–	Y	–	Y	–	Y
Observations	15,240	15,210	15.240	15,210	2,894	2,540	2,894	2,540

Notes: $*$, $**$, and $***$ indicate significance at the 90th, 95th, and 99th percent level, respectively. Standard errors are clustered at the individual level. $\mathbb{I}_{i,East,imp}$ is a dummy for the worker's home region, which is imputed using the same procedure as in the LIAB. The dummy is equal to one if the worker's home region is East Germany. $\mathbb{I}_{i,East,true}$ is a dummy for a worker's birth region (Old SOEP sample) or region of earliest non-tertiary schooling (Young SOEP sample) as read off from the SOEP survey. The sample period is 2004-2014. Male is a dummy that is equal to one if the worker is male. Age are two dummies for 30-49 years and for 50+ years. Edu are two dummies for school: i) Realschule or technical school; ii) Gymnasium or equivalent; and two dummies for post-secondary education: indicating i) at most a vocational degree; ii) a college degree.

D Proofs

D.1 Equilibrium in the Goods Market

The firm's problem is

$$\hat{\pi}_j(w) = \max_{n_h, n_c, k} pn_c + P_{h,j} \left(pn_h\right)^{1-\alpha} k^{\alpha} - \rho_j k \tag{22}$$

subject to $n_c + n_h = n_j(w)$. The first-order conditions of this problem imply

$$n_h = \frac{\rho_j}{p} \frac{1-\alpha}{\alpha} k \tag{23}$$

and assuming that both goods are supplied in equilibrium

$$P_{h,j} = \rho_j^{\alpha} \alpha^{-\alpha} \left(1-\alpha\right)^{-(1-\alpha)}. \tag{24}$$

We can plug (23) and (24) into (22) to obtain

$$\hat{\pi}_j(w) = pn_j(w) = p \sum_{i \in \mathbb{I}} \theta_j^i l_j^i(w), \tag{25}$$

where capital and labor demand for the local good has been maximized out.

The equilibrium price of the local good is determined from consumers' demand and market clearing. The aggregate demand for the local good H_j satisfies

$$P_{h,j} H_j = \left(1-\eta\right) P_j Y_j, \tag{26}$$

where, assuming that consumers own the firms and using (25), their total income is

$$P_j Y_j = \int z \left(\sum_{i \in \mathbb{I}} \theta_j^i l_j^i \left(w(z)\right) \right) v_j\left(z\right) dz + \rho_j K_j$$

and Y_j is real GDP. Aggregate supply of the local good is $H_j = (\rho_j \frac{1-\alpha}{\alpha})^{1-\alpha} K_j$, which, using the price of the local good (24), implies

$$P_{h,j} H_j \quad = \quad \frac{1}{\alpha} \rho_j K_j. \tag{27}$$

Combining demand and supply yields

$$\frac{1}{\alpha}\rho_j K_j = (1-\eta)\left\{\int p\left(\sum_{i\in\mathbb{I}}\theta_j^i l_j^i\left(w(z)\right)\right)v_j\left(z\right)dz + \rho_j K_j\right\}.$$

Given wages and the fixed K_j, this equation pins down the equilibrium price ρ_j, which in turn determines the local price P_j.

We can express the equilibrium condition in terms of ratios as follows. Starting from $P_j = (P_{h,j})^{1-\eta}$, we can substitute in with (24) and use the supply equation (27) to obtain

$$\frac{P_j}{P_x} = \left(\frac{P_{h,j}H_j}{P_{h,x}H_x}\right)^{\alpha(1-\eta)}\left(\frac{K_j}{K_x}\right)^{-\alpha(1-\eta)}.$$

Combining this expression with the demand equation (26) gives

$$\frac{P_j}{P_x} = \left(\frac{P_jY_j}{P_xY_x}\right)^{\alpha(1-\eta)}\left(\frac{K_j}{K_x}\right)^{-\alpha(1-\eta)},$$

as claimed in the main text.

D.2 Proof of Proposition 1

Firms choose the wage that maximizes profit per vacancy: they solve

$$\pi_j\left(p\right) = \max_w\left(p-w\right)\sum_{i\in\mathbb{I}}\theta_j^i l_j^i\left(w\right) \tag{28}$$

and, as shown,

$$l_j^i\left(w\right) = \frac{\mathcal{P}_j^i\left(w\right)\vartheta_j^{-\chi}\frac{\bar{a}_j^i}{\bar{a}_j}}{q_j^i\left(w\right)} \quad \text{if } w \geq R_j \tag{29}$$

which embeds the optimal behavior of workers, as described in Mortensen (2005).

The proof is constructive and it shows that firm optimality leads to the system of differential equations described. The proof relies on the insights and results of the classic Burdett-Mortensen framework, but it refines them to accommodate for multiple locations and multiple worker types.

If the function $\pi_j\left(p,w\right)$ is continuous in w for a given p, then we can take the first order condition of problem (28) and obtain

$$\frac{\left(p-w_j\left(p\right)\right)\left(\sum_{i\in\mathbb{I}}\theta_j^i\frac{\partial l_j^i\left(w_j(p)\right)}{\partial w}\right)}{\left(\sum_{i\in\mathbb{I}}\theta_j^i l_j^i\left(w_j\left(p\right)\right)\right)} = 1. \tag{30}$$

From equation (29), we find

$$\frac{\partial l_j^i(w)}{\partial w} = \frac{\frac{\partial \mathcal{P}_j^i(w)}{\partial w} q_j^i(w) - \mathcal{P}_j^i(w) \frac{\partial q_j^i(w)}{\partial w}}{q_j^i(w)^2} \vartheta_j^{-\chi} \frac{\bar{a}_j^i}{\bar{a}_j}.$$

We then define the functions in terms of p, i.e., $\tilde{x}(p) \equiv x(w(p))$ for any x, so that

$$\begin{aligned} \frac{\partial \tilde{q}_j^i(p)}{\partial p} &= \left(\frac{\partial q_j^i(w)}{\partial w}\right)\left(\frac{\partial w_j(p)}{\partial p}\right) \\ \frac{\partial \tilde{\mathcal{P}}_j^i(p)}{\partial p} &= \left(\frac{\partial \mathcal{P}_j^i(w)}{\partial w}\right)\left(\frac{\partial w_j(p)}{\partial p}\right). \end{aligned}$$

Next, we replace these equations into the above equation for $\frac{\partial l_j^i(w)}{\partial w}$ to get

$$\frac{\partial l_j^i(w)}{\partial w} = \frac{\left(\frac{\partial w_j(p)}{\partial p}\right)^{-1}}{\tilde{q}_j^i(p)^2} \left(\frac{\partial \tilde{\mathcal{P}}_j^i(p)}{\partial p} \tilde{q}_j^i(p) - \tilde{\mathcal{P}}_j^i(p) \frac{\partial \tilde{q}_j^i(p)}{\partial p}\right) \vartheta_j^{-\chi} \frac{\bar{a}_j^i}{\bar{a}_j},$$

which can itself be substituted into (30) to find a differential equation for $w_j(p)$

$$\frac{\partial w_j(p)}{\partial p} = \frac{(p - w_j(p)) \left(\sum_{i\in\mathbb{I}} \theta_j^i \frac{\frac{\partial \tilde{\mathcal{P}}_j^i(p)}{\partial p} \tilde{q}_j^i(p) - \tilde{\mathcal{P}}_j^i(p) \frac{\partial \tilde{q}_j^i(p)}{\partial p}}{\tilde{q}_j^i(p)^2} \vartheta_j^{-\chi} \frac{\bar{a}_j^i}{\bar{a}_j}\right)}{\left(\sum_{i\in\mathbb{I}} \theta_j^i \frac{\tilde{\mathcal{P}}_j^i(p)}{\tilde{q}_j^i(p)} \vartheta_j^{-\chi} \frac{\bar{a}_j^i}{\bar{a}_j}\right)}. \tag{31}$$

Since $w_j(p)$ is continuous at p by assumption, the differential equation (31), together with an appropriate boundary conditions, characterizes the optimal wage at p. Since workers can always use the backyard technology, they must be paid at least $w = R_j$. Therefore, the boundary conditions are given by

$$w_j\left(\underline{p}_j\right) = \max\left\{R_j, \arg\max_{\hat{w}} \left(\underline{p}_j - \hat{w}\right) \sum_{i\in\mathbb{I}} \theta_j^i l_j^i(\hat{w})\right\}.$$

We have thus proved that

$$w_j(p) = w_j\left(\underline{p}_j\right) + \int_{\underline{p}_j}^{p} \frac{\partial w_j(z)}{\partial z} \gamma_j(z)\, dz \tag{32}$$

as claimed. The expressions for $\tilde{q}_j^i(p)$ and $\tilde{P}_j^i(p)$ follow directly from (15) and (16).

D.3 Lemma 1

Define the real wage, adjusted for amenities, as $\hat{w} \equiv w/(\tau_j P_j)$. By assumption, $\hat{b} \equiv b_j^i/(\tau_j P_j)$ is constant across regions. Define $\hat{F}_j(\hat{w}) \equiv F_j(\hat{w}\tau_j P_j)$. Since $\theta_j^i = 1$, $\delta_j^i = \delta$, $a_{jx}^i(s_x) = 1$, and $\chi = 0$, the employed workers' value function (7) simplifies to

$$r\hat{W}(\hat{w}) = \hat{w} + \sum_{x\in\mathbb{J}} \left(\bar{v}_x \max \left[\int \hat{W}(\hat{w}')\, d\hat{F}_x(\hat{w}') - \hat{W}(\hat{w}), 0 \right] \right) + \delta \left[\hat{U} - \hat{W}(\hat{w}) \right]$$

and the unemployed worker's value function can be written as

$$r\hat{U} = \hat{b} + \sum_{x\in\mathbb{J}} \left(\bar{v}_x \max \left[\int \hat{W}(\hat{w}')\, d\hat{F}_x(\hat{w}') - \hat{U}, 0 \right] \right),$$

which no longer depend on the worker type i or the current region of the worker j. Given that $\sigma \to 0$, workers accept any offer as long as $\hat{W}(\hat{w}') \geq \hat{W}(\hat{w})$. Since $W(\hat{w})$ is increasing in $\hat{w}$, this inequality implies that workers accept any offer as long as $\hat{w}' \geq \hat{w}$.

Define $\hat{p} \equiv p/(\tau_j P_j)$. The firm's maximization problem (10) becomes

$$\hat{\pi}_j(\hat{p}) = P_j \tau_j \max_{\hat{w}} (\hat{p} - \hat{w})\, \hat{l}(\hat{w}). \tag{33}$$

From $a_{jx}^i(s_x) = 1$ and $\chi = 0$ it follows that

$$\hat{l}(\hat{w}) = \frac{\hat{\mathcal{P}}(\hat{w})}{\hat{q}(\hat{w})} \quad \text{if } \hat{w} \geq \hat{R} \tag{34}$$

where $\hat{R} \equiv R_j/(\tau_j P_j)$ is constant across regions by assumption. Since $\delta_j^i = \delta$, we have

$$\hat{q}(\hat{w}) = \delta + \sum_{x\in\mathbb{J}} \bar{v}_x \left[1 - \hat{F}_x(\hat{w}) \right] \tag{35}$$

and

$$\hat{\mathcal{P}}(\hat{w}) = \sum_{x\in\mathbb{J}} \left[\hat{E}_x(\hat{w}) + u_x \right], \tag{36}$$

where $\hat{E}_x(\hat{w}) \equiv E_x(\hat{w}\tau_j P_j)$.

The first-order condition of the wage posting problem is

$$\frac{(\hat{p} - \hat{w}) \left(\frac{\partial \hat{l}(\hat{w})}{\partial \hat{w}} \right)}{\left(\hat{l}(\hat{w}) \right)} = 1, \tag{37}$$

where

$$\frac{\partial \hat{l}\left(\hat{w}\right)}{\partial \hat{w}} = \frac{\frac{\partial \hat{\mathcal{P}}(\hat{w})}{\partial \hat{w}} \hat{q}\left(\hat{w}\right) - \frac{\partial \hat{q}(\hat{w})}{\partial \hat{w}} \hat{\mathcal{P}}\left(\hat{w}\right)}{\hat{q}\left(\hat{w}\right)^2}.$$

Plugging this latter expression into the first-order condition gives

$$\frac{\left(\hat{p} - \hat{w}\right)\left(\frac{\partial \hat{\mathcal{P}}(\hat{w})}{\partial \hat{w}} \hat{q}\left(\hat{w}\right) - \frac{\partial \hat{q}(\hat{w})}{\partial \hat{w}} \hat{\mathcal{P}}\left(\hat{w}\right)\right)}{\hat{\mathcal{P}}\left(\hat{w}\right) \hat{q}\left(\hat{w}\right)} = 1. \tag{38}$$

We next define the productivity distribution $\tilde{\Gamma}(\hat{p})$ over the $\hat{p}$ across all firms in all regions, with associated density $\tilde{\gamma}(\hat{p})$. The minimum of this productivity distribution is $\underline{\hat{p}} = \min_j \left\{\underline{\hat{p}}_j\right\}$, and the maximum $\bar{\hat{p}}$ is defined analogously. To attract any workers, the least productive firm must pay at least the reservation wage

$$\hat{w}(\underline{\hat{p}}) = \hat{R}. \tag{39}$$

From (33), firms with the same $\hat{p}$ post the same wage $\hat{w}$ and therefore attract the same number of workers. Moreover, from the usual complementarity between firm size and productivity, more productive firms post higher real wages $\hat{w}$. Define a job offer distribution across regions as a function of productivity

$$\tilde{F}(\hat{p}) = \frac{1}{\bar{\tilde{v}}} \int_{\underline{\hat{p}}}^{\hat{p}} \tilde{v}(z)\tilde{\gamma}(z)dz,$$

where

$$\bar{\tilde{v}} = \int_{\underline{\hat{p}}}^{\bar{\hat{p}}} \tilde{v}(z)\tilde{\gamma}(z)dz$$

and from the solution to problem (11) the mass of vacancies across regions, $\tilde{v}(\hat{p})$, is

$$\tilde{v}(\hat{p}) = \sum_j \left[\left(\xi_j'\right)^{-1}\left(\hat{\pi}_j\left(\hat{p}\right)\right)\right].$$

Define $\tilde{x}(\hat{p}) \equiv \hat{x}(\hat{w}(p))$ for any $\hat{x}$. We can then re-define (35) and (36) using these definitions to obtain

$$\tilde{q}\left(\hat{p}\right) = \delta + \bar{\tilde{v}}\left[1 - \tilde{F}(\hat{p})\right] \tag{40}$$

and

$$\tilde{\mathcal{P}}(\hat{p}) = \tilde{E}(\hat{p}) + u \equiv (1-u)\tilde{G}(\hat{p}) + u, \tag{41}$$

where $\tilde{E}(\hat{p}) \equiv \sum_{x \in \mathbb{J}} \tilde{E}_x\left(\hat{p}\right)$ and $u \equiv \sum_{x \in \mathbb{J}} u_x$, and $\tilde{G}(\hat{p}) \equiv \tilde{E}(\hat{p})/(1-u)$ is the distribution of workers to firms.

