Labor Misallocation Across Firms and Regions[*,†]

Sebastian Heise
Federal Reserve Bank of New York

Tommaso Porzio
Columbia University, NBER, and CEPR

March 27, 2023

Abstract

We develop a frictional labor market model with multiple regions and heterogeneous firms to study how frictions impeding labor mobility across space affect the joint allocation of labor across firms and regions. Bringing the model to matched employer-employee data from Germany, we find that spatial frictions generate large misallocation of labor across firms *within* regions. By shielding firms from competition for workers from other regions, spatial frictions allow low productivity firms to expand, reducing aggregate productivity. Overall, we show that taking into account the characteristics of the local labor market is important to quantify the aggregate losses from spatial frictions.

JEL: J6, O1, R1

[*]We thank Michael Peters for a very insightful discussion of the paper at NBER Small Growth Group. We also thank Ufuk Akcigit, Andy Atkeson, Gharad Bryan, Paco Buera, Julieta Caunedo, Lorenzo Caliendo, Kevin Donovan, Niklas Engbom, Ben Faber, Pablo Fajgelbaum, Tarek Hassan, Gregor Jarosch, Kyle Herkenhoff, Fatih Karahan, Pete Klenow, David Lagakos, Rasmus Lentz, Paolo Martellini, Mushfiq Mobarak, Ben Moll, Simon Mongey, 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.

[†]Heise: 33 Liberty Street, New York, NY 10045, email: sebastian.heise@ny.frb.org. Porzio: 665W 130th St, New York, NY 10027, email: tommaso.porzio@columbia.edu.

1 Introduction

In many countries, there are large differences in productivity and real wages across regions.[1] This fact suggests the presence of frictions, such as moving costs, that prevent worker migration towards more productive regions. Labor is thus misallocated across space, and removing barriers to labor mobility could generate large aggregate gains (Gollin, Lagakos, and Waugh (2014), Herrendorf and Schoellman (2018)).

Focusing on spatial gaps in isolation, however, misses that there are large productivity differences across firms even *within* narrow geographic markets (Lentz and Mortensen (2012), Lindenlaub, Oh, and Peters (2022)). In this paper, we argue that taking into account this within-region heterogeneity, and more broadly the local labor market frictions, is important. It changes the quantitative estimates of the costs of spatial frictions and alters the mechanisms through which they operate. Overall, we show the importance of studying labor misallocation across space and firms jointly.

The intuition behind our argument is straightforward. Given the within-region heterogeneity, the impact of reallocating workers across regions depends on the firms these workers end up at. If there are large labor market frictions limiting workers' ability to climb the local job ladder, migrants might get stuck at unproductive firms even if they move to high productivity regions. At the same time, due to the within-region heterogeneity, worker migration is not necessary to close regional gaps. Local reallocation of workers to better firms might be sufficient. Spatial frictions may nonetheless play a key role, but through a different mechanism. They limit workers' job opportunities and thus shield firms from competition from other regions, allowing low productivity firms to survive. Removing spatial frictions could thus improve the allocation of labor even within region.

To formalize these arguments, and to quantify the extent to which local heterogeneity shapes the aggregate costs of spatial frictions, we develop a general equilibrium framework that embeds frictional labor markets as in Burdett and Mortensen (1998) within a multi-region economy. We estimate the model with matched employer-employee data from Germany, a country with large regional variation in wages and productivity. We find that the aggregate gains from reducing spatial frictions hinge on the extent to which they affect the allocation of labor across firms. In particu-

[1] Examples are the Italian South versus North or the East versus West of Germany.

lar, two economies could look identical in terms of their wage or productivity gap between regions, yet the aggregate gains from removing spatial frictions could vary dramatically between the two dependent on their local labor market frictions.

In the first part of the paper, we use micro data from the German Federal Employment Agency to document three sets of facts, which motivate our focus on the joint allocation of labor across firms and regions and guide the ingredients of our model.

First, we use the Establishment History Panel (BHP), a 50% sample of all establishments in Germany, to document that there is a large wage gap between East and West Germany, but also substantial heterogeneity across firms *within* these regions. In principle, it would be possible to close the East-West wage gap by just reallocating labor within East Germany towards higher-wage firms.

Second, we use the Linked Employer-Employee Data (LIAB) to show that East Germans get very large wage increases when moving West, suggesting substantial gains from regional integration, but that workers also experience sizable wage gains for job-to-job moves within-region. Thus, frictions hindering within-region mobility could be as costly as those limiting migration towards high productivity regions.

Third, we show that workers switch jobs mostly locally and exhibit home bias (i.e., workers have a preference for their home region), leading to a job ladder characterized by frequent return migration of workers that have left their home region.

Motivated by these facts, in the second part of the paper we develop a framework to study the joint allocation of labor across firms and regions. We design a wage-posting model with heterogeneous firms, multiple regions, worker heterogeneity, and a large set of spatial frictions often considered in the migration literature: moving costs, home bias, spatial search costs, and region-specific comparative advantages. 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 move into and out of unemployment and across firms both within and between regions. A constant returns to scale matching function transforms applications and vacancies into worker-firm meetings. Search is directed across regions, but random within region, which is important for identification of the spatial frictions.

Our model allows us to identify the different 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.

First, the unobservable distribution of job offers in each region is disciplined by 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.

Second, given within-region offer distributions, the spatial frictions are identified by comparing the wage gains and job flows across regions to their within-region analogues. 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 wage gains for movers out of their home region relative to other worker types making the same move identify home preferences. In contrast, spatial search costs are disciplined by the relative frequency of job switches. Lower worker flows across regions, compared to between firms within region, indicate that workers are less able to apply for jobs in other regions.

We estimate the model with four sub-regions of Germany corresponding to the Northwest, Southwest, Northeast, and Southeast, which we refer to as *locations* to distinguish them from the regions of East and West Germany. We incorporate four worker types reflecting the four possible home locations. The model matches the data well, despite using only 21 parameters to match 305 micro and aggregate moments.

Our estimates imply large spatial barriers, mainly due to the limited ability of workers to access job opportunities that are further away, consistent with evidence that labor markets are primarily local (e.g., Manning and Petrongolo (2017)). For a given search effort, workers generate 1/20th as many job applications when searching for jobs across locations as within. We estimate a cost of moving between any two locations of 3.1%-5.3% of lifetime income (dependent on the distance of the move), and find that workers need to be paid 7.4% of their yearly income to work away from their home location and maintain the same utility.

We then turn to the main exercise of our paper and use the estimated model to quantify the aggregate and distributional costs of spatial frictions in general equilibrium. Removing all spatial frictions, including workers' home bias, would raise GDP per capita in Germany by almost 5%, and average real wages by 9%. Importantly, these gains are due to improvements in the allocation of labor *within* each location, rather than due to net migration from low to high productivity areas, as in spatial models without a frictional labor market. When spatial frictions are removed, firms are exposed to more competition from other locations, which forces unproductive firms to shrink or to exit the market and reallocates labor towards high productivity firms in each location. Additionally, workers also gain due to better job opportunities as they climb an integrated Germany-wide job ladder.

Our model also sheds light on the distributional effects of spatial frictions. When spatial frictions are removed, East Germany's output per capita rises by 17%, while output in the West increases by only 4%. Similarly, East Germans see their wage rise by almost 20%, while West Germans gain only about 7%. Both the reallocation of labor within and across locations are important for these effects. Removing spatial frictions leads to a greater labor reallocation within East Germany than in the West because the East has more unproductive firms, which are more affected by spatial frictions. Labor reallocation across locations allows East Germans to benefit from the higher wages paid in the West, and increases the average skill level of the East German labor force due to the in-migration of West Germans.

Our results remain qualitatively unchanged when we eliminate only the spatial frictions generated by technological parameters (the moving cost and the spatial search frictions), while keeping workers' preference for their home region. However, we find strong complementarities between these types of frictions: removing technological frictions and home preferences separately generates only half of the gains from removing both sources of frictions at the same time.

In the final part of the paper, we demonstrate that the gains from removing spatial frictions decline sharply as the labor mobility within each location increases. The reason is intuitive: with more within-location mobility, labor is relatively concentrated at the most productive firms, hence the marginal gains from removing spatial frictions are limited. Importantly, we show that the average wage gap between two locations does not depend in general on the level of labor market frictions. Consequently, two economies could look *identical* in terms of their wage or productivity gap between locations, yet removing spatial frictions could lead to vastly different aggregate gains dependent on the economies' local labor market frictions.

Literature. We build on a large body of work that has studied the impact of factor misallocation on aggregate productivity (e.g., Hsieh and Klenow (2009)). In particular, we add to the growing macro literature on the role of labor market frictions in misallocating labor (Lentz and Mortensen (2012); Bilal et al. (2019); Bilal (2021); Engbom (2020); Elsby and Gottfries (2022); Martellini (2022)). Our contribution is to study jointly the allocation of labor across firms and space, and to quantify how spatial frictions shape competition in the local labor market. Our analysis is motivated by recent work showing that workers' job search is mostly local.[2]

Our paper also builds on the quantitative spatial literature that has developed general equilibrium frameworks to study the aggregate and distributional impacts of barriers to the mobility of labor across space, industries, and occupations (e.g., Caliendo et al. (2019); Bryan and Morten (2019); Hsieh et al. (2019)). Our contribution to this literature is to focus on a different margin of misallocation *(across firms of different productivities)* and to show how we can quantify it using a model with labor market frictions and matched employer-employee data.[3] Typically, the quantitative spatial literature allows for rich spatial heterogeneity and individual comparative advantages towards regions and/or occupations. Barriers to labor mobility may then lead to worker-firm mismatch and misallocation of talent. This channel has a very limited role in our framework, in which misallocation is instead driven by misallocation of inputs, closer to the wedge approach of Hsieh and Klenow (2009) but generated endogenously by the interaction between labor and spatial frictions.[4]

Methodologically, we extend a wage posting model à la Burdett and Mortensen (1998) to incorporate a spatial structure. Our framework is related to job ladder models with labor mobility across sectors, such as Meghir et al. (2015), Hoffmann and Shi (2016), and Bradley et al. (2017).[5] A limitation of these models for our context, however, is that they do not consider switching costs between sectors, and therefore two workers with the same current value of employment accept the same job offers regardless

[2]For example, Manning and Petrongolo (2017) and Le Barbanchon et al. (2020).

[3]These type of data have, to the best of our knowledge, not yet been used by this literature. One way to see our contribution is that we bring to the quantitative literature on spatial frictions insights from the large labor literature that has estimated models with on-the-job search in matched employer-employee data, e.g., Lise et al. (2016); Bagger and Lentz (2019); Bonhomme et al. (2019).

[4]Krueger and Pischke (1995), Fuchs-Schündeln et al. (2010), Uhlig (2006) and Uhlig (2008), Dauth et al. (2021), Boeri et al. (2021), and Lindenlaub, Oh, and Peters (2022) focus on understanding the large wage and productivity gaps in the specific German context. We do not seek to provide a comprehensive explanation of the East-West gap in Germany. In particular, we estimate the productivity gap between East and West Germany in our model and take it as given.

[5]A large literature has estimated versions of Burdett and Mortensen (1998) (e.g Van Den Berg and Ridder (1998); Burdett and Coles (2003); Burdett et al. (2020); Moser and Engbom (2021)).

of their current sector. In our setup, instead, workers' acceptance decisions not only depend on their current value but also on their current location. To apply the existing frameworks to our context, we would need to assume that there is no cost of moving between locations. Our framework is suitable to situations in which workers' current sector or location is a state variable for employment decisions.

At a conceptual level, we contribute to the fast-growing literature on local monopsony power (e.g., Berger et al. (2022)), and in particular to work that links labor market power to spatial frictions such as commuting costs (e.g., Caldwell and Danieli (2021), Datta (2022)). Relative to this work, our paper analyzes how changes to spatial frictions affect monopsony power and endogenously reallocate workers within local labor markets.[6] The reallocation of workers towards higher productivity firms and the exit of unproductive ones in our model is similar in spirit to the within-industry reallocation in international trade when trade barriers are removed (Pavcnik (2002), Melitz (2003)). However, reallocation in our framework comes from competition for workers in the labor market, rather than for customers in the output market.

Finally, there is a large literature that studies the size of spatial frictions and the gains from migration either in partial equilibrium (e.g., Kennan and Walker (2011); Baum-Snow and Pavan (2012)) or by estimating reduced-form specifications in panel data.[7] Closest to our work, Schmutz and Sidibé (2018) build a framework in which workers receive job offers both from their current location and from other locations. Relative to these papers, we build a general equilibrium framework that provides a structural interpretation to the reduced form evidence and that can be used to study the aggregate impact of spatial frictions. We show that removing spatial frictions can entail large equilibrium effects.

Road Map. We proceed as follows. Section 2 describes the data, and Section 3 documents facts on the German labor market. Section 4 introduces the model, which we estimate in Section 5. We quantify the aggregate and distributional effects of spatial frictions in Section 6. Section 7 concludes.

[6] Similar to us, Galenianos et al. (2011) and Bachmann et al. (2021) emphasize how firms' monopsony power reduces employment at highly productive firms; however, these papers do not analyze the role played by spatial frictions.

[7] See Combes et al. (2008); Roca and Puga (2017); Hicks et al. (2017), Lagakos et al. (2020).

2 Data

We use two datasets provided by the German Federal Employment Agency (BA): 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 social security filings and exclude government employees and the self-employed. Each establishment in the BHP is a company's unit operating in a distinct county and industry.[8] For simplicity, we will refer to these units as "firms". For each firm-year pair, the dataset contains information on location, average wages, number of employees, and employee characteristics (education, age, gender).

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. These data are linked to information about the 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.[9] 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.

An important variable for our analysis is each worker's county of residence, reported in the LIAB since 1999, which together with the workplace will be used 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 start of an unemployment spell for unemployed workers and then added to all observations of that year or spell. Workers self-report their residence, and can choose which residence to report if they have multiple homes, leading some workers to report very large distances between residence and work location

[8] 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 plants (Hethey-Maier and Schmieder (2013)).

[9] We use "unemployment spell" for the period in which an individual receives unemployment benefits. After benefit expiration, individuals are not in our data until they are employed again.

even though they live in a second home closer to work. To deal with the potential measurement error, we will define several alternative measures of migration below.

We use four additional datasets. First, we obtain information on county-level cost of living from the Federal Institute for Building, Urban Affairs, and Spatial Development (BBSR (2009)), which we use to construct real wages.[10] Second, we use annual data from the German Socio-Economic Panel (SOEP) to corroborate some of our main findings. Third, we use data from the NY Fed's Survey of Consumer Expectations (SCE) to provide support for our model mechanisms. Finally, we use information on firms' profits from the ORBIS database for the model's estimation.

Sample Construction. We refer to the period 2009-2014 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 et al. (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 et al. (2013). For the LIAB, no such variable is provided and we replicate the imputation steps ourselves. We use the corrected, real 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 data in Appendix A.

3 Motivating Facts

We document three sets of facts: (i) there is substantial wage heterogeneity both across regions and across firms within these regions; (ii) workers obtain large wage gains when they change jobs, both across and within regions; (iii) workers' job flows are biased towards their home region and towards geographically closer jobs.