Using

$$\frac{\partial \tilde{x}(\hat{p})}{\partial \hat{p}} = \frac{\partial \hat{x}(\hat{w})}{\partial \hat{w}} \frac{\partial \hat{w}(\hat{p})}{\partial \hat{p}}$$

we re-write the first-order condition (38) as

$$\frac{\partial \hat{w}(p)}{\partial \hat{p}} = \frac{(\hat{p} - \hat{w}(\hat{p}))\left(\frac{\partial \tilde{\mathcal{P}}(\hat{p})}{\partial \hat{p}} \tilde{q}\,(\hat{p}) - \frac{\partial \tilde{q}(\hat{p})}{\partial \hat{p}} \tilde{\mathcal{P}}\,(\hat{p})\right)}{\tilde{\mathcal{P}}\,(\hat{p})\,\tilde{q}\,(\hat{p})}. \tag{42}$$

By definition of a steady state, inflows and outflows from unemployment must exactly balance

$$\tilde{q}\,(\hat{p})\,\tilde{E}(\hat{p}) = \bar{\tilde{v}}\tilde{F}(\hat{p})u,$$

and hence

$$\tilde{E}(\hat{p}) = \frac{\bar{\tilde{v}}\tilde{F}(\hat{p})u}{\tilde{q}\,(\hat{p})}.$$

The mass of unemployed is given from (19) by

$$u = \frac{\delta}{\bar{\tilde{v}} + \delta}.$$

Substituting these expressions into (41) gives

$$\tilde{\mathcal{P}}(\hat{p}) = \frac{\delta}{\tilde{q}\,(\hat{p})}.$$

Plugging this expression for the acceptance probability and its derivative into (42), we obtain

$$\frac{\partial \hat{w}(\hat{p})}{\partial \hat{p}} = \frac{-2\,(\hat{p} - \hat{w}(\hat{p}))\,\frac{\partial \tilde{q}(\hat{p})}{\partial \hat{p}}}{\tilde{q}\,(\hat{p})}. \tag{43}$$

Together, equations (35), (36), (39), and (43) are the functions stated in the proposition, redefined on $\hat{p}$ instead of on p, and are the same as in the standard Burdett-Mortensen model.

E AKM Decomposition

In this section, we describe how an augmented AKM specification can recover the comparative advantage of individuals across locations. We discuss here and implement in Section G.1 a specification for East and West Germany, and we show that the data do not show any evidence of comparative advantages between these two regions. Given the lack of an East-West comparative advantage, we do not extend the analysis to the level of the four finer locations we use in the estimation in Section 6. However, the same insights and identification strategy would apply and could be performed.

Specification of the Baseline Model

We can fit in the LIAB data a linear model with additive worker and firm fixed effects, following Abowd, Kramarz, and Margolis (1999) and Card, Heining, and Kline (2013), to quantify the contribution of worker-specific and firm-specific components to the real wage gap. Specifically, we estimate

$$\log(w_{it}) = \alpha_i + \psi_{J(i,t)} + \beta \mathbb{I}^{(h_i \neq R(J(i,t)))} + BX_{it} + \epsilon_{it}, \tag{44}$$

where i indexes full-time workers, t indexes time, and $J(i,t)$ indexes worker i's firm at time t.[44] In this specification, α_i is the worker component, $\psi_{J(i,t)}$ is the component of the firm j for which worker i works at time t, and $\mathbb{I}^{(h_i \neq R(J(i,t)))}$ is a dummy that is equal to one if worker i with home region h_i (either East or West Germany) is currently employed at a firm in the other region. This term picks up the comparative advantage of workers in their home region. Finally, X_{it} is a centered cubic in age and an interaction of age and college degree, as in Card, Heining, and Kline (2013). As mentioned, we focus on regions (East or West Germany) rather than finer locations since - as we show in Section G.1 - we do not find any evidence of comparative advantage even at the regional level, where we expect it to be stronger, and since the large wage gap arises between East and West Germany.

We specify, again following Card, Heining, and Kline (2013), ϵ_{it} as three separate random effects: a match component $\eta_{iJ(i,t)}$, a unit root component ζ_{it}, and a transitory error ϵ_{it},

$$\epsilon_{it} = \eta_{iJ(i,t)} + \zeta_{it} + \epsilon_{it}.$$

In this specification, the mean-zero match effect $\eta_{iJ(i,t)}$ represents an idiosyncratic wage premium or discount that is specific to the match, ζ_{it} reflects the drift in the persistent component of the individual's earnings power, which has mean zero for each individual, and ϵ_{it} is a mean-zero noise term capturing transitory factors. We estimate the model on the largest connected set of

[44] Time is a continuous variable, since, if a worker changes multiple firm within the same year, we would have more than one wage observation within the same year.

workers in our data and results are shown in Section G.1.[45]

Identification of the Model with Comparative Advantage

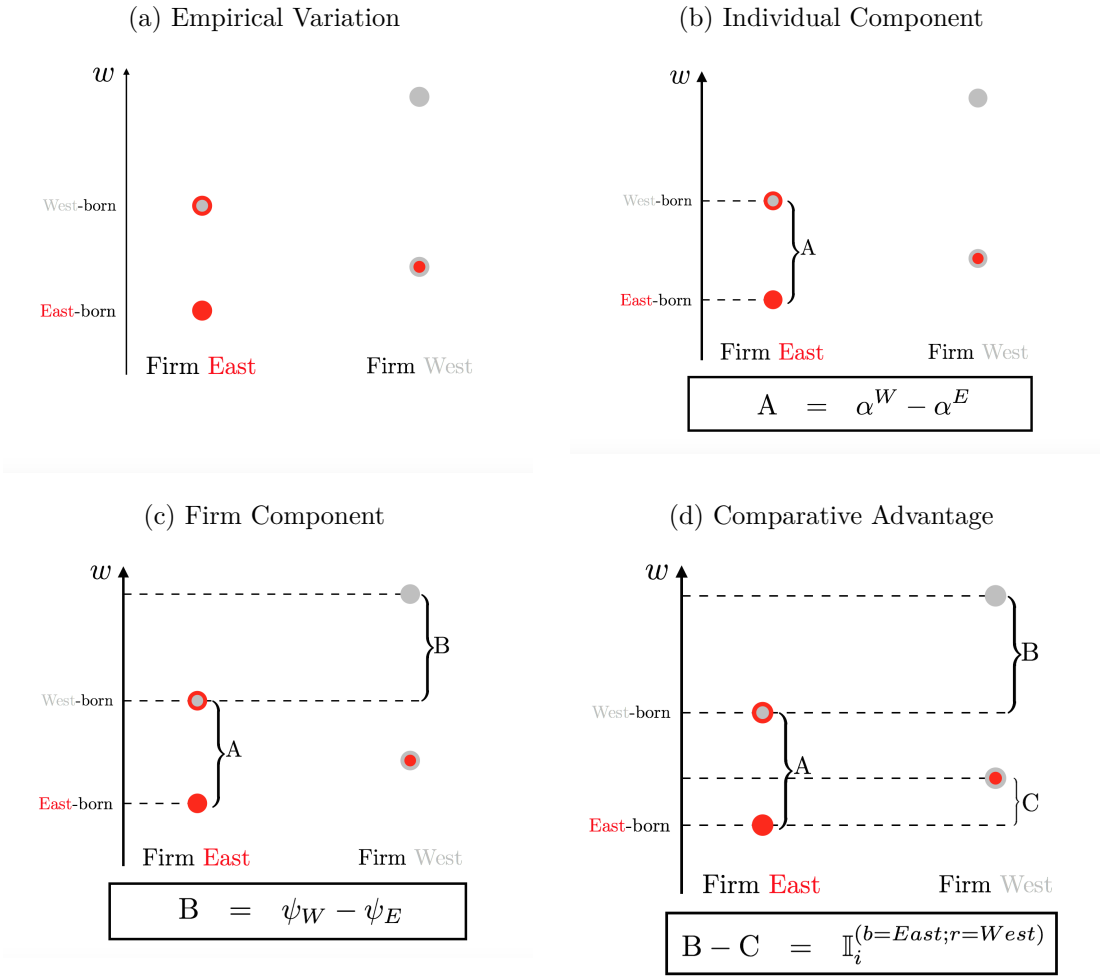
We now discuss how the specification (44) allows us to identify, through β, the comparative advantage effect by region. The same idea immediately extends to more locations.

Consider four wage observations associated with two workers: an East-born and a West-born individual working in one firm in the East, and the same two individuals working in one firm in the West. Figure A3a plots an example of these two workers' wages, where the x-axis is the identity of the firm, the y-axis is the level of the wage, the inside coloring refers to the birth region of the worker, and the outside coloring refers to the region of the firm. Figures A3b-A3d then show how these data identify the three AKM components. First, as depicted in Figure A3b, the individual components are identified from comparing the wages of the two workers when employed at the same firm. If a worker at a given firm earns a higher wage than another, this worker is identified as having a higher individual component. Second, Figure A3c highlights that the firm components are identified by comparing the same worker at two different firms. If the worker earns a higher wage at firm X than at firm Y, this difference is attributed to a higher firm component of X. Finally, Figure A3d illustrates how the comparative advantage is identified. In the absence of comparative advantages, the two workers should have an identical wage gap between them in both firms. We can thus identify the comparative advantage by comparing the wage differentials between the two workers when employed in the East- and in the West-firm, respectively.

Note that the methodology cannot separately identify whether it is the East or the West-born worker that has a comparative (dis)advantage since all that is observed is their relative wage gap. For example, if the East German worker's wage is relatively lower than the West German's wage at a firm in the West than at a firm in the East, then this difference could either arise because the East-born worker has a relative disadvantage in the West or because the West-born worker has a relative disadvantage in the East. As a result, the estimated β captures the sum of the two comparative advantages (East-born for East-Germany and West-born for West-Germany) and we need to make an arbitrary assumption in order to separately identify the two. In practice, we side-step this issue since, as we show in Section G.1, we do not find evidence of comparative advantages.

[45] While most workers (97%) are included in the sample, we miss approximately 10% of the firms included in the LIAB dataset with at least one worker during 2009-2014 in East and 11% in the West. We find that we are more likely to miss firms that pay lower wages. In fact, of the firms in the bottom decile of the average wage distribution we miss 19% in the East and 21% in the West, while of the firms in the top decile we miss 7% in the East and 5% in the West. We miss more firms than workers since – due to the nature of the exercise – large firms are more likely to be included in the connected set.

Figure A3: Identification of the AKM Components



Note: The figure illustrates the wage of two workers at two firms in East and West Germany, respectively, indexed on the x-axis. Inner coloring indicates the birth region of the worker (gray=West, red=East). Outer coloring indicates the region in which the firm is located.

F Additional Information on the Location

In this section, we provide more details on the four locations in our estimated model and the mobility between them. Figure A4 visualizes the four locations.

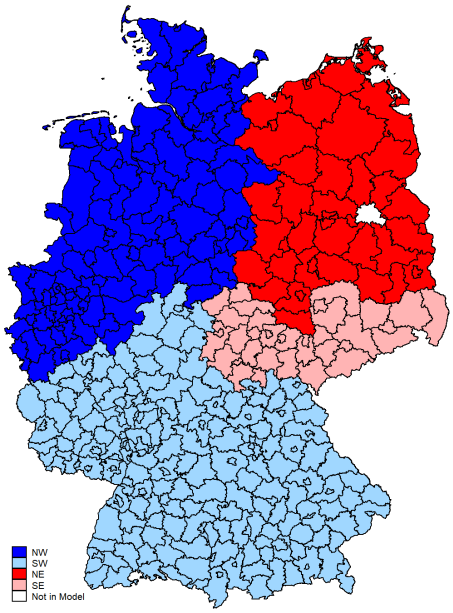
Table A11 provides some summary statistics. The first row shows the average number of individuals per year in our sample period 2009-2014 according to their work location. We include unemployed workers with their last work location prior to the unemployment spell. The Northwest location is slightly bigger than the Southwest based on the number of workers, while the Northeast and the Southeast are very similar. About 72% of workers are in West Germany. Row 2 shows the unemployment rate of each location from the German Federal Employment Agency. Unemployment is significantly higher in the East than in the West, and in both regions unemployment is higher in the North than in the South. The real GDP per capita of each location, obtained from the National Accounts of the States, mirrors this pattern, with the South of each region generating a slightly higher GDP per capita (row 3). Finally, row 4 presents the average real wage paid by firms in each location from the BHP. Real wages are very similar across the locations within East and West Germany, with a significant wage gap between the two.

Table A12 presents statistics on worker mobility across locations, analogous to the discussion of mobility between the East and West German region in Appendix B. Column 1 in the top panel of Table A12 presents the number of cross-location migrants in our core sample. Migrants are defined, as in the main text, as all workers moving job-to-job between any two locations that change their residence in the year of the move compared to one year earlier. Our sample contains about $32,000$ job-to-job migrants between locations (row 1). Column 2 of the top panel presents the same statistics using all job-to-job switchers across locations, including those that do not change their residence. Similar to the cross-region job movers in Appendix B, about 80% of cross-location job moves are done without a reported change in residence. As discussed in the main text, social security reporting regulations do not prescribe which residence to report for individuals with multiple residences, and therefore some individuals may not list the residence closest to their job. Column 3 shows our third, "intermediate" version of cross-region migration, as discussed in Section 6.2. This variable is defined as all migration moves across locations plus all cross-location job switches without a change in residence where the distance between residence and work is less than 200km at both the origin and the destination, provided that the move takes the worker further away from her current residence. We impose the upper bound on the distance between work and residence to remove workers with implausibly long commutes. Moreover, we require the distance to the residence to increase to remove job changes that take the worker closer to her current residence, since such moves do not really impose a moving cost on the worker.

The bottom panel of Table A12 shows some additional statistics for cross-location job-to-job

movers, analogous to Table A4. The columns titled "Work" show moments for the distance of the cross-location job-to-job move. The columns titled "To Live" present analogous statistics for the distance between the worker's new job after the cross-location job switch and the worker's residence. The same comments as in Appendix B apply. We note that the distances for cross-location moves are actually slightly larger than the distances for cross-region movers in Appendix B, reflecting the possibility to move large distances even within-region.

Figure A4: Locations in the Estimation



Note: The figure presents the geography of the four locations used in the estimation.

Table A11: Descriptive Statistics of the Locations

		NW	SW	NE	SE
(1)	Individuals by work location	355,907	304,158	125,377	131,959
(2)	Unemployment rate	8.8%	5.4%	12.6%	11.2%
(3)	Real GDP per capita	35,119	38,391	25,756	27,016
(4)	Average real wage	76.44	76.49	64.18	64.54

Source: BHP, LIAB, German Federal Employment Agency, National Accounts of the States, and own calculations. Notes: The table presents summary statistics for the four locations used in the estimated model. The first row shows the average number of individuals per year in our sample period 2009-2014 in each location, according to their work location. For unemployed workers, we use the last work location. Row 2 shows the average unemployment rate (Arbeitslosenquote bezogen auf abhängige, zivile Erwerbspersonen), computed as a population-weighted average across the states of each location, from the German Federal Employment Agency. Row 3 presents the real GDP per capita, computed as a population-weighted average across the states of each location, from the National Accounts of the States (Volkswirtschaftliche Gesamtrechnungen der Länder, VGRdL). The last row shows the average real wage paid by the firms in each location from the BHP.

Table A12: Number of Movers Between Locations

	Migration	All Cross-Loc	Intermediate
Number of movers	31,676	133,166	49,117
Avg. moves per year	0.006	0.022	0.009

	Migration		All Cross-Loc		Intermediate	
Distance	Work	To Live	Work	To Live	Work	To Live
Mean	322.965	81.403	292.468	144.370	244.471	87.475
P5	70.578	0	36.949	0	31.311	0
P50	323.308	14.526	295.398	49.985	199.700	38.770
P95	588.087	425.205	588.158	496.733	545.368	367.116

Source: LIAB. Notes: The first column of the top panel considers job-to-job migration moves between locations (i.e., the worker changes her residence location in the same year), the second column contains all job-to-job switches between locations, i.e., migrants plus commuters, and the third column considers migration moves plus other moves that increase the distance to the home location, as long as the distance to the residence does not exceed 200km, as described in the text. All figures are for our sample period 2009-2014. The first row of the top panel shows the number of cross-region movers between locations. The second row computes for each worker the average number of moves between locations divided by the number of years the worker is in the data and averages across all workers. The bottom panel presents some statistics on the distance of moves. The "Work" columns show the average distance between the county of the origin job and the county of the destination job for cross-location movers, as well as some selected moments of the distribution. The "To Live" present similar statistics for the distance between the work and the residence county of the worker at the destination job for cross-location movers.

G Parameters and Empirical Moments

In this section, we describe in more detail how each calibrated parameter (Section G.1) and each one of the targeted moments (Section G.2) are computed.

G.1 Calibrated Parameters

We first describe how we compute the calibrated parameters, which are shown in Table A13.

(1) Number of Firms by Region

We count in our cleaned BHP sample in each region the number of firm-year observations in the period 2009-2014. We then compute the share of firms in each region.

(2) Workers by Birth Region

To assign workers to birth regions, we obtain the population residing in each region in January 1991 from the Growth Accounting of the States (Volkswirtschaftliche Gesamtrechnung der Länder, VGRdL). This is the earliest month for which detailed population counts are available by East German states from official statistics. We do not use the LIAB data since it is not a representative sample and since it only starts in 1993. Our assumption in using residence to infer birth regions is that there was not too much net movement from East to West Germany before 1991. As a check, we obtain population estimates for the German Democratic Republic (GDR) in 1981 from Franzmann (2007), and combine these with West German population counts from the VGRdL. The population shares are, in fact, quite similar (In 1981, NW: 0.389, SW: 0.404, NE: 0.102, SE: 0.105).

(3) Separation Rate

Using the LIAB, we compute in each month the share of employed workers that become unemployed or permanently move out of the sample. We do not include workers that are temporarily out of the sample between employment spells since such workers are included in our definition of job-to-job movers. Notice that workers move out of the sample if they are either self-employed, not employed, or employed in a public sector job. We drop 2014, the last year of our sample, to avoid misclassifying workers. We then take a simple average across months for each location.

(4) Price Level

We take the price indices for each state in 2007 from the BBSR and write them forward using the inflation rate of each state obtained from the Growth Accounting of the States (Volkswirtschaftliche Gesamtrechnung der Länder, VGRdL). We aggregate the price indices in each year to the location-level by taking a population-weighted average using the population weights from the VGRdL. We then take a simple average across the years 2009-2014 for each location, and normalize Northwest to 1.

(5) Payments to Fixed Factors

We interpret the fixed factor in the model as land and set $\alpha(1-\eta)$ equal to 5%, which is the estimate of the aggregate share of land in GDP for the United States, see Valentinyi and Herrendorf (2008). As discussed in the main text, this parameter does not affect the estimation results.

(6) Elasticity of the Matching Function

We assume that the matching function has constant returns to scale - as standard in the literature, see Petrongolo and Pissarides (2001) - and puts equal weight on applications and vacancies, which gives $\chi = 0.5$. As discussed in the main text, this parameter is not separately identified from the cost of posting vacancies because we do not observe vacancy or application data.

(7) Interest Rate

We set a monthly interest rate equal to 0.5%, which accounts for both discounting and rates of retirement or death.

(8) Worker Skills

We estimate the AKM model with comparative advantage term for the worker's home area (East or West Germany)

$$\log(w_{it}) = \alpha_i + \psi_{J(i,t)} + \beta \mathbb{I}^{(h_i \neq R(J(i,t)))} + BX_{it} + \epsilon_{it}, \tag{45}$$

as discussed in Section E. As is standard, we estimate the model on the largest connected set of workers in our data, since identification of workers and firm fixed effects requires firms to be connected through worker flows.[46] This sample includes approximately 97% of West and East

[46] We use a slightly longer time period from 2004-2014 to increase the share of firms and workers that are within the connected set.

workers in the LIAB.

The estimation yields a comparative advantage estimate of $\beta = 0.019$, indicating a small *negative* comparative advantage towards the home area. Thus, a typical East-born worker is paid, controlling for firm characteristics, almost 1% more if she works in the West.[47] One possible explanation for this finding could be selection, since the workers that move to the West could be those whose skills are particularly valuable there. Since the presence of the premium would require the remaining frictions to be larger to rationalize the lack of East-to-West mobility, we conservatively set the comparative advantage to zero in our estimation.

We obtain the absolute advantage of workers from the average worker fixed effects by performing the projection

$$\hat{\alpha}_i = \eta^h \mathbb{I}_i^h + CX_i + \varepsilon_i, \tag{46}$$

where $\hat{\alpha}_i$ is the estimated worker fixed effect, $\mathbb{I}_i^h$ are dummies for the workers' home location, and X_i are dummies for worker age groups, gender, and college. We let NW be the omitted category, and obtain the η^h for the remaining three regions. We take their exponent since the AKM was estimated in logs, and present the exponentiated estimates in Table A13.

[47] We attribute half of the overall wage differential to comparative advantage of the East worker in the West and half to comparative advantage of the West worker in the East. As discussed, we cannot identify these separately.

Table A13: Calibrated Parameters

	Parameters	Source		Values	
				West	*East*
(1)	M_j: Firms by region	BHP	*North*	0.377	0.088
			South	0.445	0.090
(2)	$\bar{D}^i$: Workers by birth-region	VGRdL	*North*	0.362	0.118
			South	0.400	0.120
(3)	δ_j: Separation rate by region	LIAB	*North*	0.011	0.017
			South	0.012	0.015
(4)	P_j: Price Level by region	BBSR, VGRdL	*North*	1	0.948
			South	1.029	0.941
(5)	$\alpha(1-\eta)$: Payments to fixed factors	Valentinyi and Herrendorf (2008)		0.05	
(6)	χ: Elasticity of matching function	Assumption		0.50	
(7)	r: Monthly interest rate	Assumption		0.5 %	
(8)	θ^i: Workers' skills	AKM in LIAB	*North*	1	0.911
			South	0.986	0.896

Notes: This table reports all the parameters that are calibrated outside of the model before the estimation is run. The "Source" column provides the data source.

G.2 Moments for Estimation

Next, we turn to the 305 empirical moments targeted in the estimation and described in Table 3. Unless otherwise mentioned, all moments are constructed using the cleaned data described in the data section of the main text, for the core sample period 2009-2014.

We follow the order of the table in describing each set of moments in detail.

G.2.1 Wage Gains of Job-to-Job Movers

We compute the average wage gains of job-to-job movers between any combination of locations by estimating on all employed workers in our cleaned LIAB data the specification

$$\Delta \log(w_{it}) = \sum_{h \in \mathbb{H}} \sum_{s \in \mathbb{S}} \beta_{hs} d_{it}^s \mathbb{I}_i^h + BX_{it} + \gamma_t + \epsilon_{it}, \tag{47}$$

where $\Delta \log(w_{it})$ is the difference between a worker's log average real wage in the year after the job-to-job move and her log real wage in the job before the switch, d_{it}^s are dummies that are equal to one if worker i makes a job-to-job switch of type s at time t, and γ_t are year fixed effects. Here, $\mathbb{S}$ is the set of the 12 possible cross-location migration moves (NW-SW, NW-NE, NW-SE, SW-NW, and so on) and the 4 possible within-location moves. We define migration

moves as all job switches across locations that entail the worker updating her residence county, plus all job moves that take the worker further away from her current residence as long as the worker's residence remains within 200km of her job, as discussed in more detail in Section 6.2. We interact the move dummies with four indicator variables $\mathbb{I}_i^h$ for worker i's home location (NW, SW, NE, or SE) to identify average wage gains separately for different types of workers. Thus, in total we have 16*4=64 move-by-birth dummies of interest. The controls X_{it} contain dummies for eight age groups (26-30 years, 31-35 years, ... 56-60 years, older than 60 years, where the group of under 26 year olds is the omitted category), a dummy for whether the worker has a college degree, and a dummy for the worker's gender. The controls also include 12 dummies for non-migration cross-location job moves (for example because the worker did not change residence location and moved closer to her residence), interacted with birth location dummies. We include these latter controls so that the variables of interest, d_{it}^s, pick up wage gains of migrants relative to stayers, the omitted category. Table A14 shows the estimated coefficients on the migration dummies, and their standard errors. All coefficients are tightly estimated given the very large sample size. For each coefficient, the first column indicates the worker's home location, the second column shows the location of the worker's initial job, and the top row shows the location of the worker's new job.

Table A14: Average Log Wage Gains for Job-Job Movers by Birth and Migration Locations

Dep. var.:	New Job								
d_{it}^s	Location:	NW		SW		NE		SE	
Home	Origin Job								
Location	Location	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
NW	NW	0.109	(0.001)	0.282	(0.011)	0.136	(0.023)	0.244	(0.041)
	SW	0.195	(0.013)	0.090	(0.006)	0.048	(0.072)	0.108	(0.054)
	NE	0.127	(0.022)	0.206	(0.069)	0.051	(0.008)	0.075	(0.052)
	SE	0.164	(0.038)	0.219	(0.039)	0.202	(0.068)	0.072	(0.011)
SW	NW	0.100	(0.008)	0.169	(0.014)	0.120	(0.075)	0.134	(0.071)
	SW	0.281	(0.011)	0.107	(0.001)	0.280	(0.062)	0.186	(0.024)
	NE	0.260	(0.077)	0.138	(0.051)	0.049	(0.012)	0.029	(0.045)
	SE	0.152	(0.053)	0.161	(0.023)	0.130	(0.038)	0.085	(0.007)
NE	NW	0.081	(0.004)	0.150	(0.031)	0.031	(0.018)	0.101	(0.055)
	SW	0.177	(0.030)	0.082	(0.006)	-0.020	(0.026)	0.097	(0.043)
	NE	0.236	(0.012)	0.283	(0.027)	0.057	(0.002)	0.168	(0.015)
	SE	0.270	(0.060)	0.276	(0.038)	0.076	(0.025)	0.093	(0.008)
SE	NW	0.085	(0.008)	0.189	(0.033)	0.065	(0.056)	0.044	(0.026)
	SW	0.207	(0.032)	0.072	(0.006)	0.052	(0.077)	0.034	(0.017)
	NE	0.153	(0.060)	0.176	(0.056)	0.045	(0.010)	0.112	(0.027)
	SE	0.325	(0.024)	0.269	(0.013)	0.111	(0.014)	0.091	(0.002)

Notes: The table shows the average wage gains of job movers by origin location-destination location-home location.