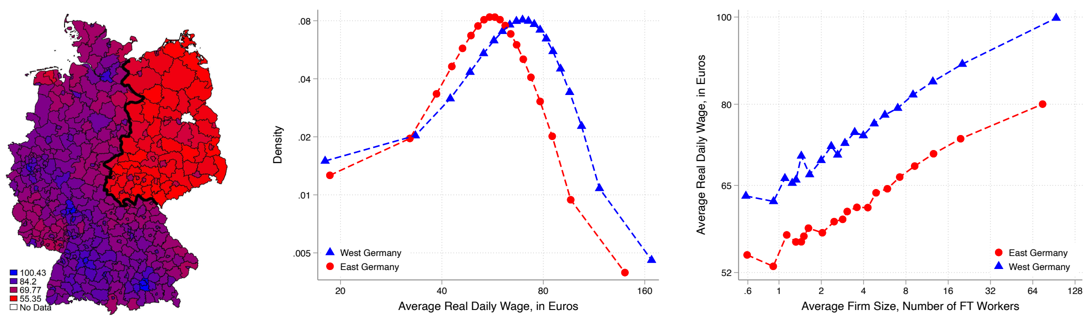
3.1 Wage Heterogeneity Between and Within Regions

Figure 1a shows that there is a large difference in the average real wage paid between counties in East and in West Germany in our baseline period. To examine whether

[10] The data cover about two thirds of the consumption basket, including housing rents. We provide further information in Appendix A. East Germany has a 7% lower average price level.

Figure 1: Real Wages Between and Within Regions

(a) Between Regions (b) Within-Region (c) Wage-Size Ladders



Source: BHP and authors' calculations. Notes: The left figure shows real daily wages in each county in 2009-2014, 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". The middle panel plots the density of the average wage of firms in each twentile of the wage distribution. Wages are residualized by regressing the log real wage on 3-digit industry dummies and time dummies, for East and West Germany separately. The cleaned wage is the residuals from this regression plus the mean of the log wage in each region, transformed back into levels. While all firms are weighted equally, only a very small share of overall employment is at the lowest wage firms. The right panel plots the average number of full-time workers for each twentile of the firm size distribution against the average real daily wage of firms in the twentile, where wages and size are residualized as before.

this wage gap is due to observables, we run in the BHP firm-level regressions

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

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 find an East-West wage gap of $\gamma = -.2609$ (s.e. .0074) without controls. Controlling for worker gender, education, and age, firm size, and industry lowers the real wage gap to $\gamma = -.2052$ (s.e. .0027); about 80% of the gap remains unexplained.

While the wage gap between East and West Germany is striking, we next show that there is even larger wage heterogeneity between firms within each region. Figure 1b plots the real wage distribution in each region (residualized by industry) and shows that the wage gap between the lowest- and highest-paying firms in each region exceeds the average wage gap between East and West.[11]

Figure 1c further plots the average firm size against the firms' average real wage

[11] In Supplemental Appendix L, available on the authors' websites, we show that there is similarly large wage dispersion across firms even *within the same county*. Hence the large dispersion is not just reflecting cross-county differences, consistent with the limited dispersion shown in Figure 1a.

for twentiles of the firm size distribution. Average real wages increase significantly with firm size in both regions, suggesting a job ladder. Additionally, East German firms pay a lower real wage than West German ones for each firm size, suggesting the presence of frictions that shield East German firms from West German competition and allow them to reach a larger size at the same wage level.

In Supplemental Appendix L[12], we show details of regression (1) and provide additional empirical results: (i) the between-region wage gap is persistent over time and similar for all industries; (ii) there are limited differences in observables between East and West Germans; (iii) there are no clear regional differences in tax rates.

3.2 Large Wage Gains of Movers Across and Within Regions

We show that workers obtain large wage gains when they change jobs, both across and within regions.

We analyze workers' wage dynamics around the time of a job-to-job move by running a standard system of local projections, consisting of one regression for each time period $\tau \in \{t-3, ..., t-1, t+1, ..., t+5\}$ around t:[13]

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

where $w_{i\tau}$ is an individual's weighted average wage across all employment spells in year τ, and we use each spell's length as its weight. We define a job-to-job move as a job switch between two firms without intermittent unemployment spell. The variable $\Delta \log(w_{i\tau})$ is the log change of the 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 payments.[14]

The variable d_{it}^{s} is 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. We distinguish between migration and commuting for moves between East and West

[12]This Supplemental Appendix is not meant for publication and includes additional material to provide context or robustness checks. It is available on the authors' websites.

[13]We pool together all data for time t from 2004 to 2014, creating an unbalanced panel. Working with an unbalanced panel could be problematic. In our case, we are less concerned because: i) we do not observe post-trends; and ii) we are focused on the wage growth on impact.

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

Germany 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, and define all other moves between East and West as commuting.[15]

The variable $\mathbb{I}_i^{East}$ is a dummy for whether an individual's birth region is East Germany. 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 more details. 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 show 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".[16]

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.

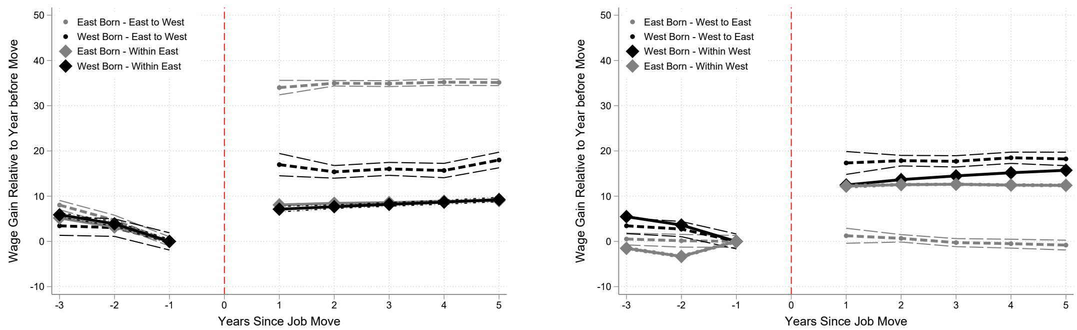
Figure 2a shows that East-born movers to the West receive on average almost a 35% real wage increase relative to their average within-firm wage growth, which is almost double the wage gain obtained by West-born workers making the same move. Figure 2b shows that moves to the East are associated with sizable wage gains for West-born

[15]We compare residence location across years since the variable is only updated at the end of each year. 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.

[16]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.

Figure 2: Wage Gains for Job-to-Job Moves

(a) East-to-West Migration vs. Within-East (b) West-to-East Migration vs. Within-West



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 $i \in$ {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 dashed lines in the left panel show the normalized coefficients for $\beta^{West}_{EW,\tau}$ and $\beta^{East}_{EW,\tau}$, and the solid lines with diamonds show $\beta^{East}_{EE,\tau}$ and $\beta^{West}_{EE,\tau}$. The dashed lines in the right panel show the normalized coefficients for $\beta^{West}_{WE,\tau}$ and $\beta^{East}_{WE,\tau}$, and the solid lines with diamonds show $\beta^{West}_{WW,\tau}$ and $\beta^{East}_{WW,\tau}$.

workers and almost no effect for East-born ones. The figures highlight that cross-region movers obtain significant wage increases, in particular those moving out of their home region, suggesting that workers need to be compensated to leave (home bias). Moreover, average wage gains for moves to the East tend to be smaller than for moves to the West, consistent with the lower average wage level in the East, and suggesting the presence of large gains from regional integration.

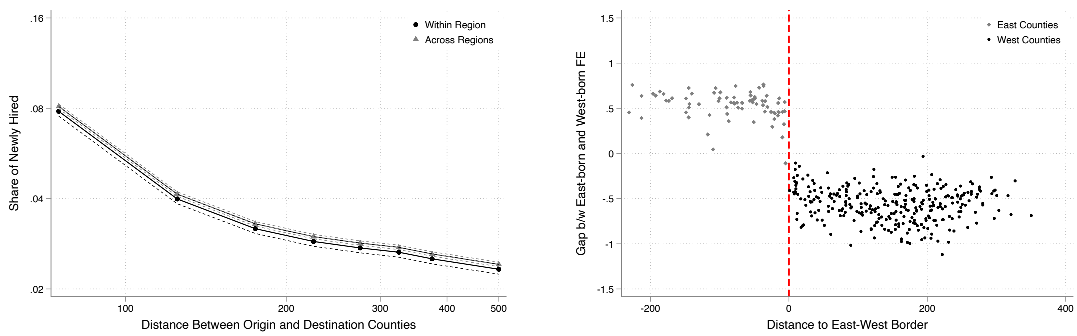
Workers obtain large wage gains not only from cross-region migration, but also from within-region job switches. Figures 2a and 2b show that workers experience wage gains of on average 10% from within-region job moves, consistent with the notion that they are climbing a job ladder in the presence of labor market frictions.

These observational returns from migration and job-to-job moves should not be interpreted as causal effects. Movers are selected: they are the ones that received sufficiently appealing job offers. Nonetheless, these large wage gains highlight the importance of labor mobility, both within and between regions, for aggregate productivity and they will offer relevant empirical targets to which our model is going to provide a structural interpretation.

In Supplemental Appendix M, we list the full estimates from specification (2), and show robustness to alternative definitions of job switches and migration.

Figure 3: Results from the Gravity Equation: Geography versus Home Bias

(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 border. We normalize the fixed effect coefficients for each worker type by their mean, and plot counties in the East with a negative distance.

3.3 Distorted Job Ladder

Finally, we study job flows and show that workers climb a country-wide job ladder, which is distorted by spatial frictions.

Let $n^h_{o,d,t}$ be the number of workers with home region h (either East or West Germany) that were in a job in county o in year $t-1$ and that have made a job-to-job move to a new job in county d in year t. We compute the share $s^h_{o,d}$ 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 core period.[17] We then 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 $\mathbb{X}$ contains seven 50km intervals from 50km-99km onward to 350km-399km and an eighth group for counties that are

[17] 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.

Table 1: Summary Statistics on Mobility

		Home: West	Home: East
	Workers moving job-to-job per month...		
(1)	- ... within region	1.13%	1.04%
(2)	- ... across regions	0.01%	0.06%
(3)	Ever crossed border	4.6%	23.9%
(4)	Returned movers	46.3%	36.1%
(5)	Mean years away (returners)	2.90	2.41

Source: LIAB. Notes: The table shows statistics for workers with at least one full-time employment spell in 2009-2014. Row 1 presents the share of these workers moving job-to-job per month within-region, defined as the number of job-to-job switchers whose new job is in the same region as the old one divided by all employed workers in the initial month, and averaged across months. Row 2 presents the average monthly share of movers across regions, defined analogously and taking all job movers across regions. Row 3 shows the share of the workers in our sample that have ever had a full-time job in their non-home region over the entire sample since 1993. Row 4 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 5 presents the average number of years away.

further than 399 km apart. The term $\mathbb{I}_{R(o)\neq R(d)}$ captures any geographical barriers beyond distance affecting mobility between East and West Germany. The 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, due to preferences, comparative advantage, or possibly due to a social network that allows them to find job opportunities.

Figure 3a shows that workers move mostly locally, and job switches become less likely for counties that are further apart.[18] Conditional on the origin and destination effects, we do not find a role for geographical barriers at the East-West border, since the gray line (the coefficients $\hat{\phi}_x + \hat{\rho}$) is almost on top of the black one.

Figure 3b shows that East individuals have significantly higher destination fixed effects for counties in East Germany. This result implies that East Germans are more likely to move to counties in the East than West German workers regardless of their current location.[19] Conversely, East-born workers are less likely to move to counties in the West. Supplemental Appendix O provides additional robustness checks for different sub-groups of the population and for different definitions of migration.

Despite the strong effects of distance and home bias on worker mobility, the labor markets of East and West Germany are, in fact, tightly connected. Table 1 shows

[18] We show the full list of estimated coefficients of regression (3) in Supplemental Appendix O.

[19] 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. This normalization is without loss of generality since we are interested only in the relative fixed effects across counties.

that on average 1% of all employed West and East Germans switch jobs within-region in an average month (row 1). For East Germans, the job-to-job transition rate across regions is about one twentieth as high as the transition rate within region (row 2). Row 3 illustrates that 4.6% of West-born and 23.9% of East-born in our sample have ever had a full-time job in the other region over the entire period since 1993. However, between one third and one half of the workers taking a job in the other region return to a job at home, after spending on average only 2-3 years away (rows 4-5).[20] Overall, workers climb a country-wide job ladder, but this ladder is distorted by spatial frictions since workers change jobs mostly locally and frequently return home. The substantial return migration implies that the gains from cross-region migration may be short-lived if workers, when returning home, move to relatively low productivity firms. This possibility highlights the importance of studying worker allocation to firms both within and between regions.

In Appendix B we present additional statistics on movers and show that the share of workers away from their home region has been relatively stable recently. This fact, together with the stable wage gap, motivates our analysis in steady state below.

4 Model

We now develop a model to quantify how spatial barriers and labor market frictions jointly affect worker mobility across space and firms. The model's ingredients are tied to the empirical facts shown above: (i) the wage dispersion and wage gains within-region call for a model with heterogeneous firms and labor market frictions; (ii) the spatial wage gaps and the asymmetries in wage gains and job flows necessitate a model with mobility costs and home bias; (iii) the presence of repeated moves across East and West suggests a framework in which individuals draw (infrequently) jobs from different regions. To capture these facts, our framework embeds the on-the-job search model of Burdett and Mortensen (1998) into a multi-region economy inhabited by heterogeneous firms and workers, subject to different types of spatial frictions commonly used in the literature: moving costs, home preferences, regional comparative advantages, and spatial search frictions.[21]

[20] The average non-returner is employed in the other region, until her employment history ends, for more than three times as long: 9.4 years for West Germans and 7.5 years for East Germans.

[21] We introduce spatial search frictions that make it easier for workers to find jobs locally, building on a recent literature which uses job application data to show that workers' number of applications

Table 2: Economic Environment

Workers of types i	$\bar{D}^i$, θ^i_j
Preferences for each location j	τ^i_j
Workers' indirect utility	$\mathcal{V}^i_j = w\theta^i_j\tau^i_j/P_j$
Moving costs between locations	κ^i_{jx}
Applications and spatial search frictions	$a^i_{jx}(s_x) = z^i_{jx}s_x$
Application cost (employed and unemployed)	$\psi(s_x) = \frac{s_x^{1+\epsilon}}{1+\epsilon}$, $\psi_u(s_x) = \nu^{-\epsilon}\frac{s_x^{1+\epsilon}}{1+\epsilon}$
Extreme value shocks upon an offer	$(\varepsilon_j, \varepsilon_x) \sim EV(0, \sigma)$
Firms' distribution within location j	M_j, $p \sim \gamma_j(.)$
Firm output net of vacancy cost (in steady state)	$vp\sum_{i\in\mathbb{I}}\theta^i_j l^i_j(w) - \xi_j(v)$
Random matching in each location j	$M(\bar{a}_j, \bar{v}_j) = \bar{a}_j^{\chi}\bar{v}_j^{1-\chi}$
Law of motion of labor per vacancy	$\dot{l}^i_j(w) = \vartheta_j^{-\chi}\frac{\bar{a}^i_j}{\bar{a}_j}\mathcal{P}^i_j(w) - q^i_j(w)l^i_j(w)$
Relative local price	$\frac{P_j}{P_x} = \left(\frac{P_jY_j}{P_xY_x}\right)^{\alpha(1-\eta)}\left(\frac{K_j}{K_x}\right)^{-\alpha(1-\eta)}$

We solve the model in general equilibrium to study the effects of removing spatial barriers on the allocation of workers to firms. The model is dynamic, but we focus on the tractable stationary equilibrium since the East to West wage gap is persistent and the number of workers away from home has been stable in recent years.