G.2.2 Flows of Job-to-Job Movers

We compute in our cleaned LIAB data in each month the number of workers making a job-to-job move between any combination of locations. There are 12 possible migration moves (NW-SW, NW-NE, NW-SE, SW-NW, and so on) and 4 possible within-location job moves. We define migration moves as all job switches across locations that entail the worker updating her residence county, plus all job moves that take the worker further away from her current residence as long as the worker's residence remains within 200km of her job, as discussed in more detail in Section 6.2. We compute these movers by worker home location (i.e., their type). In total, there are thus 16*4=64 worker flows. We translate these raw flows into shares by dividing them in each month by the total number of employed workers of the given type in the location of the origin job. We exclude workers that leave the sample in the next month from this calculation, since we do not have information on whether they move or stay within the location. We also exclude the last month in our data, December 2014, for the same reason. We then take the average of these shares across months.

Table A15 shows the resulting shares. For each worker home location (first column) and location of the current job (second column), we show the share of workers changing jobs to a given destination location (indicated in the top row) in an average month, as a fraction of all employed workers of the given home location and current location.

Table A15: Job-to-Job Migration Flows Between Locations by Birth Location

	Move to Location:	NW	SW	NE	SE
Birth Location	Current Work Location				
NW	NW	0.977%	0.020%	0.004%	0.002%
	SW	0.208%	1.094%	0.006%	0.009%
	NE	0.194%	0.030%	0.948%	0.028%
	SE	0.133%	0.068%	0.041%	1.057%
SW	NW	0.983%	0.215%	0.007%	0.007%
	SW	0.025%	1.244%	0.001%	0.006%
	NE	0.084%	0.133%	0.881%	0.074%
	SE	0.033%	0.159%	0.027%	1.111%
NE	NW	1.054%	0.032%	0.077%	0.011%
	SW	0.073%	1.247%	0.069%	0.029%
	NE	0.043%	0.010%	0.911%	0.031%
	SE	0.038%	0.047%	0.124%	1.006%
SE	NW	1.031%	0.089%	0.019%	0.094%
	SW	0.043%	1.179%	0.010%	0.117%
	NE	0.031%	0.030%	0.608%	0.138%
	SE	0.011%	0.033%	0.020%	1.080%

Notes: The table presents the share of employed workers that make a job-to-job move for each triplet of home location, current location, destination location in an average month.

G.2.3 Employment Share

We count in our cleaned LIAB data in each month the number of employed workers of a given type (home location) living in each location, and we divide by the total number of employed workers of that type in our LIAB data to obtain shares. We then average across months. We similarly compute the share of employed workers working in each location. Table A16 presents these worker shares. The first column indicates the home location of the worker, and the second column indicates the residence/work location. Columns 3 and 4 show the shares of employed workers of the given home location that live in a given location (column 3) and work in a given location (column 4). In our baseline estimation, we use the residence location as target for the distribution of labor since it more closely reflects the way in which we define a cross-location move. We use the work location in some of the robustness checks in Supplementary Appendix P.

Table A16: Share of Employed Workers by Location of Residence or Work Location

Home Location	Location of...	...Residence	...Work
NW	NW	92.7%	92.0%
	SW	4.4%	5.6%
	NE	2.0%	1.6%
	SE	0.8%	0.8%
SW	NW	4.3%	6.1%
	SW	92.5%	90.9%
	NE	0.8%	0.8%
	SE	2.3%	2.2%
NE	NW	7.6%	12.8%
	SW	4.3%	5.8%
	NE	84.7%	77.1%
	SE	3.4%	4.4%
SE	NW	3.0%	4.4%
	SW	6.7%	9.8%
	NE	2.5%	3.9%
	SE	87.7%	81.9%

Notes: The table shows the fraction of employed workers of the home location indicated in column 1 that live in the location indicated in column 2 and that work the location indicated in column 2, respectively.

G.2.4 Unemployment Share

We count in our cleaned LIAB data in each month the number of unemployed workers of a given type (home location) living in each location, and we divide by the total number of unemployed workers of that type to obtain shares. We then average across months. We similarly compute the share of unemployed workers by last work location of the worker. We obtain the last work location as the location of the most recent job before the unemployment spell, and we exclude unemployed workers whose last job was in Berlin and workers that do not have a prior employment spell. Table A17 presents these worker shares. In our baseline estimation, we use the residence location as target for the distribution of labor since it more closely reflects the way in which we define a cross-location move. We use the work location in some of the robustness checks in Supplementary Appendix P.

Table A17: Share of Unemployed Workers by Location of Residence or Location of Last Job

	Location of...	Residence	Last Job
Home Location			
NW	NW	90.9%	89.1%
	SW	4.5%	6.5%
	NE	3.3%	3.1%
	SE	1.3%	1.4%
SW	NW	4.7%	7.4%
	SW	90.2%	87.5%
	NE	1.5%	1.5%
	SE	3.6%	3.6%
NE	NW	4.9%	10.6%
	SW	2.9%	5.5%
	NE	89.5%	78.8%
	SE	2.7%	5.2%
SE	NW	2.4%	4.2%
	SW	4.8%	9.2%
	NE	2.9%	4.2%
	SE	90.0%	82.4%

Notes: The table shows the fraction of unemployed workers of the home location indicated in column 1 that live in the location indicated in column 2 and whose last job was in the location indicated in column 2, respectively.

G.2.5 Firm Component of Wages by Location and Worker Type

We perform in our cleaned LIAB data a regression of the firm fixed effects from our AKM model on dummies for an employed worker's residence location, by worker type, and controls

$$fe_{it} = \sum_{h \in \mathbb{H}} \sum_{l \in \mathbb{L}} \beta_{hl} \mathbb{I}_{it}^{l} \mathbb{I}_{i}^{h} + BX_{it} + \epsilon_{it}, \tag{48}$$

where fe_{it} is the firm fixed effect of the firm at which worker i is employed at time t, obtained from the AKM estimated in Section G.1, $\mathbb{I}_{it}^{l}$ are dummies that are equal to one if worker i lives in location l at time t, $\mathbb{L} = \{NW, SW, NE, SE\}$, and $\mathbb{I}_{i}^{h}$ are dummies that are equal to one if worker i's home location is location h. Here, $\mathbb{H}$ is the set of the 4 possible birth locations (NW, SW, NE, and SE). The controls X_{it} contain dummies for eight age groups (26-30 years, 31-35 years, ... 56-60 years, older than 60 years, where the group of under 26 year olds is the omitted category), a dummy for whether the worker has a college degree, and a dummy for the worker's gender. In a second specification, we run an analogous regression using dummies for a worker's work location rather than her residence location.

Table A18 shows the estimated coefficients. The first two columns with data show the estimated coefficients β_{hl} for workers with home location h indicated in column 1 and residence

location l indicated in column 2, together with their standard errors. Each of the coefficients is relative to the coefficient of workers born in the Northwest and living in the Northwest. The last two data columns show the analogous estimates for workers with home location h indicated in column 1 and work location l indicated in column 2. In our baseline estimation, we use the moments related to the residence location as target since they more closely reflect the way in which we define a cross-location move. We use the moments related to the work location in some of the robustness checks in Supplementary Appendix P.

Table A18: Firm Fixed Effects by the Birth and Current Location of Workers

Dep. var.: fe_{it}	Location of...	Live		Work	
Home Location		Coefficient	SE	Coefficient	SE
NW	SW	-0.064	0.001	-0.060	0.001
	NE	-0.141	0.001	-0.210	0.001
	SE	-0.139	0.002	-0.147	0.002
SW	NW	-0.036	0.001	-0.038	0.001
	SW	-0.046	0.000	-0.046	0.000
	NE	-0.193	0.002	-0.213	0.002
	SE	-0.165	0.001	-0.187	0.001
NE	NW	-0.090	0.001	-0.070	0.001
	SW	-0.104	0.001	-0.113	0.001
	NE	-0.198	0.000	-0.211	0.000
	SE	-0.119	0.001	-0.163	0.001
SE	NW	-0.056	0.001	-0.062	0.001
	SW	-0.090	0.001	-0.088	0.001
	NE	-0.171	0.002	-0.163	0.001
	SE	-0.169	0.000	-0.177	0.000

Notes: The table shows the estimated coefficients β_{hl} in specification (48). The first two columns with data show the coefficients for workers with home location h indicated in column 1 and residence location l indicated in column 2, together with their standard errors. Each of the coefficients is relative to the coefficient of workers born in the Northwest and living in the Northwest. The last two data columns show the analogous estimates for workers with home location h indicated in column 1 and work location l indicated in column 2.

G.2.6 Firm Component of Wages by Firm Location

We collapse the cleaned LIAB data to the firm-level and perform a regression of the firm fixed effects from our AKM model on dummies for each firm's location:

$$fe_j = \sum_{l \in \mathbb{L}} \beta_l \mathbb{I}_j^l + \epsilon_j, \tag{49}$$

where fe_j is the estimated firm fixed effect of firm j, and $\mathbb{I}_j^l$ are dummies that are equal to one if firm j is in location l. Using the firm fixed effects instead of actual real wages isolates the firm component of wages and removes differences in wages due to worker composition. We do

not include industry controls since we want our model to be consistent with the aggregate wage gaps between locations, which could partially be due to differences in industry composition. Our estimated productivity shifters therefore also reflect industry differences across locations, although they are not quantitatively important, as shown in Supplemental Appendix M. For similar reasons, we do not include demographic controls. Table A19 presents the estimated coefficients β_l for firm location l indicated in column 1, where NW is the omitted category.

While in our baseline specification we do not include controls since we simply want to capture the differences in average firm productivity across locations, we also computed an alternative specification with a vector of controls X_j. We control for firm-level averages, averaged across all workers at the firm, of dummies for eight age groups (26-30 years, 31-35 years, ... 56-60 years, older than 60 years, where the group of under 26 year olds is the omitted category), a dummy for whether a worker has a college degree, and a dummy for workers' gender. The results barely change.[48]

Table A19: Firm Fixed Effect by Location

Dep. var.: fe_j	Coef on Firm FE	SE
Location		
SW	.001	.002
NE	-.166	.002
SE	-.141	.003

Notes: The table presents the estimated coefficients β_l from specification (49) for firm location l indicated in column 1, where NW is the omitted category.

G.2.7 GDP per Capita

We obtain nominal GDP per capita for each federal state from the National Accounts of the States (Volkswirtschaftliche Gesamtrechnungen der Länder, VGRdL) for each year. To translate the nominal figures into real ones, we compute the price level in each state in 2007 as a population-weighted average across the county-level prices reported by the BBSR. We then extend the resulting state-level prices in 2007 forward to 2014 using the state-level deflators available in the VGRdL. We deflate each state's nominal GDPpc with the resulting prices in each year to obtain state-level real GDPpc in each year, and we aggregate to the location level using each state's population in each year, also reported in the VGRdL. We take a simple average over the years in our core sample period and normalize real GDP per capita in NW to 1. Table A20 presents the results.

[48]Specifically, the three coefficients for SW, NE, and SE become: -0.001, -0.154, -.144.

Table A20: Average GDP per capita by Location

Location	Avg. GDP pc	Normalized to 1
NW	35,119	1
SW	38,391	1.09
NE	25,756	0.73
SE	27,016	0.77

Notes: The table shows a simple average over the GDPpc of each location in the period 2009-2014. We obtain nominal GDPpc from the National Accounts of the States (Volkswirtschaftliche Gesamtrechnungen der Länder, VGRdL), and construct price deflators from the inflation rates in the VGRdL and the price data from the survey of the Federal Institute for Building, Urban Affairs and Spatial Development (BBSR).

G.2.8 Unemployment Rate

We obtain the unemployment rate (Arbeitslosenquote bezogen auf abhängige, zivile Erwerbspersonen) of each federal state in each month from the official unemployment statistics of the German Federal Employment Agency. We compute this moment from the official statistics rather than from the smaller LIAB sample since the latter is not representative and includes unemployed individuals only for as long as they are receiving unemployment benefits. We aggregate across states to locations using each state's labor force as weight, and take a simple average across the months in our core sample period. Table A21 shows the estimates.

Table A21: Unemployment Rate by Location

Location	Unemployment Rate
NW	8.82%
SW	5.35%
NE	12.58%
SE	11.16%

Note: The table shows the average unemployment rate in each location in the period 2009-2014, computed from the official unemployment statistics of the German Federal Employment Agency.

G.2.9 Labor Share for Each Decile of Firm Size Distribution

We obtain in our cleaned BHP data the number of full-time workers employed at each firm in each year in our core sample period. We then remove variation due to observables that are not present in our model by performing, for each work location, the following regression

$$\ln(y_{jlt}^{size}) = B_l X_{jlt} + \gamma_t + \epsilon_{jlt}, \tag{50}$$

where y_{jlt}^{size} is the number of full-time workers of firm j in location l in year t and γ_t are year fixed effects. The controls X_{jlt} include the share of male full-time workers, the share of young full-time workers (of age less than 30 years old), and the share of full-time workers of medium age (30-49 years old). The controls also include the share of full-time workers of low qualifications (individuals with a lower secondary, intermediate secondary, or upper secondary school leaving certificate but no vocational qualifications) and the share of full-time workers of medium qualifications (individuals with a lower secondary, intermediate secondary, or upper secondary school leaving certificate and a vocational qualification). Finally, we include 3-digit time-consistent industry dummies based on Eberle, Jacobebbinghaus, Ludsteck, and Witter (2011) (WZ93 classification).

Based on the four regressions (one for each work location l) we obtain residuals for the log number of workers at each firm j, $\hat{\epsilon}_{jlt}^{size}$. We add back the mean log number of workers in each location, $\overline{\ln(y_{jlt}^{size})}$, to obtain a cleaned number of workers, $\hat{y}_{jlt}^{size} = \exp[\overline{\ln(y_{jlt}^{size})} + \hat{\epsilon}_{jlt}^{size}]$. We then construct deciles of the distribution of residualized firm size in each location and compute the share of residualized workers employed in each decile. Table A22 presents the resulting shares. Each column of the table shows the share of the location's workers employed at firms in the decile of the location's residualized firm size distribution indicated in column 1.