4.1 Environment

Let time be continuous and all agents discount future income at rate r. We consider an economy partitioned into $\mathbb{J} = \{1, ..., J\}$ sites, which we refer to as *locations*.[22] Table 2 shows the key features of the environment, which we now describe.

Workers. There is a continuum of workers of types $i \in \mathbb{I}\{1, .., I\}$ with mass $\bar{D}^i$, where $\sum_{i\in\mathbb{I}}\bar{D}^i = 1$. Throughout the text, we will use superscripts for worker types and subscripts for locations. Workers of type i have a preference parameter τ^i_j for being at location j, and consume both a tradable and a local good, such as housing. Their utility is $\mathcal{U}^i_j = \tau^i_j c^\eta h^{1-\eta}$, where c and h are the amounts of tradable good and local good, respectively. A worker of type i produces θ^i_j units of output per time

declines sharply with the distance of the vacancy (Manning and Petrongolo (2017); Le Barbanchon et al. (2020)). Schmutz and Sidibé (2018) also incorporate similar frictions into their *partial equilibrium* model to capture the lack of migration between areas with different unemployment rates. In Bilal (2021), instead, unemployed workers search for jobs only in their local market.

[22] 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: two in the East and two in the West.

unit in location j. Her indirect utility from receiving wage rate w in location j is $\mathcal{V}_j^i = w\theta_j^i\tau_j^i/P_j$, 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 of the local good.[23] We normalize $P_c = 1$.

Workers choose search effort s_x for each location x, file applications, and randomly and infrequently receive wage offers from firms. Search effort s_x directed by worker i in location j to location x generates $a_{jx}^i(s_x) = z_{jx}^i s_x$ job applications, where z_{jx}^i is the worker's search efficiency. Search effort is subject to a cost, to be paid separately for each location x in which the worker files applications, given by $\psi(s_x) = \frac{s_x^{1+\epsilon}}{1+\epsilon}$ for employed workers and $\psi_u(s_x) = \nu^{-\epsilon}\frac{s_x^{1+\epsilon}}{1+\epsilon}$ for unemployed ones. Here, $\nu \geq 1$ modulates the higher search intensity of unemployed workers along the lines of Moscarini and Postel-Vinay (2016). Upon receiving an offer from location x, workers draw idiosyncratic preference shocks for current location j (ε_j) and destination location x (ε_x) from a type-I extreme value distribution with zero mean and standard deviation σ.[24] We assume that workers can always separate into unemployment keeping the same shocks, which allows us to pin down the lower bound for wages in each location, as in the original formulation of Burdett and Mortensen (1998). 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.

Workers accept an offer if it provides higher value than the current one, solving

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

where $W_j^i(w)$ and $W_x^i(w')$ are the values of employment at wage w and w' in locations j and x, respectively, and $\kappa_{jx}^i = 0$ if $j = x$. The value $W_j^i(w)$ solves

$$\begin{aligned} rW_j^i(w) &= \frac{w\theta_j^i\tau_j^i}{P_j} + \delta_j^i\left[U_j^i - W_j^i(w)\right] \\ &+ \max_{\{s_x\}_{x\in\mathbb{J}}}\sum_{x\in\mathbb{J}}\left(a_{jx}^i(s_x)\,\vartheta_x^{1-\chi}\left[\int V_{jx}^{E,i}(w,w')\,dF_x(w') - W_j^i(w)\right] - \psi(s_x)\right). \end{aligned} \tag{4}$$

The first term, $w\theta_j^i\tau_j^i/P_j$, is the real flow value of employment. The second term

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

[24] These are shocks to workers' preferences for being in a specific firm and location. The problem is isomorphic to an alternative formulation in which workers only draw a shock for the value of accepting the offer, where that shock follows a logistic distribution.

is the continuation value for separating into unemployment, which occurs at rate δ_j^i. The third term is the continuation value for drawing new job offers from all the locations x, where $V_{jx}^{E,i}(w, w') \equiv \sigma \log\left(\exp\left(W_j^i(w)\right)^{\frac{1}{\sigma}} + \exp\left(W_x^i(w') - \kappa_{jx}^i\right)^{\frac{1}{\sigma}}\right)$ due to properties of the the type-I extreme value distribution. Applications become jobs at the equilibrium rate $\vartheta_x^{1-\chi}$. Last, $\{F_j\}_{j\in\mathbb{J}}$ are the endogenous distributions of wage offers in each region, generated by the firm's problem as we describe below.

Unemployed workers receive a benefit rate b_j^i and their value thus solves

$$rU_j^i = \frac{b_j^i \theta_j^i \tau_j^i}{P_j} + \max_{\{s_x\}_{x\in\mathbb{J}}} \sum_{x\in\mathbb{J}} \left(a_{jx}^i(s_x)\, \vartheta_x^{1-\chi} \left[\int V_{jx}^{U,i}\left(b_j^i, w'\right) dF_x(w') - U_j^i \right] - \psi_u(s_x) \right), \tag{5}$$

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

Firms and Goods Market. In each location $j \in \mathbb{J}$ there is a mass M_j of firms, with $\sum_{j\in\mathbb{J}} M_j = 1$. Firms are distributed over productivity p with location-specific density $\frac{\gamma_j(p)}{M_j}$ with support on a closed set $[\underline{p}_j, \bar{p}_j] \subseteq \mathbb{R}^+$.[25] Firms cannot change location.[26]

Each firm decides how many vacancies $v_j(p)$ to post and what wage rate $w_j(p)$ to offer, which jointly determines firm size, and then allocates labor to the production of the tradable and the local good. Denote by l_j^i the measure of workers of type i employed per vacancy of a firm, so that $n_j \equiv \sum_{i\in\mathbb{I}} \theta_j^i l_j^i$ is the total 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 must satisfy $n_c + n_h = \sum_{i\in\mathbb{I}} \theta_j^i l_j^i$. The term k is a factor in fixed supply, such as land, with aggregate supply in location j of K_j and equilibrium price ρ_j.

We first take as given the measure of workers hired and solve the firm's problem of allocating workers between the two goods. A firm that has hired n_j units of labor per vacancy maximizes profits excluding labor costs:

[25] Thus, $\gamma_j(p)$ will integrate to the mass of firms in location j, M_j.

[26] This assumption is motivated by the fact that, as mentioned, our data is at the establishment level, and thus we cannot see firms relocating or deciding where to open establishments. The model, nonetheless, could easily be adapted to allow entrepreneurs to make a location choice. Note that we allow firms to change their size by changing their number of vacancies, and to effectively enter or exit (across steady states) by going from zero to positive vacancies or vice versa.

$$\hat{\pi}_j(n_j) = \max_{n_h,n_c,k} \left\{ pn_c + P_{h,j} \left(pn_h \right)^{1-\alpha} k^{\alpha} - \rho_j k \right\} \tag{6}$$

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

$$\frac{P_j}{P_x} \;=\; \left(\frac{P_j Y_j}{P_x Y_x} \right)^{\alpha(1-\eta)} \left(\frac{K_j}{K_x} \right)^{-\alpha(1-\eta)}, \tag{7}$$

where $P_j Y_j$ is the nominal output of location j, and recall that $P_{h,j}^{1-\eta} = P_j$.[27] Substituting in the optimal choices and equilibrium price, we can simplify $\hat{\pi}_j(n_j)$ to

$$\hat{\pi}_j(n_j) = pn_j = p \sum_{i \in \mathbb{I}} \theta_j^i l_j^i(w). \tag{8}$$

The firm's profits thus become a linear expression in n_j, as in the standard Burdett-Mortensen framework. We provide details in Appendix D.1.

We now turn to the firm's wage posting problem. Since the firms' production functions are linear, the firm-level problem of posting vacancies and choosing wages can be solved separately. Total profits per vacancy are $\pi_j(n_j) = \max_w \{\hat{\pi}_j(n_j) - wn_j\}$. Using (8), the wage rate maximizes

$$\pi_j\left(p\right) = \max_w \left(p - w\right) \sum_{i \in \mathbb{I}} \theta_j^i l_j^i\left(w\right). \tag{9}$$

In choosing which wage to offer, firms take into account that a higher wage reduces their profit margin, but allows them to hire and retain more workers since, as we discuss below, the equilibrium function $l_j^i\left(w\right)$ is increasing in w. We assume that firms are competing for all worker types in one unified labor market, hence they must offer only one wage per efficiency unit to all workers.[28]

[27] 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)).

[28] This 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 (2021)) assumes that the labor market is segmented by type. In our framework, each firm p located in j posts a single wage rate $w_j(p)$, which determines the composition of worker types it attracts.

The number of vacancies per firm then maximizes total profits

$$\varrho_j(p) = \max_v \pi_j(p)\,\vartheta_j^{-\chi} v - \xi_j(v), \tag{10}$$

where $\xi_j(v)$ is a vacancy posting cost. The size of a firm p in location j is thus given by $l_j(w_j(p))v_j(p)$, where $w_j(p)$ is the profit-maximizing wage.

Matches. Matches in location j are created as a function of the total mass of applications filed by workers of all types i towards j, $\bar{a}_j = \sum_{i\in\mathbb{I}} \bar{a}_j^i$, and vacancies posted by firms, $\bar{v}_j$. Matching takes place according to a matching function $M(\bar{a}_j, \bar{v}_j) = \bar{a}_j^{\chi}\bar{v}_j^{1-\chi}$ as in Diamond-Mortensen-Pissarides models (e.g., Pissarides (2000)), where

$$\bar{a}_j^i = \sum_{x\in\mathbb{J}} \left[\int a_{xj}^{E,i}(w)\,dE_x^i(w) + a_{xj}^{U,i}(b)\,u_x^i\right], \tag{11}$$

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

Here, $a_{xj}^{E,i}(w)$ and $a_{xj}^{U,i}(b)$ are the equilibrium measures of applications sent by employed and unemployed workers of type i from x to j, and $E_j^i(w)$ is the mass of employed workers of type i at firms in location j receiving at most w. The matching function implies market tightness in location j of $\vartheta_j \equiv \frac{\bar{v}_j}{\bar{a}_j}$. The rate at which a vacancy is filled is $\vartheta_j^{-\chi}$, and the rate at which an application becomes a job is $\vartheta_j^{1-\chi}$.

Workers can direct applications towards each location, but search is random within location. Therefore, offers in location j are drawn from the following wage distribution

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

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. The law of motion for $l_j^i(w)$ is

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

where $l_j^i(w) = 0$ if $w < R_j^i$, and R_j^i is the reservation wage which solves $rW_j^i\left(R_j^i\right) = rU_j^i$. The first term in (14) 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; ii) $\frac{\bar{a}_j^i}{\bar{a}_j}$, the share of applications going towards location j filed by workers of type i; and iii) $\mathcal{P}_j^i(w) \in [0,1]$, the probability that an offer w 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,

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

where $\mu_{xj}^{E,i}(w', w)$ is the probability that an offer w is accepted by an individual currently employed in region x at wage w' and $\mu_{xj}^{U,i}(b, w)$ is the corresponding probability for an unemployed. Their closed form expressions are in Appendix D.2.

The second term in (14) is the separation rate

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

which consists of the exogenous separation rate into unemployment plus the rate at which workers receive and accept offers from other firms. As usual, we can use the law of motion (14) to solve for the steady state mass of workers per vacancy (which is zero if $w < R_j^i$)

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

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

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

The law of motion for unemployed workers is $\dot{u}_j^i = \delta_j^i e_j^i - \varphi_j^i u_j^i$, where the rate at which workers leave unemployment is $\varphi_j^i \equiv \sum_{x \in \mathbb{J}} \vartheta_x^{1-\chi} a_{jx}^{U,i}(b) \int \mu_{jx}^{U,i}(b, w') \, dF_x(w')$. Thus, in steady state, the mass of unemployed workers is

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

where $\bar{D}_j^i = E_j^i(w(\bar{p}_j)) + u_j^i$ is the total mass of workers i in region j.

4.2 Stationary Equilibrium

As discussed, we focus on the stationary equilibrium, which we now define.

Definition 1: Stationary Labor Market Equilibrium. *A stationary equilibrium 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 probabilities $\left\{\mu_{jx}^{E,i}(w, w'), \mu_{jx}^{U,i}(b, w')\right\}_{j \in \mathbb{J}, x \in \mathbb{J}, i \in \mathbb{I}}$, labor per vacancy $\left\{l_j^i(w)\right\}_{j \in \mathbb{J}, i \in \mathbb{I}}$, 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 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 profits, taking as given the function $\left\{l_j^i(w)\right\}_{j \in \mathbb{J}, i \in \mathbb{I}}$;*
3. *arrival rates of offers and wage distributions are consistent with wage policies, applications, and vacancy posting, according to (9), (11), and (12);*
4. *firm sizes and worker distributions satisfy (17), (18), and (19).*

The following proposition shows that the wage policies follow a system of differential equations, facilitating the computation of the model.

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

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

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

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

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

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

Proof. See Appendix D.3. □

In Supplemental Appendix P, we show that our model collapses to the Mortensen (2005) framework if we shut down the spatial heterogeneity and the preference shocks.

5 Bringing the Model to the Data

We bring the model to our German data and quantify the different spatial and labor market frictions that limit the ability of workers to reallocate across firms and regions.

5.1 Estimation

To estimate the model, we impose a few assumptions to reduce dimensionality, calibrate outside of the model all the parameters that have a corresponding empirical moment, and jointly estimate the remaining ones within the structure of the model.

Parametrization and Functional Forms. To keep the estimation time feasible, we set the number of locations to four, two in the West and two in the East − Northwest ($j = NW$), Southwest ($j = SW$), Northeast ($j = NE$), and Southeast ($j = SE$), and choose four worker types, which are distinguished by their home location.[29] We can thus distinguish the role of the former East-West border from other spatial frictions between locations, and we will continue to refer to East and West Germany overall as "regions".

[29]This parametrization implies that we need to match $4 \times 4 \times 4 = 64$ wage gains and 64 worker flows. Appendix A provides details on the locations. In robustness checks, we analyze the effect of increasing the number of locations.

We set a unit interval of time to be one month.[30] Firms' log productivity is drawn from a log-normal distribution with equal variance in all locations, Σ, and mean A_j, normalized with $A_{NW} = 1$. The vacancy cost function is $\xi_j(v) = \xi_{0,j} v^{1+\xi_1} \bar{\pi}_j(p)$, where the curvature ξ_1 is constant across locations but $\xi_{0,j}$ is specific to the overall region – i.e. we estimate $\xi_{0,W}$ and $\xi_{0,E}$.

Rather than pinning down the unemployment benefits b_j^i, we set the reservation wages R_j^i directly to be a fraction ι of the lowest firm productivity, $R_j^i = \iota \underline{p}_j$ and assume that all worker types have identical reservation wage (but not utility) within location.[31]

We restrict the different sources of spatial frictions – τ_j^i, κ_{jx}^i, and z_{jx}^i – as follows

$$
\begin{aligned}
\tau_j^i &= \tau_j \left(1 - \tau_l \mathbb{I}_{(i \neq j) \cap (r(i) = r(j))}\right) \left(1 - \tau_r \mathbb{I}_{r(i) \neq r(j)}\right) \\
\kappa_{jx}^i &= \begin{cases} 0 & \text{if } j = x \\ \kappa_0 e^{\kappa_1 dist_{jx}} \bar{W}^i & \text{if } j \neq x \end{cases} \\
z_{jx}^i &= \begin{cases} \left(1 - z_{l,1} \mathbb{I}_{i \neq j}\right) & \text{if } j = x \\ \left(z_0 e^{-z_1 dist_{jx}}\right) \left(1 + z_{l,2} \mathbb{I}_{i = x}\right) \left(1 + z_r \mathbb{I}_{(r(i) = r(x)) \cap (i \neq x)}\right) & \text{if } j \neq x \end{cases}.
\end{aligned}
$$

Preferences τ_j^i are the product of general amenities of location j (τ_j), the worker's utility cost to live outside of her home location but inside her home region (τ_l), and the cost to live outside the home region (τ_r). The index function $r(i)$ maps location i to its region. The moving cost κ_{jx}^i is symmetric (which is important for identification, as we discuss below) and proportional to the average value for each worker, $\bar{W}^i \equiv \frac{1}{e^i} \sum_{j \in \mathbb{J}} \int W_j^i(w) \, dE_j^i(w)$, where $e^i \equiv \sum_{j \in \mathbb{J}} E_j^i(w(\bar{p}_j))$. The search efficiency is a function of whether the worker searches within the current location ($z_{l,1}$), of the distance of the destination location (z_0,z_1), and of whether the search is directed towards the home location or region ($z_{l,2}$, z_r).

Finally, we restrict $A_{NE} = A_{SE}$ since average wages are similar in the Northeast and the Southeast.[32] We also assume that local amenities are the same, $\tau_{NE} = \tau_{SE} = \tau_E$. We show below that despite these restrictions, we match well the location-specific

[30] 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 as $Prob_u = 1 - e^{-\delta}$.

[31] While R_j is endogenous, setting its value directly is the same as choosing a set of unemployment benefits b_j^i that solve $U_j^i = W_j^i(R_j)$.

[32] See Supplemental Appendix K.

moments of the Northeast and Southeast.

Calibrated Parameters. We calibrate eight sets of parameters listed in Table 3. We describe how we set their values in Appendix E, and focus here on how we set workers' relative productivity, θ_j^i (row 1). Due to wage posting and since firms post the same wage to all workers, the model yields a log additive wage equation

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

This equation is similar to the specification by Abowd et al. (1999), with the main difference that in our specification θ_j^i is both individual- and location-specific. This allows for *comparative advantage*, i.e., a worker employed in her home location could have higher productivity there. We show in Appendix F that we can run a modified AKM regression to identify the strength of the comparative advantage. This effect, however, turns out to be close to zero. Therefore, we can focus on workers' average skills, θ^i, which we obtain as the the average worker fixed effects from an AKM regression. We find that the average East German worker's unobserved skills are about 9 percentage points below those of a West German worker.[33]

Estimated Parameters and Targeted Moments. We are left with 21 parameters that we estimate through simulated method of moments, and target the 305 moments summarized in Table 4.[34] Appendix E summarizes how each moment is computed.

Overall, we target all the key moments presented in the motivating evidence of Section 3, but for locations rather than regions. We then add a few more specific moments to discipline as well as possible the extent of labor market frictions.

The mapping between model and data is straightforward since we can compute exactly the same objects in both. The main complication is to define worker flows across locations in the data consistently with the model. A sizable share of individuals in our data report to be working in a location different from their residence, while in the model we do not distinguish between migration and commuting. As our baseline,

[33]A recent literature has shown several concerns related to the estimation of second moments in AKM regressions (see Andrews et al. (2008), Andrews et al. (2012), and Bonhomme et al. (2019)). For our application, these concerns do not apply since we focus on first moments, which are unbiased (Andrews et al. (2008)).

[34]We estimate the model using a standard indirect inference approach and provide more details on our estimation algorithm in Appendix G. Figure A4 shows that the model likelihoods are locally single-peaked around each parameter estimate.

Table 3: Calibrated Parameters

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

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.

we therefore count as migrants in the data all individuals that change their work location and satisfy either one of these conditions: i) they update their residence; ii) their new job is farther away from their residence than the old one and both jobs are within 200km of their residence (otherwise, we suspect that the residence is simply misreported).[35] We next describe how the moments identify our parameters.

Identification. The flows and wage gains of movers within and between regions, and how they differ by birth-place, allow us to identify the spatial frictions (κ^i_{jx},

[35]About 7% of workers work in a location different from their residence. Defining cross-location movers as only those workers that change the location of their job and update their residence could overestimate spatial frictions since some job offers lead workers to commute, and hence these workers do not update their residence. However, since the living location is self-reported as discussed in Section 2, we do not want to include individuals that report to be living very far away from their job as these observations are likely misreported. We also do not want to define cross-location moves as all changes in work location regardless of residence since that could underestimate the frictions, as commuters, especially the ones moving back closer to their home, most likely do not pay the same fixed costs of relocating as migrants. Our definition strikes a balance between these concerns. In Supplemental Appendix T, we re-estimate our model with a broader and a narrower definition of cross-location moves and show that this mainly affects our estimates of the moving cost, while keeping most results unchanged.

Table 4: Targeted Moments

	Moments	N	Source	Model Fit	Key Parameters
(1)	Wage gains w/i locations, by (i, j)	16	Q.2.1	Fig 5	Σ, σ
(2)	Wage gains b/w locations, by (i, j, x)	48	Q.2.1	Fig 5	κ, τ_j^i, Σ
(3)	Job flows w/i locations, by(i, j)	16	Q.2.2	Fig 5	ϵ, ξ_0
(4)	Job flows b/w locations, by (i, j, x)	48	Q.2.2	Fig 5	z_{jx}^i, κ
(5)	Employment shares, by (i, j)	16	Q.2.3	Fig A6	κ, z_{jj}^i, τ_j, τ_j^i, z_{jx}^i
(6)	Unemployment shares, by (i, j)	16	Q.2.4	Fig A6	κ, z_{jj}^i, τ_j, τ_j^i, z_{jx}^i
(7)	Firm component of wages, by (i, j)	15	Q.2.5	Fig A6	A_j, τ_j
(8)	Average firm component of wages, by j	3	Q.2.6	Fig A6	A_j, τ_j ,z_{jj}^i, τ_j^i
(9)	Relative output per worker, by j	3	Q.2.7	Fig A6	A_j, ν
(10)	Unemployment rates, by j	4	Q.2.8	Fig A6	ν
(11)	Deciles of firm-size distributions, by j	40	Q.2.9	Fig A7	ξ_1
(12)	Slope of wage vs firm size relationship, by j	4	Q.2.10	Fig A8	ξ_1 , ι
(13)	Slope of J2J wage gain vs initial wage, by j	4	Q.2.11	Fig A8	Σ, σ
(14)	Slope of separation rate vs firm wage, by j	4	Q.2.12	Fig A8	ξ_0 , σ, ϵ
(15)	Std of job-job wage gains, by (i, j, x)	64	Q.2.13	Fig A9	Σ, ξ_0 , ϵ, ι, σ
(16)	Profit to labor cost ratio, by j	4	Q.2.14	Table A8	σ, ξ_1 , ι,ξ_0, τ_j^i

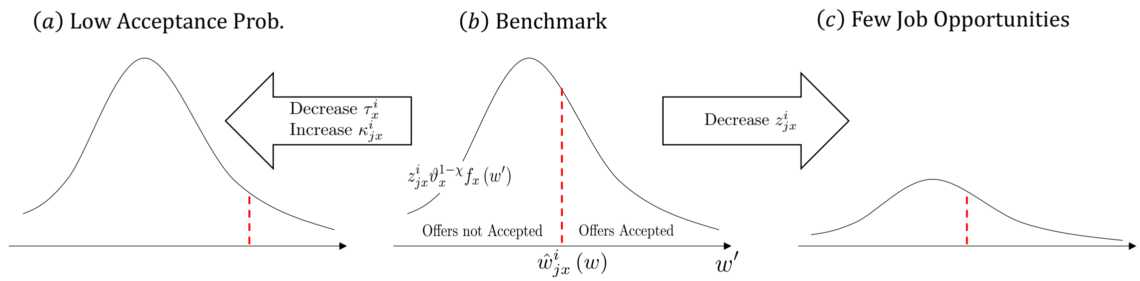
Notes: The table reports the moments used in the estimation. The column titled "N" lists the number of moments in the group. Column "Source" links to the appendix section where the moment is computed, and column "Model fit" lists the table or figure that compares the empirical moment to the model-computed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

τ_j^i, and z_{jx}^i) and to separate them from the labor market frictions (rows 1-4 of Table 4).[36] Labor market frictions mainly affect the allocation of labor within each location, and can therefore be identified from within-location moments, as quite standard in the Burdett-Mortensen framework (see, e.g., Bontemps et al. (2000)). The spatial frictions can then be inferred from cross-location moments. Figure 4 illustrates our identification argument, highlighting the importance of using both worker flows and wage gains. 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, $z_{jx}^i \vartheta_x^{1-\chi} f_x(w')$.

Assume that $\sigma \to 0$ so that an offer w' from location x is accepted by worker type i employed in region j at wage w if and only if $W_x^i(w') - \kappa_{jx}^i \geq W_j^i(w)$. Let $\hat{w}_{jx}^i(w)$ be the cutoff wage offer such that $W_x^i\left(\hat{w}_{jx}^i(w)\right) - \kappa_{jx}^i = W_j^i(w)$. The accepted offers are

[36] Since the job flows are also tied in steady state to the allocation of labor across locations, we target those moments as well (rows 5-6).

Figure 4: Identifying Spatial Frictions



the ones to the right of $\hat{w}_{jx}^i(w)$, and the mass of worker flows per unit of search effort is the integral under the wage offer density to the right of $\hat{w}_{jx}^i(w)$. Starting from the benchmark (b), we consider a decrease in the search efficiency z_{jx}^i in panel (c). This decrease reduces the mass of offers received, and hence the worker flows. However, it does not affect the cutoff $\hat{w}_{jx}^i(w)$, and hence has no effect on the average wage gain of workers that accept an offer.

In contrast, a decline in workers' preference for location x, τ_x^i, shifts the acceptance threshold to the right in panel (a) (a similar argument applies to a rise in the moving cost, κ_{jx}^i). Therefore, a worker accepts only relatively better offers. Hence, the expected wage gain of a move increases in κ_{jx}^i and decreases in τ_x^i. Wage gains between locations are thus informative of κ_{jx}^i and τ_x^i, while flows between locations primarily help identify z_{jx}^i.

Without further restrictions, we cannot separate moving costs from location preferences. We thus assume that moving costs are identical for all worker types. This assumption allows us to identify the location preferences using differences in wage gains for individuals of different types that make the same migration move, e.g., East versus West Germans that move from East to West.[37]

The productivity shifters (A_j) and the search efficiency of the unemployed (ν) are mainly related to the average firm wages, output per worker and unemployment by location (rows 7-10). When productivity is higher, firms offer a higher wage, everything else equal. A greater search efficiency of the unemployed reduces the unemployment rate and increases output per capita. The moments are also related to the location's amenity (τ_j), which leads to lower wages due to compensating differentials.

[37] Supplemental Appendix R provides further details on the identification, and Appendix G uses simulations of the model to verify that the identification argument holds in practice.

The variance of firm productivity (Σ), the labor market friction parameters ($\xi_0, \xi_1, \epsilon, \sigma$), and the level of the reservation wage relative to firm productivity (ι) are linked to the efficiency of the job ladder (rows 11-15). This is expected: a higher variance of productivity raises the variance of wages; ξ_0 and ξ_1 determine the intensity of vacancy posting and how it varies across firms; the cost of search effort ϵ modulates the relationship between workers' search intensity and the value of employment at their current firm; and σ determines how much the job moves are directed towards higher wage offers.

Finally, the labor market friction parameters (σ, ξ_0, ξ_1) and the reservation wage (ι) are related to firm profitability (row 16). Greater labor market frictions increase firms' local monopsony power and hence profits, while a higher reservation wage decreases profitability. The home preference τ_j^i also plays a role: when workers are more attached to a location, firms face less competition from other locations.

To verify our heuristic identical argument, we analyze the connection between all parameters and moments via model simulations and show the Jacobian matrix illustrating the elasticity of each (model generated) moment to each parameter in Appendix G. The last column of Table 4 reports the most important parameters for each moment based on this exercise.[38]

Importantly, the Jacobian matrix confirms the importance of targeting both wage gains and workers flows across locations to identify the spatial frictions. The wage gains between locations (row 2 of Table 4) are crucial for the moving costs κ and the preference τ_j^i, while the job flows between locations (row 4) are important for the relative search efficiencies z_{jx}^i. The spatial frictions are also key for the steady state allocation of employed and unemployed workers (rows 5-6). As expected, the within-location wage gains and flows (rows 1 and 3) are not relevant for the spatial frictions. Large within-location wage gains are driven by either a large variance of productivity (Σ), or a low variance of the taste shock (σ) so that workers only accept

[38] Since the full Jacobian matrix includes 6,405 (305×21) cells, in our exposition we take averages of the 16 blocks of moments shown in Table 4 and show these averages rather than each moment separately. In the table and graph, we bundle together a few sets of closely related parameters and refer to them jointly as follows: i. the two relative amenities τ_{SW} and τ_E (we refer to them jointly as $\tau_j \equiv \{\tau_{SW}, \tau_E\}$); ii. the two home biases τ_l and τ_r $\left(\tau_j^i \equiv \{\tau_l, \tau_r\}\right)$; iii. the relative search efficiencies between regions z_0, z_1, $z_{l,2}$ and z_r $\left(z_{jx}^i \equiv \{z_0, z_1, z_{l,2}, z_r\}\right)$; iv. the cost of moving κ_0 and κ_1 ($\kappa \equiv \{\kappa_0, \kappa_1\}$); v. the two relative productivities A_{SW} and A_E ($A \equiv \{A_{SW}, A_E\}$); vi. the two costs of vacancy posting $\xi_{0,W}$ and $\xi_{0,E}$ ($\xi_0 \equiv \{\xi_{0,W}, \xi_{0,E}\}$).

job offers with an associated wage increase. Within-location flows are related to the parameters of the costs of applying to jobs (ϵ) and of posting vacancies (ξ_0).

5.2 Results

Next, we turn to the results, show the model fit, and discuss the parameter estimates.

Model Fit. The left panel of Figure 5 plots the wage gains of job-to-job movers in the data against those in the model, while the right panel shows the labor flows.[39] Each dot is for one of the 64 different types of moves by origin-destination-home location. The model matches the data well. For example, it generates larger wage gains for moves towards the West (blue symbols). Individuals are also more likely to move within-location (gray circles) and to move back to their home location and region (diamonds) than away from home (stars).

We discuss the fit of all other moments in Appendix H, and summarize here the takeaways. The model matches well the steady state distributions of workers and the average GDP, wages, and unemployment rates, consistent with the hypothesis that the German labor market is in steady state. The model's job ladder mechanism implies that more productive firms offer higher wages and have a lower rate of quits, which 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.[40]

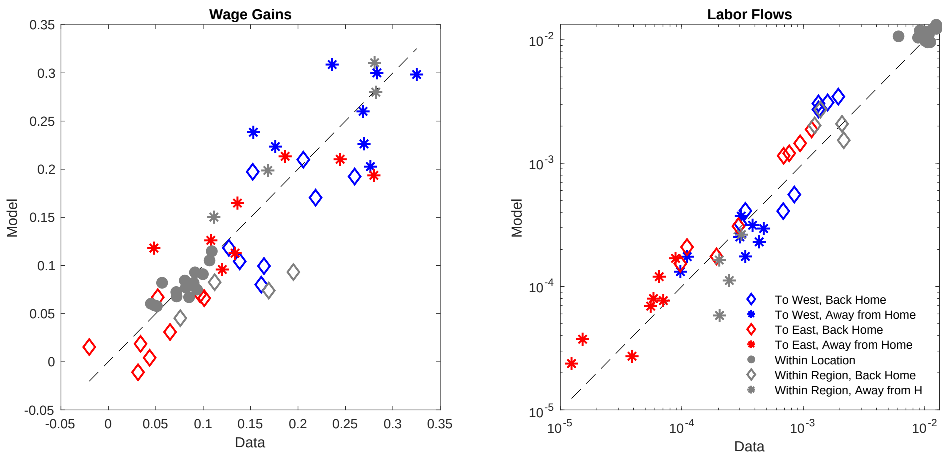
Overall, the fit is good considering that we estimate 21 parameters to target 305 moments.[41] Several structural restrictions imposed by the model on the joint distri-

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

[40]In Figure A8 we show the non-parametric relationships for the moments in rows 12, 13, and 14 of Table 4. In Figure A9, we show that adding individual fixed effects in wage growth brings the empirical estimates for the standard deviations of wage growth very close to the model's ones.