Table A22: Share of Workers by Firm Size Decile and Location

Firm Size Decile	NW	SW	NE	SE
1	0.009	0.008	0.010	0.009
2	0.013	0.013	0.015	0.015
3	0.017	0.016	0.019	0.019
4	0.022	0.021	0.024	0.024
5	0.029	0.028	0.034	0.033
6	0.038	0.036	0.043	0.042
7	0.052	0.050	0.058	0.057
8	0.074	0.071	0.083	0.081
9	0.124	0.119	0.136	0.135
10	0.622	0.636	0.578	0.584

Notes: Each column of the table shows the share of the location's workers employed at firms in the decile of the location's firm size distribution indicated in column 1. The number of workers used in the table is residualized using firms' share of male workers, share of workers with low and medium skills, share of young and medium-aged workers, and industry dummies, as described in the text.

G.2.10 Relationship between Firm Wage and Firm Size

We obtain in our cleaned BHP data the number of full-time workers and their average wage at each firm, where top coded wages are imputed as in Card, Heining, and Kline (2013). We then remove variation due to observables that is not present in our model by performing, for each work location l, the following regression

$$\ln(y_{jlt}) = B_l X_{jlt} + \gamma_t + \epsilon_{jlt},$$

where y_{jlt} is either the number of full-time workers of firm j in location l in year t or their average wage, and γ_t are year fixed effects. The controls X_{jlt} include the share of male full-time workers, the share of young full-time workers (of age less than 30 years old), and the share of full-time workers of medium age (30-49 years old). The controls also include the share of full-time workers of low qualification (individuals with a lower secondary, intermediate secondary, or upper secondary school leaving certificate but no vocational qualifications) and the share of full-time workers of medium qualification (individuals with a lower secondary, intermediate secondary, or upper secondary school leaving certificate and a vocational qualification). Finally, we include 3-digit time-consistent industry dummies based on Eberle, Jacobebbinghaus, Ludsteck, and Witter (2011) (WZ93 classification).

We obtain from these four regressions (one for each location l) residuals for the log real wage, $\hat{\epsilon}^{wage}_{jlt}$, and for the log number of workers, $\hat{\epsilon}^{size}_{jlt}$. We add back the mean of each variable in each location, $\overline{\ln(y^{wage}_{jlt})}$ and $\overline{\ln(y^{size}_{jlt})}$, to obtain a cleaned log real wage, $\ln(\hat{y}^{wage}_{jlt}) = \overline{\ln(y^{wage}_{jlt})} + \hat{\epsilon}^{wage}_{jlt}$ and a cleaned log number of workers, $\ln(\hat{y}^{size}_{jlt}) = \overline{\ln(y^{size}_{jlt})} + \hat{\epsilon}^{size}_{jlt}$ for each firm. We then regress the residualized log real wage on the residualized log number of workers in each location

$$\ln(\hat{y}^{wage}_{jlt}) = \beta_{0,l} + \beta_{1,l} \ln(\hat{y}^{size}_{jlt}) + \varepsilon_{jlt}, \tag{51}$$

and report the slope coefficients $\beta_{1,l}$ in Table A23. We also plot the non-parametric relationships between $\ln(\hat{y}^{wage}_{jlt})$ and $\ln(\hat{y}^{size}_{jlt})$ in Figure A8, panel (a).

Table A23: Log Wage on Log Firm Size by Location

Dep. var.: $\ln(\hat{y}_{jlt}^{wage})$	Coefficient	SE
Location		
NW	.124	.000
SW	.124	.000
NE	.110	.001
SE	.109	.001

Notes: The table presents the coefficients $\beta_{1,l}$ of regression (51), by location of the firm, indicated in the first column. The residualization procedure is described in the text.

G.2.11 Wage Gains of Job-to-Job Movers by Origin Firm Wage

We identify in our cleaned LIAB data all job-to-job moves and determine for each move the origin location of the worker (NW, SW, NE, or SE). We restrict the dataset to only these observations. We compute the log real wage gain associated with each job-to-job move, defined as the difference between a worker's log average real wage in the year after the job-to-job move and her log real wage in the job before the switch. We then residualize these wage gains to take out observable heterogeneity not present in our model by running, separately for each location l of the initial job, the regression

$$\Delta \ln(w_{ilt}) = B_l X_{ilt} + \gamma_t + \epsilon_{ilt}, \tag{52}$$

where $\Delta \ln(w_{ilt})$ is the log real wage gain associated with the move and γ_t are year fixed effects. The controls X_{ilt} contain dummies for eight age groups (26-30 years, 31-35 years, ... 56-60 years, older than 60 years, where the group of under 26 year olds is the omitted category), a dummy for whether the worker has a college degree, a dummy for the worker's gender, and 3-digit time-consistent industry (of the origin firm) dummies based on Eberle, Jacobebbinghaus, Ludsteck, and Witter (2011) (WZ93 classification). From these four regressions (one for each location l), we construct residuals for the log real wage gain, $\hat{\epsilon}_{ilt}^{gain}$. We add back the mean of the log real wage gain in each location, $\overline{\Delta \ln(w_{ilt})}$, to obtain a cleaned log real wage, $\Delta \ln(\hat{w}_{ilt}) = \overline{\Delta \ln(w_{ilt})} + \hat{\epsilon}_{ilt}^{gain}$. We similarly residualize the log real wage of the worker at the origin firm, $\ln(w_{ilt-1})$, to obtain the residualized initial log real wage, $\ln(\hat{w}_{ilt-1})$. We then regress the residualized log real wage gains on the residualized log initial real wages in each location

$$\Delta \ln(\hat{w}_{ilt}) = \beta_{0,l} + \beta_{1,l} \ln(\hat{w}_{ilt-1}) + \varepsilon_{ilt} \tag{53}$$

and report the slope coefficients $\beta_{1,l}$ in Table A24. In this table, each row shows the estimated regression coefficient on the residualized log initial wage for job-to-job moves originating in the location indicated in the first column. We also plot the non-parametric relationships between $\Delta \ln(\hat{w}_{ilt})$ and $\ln(\hat{w}_{ilt-1})$ in Figure A8, panel (b).

Table A24: Log Wage Gain of Movers by Initial Wage

Dep. var.: $\Delta \ln(\hat{w}_{irt})$	Coefficient	SE
Location		
NW	-.549	.001
SW	-.577	.000
NE	-.562	.003
SE	-.561	.002

Note: The table presents the coefficients $\beta_{1,l}$ of regression (53), by location of the origin firm. The residualization procedure is described in the text.

G.2.12 Separation/Quit Rate by Initial Wage

We identify in our cleaned LIAB data in each month the workers moving job-to-job, from a job into unemployment, or from a job to permanently out of the sample. We construct a dummy that is equal to one if worker i with current job in location l at time t makes such a move, d^{sep}_{ilt}. We also obtain the log real wage of each worker in the job prior to the move, $\ln(w_{ilt})$. We then residualize these two variables to take out observable heterogeneity not present in our model by running, separately for each location of the initial job, the regression

$$y_{ilt} = B_l X_{ilt} + \gamma_t + \epsilon_{ilt}, \tag{54}$$

where y_{ilt} is either the dummy indicating a separation or the worker's log real wage in the job prior to the move. The controls X_{ilt} contain dummies for eight age groups (26-30 years, 31-35 years, ... 56-60 years, older than 60 years, where the group of under 26 year olds is the omitted category), a dummy for whether the worker has a college degree, a dummy for the worker's gender, and 3-digit time-consistent industry (of the origin firm) dummies based on Eberle, Jacobebbinghaus, Ludsteck, and Witter (2011) (WZ93 classification). From these four regressions (one for each location l), we construct residuals for the log initial real wage, $\hat{\epsilon}^{wage}_{ilt}$, and for the separation dummy, $\hat{\epsilon}^{sep}_{ilt}$, and add back the mean of each variable in each location, $\overline{\ln(w_{ilt})}$ and $\overline{d^{sep}_{ilt}}$, to obtain a cleaned log wage, $\ln(\hat{w}_{ilt}) = \overline{\ln(w_{ilt})} + \hat{\epsilon}^{wage}_{ilt}$ and a cleaned separation dummy $\hat{d}^{sep}_{ilt} = \overline{\ln(d^{sep}_{ilt})} + \hat{\epsilon}^{sep}_{ilt}$. We then regress the residualized separation dummy on the residualized log wages for each location

$$\hat{d}^{sep}_{ilt} = \beta_{0,l} + \beta_{1,l} \ln(\hat{w}_{ilt}) + \varepsilon_{ilt} \tag{55}$$

and report the slope coefficients $\beta_{1,l}$ in Table A25. In this table, each row shows the estimated regression coefficient on the residualized log initial real wage for separations from jobs in the location indicated in the first column. We also plot the non-parametric relationships between $\hat{d}^{sep}_{ilt}$ and $\ln(\hat{w}_{ilt})$ in Figure A8, panel (c).

Table A25: Avg. Separation Rates of Workers by Initial Wage

Dep. var.: $\hat{d}^{sep}_{irt}$	Coefficient	SE
Location		
NW	-.029	.000
SW	-.033	.000
NE	-.037	.000
SE	-.036	.000

Notes: The table presents the coefficients $\beta_{1,l}$ of regression (55), by location of the firm. The residualization procedure is described in the text.

G.2.13 Standard Deviation of Wage Gains

We identify in our cleaned LIAB data all migration moves between any combination of locations (NW-SW, NW-NE, NW-SE, SW-NW, and so on). We define migration moves as all job switches across locations that entail the worker updating her residence county, plus all job moves that take the worker further away from her current residence as long as the worker's residence remains within 200km of her job, as discussed in more detail in Section 6.2. We also identify job-to-job moves within-location, for each of the four locations. We indicate for each move the home location of the worker making the move. We restrict the dataset to these job-to-job moves and compute the log real wage gain associated with each move, defined as the difference between a worker's log average real wage in the year after the job-to-job move and her log real wage in the job before the switch. We then residualize these wage gains to take out observable heterogeneity not present in our model by running, separately for each location of the initial job, the regression

$$\Delta \ln(w_{ilt}) = B_l X_{ilt} + \gamma_t + \epsilon_{ilt}, \tag{56}$$

where $\Delta \ln(w_{ilt})$ is the log real wage gain associated with the move of worker i with initial job in location l at time t. The controls X_{ilt} contain dummies for eight age groups (26-30 years, 31-35 years, ... 56-60 years, older than 60 years, where the group of under 26 year olds is the omitted category), a dummy for whether the worker has a college degree, and a dummy for the worker's gender. From these four regressions (one for each location of the initial job l), we construct residuals for the log real wage gain, $\hat{\epsilon}^{gain}_{ilt}$. We then compute the standard deviation of these residualized wage gains for each home location-origin-destination combination. These coefficients are in Table A26. For each worker home location (first column) and location of the

current job (second column), we show the standard deviation of wage gains for workers changing jobs to a given destination location (indicated in the top row).

Table A26: Standard Deviation of the Residual Wage Gains for Job Movers

	New Job Location:	NW	SW	NE	SE
Home Location	Current Job Location				
NW	NW	0.564	0.763	0.640	0.772
	SW	0.656	0.546	0.655	0.546
	NE	0.545	0.671	0.442	0.486
	SE	0.562	0.435	0.589	0.435
SW	NW	0.558	0.660	0.652	0.644
	SW	0.743	0.543	0.948	0.734
	NE	0.834	0.682	0.413	0.463
	SE	0.625	0.589	0.392	0.437
NE	NW	0.445	0.587	0.522	0.584
	SW	0.573	0.457	0.473	0.520
	NE	0.651	0.752	0.455	0.684
	SE	0.695	0.503	0.525	0.472
SE	NW	0.477	0.613	0.485	0.499
	SW	0.661	0.470	0.691	0.530
	NE	0.640	0.628	0.424	0.578
	SE	0.729	0.645	0.526	0.471

Notes: The table shows the standard deviation of the residualized wage gains of job-to-job movers, $\hat{\epsilon}_{ilt}^{gain}$, for workers of a given home location (column 1) and current job location (column 2) that move jobs to a given destination location (top row). The residualization procedure is described in the text.

G.2.14 Ratio of Profits to Labor Costs

We obtain the pre-tax profits of all firms in Germany from the ORBIS database provided by the company Bureau van Dijk. We allocate firms to our four locations based on the ZIP code of their address, and drop firms with fewer than 5 employees since their profits are very noisy. We then construct the ratio of profits to labor costs by dividing pre-tax profits by total labor costs reported in ORBIS, and average across firms and years to compute the average ratio in each location. We drop outlier profit ratios below the 5th and above the 95th percentile of the distribution of profit ratios in each location, and compute the average over the remaining ratios. Table A27 presents the estimates.

Table A27: Average Ratio of Firm Profits to Labor Costs by Location

Location	Avg. Profit Share
NW	27.44%
SW	25.87%
NE	29.87%
SE	26.26%

Source: ORBIS database. Notes: The table presents the average ratio of pre-tax profits to total labor costs for firms in the location indicated in the first column.

H Model's Computation and Estimation

We here provide a brief explanation of the solution algorithm and more details on the estimation approach and outcomes.

H.1 Solution Algorithm

To solve the model, we follow a nested iterative procedure. Leveraging Proposition 1, we solve the model in the one-dimensional productivity space. In other words, rather than keep track of both wages and productivity, we simply solve for all the functions directly on the productivity support. Our procedure is as follows:

1. Make an initial guess for wage offer distributions, $\{w_j(p)\}_{j\in\mathbb{J}}$, firm vacancies $\{v_j(p)\}_{j\in\mathbb{J}}$, market tightness $\{\vartheta_j\}_{j\in\mathbb{J}}$, and vacancy sizes $\left\{\tilde{l}_j^i(p)\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$, which gives

$$\left\{w_j(p;k), v_j(p;k), \vartheta_j(k), \tilde{l}_j^i(p;k)\right\}_{j\in\mathbb{J},i\in\mathbb{I},k=0},$$

   where k indexes the external iteration loop.