[41]Given the arbitrary distinction between targeted and not targeted moments, we decided to simply include as targets all the key relevant moments. The model performance is thus evaluated by its ability to simultaneously match several features of the data despite its relatively limited flexibility.

Figure 5: 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 E.

butions of firm wages, employment, wage gains, and labor flows are satisfied in the data, building confidence in our estimated frictions.

Parameter Estimates. We present the estimated spatial frictions in Table 5, and include the remaining parameters in Appendix G. Row 1 reports the one-time moving costs, κ_{jx}, as a fraction of the present discounted value of income. Since these costs vary with distance, we present a range for moves between the closest two locations and moves between the farthest two locations. We find moving costs in the range of $3-5\%$ of the PDV of income. Rows 2 and 3 show a strong preference for birth-location: workers need to be paid about 7.4% more than in her home location to obtain the same flow utility, and moving towards the non-home region would require a yearly compensation of almost 10%.[42,43]

[42] In Supplemental Appendix V, we 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.

[43] Our estimated moving and preference costs are consistent with the findings in Schmutz and Sidibé (2018), who estimate moving costs between 13,700 € and 16,900 € between cities in France. The moving costs we estimate are smaller than in work that does not account for a frictional labor market, for two reasons: first, since any cross-location move is also a move between firms, part of the wage gain from migration reflects general labor market frictions that are also present within region, rather than moving costs; second, the search frictions across locations in our model allow us to match a low cross-regional migration rate without the need of a very large moving cost.

Table 5: 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, 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 A7. 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_{jj}^{j}, 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.

Rows 4 and 5 report the search efficiencies, relative to the within-home location level, which is normalized to 100%. Search efficiency is much lower across locations, consistent with evidence that workers search for jobs primarily locally.[44] For example, row 5.i shows that one unit of search effort expended across locations to the non-home region translates into filing only about 1/20th as many applications. Search is also more efficient towards the home location, (compare 5.i and 5.iii) possibly reflecting social connections (Burchardi and Hassan (2013), Bailey et al. (2020)).

We discuss the remaining parameters in Appendix G. We note that our model infers an amenity value in the East that is 11% higher than in the West. This additional amenity is consistent with the large fiscal transfers towards East Germany (Henkel et al. (2021)) and it could additionally reflect remaining cost of living differences that are not picked up by our price indices.

Discussion. Our results hinge on two core assumptions of the Burdett-Mortensen 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 the

[44] Manning and Petrongolo (2017), Le Barbanchon et al. (2020), Datta (2022).

outside option has a limited effect on workers' wages (Jäger et al. (2020)) and that, conditional on the current firm, a worker's previous firm has almost no effect on current wages (Kline et al. (2019)). Nonetheless, we note that under a different wage setting method larger wage gains for movers between locations could be driven by firms offering wage premia to compensate workers that have to migrate to take a job. 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 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 meaning 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.

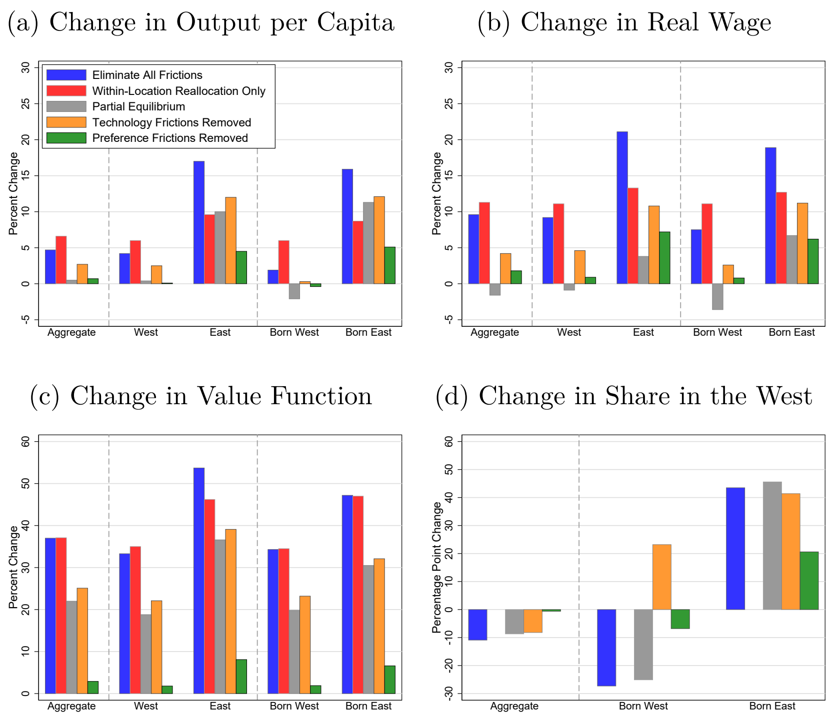
6 Labor Misallocation across Firms and Regions

We now study the role of spatial frictions in our framework with labor reallocation both across firms and regions. First, we analyze the aggregate effects of spatial frictions. Then, we turn to the distributional effects across regions and workers' types. Finally, we study the extent to which the results are affected by the strength of the labor market frictions. Since the heterogeneity across locations within regions is small relative to the East-West differences, we aggregate the results by region rather than showing individual locations throughout this section.

6.1 Aggregate Effects of Spatial Frictions

We recompute the equilibrium keeping all the parameters at their estimated values, but removing the spatial frictions: the moving cost ($\kappa_0 = 0$), the home preferences

Figure 6: Aggregate and Distributional Effects of Removing Spatial Frictions



(a) Change in Output per Capita (b) Change in Real Wage

(c) Change in Value Function (d) Change in Share in the West

Notes: Figure shows the effects of various exercises, shown with the different-colored bars, on four outcomes: output per capita (top-left), real wage (top-right), average value (bottom-left), and the share of workers in the West (bottom-right). Bars show percentage change relative to the baseline economy.

($\tau_l = \tau_r = 0$), and the differences in search efficiency across and within locations ($z_{l,1} = z_{l,2} = z_r = z_1 = 0$ and $z_0 = 1$). We then compute four core statistics for the baseline and the counterfactual long-run steady state equilibrium: (i.) output per capita; (ii.) the average of workers' value functions across all employed and unemployed workers; (iii.) average real wage, $w_j(p)\,\theta_j^i/P_j$; and (iv.) the share of the overall employment in West Germany.

The results are shown by the first set of blue bars in Figure 6. Removing all spatial frictions leads to an increase in output per capita of slightly less than 5%.[45] Despite these relatively modest output gains, the increase in the average worker's value is much larger.[46] The reason is twofold. First, without spatial frictions workers no

[45]The aggregate productivity cost of spatial frictions is smaller in our model than in other contexts (e.g., Bryan and Morten (2019)), which is likely due to the different context (developed versus developing country) and due to the fact that our model does not contain a key mechanism in their work, namely that each individual draws a vector of location-specific comparative advantages.

[46]We use the term workers' value rather than welfare since we are, in the counterfactual, effectively changing preferences through the taste spatial friction τ_j^i.

longer incur the moving cost κ_{jx}^{i} or the utility cost $(\tau_l,\ \tau_r)$ when they cross locations. Moreover, workers' search efficiency across locations rises. These factors lead to a higher continuation value. Second, eliminating spatial frictions exposes firms to stronger competition for workers from other locations, which raises wages more than the increase in labor productivity due to a reduction in firms' monopsony rents.

The bottom right panel illustrates that there is net reallocation of labor towards the East, hence, towards the region with, on average, lower productivity. This result could seem counterintuitive: in a neoclassical framework we would have expected labor to reallocate towards the West. However, it is a direct implication of an inherent asymmetry in our frictional setting. In the data, and in our baseline estimation, there are only about a third as many East Germans as West Germans. Therefore, more workers have a strong attachment to the West than to the East. Once we remove spatial frictions, even though a relatively smaller share of West Germans than East Germans migrate, there is a relatively larger positive labor supply shock in the East, as it is opening up to a larger labor market.

We now further investigate the mechanisms behind these findings in our model with worker allocation both across firms and locations. First, we analyze the importance of within-location reallocation of labor compared to worker reallocation across locations. Second, we discuss the role of the equilibrium response of firms. Third, we separately analyze the different types of frictions.

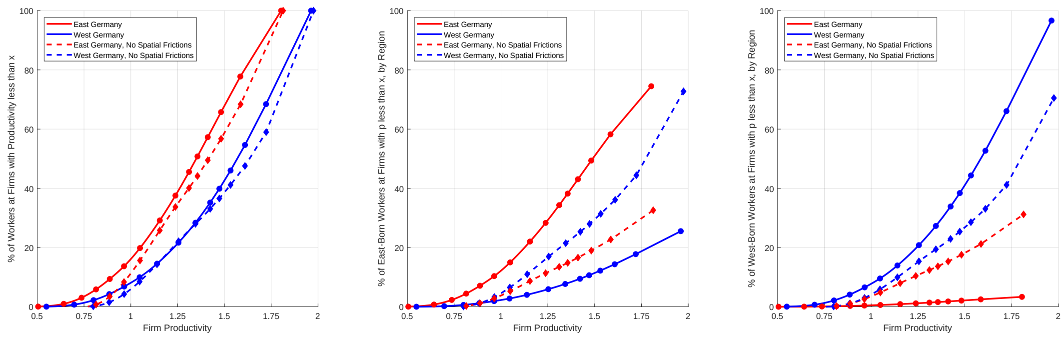
The Importance of the Within-Location Allocation of Labor. Our first key result is that the aggregate gains we find arise from a better allocation of labor *within* locations, rather than from migration of workers towards high productivity locations. To reach this result, we recompute the aggregate gains holding fixed the share of workers in each location at the baseline level. We continue to change the within-location distribution of workers and firms' policy functions as in the full counterfactual. The results, in the red bars in Figure 6, show that shutting down the migration across locations actually *raises* output and wages, since in the full counterfactual workers migrate towards the lower productivity East.

Panel (a) of Figure 7a analyzes how the within-region reallocation of workers generates the aggregate gains by presenting CDFs of employment to firms of different productivity within East and West Germany.[47] Removing spatial frictions shifts the

[47] Since the baseline was estimated from the data moments, it is consistent with the within-region

Figure 7: Labor Allocation Across Firms and Regions

(a) All Workers (b) East Germans (c) West Germans



Notes: The left panel shows the CDF of workers over firm productivity within East (in red) and West Germany (in blue). The solid line is our benchmark estimation, while the dashed one the counterfactual without spatial frictions. The middle panel is a semi-CDF that shows the distribution of employment for East German workers across the whole Germany. To interpret the figure, consider that, at baseline, more than 75% of employment is in East Germany, and the remaining employment is in the West (i.e., the two last points of the solid lines add up to one, and similar for the dashed lines). The right panel shows the semi-CDF for West Germans.

distributions to the right as labor reallocates towards the more productive firms. In the baseline economy, spatial frictions partially shield low productivity firms from competition through two margins: i.) by reducing the value of unemployment, thus allowing firms to hire workers at a relatively low wage; ii.) by limiting the rate at which workers are poached, as they are only rarely poached from firms in the other region. As spatial frictions are removed, these protections are eliminated. Therefore, it becomes harder for unproductive firms to hire and to retain workers, forcing them to shrink or to stop posting vacancies altogether. While removing spatial frictions also makes it easier for unproductive firms to hire from the other region, on net the negative effect dominates. This effect is stronger in the East because it has the lowest productivity firms overall.

We can decompose the total labor of a firm of productivity p in region j as

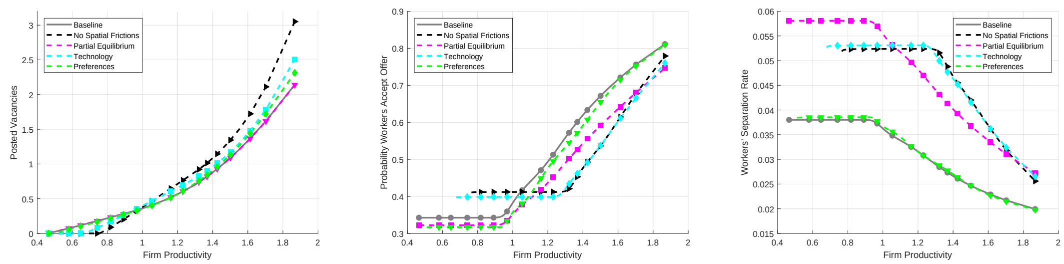
$$e_j\left(p\right) = \underbrace{\vartheta_j^{-\chi}}_{\text{Tightness}} \underbrace{v_j\left(p\right)}_{\text{Vacancies}} \sum_{i \in \mathbb{I}} \left(\frac{\bar{a}_j^i}{\bar{a}_j} \underbrace{\mathcal{P}_j^i\left(w(p)\right)}_{\text{Accept Probability}} \underbrace{\left(q_j^i\left(w(p)\right)\right)^{-1}}_{\text{Separation Rate}} \right).$$

The first term captures the local market tightness and thus only affects the allocation of labor between regions. The other three terms could, in principle, explain the

wage distributions shown in Figure 1b if wages are increasing in productivity, as in our model.

Figure 8: Margins of Employment in East Germany

(a) Posted Vacancies (b) Acceptance Probability (c) Separation Rate



Notes: All panels are for firms in East Germany and show outcomes as a function of firm productivity. The left panel shows the change in the number of posted vacancies. The middle panel shows the probability that a given wage is accepted by the worker it matches with. The right panel shows the monthly rate at which workers separate towards either other firms or unemployment. We consider four possible counterfactuals, described in text.

reallocation of labor towards more productive firms. Removing spatial frictions might allow high productivity firms to post relatively more vacancies (high $v_j(p)$), make it easier for them to attract workers upon meeting them (high $\mathcal{P}^i_j(w)$), or facilitate worker retention (low $q^i_j(w)$).

In Figure 8, we plot these three objects as a function of firm productivity.[48] Panel (a) shows that the number of posted vacancies contributes positively to the reallocation of labor from low- to high-productivity firms. Going from the baseline to the counterfactual, more productive firms increase their number of vacancies while unproductive firms shrink. The separation rate also contributes to the improved allocation of labor (panel (c)): in the counterfactual equilibrium all workers search more intensively, leading to a higher separation rate than in the baseline, but this effect is larger at lower productivity firms. The acceptance probability, instead, mitigates the reallocation gains (panel (b)). Workers are relatively more likely to accept offers at lower productivity firms in the economy without spatial frictions. This result is driven by the fact that access to the country-wide pool of unemployed workers, as previously noted, has a larger relative impact on the lower productivity firms.[49]

Large Equilibrium Effects due to Lower Monopsony Power. Our second key result is that the majority of the aggregate gains arise from the change in labor market competition and the resulting decline in firms' local monopsony power, rather

[48] The plots are for East Germany. The ones for the West are similar and are in Supplemental Appendix W.

[49] For the higher productivity firms, instead, the probability that an offer is accepted decreases due to the overall improvement in the allocation of labor and the increased effective competition.

than from changes in workers' behavior. In the counterfactual equilibrium, both workers and firms adjust. Workers search more intensively across locations and are more willing to accept job offers that are further away. Firms adjust their wages and vacancies to more competition. To disentangle these two effects, we recompute the steady state holding fixed firms' wages and posted vacancies at their baseline values while allowing workers to adjust their search and acceptance behavior.

The first set of gray bars in Figure 6 show that when firms' wage and vacancy policies are held fixed, the output per capita increases by only 0.5%. Thus, firms' equilibrium response to more competition is the main driver of the aggregate gains. Intuitively, when firms are not able to adjust vacancies, one of the key drivers of the improved within-allocation is muted, as illustrated by the dashed pink line on top of the gray line in Panel (a) of Figure 8.[50] While the separation rate still rises more for low-productivity firms than for high-productivity ones (Panel (c)), this channel alone has only a modest effect.

Effects of Individual Frictions. Our third finding is that there are strong complementarities between the *technological* spatial frictions imposed by the moving cost κ and the search productivity z, and *preference* spatial frictions due to home preferences τ. Technological frictions could be affected by policy, for example by a faster railway system or an integrated online job portal. Instead, preference frictions are plausibly harder to affect, as they are typically a slow moving object (Alesina and Fuchs-Schündeln (2007)). To analyze their effects separately, we recompute the equilibrium of the economy when we remove either only the technological spatial frictions or the home preferences. The yellow and green bars in Figure 6 show that removing technological barriers alone generates aggregate gains that are about a third to one half as large as the baseline. In contrast, removing home preferences generates only modest gains. Addressing both types of frictions jointly is important: summing over the aggregate gains from both separate exercises yields only about half the effect of removing both sources of frictions at the same time.

Robustness. We explore the quantitative role of two features of our model in Appendix I: (i.) there are only two locations in each region; (ii.) the locations in East Germany are smaller, hence have fewer firms and workers. We find that there are still large gains from the within-location reallocation of labor even as we increase the

[50] In Supplemental Appendix W we replicate Figure 7 for this alternative counterfactual.

number of locations to 24, and that there are only relatively small effects of the size of the labor force in each location.

Evidence of Our Mechanism. Our findings rely on a key mechanism: lower spatial frictions increase competition, which improves allocative efficiency by forcing inefficient firms to shrink or to exit the market. While we do not have direct evidence on the response of firms' wage and vacancy policies, we can bring indirect evidence to test our mechanism. Specifically, we study the relationship between spatial frictions and workers' reservation wage and search effort, which can be observed in the U.S. Survey of Consumer Expectations (SCE). See Appendix J for details.

We verify three implications of the model. First, workers that commute longer distances conditional on a given wage, hence are effectively more exposed to spatial frictions, spend more time searching for a new job and send more applications, consistent with a "local firm advantage".[51] Second, workers at the bottom of the local pay distribution search more for new jobs for a given commuting time, supporting the notion that low-wage firms are most affected by spatial frictions. Third, conditional on their current wage workers have a higher reservation wage when their local labor market has more job opportunities, and hence local economic conditions matter.

6.2 Distributional Effects of Spatial Frictions

We next focus on the distribution of resources across regions and worker types.

Differences by Region. The second and third set of bars in Figure 6 examine the effects of removing all spatial frictions separately for individuals in the West and in the East of Germany (blue bars).[52] The baseline gains are larger in the East than in the West, for two reasons. First, despite similar observables, we estimated a large gap between East and West workers in unobservable skills (see row 1 of Table 3). Therefore, as West workers move East and East workers move West, relative human capital improves in the East. Second, the reallocation of labor away from lower productivity firms is stronger in the East since there are more low productivity firms in that region.

[51] Workers also search more intensively for a new job when they are "less satisfied" with the current one. This variable captures potential non-pay amenities of the job ladder.

[52] In the model, individuals move continuously across locations. Nonetheless, we can compute the outcomes for the individuals that are, in our long-run steady state, in either East or West Germany. The statistics account for the possibility that individuals move across locations and regions.

While East Germany gains the most, output and wages also rise in West Germany. This outcome differs from a neoclassical benchmark model with one representative firm in each region, where eliminating barriers to labor mobility would lead to net worker flows towards the West until marginal labor productivity is equalized across regions. In our model with heterogeneous firms, workers in the West gain since there is net reallocation of labor towards the East, hence a less tight West German labor market, and, moreover, an improvement in the within-region allocation of labor.

In our second exercise (red bars), we find that when migration across regions is shut down, the output and wage gains in East Germany fall by nearly half. Intuitively, a large part of the East German gains is due to the increase in average human capital due to the in-migration of West German workers. Instead, in West Germany, migration has a negative effect on output and wages. Due to the importance of migration, the partial equilibrium gains (gray bars) in East Germany are also relatively large.

Differences by Worker Type. The fourth and fifth set of bars in Figure 6 compare the effects of removing spatial frictions for East and West Germans. While everyone benefits, East Germans see a larger increase in their output per capita and wages since a sizable share of them move from the East to the high productivity West. Panel (b) of Figure 7 illustrates this move by plotting the semi-CDF of East Germans in each region.[53] The share of East Germans in the West rises significantly.

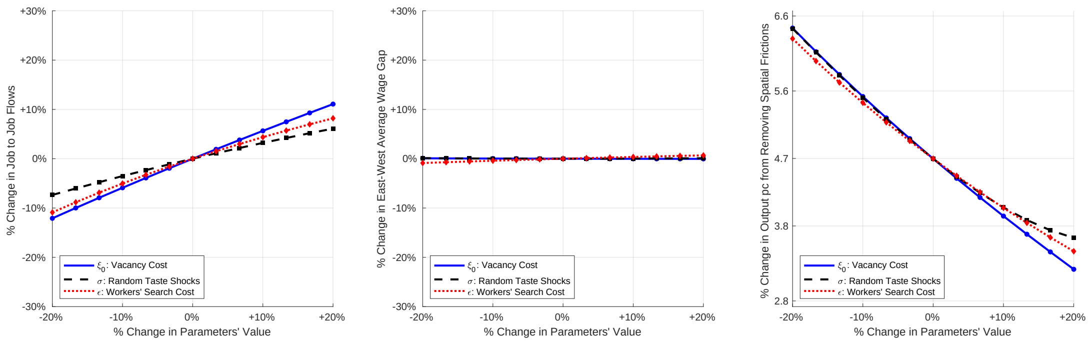
West Germans, as shown in Panel (c) of Figure 7, migrate on net towards the less productive East. Nonetheless, their wage rises because of the equilibrium increase in average wage in both regions, and from the overall improvement in the allocation of labor within region.

Implications for the West-East Gaps. We show in Appendix I that eliminating spatial frictions shrinks the gaps in output, value, and real wages between East and West Germany to 16%, 0.4%, and 14%, respectively. The remaining gaps are due to the average higher productivity of firms in the West, the higher estimated amenity in the East, and the presence of labor market frictions. The higher amenity in the East allows firms there to still retain workers while paying a lower real wage.

[53] Note that each line does not end at one but at the share of East German workers in each region. Adding up the last points on the two solid lines or on the two dashed lines gives one.

Figure 9: Sensitivity of Micro and Macro Moments to Labor Market Parameters

(a) Within-Location Job Flows (b) East-West Wage Gap (c) Output per Capita Gains



Notes: We vary three different parameters modulating the labor market frictions, recompute selected targeted moments, and compare them with the baseline economy. The left panel shows the job to job flows (the lines marked with a cross are the job flows within region) relative to the baseline. The middle panel shows the change in the gap in average wage between West and East Germany relative to the baseline. The right panel shows the overall effect on GDP per capita.

6.3 The Role of the Local Labor Market for Aggregate Gains

Our final key result is that the micro-level details of the labor market matter for the gains of removing spatial frictions due to their impact on the allocation of labor across firms. To show this result we vary, one at a time, three core labor market frictions: (i.) the vacancy cost (ξ_0), which affects the overall mass of vacancies posted by firms; (ii.) the variance of the preference taste shocks (σ), which affects the allocative power of wages; (iii.) the elasticity of workers' search cost (ε), which modulates the ability of workers to move up the job ladder.

Panel (a) of Figure 9 shows that within-location job-to-job flows increase relative to the baseline as we raise each parameter, leading to more within-location reallocation. However, as Panel (b) highlights, changing the labor market frictions has no significant effect on the aggregate wage gap between East and West Germany, consistent with the idea that labor market frictions mainly affect the distribution of labor within, rather than between, regions. In Panel (c), we compute, just as in Section 6.1, the aggregate gains from removing spatial frictions as we vary the degree of labor market frictions in the economy. The aggregate gains in output per capita decline substantially as within-location labor mobility increases. For example, in an economy with 10% higher vacancy costs, the aggregate gains are reduced by a quarter compared to the baseline, from 4.7% to 3.9%.[54] This result is intuitive: higher labor mobility implies smaller potential gains from improving the within-region allocation of labor. This result is also important: ignoring within-location reallocation can lead to wrong assessments about the importance of spatial frictions. Two economies could look identical in terms of their regional wage gaps, yet removing spatial frictions could lead to very different aggregate outcomes dependent on labor market frictions.[55]

7 Conclusion

In this paper, we have shown that taking into account the within-region heterogeneity across firms, and the extent of local labor market frictions, is important to quantify the costs of spatial barriers on the aggregate economy and to understand the mechanisms through which they operate. To reach this conclusion, we design a model which encompasses both spatial and labor market frictions, allowing us to study the joint allocation of labor across firms and locations. Bringing the model to data from Germany, we learn four insights that are relevant beyond this specific context.

First, removing spatial frictions can improve the allocation of workers within locations, leading them to concentrate towards more productive firms and generating aggregate gains. Second, spatial frictions provide firms with local monopsony power and allow unproductive firms to grow. When spatial frictions are removed, the additional competition for workers reallocates jobs towards the most productive firms, possibly generating large aggregate gains, as we find in our context. Third, the aggregate gains from removing spatial frictions can vary substantially across economies dependent on their local labor market frictions, even when these economies have the same wage gap between locations. Analyzing spatial wage gaps without firm-level data may therefore give an incomplete picture. Finally, even in a context, such as ours, in which the within-location reallocation of workers is important for the aggregate gains, reallocation across regions is still important for the distributional effects, as workers born in a low productivity locations are trapped there by spatial frictions.

[54] It is possibly surprising that the effect of varying each source of labor market frictions in panel (c) is similar. There is no fundamental reason why this should be the case, and it is due to the fact that all three frictions have similar impact on labor mobility as shown in Panel (a).

[55] In Supplemental Appendix W we present additional plots of job-to-job movers' average wage gains, the change in workers' value, and East Germans' real wage increase.

References

Abowd, John, Francis Kramarz, and David Margolis (1999). “High Wage Workers and High Wage Firms”. In: *Econometrica* 67.2, pp. 251–333.

Alesina, Alberto and Nicola Fuchs-Schündeln (2007). “Good-bye Lenin (or not?): The effect of communism on people’s preferences”. In: *American Economic Review* 97.4, pp. 1507–1528.

Allen, Treb and Costas Arkolakis (2014). “Trade and the Topography of the Spatial Economy”. In: *The Quarterly Journal of Economics* 129.3, pp. 1085–1140.

Andrews, Martyn J, Len Gill, Thorsten Schank, and Richard Upward (2008). “High wage workers and low wage firms: negative assortative matching or limited mobility bias?” In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171.3, pp. 673–697.

Andrews, Martyn J, Leonard Gill, Thorsten Schank, and Richard Upward (2012). “High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias”. In: *Economics Letters* 117.3, pp. 824–827.

Bachmann, Rüdiger, Christian Bayer, Heiko Stüber, and Felix Wellschmied (2021). “Monopsony Makes Firms not only Small but also Unproductive: Why East Germany has not Converged”. mimeo.

Bagger, Jesper and Rasmus Lentz (2019). “An empirical model of wage dispersion with sorting”. In: *The Review of Economic Studies* 86.1, pp. 153–190.

Bailey, Michael, Patrick Farrell, Theresa Kuchler, and Johannes Stroebel (2020). “Social connectedness in urban areas”. In: *Journal of Urban Economics*, p. 103264.

Baum-Snow, Nathaniel and Ronni Pavan (2012). “Understanding the city size wage gap”. In: *The Review of economic studies* 79.1, pp. 88–127.

BBSR (2009). “Regionaler Preisindex”. In: *Berichte*. Vol. 30. Bonn: BBSR.

Berger, David, Kyle Herkenhoff, and Simon Mongey (2022). “Labor market power”. In: *American Economic Review* 112.4, pp. 1147–93.

Bilal, Adrien (2021). *The geography of unemployment*. Tech. rep. National Bureau of Economic Research.

Bilal, Adrien G, Niklas Engbom, Simon Mongey, and Giovanni L Violante (2019). *Firm and worker dynamics in a frictional labor market*. Tech. rep. National Bureau of Economic Research.

Bird, Richard and Enid Slack (2002). “Land Taxation in Practice: Selected Case Studies. Germany.” World Bank Conference on Innovations in Local Revenue Mobilization.

Boeri, Tito, Andrea Ichino, Enrico Moretti, and Johanna Posch (2021). “Wage equalization and regional misallocation: evidence from Italian and German provinces”. In: *Journal of the European Economic Association* 19.6, pp. 3249–3292.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2019). “A distributional framework for matched employer employee data”. In: *Econometrica* 87.3, pp. 699–739.

Bontemps, Christian, Jean-Marc Robin, and Gerard J Van den Berg (2000). “Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation”. In: *International Economic Review* 41.2, pp. 305–358.

Bradley, Jake, Fabien Postel-Vinay, and Hélène Turon (2017). “Public Sector Wage Policy and Labor Market Equilibrium: A Structural Model”. In: *Journal of the European Economic Association* 15.6, pp. 1214–1257.

Bryan, Gharad and Melanie Morten (2019). “The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia”. In: *Journal of Political Economy*.

Burchardi, Konrad B and Tarek A Hassan (2013). “The economic impact of social ties: Evidence from German reunification”. In: *The Quarterly Journal of Economics* 128.3, pp. 1219–1271.

Burdett, Kenneth, Carlos Carrillo-Tudela, and Melvyn Coles (2020). “The cost of job loss”. In: *The Review of Economic Studies* 87.4, pp. 1757–1798.

Burdett, Kenneth and Melvyn Coles (2003). “Equilibrium wage-tenure contracts”. In: *Econometrica* 71.5, pp. 1377–1404.

Burdett, Kenneth and Dale T. Mortensen (1998). “Wage Differentials, Employment Size, and Unemployment”. In: *International Economic Review* 39.2, pp. 257–273.

Caldwell, Sydnee and Oren Danieli (2021). “Outside options in the labor market”. mimeo.

Caliendo, Lorenzo, Maximiliano A. Dvorkin, and Fernando Parro (2019). "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock". In: *Econometrica* 87.3, pp. 714–835.

Card, David, Jörg Heining, and Patrick Kline (2013). "Workplace Heterogeneity and the Rise of West German Wage Inequality". In: *Quarterly Journal of Economics* 128.3, pp. 967–1015.