2. Given $\left\{w_j(p;k), v_j(p;k), \vartheta_j(k), \tilde{l}_j^i(p;k)\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$, solve the workers' problem through value function iteration, which yields the value functions, and most importantly, the acceptance probabilities for every pair of firms (p,p') and each worker's type i, and the job applications:

$$\begin{aligned}&\left\{\tilde{\mu}_{jx}^{E,i}(p,p';k), \tilde{\mu}_{jx}^{U,i}(b,p';k)\right\}\\&\left\{\tilde{a}_{jx}^{E,i}(p;k), \tilde{a}_{jx}^{U,i}(b;k)\right\}_{j\in\mathbb{J},x\in\mathbb{J},i\in\mathbb{I}}.\end{aligned}$$

3. Given $\left\{\tilde{\mu}_{jx}^{E,i}(p,p';k), \tilde{\mu}_{jx}^{U,i}(b,p';k), \tilde{a}_{jx}^{E,i}(p;k), \tilde{a}_{jx}^{U,i}(b;k)\right\}_{j\in\mathbb{J},x\in\mathbb{J},i\in\mathbb{I}}$, we use equation (16) to solve for $\left\{\tilde{q}_j^i(p;k)\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$ and then iterate through equations (15), (17), and (18) until convergence to get a new guess for the firm size per vacancy $\left\{\tilde{l}_j^i(p;k+1)\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$ that is consistent with the steady state employment distributions $\tilde{E}_j^i(p;k)$ and the probability of accepting offers $\tilde{\mathcal{P}}_j^i(p;k)$.

4. Finally, using $\left\{\tilde{l}_j^i(p;k), \tilde{q}_j^i(p;k)\right\}_{j\in\mathbb{J},x\in\mathbb{J},i\in\mathbb{I}}$, and solving for the boundary conditions at $w_j\left(\underline{p}_j\right)$ we can solve for a new guess for firm wages $\{w_j(p;k+1)\}_{j\in\mathbb{J}}$ using the system of differential equations in Proposition 1. Then, using the equations shown in the model section, we can get new guesses for vacancies and market tightness. We thus have a new vector

$$\left\{w_j(p;k+1), v_j(p;k+1), \vartheta_j(k+1), \tilde{l}_j^i(p;k+1)\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$$

and can go back to point 2.

5. We iterate the external loop 2-4 until there is convergence within each iterative loop, namely the ones for value functions, vacancy sizes, and firm wages.

In order to compute the general equilibrium counterfactuals, we follow the same algorithm, but with two differences. First, as mentioned in the main text, during the estimation of the model, we solve - within each loop - for the unemployment benefits that make the value of unemployed workers identical to the value of workers at the lowest productivity firm in our estimated economy. In the counterfactuals, instead, we keep the unemployment benefits fixed at their estimated value. Second, while during the estimation we can keep each location's prices fixed at their observed values, in the counterfactual we must solve for the new equilibrium prices. Therefore, within each loop, we calculate each location's GDP and then we use it to calculate the new aggregate equilibrium prices.

H.2 Estimation Algorithm and Outcomes

The objective is to find a parameter vector ϕ^* that solves

$$\phi^* = \arg\min_{\phi \in \mathbb{F}} \mathcal{L}(\phi) \tag{57}$$

where

$$\mathcal{L}(\phi) \equiv \sum_x \left[\omega_x \left(T_x \left(m_x(\phi), \hat{m}_x\right)\right)^2\right]$$

and $\mathbb{F}$ is the set of admissible parameter vectors, which is bounded to be strictly positive (or negative for search distance) and finite. In the choice of the function $T_x(\cdot)$, for most moments we follow Jarosch (2016) and Lise, Meghir, and Robin (2016) and minimize the sum of the percentage deviations between model-generated and empirical moments; for others, instead, we use log differences. Specifically, for the moments that are already expressed in logs – rows (1), (5), (6), (7), (10), (11), (12), (14) of Table 3 – $T_x(\cdot)$ is the percentage deviation: $T_x(m_x(\phi), \hat{m}_x) = \frac{m_x(\phi)-\hat{m}}{\hat{m}}$. For the other moments, $T_x(\cdot)$ is the log difference: $T_x(m_x(\phi), \hat{m}_x) = \log m_x(\phi) - \log \hat{m}$. Using the log difference is important especially for job flows to avoid giving excessive weight to deviations between model and data for flows that have very small magnitudes. Nonetheless, we have re-estimated the model using percentage deviations for all moments, and the results are broadly consistent, although the estimation procedure is less effective. We also introduce an additional weighting factor ω_x to give equal weight to each one of the 14 groups of parameters that we target, shown in Table 3.

The minimization algorithm that we use to solve the problem (57) combines the approaches of Jarosch (2016) and Lise, Meghir, and Robin (2016), and Moser and Engbom (2018), both

adapted to our needs.

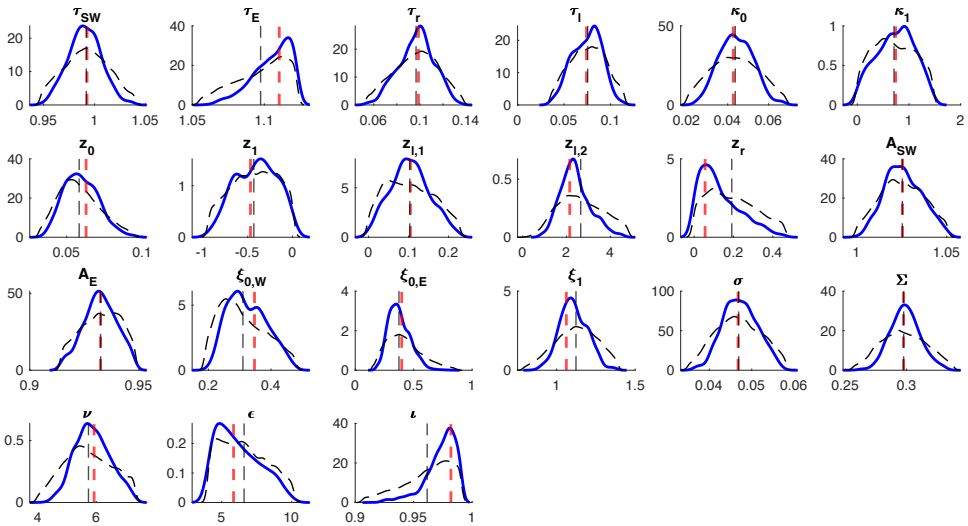
We simulate, using Markov Chain Monte Carlo for classical estimators as introduced in Chernozhukov and Hong (2003), 200 strings of length 10,000 (+ 1,000 initial scratch periods used only to calculate posterior variances) starting from 200 different guesses for the vector of parameters ϕ_0. In the first run, we choose the initial guesses to span a large space of possible parameter vectors. In updating the parameter vector along the MCMC simulation, we pick the variance of the shocks to target an average rejection rate of 0.7, as suggested by Gelman, Carlin, Stern, Dunson, Vehtari, and Rubin (2013). The average parameter values across the 200 strings for the last 1,000 iterations provide a first estimate of the vector of parameters. We then repeat the same MCMC procedure, but we start each string from the parameter estimates of the first step. We pick our final estimates as the average across the parameter vectors, picked from all strings, that are associated with the 100 smallest values of the likelihood functions.

Figure A5 illustrates our approach and how it slightly differs from Jarosch (2016) and Lise, Meghir, and Robin (2016). The black dotted line shows the density function of the last 1,000 iterations across all strings. The usual approach is to pick the average across all these draws, which we highlight in the picture with a vertical black dotted line. However, this approach could be problematic if the parameter space is bounded, hence the estimated densities are not symmetric, as in our case for some parameters. Therefore, given our vector of parameters and likelihoods, we pick the optimal parameter following Moser and Engbom (2018), and simply select the vector of parameters that minimizes the objective function among all our draws.[49] Our estimates are shown with red dotted lines in the figure. For most parameters, they are almost identical to the alternative approach. Finally, the blue density functions shows the density, across all strings, of the 10 best outcomes within each string. This density provides a visual representation of the tightness of our estimates, which are, in general, quite good – especially for the key parameters that determine the spatial frictions. It is also relevant to notice that all the densities are single-peaked, which suggests that the model is, at least locally, tightly identified.

All the estimated parameters, corresponding to the vertical dotted red lines, are included in Table A28 below.

[49]More precisely, we take the average across the 100 best outcomes across all the 2,000,000 draws.

Figure A5: Estimation Outcomes



Notes: The figure shows the outcomes of the estimation. Each panel shows a different one of the 21 estimated parameters. As described in the text, the black dashed and blue lines show the densities for different sub-sets of parameter draws. The red vertical lines are our estimated parameters, while the black vertical lines show the estimates that we would obtain with the alternative approach, described above. The top row shows the estimation results for τ_{SW}, τ_E, τ_r, τ_l, κ_0 and κ_1. The second row shows the results for z_0, z_1, $z_{l,1}$, $z_{l,2}$, z_r, and A_{SW}. The third row shows the estimates for A_E, $\xi_{0,W}$, $\xi_{0,E}$, ξ_1, σ, and Σ. The last row shows the estimates for ν, ϵ, and ι.

Table A28: All Estimated Parameters

(1)	τ_{SW}: amenity SW	0.993	(12)	A_{SW}: productivity SW	1.025
(2)	τ_E: amenity East	1.110	(13)	A_E: productivity East	0.932
(3)	τ_r: region preference	0.099	(14)	$\xi_{0,W}$: vacancy cost West	0.347
(4)	τ_l: location preference	0.074	(15)	$\xi_{0,E}$: vacancy cost East	0.398
(5)	κ_0: move cost out of location	0.043	(16)	ξ_1: vacancy curvature	1.062
(6)	κ_1: move cost distance	0.742	(17)	σ: variance of taste shocks	0.047
(7)	z_0: search out of location	0.063	(18)	Σ: variance p distribution	0.297
(8)	z_1: search distance	-0.469	(19)	ν: search intensity of unemployed	5.926
(9)	$z_{l,1}$: search in home location	0.105	(20)	ϵ: curvature search cost	5.841
(10)	$z_{l,2}$: search to home location	2.146	(21)	ι: workers' outside option	0.982
(11)	z_r: search to home region	0.055			

Notes: The table reports the 21 parameters estimated from our model, estimated according to the procedure described above.

I Further Details on Model Fit

This section presents additional figures and tables to describe the model fit with the data. While all the moments are included in this section, for completeness we present in Supplemental Appendix Q the numerical values of each one of the 305 moments in the model and data. All these moments are included already in this section, but in figures rather than tables.

Figure A6 shows that the model fits the empirical moments well in several dimensions. Each panel plots a set of moments in the data (x-axis) against their values in the model (y-axis), with the 45-degree line indicating a perfect fit. In each of the top three panels, moments relating to West German workers are in blue and moments for East German workers are in red. The top left panel presents the share of employed workers of each type in a given location. As highlighted in the text, we use each worker's residence to determine her location since our definition of a cross-location move is based on the worker's residence. The empirical values of these moments were computed in Section G.2.3. The panel shows that in our model, as in the data, most workers are in their home location (circles). Moreover, East-born workers are more likely to be in the West than West-born workers in the East (stars), consistent with the fact that the West has higher productivity and a larger ratio of firms to workers. The top middle panel shows the share of unemployed workers in each region, which is similar to the distribution of employed workers. The top right panel presents the average firm component of wages paid to workers of a given type in each region (as computed in Section G.2.5). Consistent with the data, workers in the East earn lower wages than workers in the West. Furthermore, the relative wage gaps differ by workers' location, in a similar way in the model and in the data. In particular, the average wage gap between East and West German workers is smaller for the group of workers that are away from their home region (red versus blue stars) than for the group of workers that are in their home location (red versus blue circles).

The bottom three panels of Figure A6 present the average firm component of wages (the empirical moments were computed in Section G.2.6), GDP per worker (Section G.2.7), and the unemployment rate (Section G.2.8) in each of the four locations. The model matches the data well, generating lower wages, lower GDP per worker, and higher unemployment in the East.

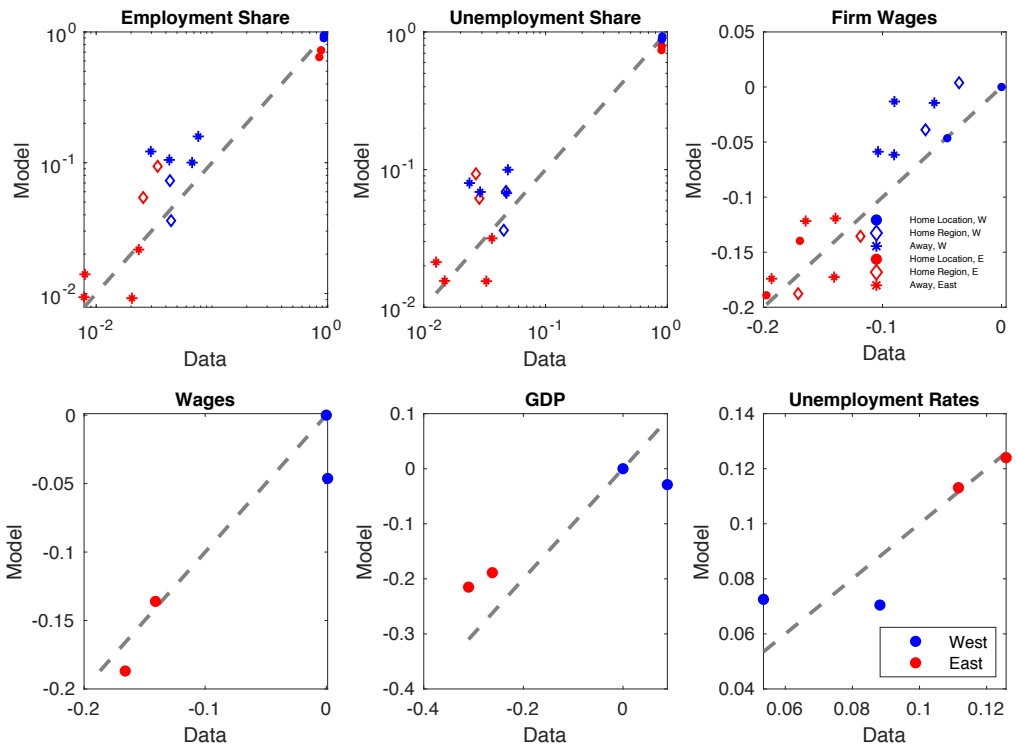
Figure A7 plots the firm size distributions in each location in the model and in the data, computed as described in Section G.2.9. The model matches almost perfectly the share of employment in the middle of the size distribution, and only slightly underestimates the mass of employment at the bottom and top deciles. In each location, approximately half of the overall employment is accounted for by the largest decile of firms.