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

Chernozhukov, V. and H. Hong (2003). "An MCMC Approach to Classical Estimation". In: *Journal of Econometrics* 115.2, pp. 293–346.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon (2008). "Spatial wage disparities: Sorting matters!" In: *Journal of Urban Economics* 63.2, pp. 723–742.

Datta, Nikhil (2022). "Local Monopsony Power". mimeo, University College London.

Dauth, Wolfgang, Sang Yoon Tim Lee, Sebastian Findeisen, and Tommaso Porzio (2021). "Transforming Institutions: Labor Reallocation and Wage Growth in a Reunified Germany". mimeo.

Eberle, Johanna, Peter Jacobebbinghaus, Johannes Ludsteck, and Julia 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.

Elsby, Michael WL and Axel Gottfries (2022). "Firm dynamics, on-the-job search, and labor market fluctuations". In: *The Review of Economic Studies* 89.3, pp. 1370–1419.

Engbom, Niklas (2020). "Misallocative Growth". mimeo.

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

Fuchs-Schündeln, Nicola, Dirk Krueger, and Mathias Sommer (2010). "Inequality Trends for Germany in the Last Two Decades: A Tale of Two Countries". In: *Review of Economic Dynamics* 13.1, pp. 103–132.

Galenianos, Manolis, Philipp Kircher, and Gabor Virag (2011). "Market power and efficiency in a search model". In: *International economic review* 52.1, pp. 85–103.

Gelman, Andrew, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin (2013). *Bayesian data analysis.* Chapman and Hall/CRC.

Gollin, Douglas, David Lagakos, and Michael E. Waugh (2014). "The Agricultural Productivity Gap". In: *Quarterly Journal of Economics* 129.2, pp. 939–993.

Henkel, Marcel, Tobias Seidel, and Jens Suedekum (2021). "Fiscal transfers in the spatial economy". In: *American Economic Journal: Economic Policy* 13.4, pp. 433–68.

Herrendorf, Berthold and Todd Schoellman (2018). "Wages, Human Capital, and Barriers to Structural Transformation". In: *American Economic Journal: Macroeconomics* 10.2, pp. 1–23.

Hethey-Maier, Tanja and Johannes 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, Joan Hamory, Marieke Kleemans, Nicholas Y Li, and Edward Miguel (2017). *Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata.* Tech. rep. National Bureau of Economic Research.

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

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow (2019). "The allocation of talent and us economic growth". In: *Econometrica* 87.5, pp. 1439–1474.

Hsieh, Chang-Tai and Peter J. Klenow (2009). "Misallocation and Manufacturing TFP in China and India". In: *Quarterly Journal of Economics* 124.4, pp. 1403–1448.

Jäger, Simon, Benjamin Schoefer, Samuel Young, and Josef Zweimüller (2020). "Wages and the Value of Nonemployment". In: *The Quarterly Journal of Economics* 135.4, pp. 1905–1963.

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

Kennan, John and James R Walker (2011). "The effect of expected income on individual migration decisions". In: *Econometrica* 79.1, pp. 211–251.

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

Krueger, Alan B. and Jörn-Steffen 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 Richard B. Freeman and Lawrence F. Katz. University of Chicago Press, pp. 405–446.

Lagakos, David, Samuel Marshall, Ahmed Mushfiq Mobarak, Corey Vernot, and Michael E Waugh (2020). "Migration costs and observational returns to migration in the developing world". In: *Journal of Monetary Economics*.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet (Oct. 2020). "Gender Differences in Job Search: Trading off Commute against Wage". In: *Quarterly Journal of Economics* 136.1, pp. 381–426. DOI: 10.1093/qje/qjaa033.

Lentz, Rasmus and Dale T Mortensen (2012). "Labor Market Friction, Firm Heterogeneity, and Aggregate Employment and Productivity". In: mimeo, University of Wisconsin.

Lindenlaub, Ilse, Ryungha Oh, and Michael Peters (2022). *Firm Sorting and Spatial Inequality*. Tech. rep. National Bureau of Economic Research.

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin (2016). "Matching, sorting and wages". In: *Review of Economic Dynamics* 19, pp. 63–87.

Manning, Alan and Barbara Petrongolo (2017). "How local are labor markets? Evidence from a spatial job search model". In: *American Economic Review* 107.10, pp. 2877–2907.

Martellini, Paolo (2022). *Local labor markets and aggregate productivity*. Tech. rep. Working Paper.

Meghir, Costas, Renata Narita, and Jean-Marc Robin (2015). "Wages and informality in developing countries". In: *American Economic Review* 105.4, pp. 1509–46.

Melitz, Marc J (2003). "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity". In: *econometrica* 71.6, pp. 1695–1725.

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

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

Moser, Christian and Niklas Engbom (2021). "Earnings Inequality and the Minimum Wage: Evidence from Brazil". mimeo, Columbia University.

Paetzold, Jörg and Markus Tiefenbacher (2018). "Distributional and Revenue Effects of a Tax Shift from Labor to Property". In: *International Tax and Public Finance* 25.5, pp. 1215–1251.

Pavcnik, Nina (2002). "Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants". In: *The Review of economic studies* 69.1, pp. 245–276.

Petrongolo, Barbara and Christopher A Pissarides (2001). "Looking into the black box: A survey of the matching function". In: *Journal of Economic literature* 39.2, pp. 390–431.

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

Roca, Jorge De La and Diego Puga (2017). "Learning by working in big cities". In: *The Review of Economic Studies* 84.1, pp. 106–142.

Schmutz, Benoît and Modibo Sidibé (2018). "Frictional Labour Mobility". In: *The Review of Economic Studies* 86.4, pp. 1779–1826.

Uhlig, Harald (2006). "Regional Labor Markets, Network Externalities, and Migration: The Case of German Reunification". In: *American Economic Review* 96.2, pp. 383–387.

— (2008). "The Slow Decline of East Germany". In: *Journal of Comparative Economics* 36.4, pp. 517–541.

Valentinyi, Akos and Berthold Herrendorf (2008). "Measuring factor income shares at the sectoral level". In: *Review of Economic Dynamics* 11.4, pp. 820–835.

Van Den Berg, Gerard J. and Geert Ridder (1998). "An Empirical Equilibrium Search Model of the Labor Market". In: *Econometrica* 66.5, pp. 1183–1221.

Online Appendix

A Further Details on Data and Data Construction

In this section, we discuss the main variables and data construction steps of the paper. We provide detail on other variables in Supplementary Appendix K.[56]

We use the Establishment History Panel (BHP) version 7514, covering the years 1975-2014. We use the longitudinal model of the Linked Employer-Employee Data (LIAB), version 9314, covering 1993-2014. IAB regulations do not allow us to merge these datasets. 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.

BHP Data. The BHP is a 50% sample of all establishments in Germany with at least one employee subject to social security as of 30th of June of a given year. The data are reported as a panel dataset at the establishment-year level. As in the main text, we refer to establishments as "firms" going forward.

We obtain for each firm the location at the county level and the number of full-time workers, as well as the share of full-time workers by gender, education, and age. We create a dummy for whether a firm is in East Germany, and we code the dummy as missing if the firm is in Berlin. We obtain the mean gross daily wage paid to full-time employees by each firm in each year. Since the data contain earnings only 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, which we use for all our analyses. The imputation procedure follows Card et al. (2015). We use the time-consistent 3-digit industry codes at the WZ93 level for each firm. These time-consistent codes were constructed by Eberle et al. (2011) and are provided to us by the IAB.

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 and remove firms located in Berlin, which reduces the sample size by 3.8 million and 200,000, respectively. We

[56] This Supplemental Appendix is not meant for publication and includes additional material to provide context or robustness checks. It is available on the authors' websites.

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. We present some summary statistics in Supplementary Appendix K.

LIAB Data. The LIAB data provide matched employer-employee data that link more than 1.9 million individuals to about 400,000 firms. 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.

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. About 10% of worker-start date-end date episodes are associated with multiple spells (7% if we exclude part-time work, which is our sample below). We replace partially overlapping employment spells with artificial observations with new dates so that completely parallel and completely non-overlapping periods are created. We keep only the worker's highest-paying job in cases of completely overlapping spells. This main job, on average, accounts for 81% of the worker's period income (median: 86%).

We obtain an individual's daily wage or unemployment benefit. As in the BHP, earnings are only reported up to the upper earnings limit for statutory pension insurance contributions. 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 et al. (2015).

We obtain each worker's county of residence, which is available since 1999, and for employed workers the county of their job. We set each individual's birth county as the earliest available county of residence or county of work recorded for the worker, from any record, including part-time or unemployed. If the earliest county of work and county of residence are from the same spell, we use the county of residence. We compute the distance between any county pair from Google maps, using the mid point of the counties.

We 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), as well as a gender dummy and a dummy for whether the worker has a college education. The education variable is less than 85% complete for employed workers and unavailable for unemployed workers. We therefore set the dummy to zero if education is missing and include in our analyses an additional dummy to capture missing cases.

Our baseline analysis contains 15.1 million employment or unemployment spells for our baseline period 2009-2014. We drop part-time workers, which removes 5.0 million spells. 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 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. Supplementary Appendix K provides some summary statistics.

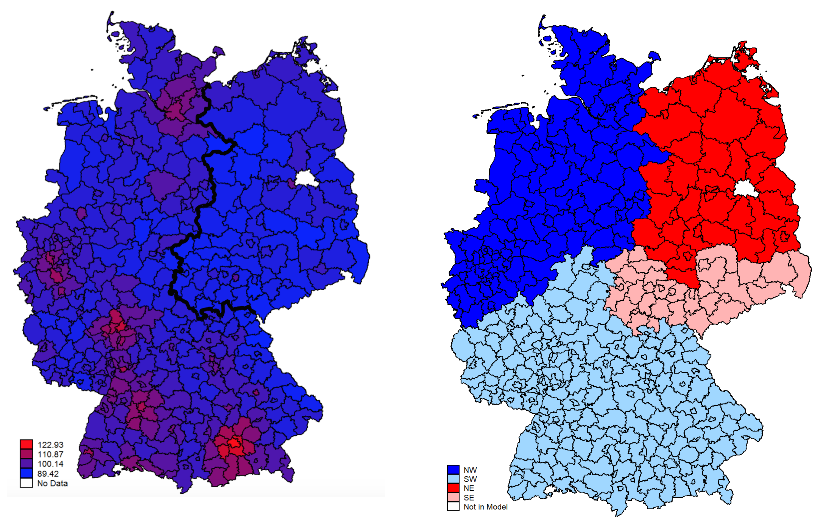
We obtain firm-level information from the matched BHP data for firms in which at least one worker in the LIAB has an employment spell. As in the BHP above, we keep only firms with at least one full-time worker, which reduces the number of firm observations from originally 2.4 million to 2.0 million. The matched sample contains about 40% of the firm-year observations of our BHP sample above. We present some summary statistics in Supplementary Appendix K. Due to the smaller size of the LIAB-BHP sample, we rely on the BHP sample to compute the firm-level moments we use in our model estimation.

Price Deflators. We obtain regional price data 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. We obtain 2007 prices for each of the 402 counties in the IAB data (shown in Figure A1a), and then apply the price deflator of the corresponding state from the growth accounting of the states to each county to obtain a county-level price index for each year in 2009-2014.

Locations for the Quantitative Estimation. In the quantitative estimation of the model we divide Germany into four locations. Figure A1b visualizes the locations. In Supplemental Appendix K, we provide further summary statistics.

Figure A1: Price Level and Locations

(a) Price Level, 2007 (b) Locations in the Estimation



Source: BBSR, authors' calculations. Note: The left figure plots the price level in 2007 for each county, in euros valued in Bonn, the former capital of West Germany, from the BBSR. The right figure presents the geography of the four locations used in the estimation.

B Statistics on Worker Mobility

Mobility Across Regions. We provide some additional statistics on worker mobility across regions. The top part of Table A1 presents the number of cross-region movers in our core sample for migrants, which, as defined in the main text, change their residence (column 1). We also show statistics for all job-to-job switchers across regions (column 2). We find that about 80% of cross-region job moves are done without a reported change in residence. We refer to such moves as "commuting".

Since individuals may not list the residence closest to their job in the case of multiple residences, there may be mismeasurement in commuting. We therefore define a third, "intermediate" version of cross-region migration as all migration moves plus all cross-region job switches without a change in residence that take the worker further away from her current residence, as long as the work county is within 200km of the residence county both before and after the move. We impose this upper bound on distance to remove workers with implausibly long commutes. Moreover, we require the distance to the residence to increase since moves that decrease the distance do not really impose a moving cost on the worker.

The bottom panel of Table A1 shows percentiles of the distance between the origin and the destination job for cross-region job-to-job movers ("Work") and between the worker's new job and her residence after the move ("Live"). Workers move on average about 300km between jobs, with some job switchers moving more than 500km. Most workers live close to their job; however, some workers in the tail report distances to their residence of 400km or more. These workers likely have a misclassified residence county. We alleviate this issue somewhat with our intermediate definition.

Figure A2 presents the time series of the share of workers that is employed or unemployed away from their home region (circles) and the share of workers that are living away from home (triangles). We find 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.

Supplemental Appendix N contains additional statistics on the characteristics of migrants (analogous to Table 1). It also contains the distribution of moves throughout workers' life time and mobility by cohort.

Worker Mobility Across Locations. Table A2 presents statistics on worker mobility across the locations used in the quantitative estimation of the model. Similar to above, we distinguish between migrants, which change their residence location (column 1), all job-to-job switchers across locations (column 2), and an "intermediate version" of cross-location migration (column 3), which contains all migration moves plus all cross-location job switches without a change in residence that take the worker further away from her current residence, as long as the work county remains within 200km of the worker's residence county. We use this intermediate definition of a cross-location move in our estimation in Section 5.

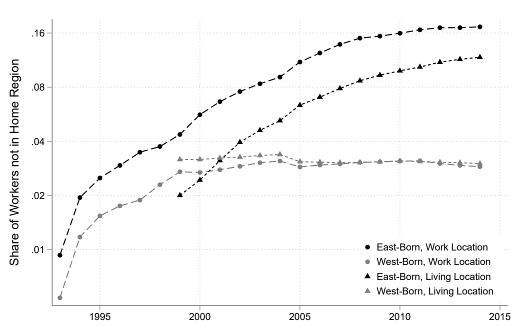
Table A1: Number of Movers Between East and West Germany

	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 region 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 cross-region moves that increase the distance to the residence county, as long as the distance to the residence county 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.

Figure A2: Stock of Individuals Away from Home Region



Source: LIAB. Notes: The circles plot the share of workers of a given type that are 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.

Table A2: 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 cross-location moves that increase the distance to the residence county, as long as the distance to the residence county 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.

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 birth region. The SOEP data consist of samples drawn in different "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, identifying the birth region 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, we use information on the individuals' schooling. We code the home region as the location of the individual's earliest observed non-tertiary schooling. We refer to individuals where we have this information as the "Young SOEP Sample". While the SOEP also asks some individuals about their place of residence in 1989, that variable is only available for about 0.5% of individual-year observations.

We construct an imputed home region in the same way and subject to the same restrictions as in the LIAB. Table A3 compares the imputed and actual home region for individuals that are in the labor force in 2009-2014. We find that in both samples the imputed and the actual home region match closely.

Table A3: Fraction of Individuals Where Imputed Home Region Matches Actual

	Old SOEP Sample		New SOEP Sample	
	East	West	East	West
	(1)	(2)	(3)	(4)
Imputed = Actual	.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 onward, exclude Berlin, and drop residence information prior to 1999. We then use the worker's region 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 region prior to 1999, to assign an imputed home region. We compare this imputed home region to the actual 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.