Table A29 shows that the model also does a reasonable job in matching the joint distributions of firm wages, sizes, and separation rates, the standard deviation of wage gains, and the profit shares (the empirical moments were computed in Sections G.2.10 to G.2.14). The core mechanism of the model generates a positive relationship between firm size and the firm wage

(row 1 of Table A29), since higher productivity firms offer higher wages to increase their size. As a result, workers climb a job ladder across firms and are more likely to separate at the bottom rungs (row 2), also facing, on average, larger wage gains when separating from firms at the bottom (row 3). These core features of the model are consistent with the data. We further explore these relationships in Figure A8, where we plot these variables in the model and in the data nonparametrically, for each of the four locations. The top panels show the relationship between firm log size and log average wage in each location, the middle panels present the expected wage gains as a function of a worker's current firm's log average wage, and the bottom panels show the relationship between the separation rate and a worker's current firm's log average wage. In both the model and data, these relationships are roughly linear.

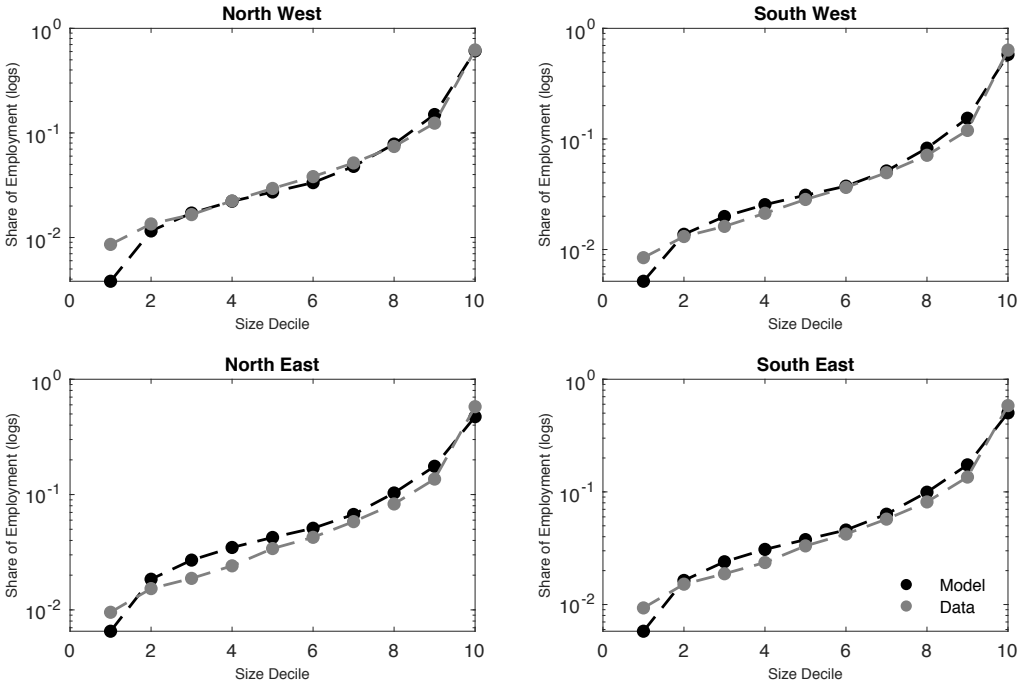
As noted in the main text, the model overestimates the relationship between job movers' expected wage gains and their current firm's average wage. Moreover, the model underestimates the standard deviation of wage gains of movers (row 4 of Table A29). This result is somewhat expected since in the model wage dispersion across firms is purely generated by labor market frictions, while in the data there may be other sources of wage dispersion that our empirical controls are not capturing. For further analysis, Figure A9 plots the distribution of the standard deviation of wage gains in the model and data for all 64 origin-destination-home location tuples. The standard deviations in the data are higher than in the model for nearly all combinations of moves. For comparison, we also plot in the figure an alternative empirical moment: the standard deviation of wage gains controlling for individual fixed effects (light gray). As expected, controlling for individual fixed effects reduces significantly the empirical variance (some individuals have persistently higher wage gains than others, as shown in the literature). Relative to this alternative target, our model slightly overestimates the standard deviation of wage gains.

Figure A6: Employment, Wages, and GDP by Location and Worker-Type



Notes: The figure graphs the value of various moments in the model against the same moments in the data. The construction of these moments is described in Sections G.2.3 to G.2.8. Each dot corresponds to one moment. The top left panel shows the share of employed workers residing in each location, by worker type. The top middle panel shows the share of unemployed workers residing in each location, again by worker type. The top right panel shows the average log firm component of wages for each worker type residing in each location, normalized relative to workers whose home location is North-West and that are currently residing in the North-West. In each panel, moments relating to West German workers are in blue and moments for East German workers are in red. Circles are for workers currently residing in their home location, squares for workers residing in their home region but not location, and stars are for workers currently out of their home region. The bottom left panel shows the average log firm component of wages by location, relative to the North-West. The bottom middle panel shows the GDP per capita of each location relative to the North West. Last, the bottom right panel shows the unemployment rates. In each of these panels, West locations are in blue and East locations are in red.

Figure A7: Within-Location Firm-Size Distributions



Notes: The figure compares the firm size distribution in the model and in the data. Each panel graphs the share of total employment that is working at each decile of the firm size distribution for each of the four locations. Model moments are in black and data moments are in gray. The construction of these moments is described in Section G.2.9.

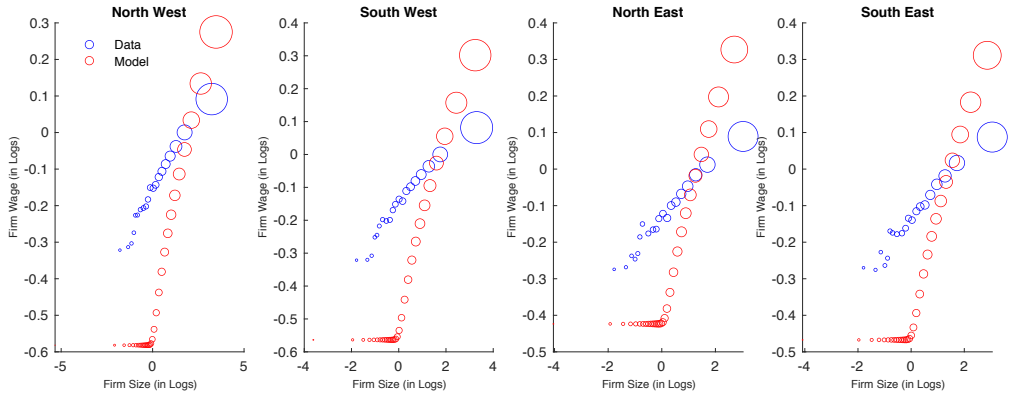
Table A29: Model Fit for Additional Moments

	Parameters		Model		Data	
			West	*East*	*West*	*East*
(1)	Slopes wage vs firm's size, by j	*North*	0.126	0.135	0.124	0.110
		South	0.161	0.140	0.124	0.109
(2)	Slopes separation vs firm's wage, by j	*North*	-0.024	-0.019	-0.029	-0.037
		South	-0.024	-0.020	-0.033	-0.036
(3)	Slopes wage gain vs firm's wage, by j	*North*	-0.805	-0.889	-0.549	-0.561
		South	-0.827	-0.870	-0.577	-0.562
(4)	Average Std of job-job wage gains, by j	*North*	0.392	0.377	0.591	0.584
		South	0.399	0.378	0.609	0.539
(5)	Profit shares, by j	*North*	0.285	0.360	0.274	0.259
		South	0.303	0.342	0.259	0.263

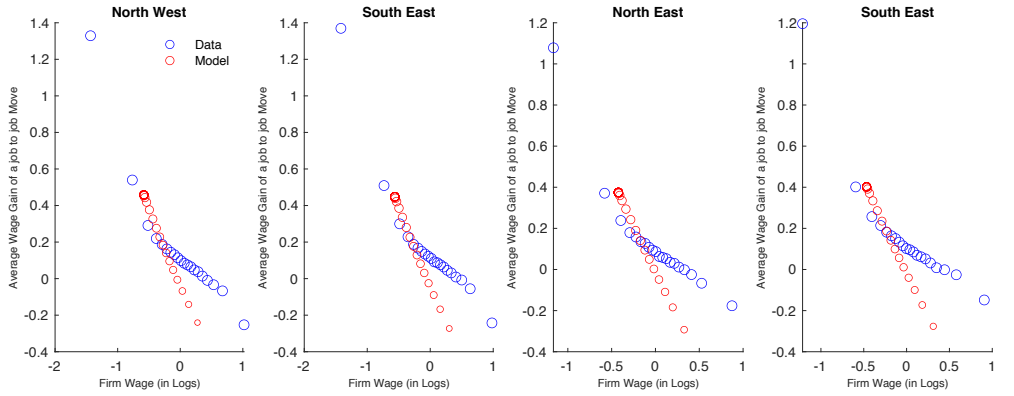
Notes: The table compares several moments in the model to their data analogues, by location of the firm. The construction of these moments is described in Sections G.2.10 to G.2.14. The first row shows the slope of the wage function with respect to firm size. The second row presents the slope of the separation rate with respect to firms' wage. The third row shows the slope of the average wage gain from a job-to-job move as a function of the origin firm's wage. The fourth row presents the standard deviation of wage gains from a job-to-job move by location of the origin firm. We take the average across all the 16 possible job-to-job moves that originated in each region. All the 64 disaggregated moments are included in Supplemental Appendix Q. The last row shows the average ratio of profits to labor costs in each location.

Figure A8: Model Fit for Joint Distribution of Firm Wages, Sizes, and Separation Rates

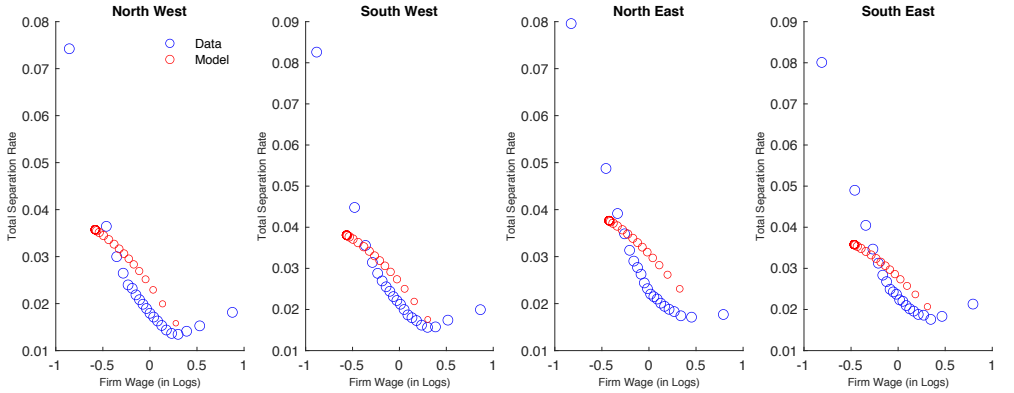
(a) Relationships between Firm Sizes and Average Wages



(b) Relationships between Firm Wages and Expected Wage Gains of Job-to-Job Moves

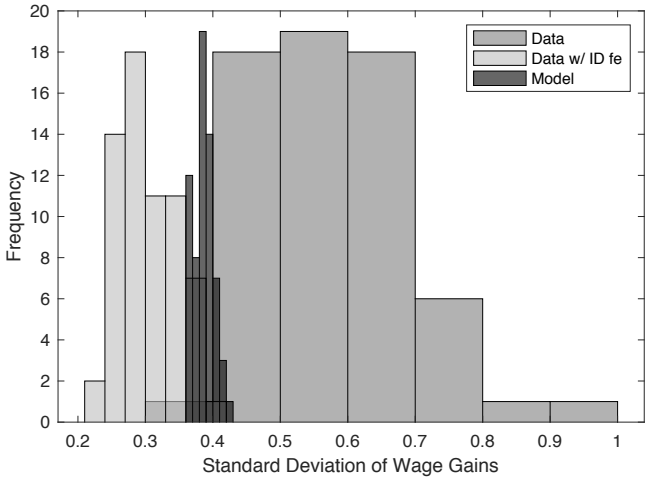


(c) Relationships between Firm Wages and Separation Rates



Notes: The figure compares various moments in the model (red) and in the data (blue), for each location. The empirical moments are computed as described in Sections G.2.10 to G.2.12. In both the data and the model, we cut the firm distribution into twentiles based on the variable on the x-axis and then compute the summary statistic within each twentile. The size of each circle represents the number of observations. Wages and sizes are normalized relative to their average in both model and data without loss of generality since they are not targeted. The top panels show the relationship between firms' average wage and their size (number of workers). The middle panels show the relationship between the average wage gain of a job-to-job move, across all possible moves, and the average wage of the worker's firm prior to the move. The bottom panels show the relationship between the rate at which workers separate, either towards a new firm, unemployment, or permanent non-employment, and the average wage of the firm prior to the move.

Figure A9: Standard Deviation of Wage Gains



Notes: The figure shows the distribution of the standard deviation of wage gains for all the 64 possible tuples of origin-destination-home location (j, x, i). The empirical moments are computed in Section G.2.13. The histogram counts the frequency with which a standard deviation of wage gains of the given value is observed. The count in the model is depicted by the black bars and the count in the data in dark gray. The light gray bars present an alternative empirical specification where in addition to the controls in Section G.2.13 we include individual fixed effects in the regression that residualizes the wage gains. The width of the bars is chosen so that each alternative has the same number of bars. It varies across alternatives dependent on how dispersed the standard deviations are. The height of the bars is comparable across alternatives and indicates the number of observations falling into the given range of standard deviations.

J Additional Quantitative Results

We provide additional results complementing the quantitative analysis of Section 6.

Table A30 includes similar statistics as Table 5, but we report the level of each variable separately for East and West Germany. Most of the results have already been described in Section 6. The table shows that most of the reduction of the West-East gap is driven by an increase in East German wages and GDP rather than by a decrease in the West German analogues.

Figure A10 shows the distributions of labor across wage levels per efficiency unit. More specifically, it shows semi-CDFs, since it depicts the CDFs of wages normalized so that the point furthest to the right is not at one but at the total mass of employed workers of the specific type in a given region. The top panel shows the distribution of labor in Germany overall, the middle panel presents the distribution in West Germany, and the bottom panel shows the distribution in East Germany. In each panel, we distinguish between West-born workers (blue), East-born workers (red), and all employed workers (black/gray). Similarly to the figures in the main text, we compare the benchmark economy to the four counterfactuals as we vary the level of spatial frictions. Figure A11 is analogous to Figure A10, but studies the distribution of labor across firm productivity levels. The figures clarify that, in the benchmark and in both the preferences and moving costs counterfactuals, workers are mostly working in their home region. Instead, once we shut down differences in search efficiency we see that workers are significantly more likely to work outside of their home region, consistent with the results shown in Table 6.

Finally, Figure A12 shows simulated employment histories, as described in the main text and shown in Figure 11, for two additional statistics. The top panel shows the share of employment in West Germany. We start workers in period zero from unemployment in their home region and then follow them over time. In the benchmark economy, the share of workers away from home slowly increases over time, but most workers remain within their home region. Removing the home preferences or moving costs, weakens the sorting across areas, but overall the share of workers in their home region remains high. Eliminating the search efficiency frictions, on the other hand, has a large effect. Many workers immediately take a job in the other region, with a large share of East-born workers ending up employed in the West, and vice versa. The bottom panel shows the monthly probability of a job-to-job move across regions – i.e. from the East to the West of Germany or vice versa. Reducing the spatial frictions significantly increases the frequency of cross-region moves throughout an individual's life-cycle. In particular, shutting down differences in search efficiency across regions increases by more than an order of magnitude the cross-regional mobility of workers. This result formalizes the idea that removing spatial frictions allows workers to climb a country-wide job ladder.

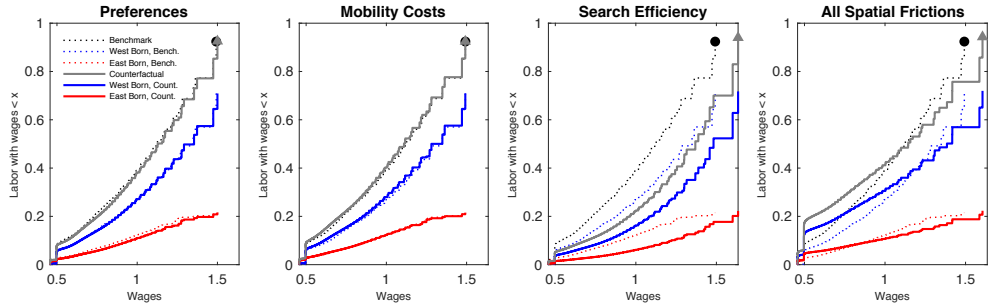
Table A30: Aggregate Effects of Spatial Frictions

		West Germany					East Germany				
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Base	τ	κ	z	*All*	*Base*	τ	κ	z	*All*
(1)	Wage per efficiency unit	1	1.008	0.993	1.137	0.972	0.869	0.898	0.856	1.021	0.825
(2)	% of West-born Labor	93.0%	88.1%	90.2%	80.1%	79.0%	10.7%	29.9%	24.4%	63.8%	68.6%
(3)	Efficiency units pc	1	0.996	0.997	0.988	0.987	0.925	0.942	0.938	0.973	0.978
(4)	Average wage paid	1	1.004	0.990	1.125	0.956	0.804	0.848	0.804	1.008	0.820
(5)	GDP per capita	1	1	0.996	1.058	0.997	0.830	0.861	0.839	0.965	0.903
(6)	Average value	1	1.010	0.999	1.110	1.167	0.880	0.927	0.890	1.074	1.141

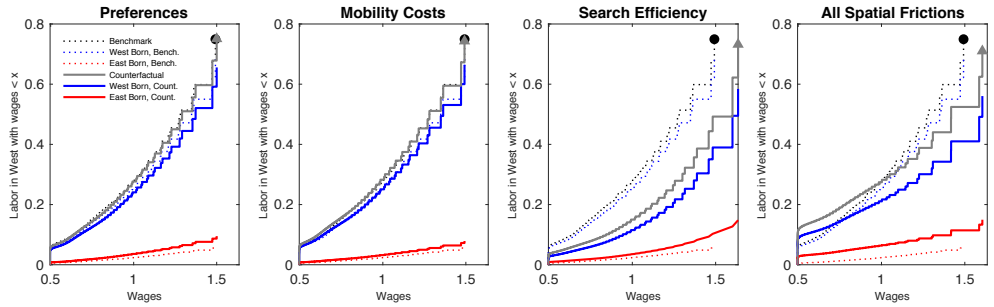
Notes: The table shows the aggregate effects of spatial frictions by comparing the benchmark (Base) with four hypothetical economies: i) No home preference (τ); ii) No moving costs (κ); iii) No differences in search efficiency across regions (z); iv) No spatial frictions (All). Columns 1-5 present statistics for West Germany. Columns 6-10 present statistics for East Germany. Statistics in rows 1, 3, 4, 5, and 6 are normalized relative to the benchmark in West Germany. Row 1 presents the average wage per efficiency unit, $w_j(p) P_j^{-1}$, averaged across all employed workers in j. Row 3 shows efficiency units per capita, the average of θ_j^i across all workers in j. Row 4 displays the average wage paid, $w_j(p) P_j^{-1} \theta_j^i$. Row 5 presents the average output per capita, $p \sum_{i \in \mathbb{I}} \theta_j^i l_j^i(w)$. Row 6 shows the average value, obtained by averaging across U_j^i and $W_j^i(w)$ using the distribution of labor across firms, regions, and employment status.

Figure A10: Employment and Wage Distributions

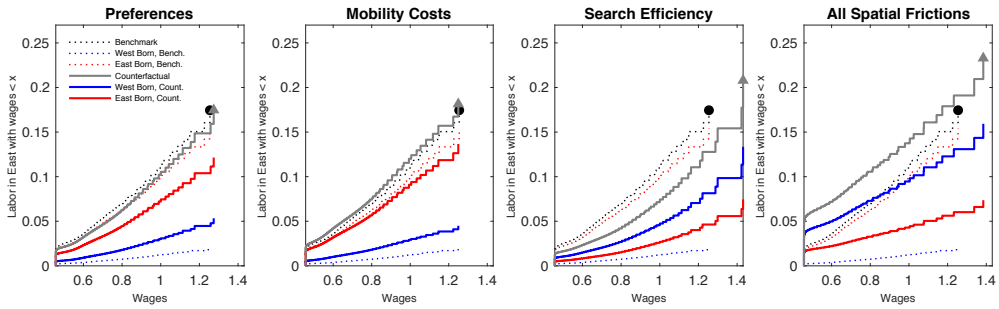
(a) Germany



(b) West Germany



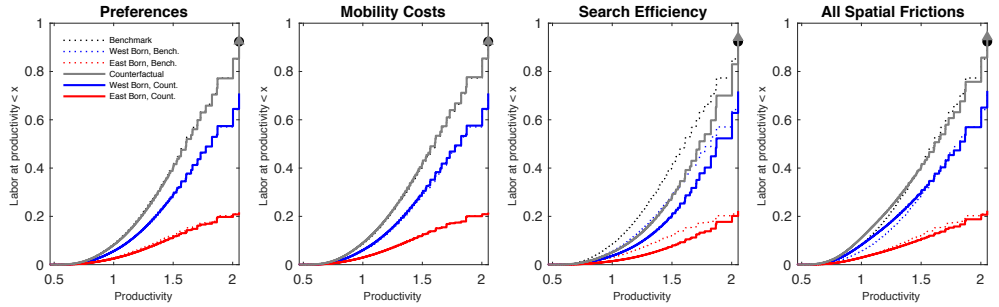
(c) East Germany



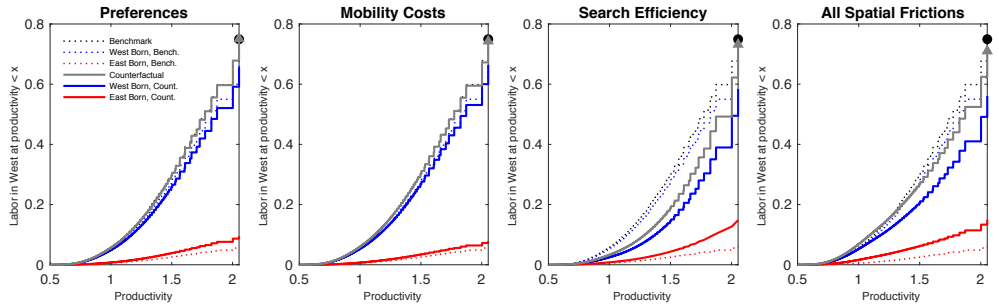
Notes: The figure reports the CDFs of real wages per efficiency unit by firms' region and workers' birth region. The CDFs are scaled such that their point furthest to the right is not one but instead the mass of workers of the given type in the given region. As usual, East-born workers are in red and West-born workers in blue. In each panel, we compare the benchmark economy (dotted lines) with one counterfactual economy (solid line). We consider four counterfactuals, left to right: i) no home preferences; ii) no moving costs; iii) equal search efficiency towards each region; iv) no spatial frictions.

Figure A11: Employment and Productivity Distributions

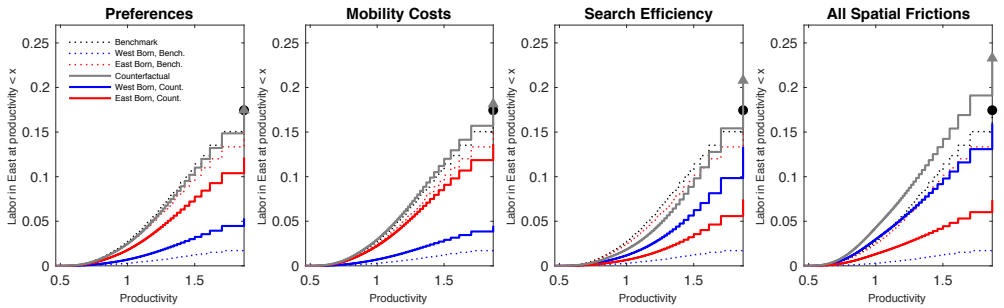
(a) Germany



(b) West Germany



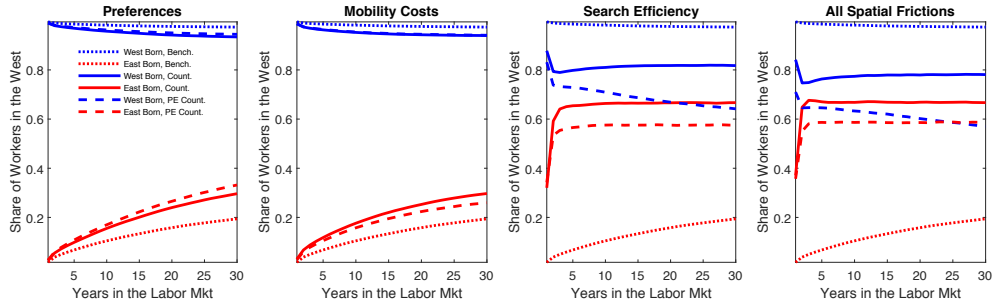
(c) East Germany



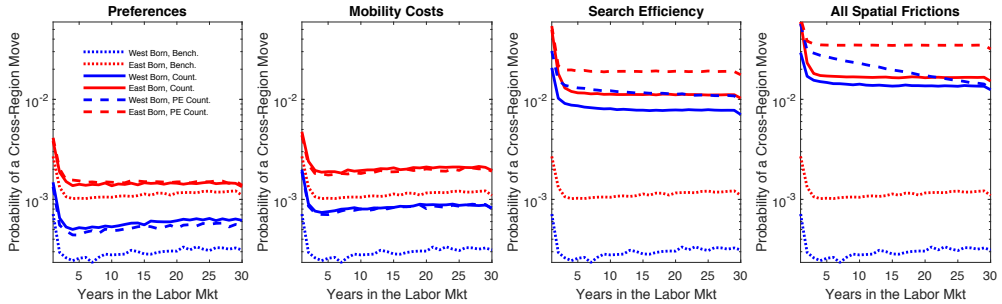
Notes: The figure reports the CDFs of workers by firm productivity by firms' region and workers' birth region. The CDFs are scaled such that their point furthest to the right is not one but instead the mass of workers of the given type in the given region. As usual, East-born workers are in red and West-born workers in blue. In each panel, we compare the benchmark economy (dotted lines) with one counterfactual economy (solid line). We consider four counterfactuals, left to right: i) no home preferences; ii) no moving costs; iii) equal search efficiency towards each region; iv) no spatial frictions.

Figure A12: Simulated Employment Histories: Employment Region and Cross-Regional Moves

(a) Share of Employment in West Germany



(b) Frequency of Cross-Area Moves



Notes: The figures show results from 100,000 simulated job histories, each one starting with individuals unemployed in their home region. Red is East and blue is West. The top panel shows the share of workers in the West. The bottom panel shows the share of workers that make, in that month, a cross-regional move. Dotted lines are the benchmark economy. Solid lines represent each one of our four counterfactuals. The dashed lines are partial equilibrium counterfactuals, for which we keep constant the distribution of wage offers, prices, vacancies, and aggregate applications, but we solve for individual decision rules based on the relevant set of spatial frictions.