As a more rigorous test, 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.

Table A4 shows the results. The wage gap is similar under both the true and the imputed region 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 A4: 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)
$\mathbb{I}_{i,East,imp}$	$-\mathbf{.346}^{***}$	$-\mathbf{.404}^{***}$			$-\mathbf{.160}^{***}$	$-\mathbf{.163}^{***}$		
	(.0212)	(.0196)			(.0325)	(.0309)		
$\mathbb{I}_{i,East,true}$			$-\mathbf{.338}^{***}$	$-\mathbf{.406}^{***}$			$-\mathbf{.133}^{***}$	$-\mathbf{.127}^{***}$
			(.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 and Additional Formulas

D.1 Equilibrium in the Goods Market

The firm's problem in the goods market is

$$\hat{\pi}_j(w) = \max_{n_h, n_c, k} \left\{ pn_c + P_{h,j} \left(pn_h \right)^{1-\alpha} k^{\alpha} - \rho_j k \right\} \tag{20}$$

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

$$n_h = \frac{\rho_j}{p} \frac{1-\alpha}{\alpha} k \tag{21}$$

and assuming that both goods are supplied in equilibrium

$$P_{h,j} = \rho_j^{\alpha} \alpha^{-\alpha} \left(1-\alpha\right)^{-(1-\alpha)}. \tag{22}$$

We can plug (21) and (22) into (20) to obtain

$$\hat{\pi}_j(w) = pn_j(w) = p \sum_{i \in \mathbb{I}} \theta_j^i l_j^i(w), \tag{23}$$

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

The equilibrium price of the local good is determined from consumers' demand and market clearing. Due to the Cobb-Douglas utility, the aggregate demand for the local good H_j satisfies

$$P_{h,j} H_j = (1 - \eta) P_j Y_j, \tag{24}$$

where, assuming that consumers own the firms and using (23), their total income is

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

and Y_j is real GDP.

Using the production function $h = (pn_h)^{1-\alpha} k^\alpha$, and plugging in (21), aggregate supply of the local good in location j is $H_j = (\rho_j \frac{1-\alpha}{\alpha})^{1-\alpha} K_j$, which, using the price of the local good (22), implies

$$P_{h,j} H_j = \frac{1}{\alpha} \rho_j K_j. \tag{25}$$

Combining demand and supply yields

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

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 (22) and use the supply equation (25) 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 (24) gives

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

as claimed in the main text.

D.2 Additional Formulas

The probability that a worker of type i employed at wage w in region j accepts an offer w' from region x is

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

The corresponding probability for an unemployed worker is

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

D.3 Proof of Proposition 1

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

$$\pi_j\left(p\right) = \max_w \left(p - w\right) \sum_{i \in \mathbb{I}} \theta_j^i l_j^i\left(w\right) \tag{26}$$

and, as shown in equation (17),

$$l_j^i\left(w\right) = \frac{\mathcal{P}_j^i\left(w\right) \vartheta_j^{-\chi \frac{\bar{a}_j^i}{\bar{a}_j}}}{q_j^i\left(w\right)} \quad \text{if } w \geq R_j^i \tag{27}$$

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 $\pi_j(p, w)$ is continuous in w for a given p, then we can take the first order condition of problem (26) and obtain

$$\frac{(p - w_j(p))\left(\sum_{i\in\mathbb{I}} \theta_j^i \frac{\partial l_j^i(w_j(p))}{\partial w}\right)}{\left(\sum_{i\in\mathbb{I}} \theta_j^i l_j^i(w_j(p))\right)} \;=\; 1. \tag{28}$$

From equation (27), 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 (28) to find a differential equation for $w_j(p)$

$$\frac{\partial w_j(p)}{\partial p} = \frac{(p - w_j(p))\left(\sum_{i\in\mathbb{I}} \theta_j^i \frac{\frac{\partial \tilde{\mathcal{P}}_j^i(p)}{\partial p}\tilde{q}_j^i(p) - \tilde{\mathcal{P}}_j^i(p)\frac{\partial \tilde{q}_j^i(p)}{\partial p}}{\tilde{q}_j^i(p)^2}\vartheta_j^{-\chi}\frac{\bar{a}_j^i}{\bar{a}_j}\right)}{\left(\sum_{i\in\mathbb{I}} \theta_j^i \frac{\tilde{\mathcal{P}}_j^i(p)}{\tilde{q}_j^i(p)}\vartheta_j^{-\chi}\frac{\bar{a}_j^i}{\bar{a}_j}\right)}. \tag{29}$$

Since $w_j(p)$ is continuous at p by assumption, the differential equation (29), together with an appropriate boundary conditions, characterizes the optimal wage at p. Since workers can always voluntarily separate into unemployment while keeping their preference shocks, they must be paid at least $w = R_j^i$. Therefore, the boundary conditions are given by

$$w_j\left(\underline{p}_j\right)=\max\left\{\min_{i\in\mathbb{I}} R^i_j, \arg\max_{\hat{w}}\left(\underline{p}_j-\hat{w}\right)\sum_{i\in\mathbb{I}}\theta^i_j l^i_j\left(\hat{w}\right)\right\}.$$

We have thus proved that

$$w_j\left(p\right) \;=\; w_j\left(\underline{p}_j\right)+\int_{\underline{p}_j}^{p}\frac{\partial w_j\left(z\right)}{\partial z}\gamma_j\left(z\right)dz \tag{30}$$

as claimed.

E Parameters and Empirical Moments

We describe how the parameters and targeted moments are computed. We provide extensive details in Supplemental Appendix Q.

Calibrated Parameters

We estimate workers' skills θ^i from an AKM model with comparative advantage term, building on Abowd et al. (1999) and Card et al. (2013). We estimate in the LIAB data the following model with two regions, East and West Germany:

$$\log(w_{it})=\alpha_i+\psi_{J(i,t)}+\beta\mathbb{I}^{(h_i\neq R(J(i,t)))}+BX_{it}+\epsilon_{it}, \tag{31}$$

where i indexes full-time workers, t indexes time, and $J(i,t)$ indexes worker i's firm at time t.[57] 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 et al. (2013). We discuss the identification of this model in the dedicated appendix F below.

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

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

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

We estimate the model on the largest connected set of workers in our data.[58]

The estimation yields a comparative advantage estimate of $\beta = 0.019$, indicating a small *negative* comparative advantage towards the home region. Given the lack of comparative advantage at the regional level where we would expect to find the largest effect, we do not extend the analysis to the level of the four finer locations we use in the estimation in Section 5. However, the same insights and identification strategy would apply and could be performed. 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 workers' absolute advantage from the average worker fixed effect for each worker type, see Supplemental Appendix Q.1.

Table A5: Calibrated Parameters

	Parameter	Computed from...
(2)	M_j : Firms by location	Number of firm-year observations in the BHP
(3)	$\bar{D}^i$: Workers' home location	Population residing in each location in January 1991 from the Growth Accounting of the States
(4)	δ_j : Separation rate by location	Workers' monthly probability of separating into unemployment or permanent non-employment from LIAB
(5)	P_j : Price level by location	Weighted average of state-level price indices from BBSR
(6)	$\alpha(1-\eta)$: Pay to fixed factors	Interpret fixed factor as land, use share of land in GDP for the U.S. from Valentinyi and Herrendorf (2008)
(7)	χ : Matching elasticity	Assume CRS, set by assumption.
(8)	r : Interest rate	Assume infinitely lived individuals, set by assumption.

Notes: This table provides a brief summary how each calibrated parameter is computed. Details are in Supplemental Appendix Q.

[58] 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 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 large firms are more likely to be included in the connected set.

Table A5 contains a brief discussion of the remaining parameters.

Moments for the Estimation

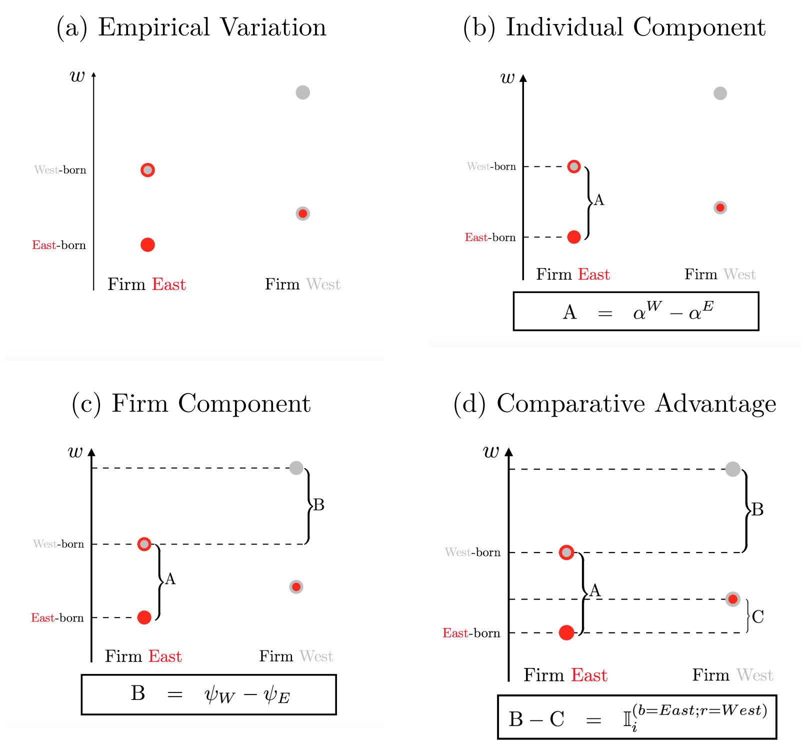
Unless otherwise mentioned, all moments are constructed using the cleaned data described in the data section of the main text, for the core sample period 2009-2014. We briefly summarize the construction of the moments in Table A6, and delegate details to Supplemental Appendix Q.2.

Table A6: Targeted Moments

	Parameter	Computed from...
(1)+(2)	Wage gains w/i + b/w locations	Regression in LIAB of movers' wage gains on dummies for moves w/i and b/w locations and conrols.
(3)+(4)	Job flows w/i + b/w locations	Monthly worker flows w/i and b/w locations from LIAB.
(5)+(6)	(Un)employment shares	Share of (un)employed workers living in each location in LIAB.
(7)	Firm component of wages	Regression of AKM firm fixed effects on worker location and home location dummies
(8)	Avg. firm component of wages	Regression of AKM firm fixed effects on firm location dummies
(9)+(10)	Output p.c. and unemployment	National Accounts of the States and official unemployment statistics
(11)	Deciles of firm-size distribution	Share of workers employed at each firm size decile from BHP, with firm size residualized by age, education, gender, industry.
(12)	Slope of wage vs firm size	Coefficient of regression of firms' (residualized) average wage on (residualized) firm size from BHP.
(13)	Slope of J2J wage gain vs wage	Coefficient of regression of workers' (residualized) job-job wage gain on their (residualized) initial wage in LIAB.
(14)	Slope of separation rate vs firm wage	Coefficient of regression of (residualized) dummy for worker separation on their (residualized) initial wage in LIAB.
(15)	Std of job-job wage gains	Std of (residualized) wage gains of job movers in LIAB.
(16)	Profits to labor cost ratio	Pre-tax profits divided by total labor costs from ORBIS.

Notes: Table provides a summary of how each moment is computed. Details are in Supplemental Appendix Q.

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 Identification of Workers' Skills

We now discuss how the specification (31) allows us to identify, through β, the comparative advantage effect by region. The same idea 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, she 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. As a result, the estimated β captures the sum of the two comparative advantages and we need to make an assumption in order to separately identify the two. In practice, we side-step this issue since we do not find evidence of comparative advantages as described above.

G Model's Computation and Estimation

We here provide a brief explanation of the solution algorithm and more details on the estimation approach and outcomes. Additional details, with a complete description of the solution and estimation algorithm are found in Supplemental Appendix S.

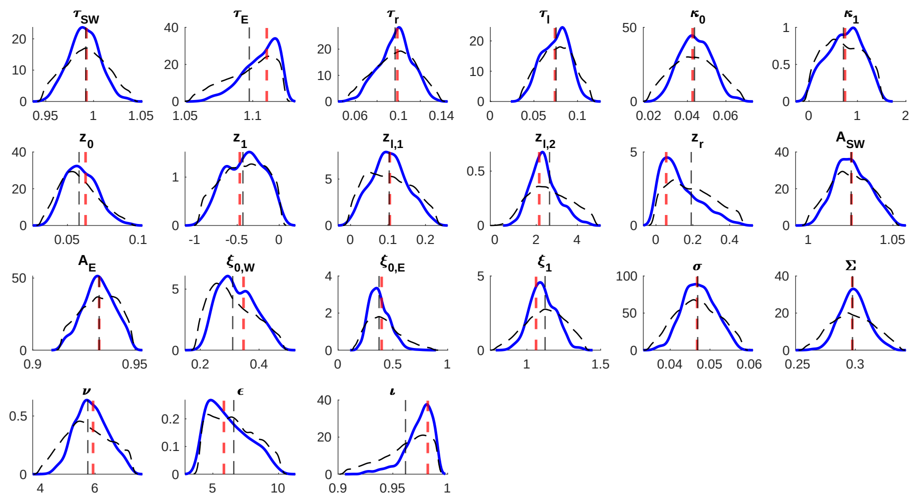
Computation. To solve the model, we follow a nested iterative procedure which leverages Proposition 1 to solve the model in the one-dimensional productivity space. For the counterfactuals, we follow the same algorithm, but keep the unemployment benefit constant (rather than estimating it to match the reservation wage), and allow equilibrium prices to respond to local GDP (rather than simply using their empirical values, which we do in the estimation).

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

$$\phi^* \quad = \quad \arg\min_{\phi \in \mathbb{F}} \sum_x \left[\omega_x \left(T_x \left(m_x \left(\phi \right), \hat{m}_x \right) \right)^2 \right] \tag{32}$$

and $\mathbb{F}$ is the set of admissible parameter vectors. The choice of the function $T_x(\cdot)$ minimizes either the sum of the percentage deviations between model-generated and empirical moments or log differences, as explained in the Supplemental Appendix S.

Figure A4: 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.

We introduce a weighting factor ω_x to give equal weight to each one of the 14 groups of parameters that we target, shown in Table 4.

We then use a fairly standard simulation-based minimization routine to solve the minimization problem. Figure A4 illustrates our approach and how it slightly differs from others such as Jarosch (2016) and Lise et al. (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. We instead pick the parameters following Moser and Engbom (2021), and thus select the vector of parameters that minimizes the objective function among all our draws. 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,

Table A7: 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.

tightly identified.

All the estimated parameters, corresponding to the vertical dotted red lines, are included in Table A7.

Jacobian Matrix and Identification. To formally explore the connection between parameters and moments, we compute the elasticity of each (model-generated) moment to each model parameter.

Specifically, we start from the estimated vector of parameters ϕ^*, and we create 42 alternative vectors, two for each parameter j, as follows: $\underline{\phi}(j) = \left\{\phi^*_{-j}, 0.95\phi^*_j\right\}$ and $\overline{\phi}(j) = \left\{\phi^*_{-j}, 1.05\phi^*_j\right\}$, where $\underline{\phi}(j)$ keeps all parameters except for j constant and decreases j by 5%, while $\overline{\phi}(j)$ does the same, but increasing j by 5%.

We then compute with our model the vectors of moments corresponding to each vector of parameters and use them to compute

$$\Delta_{jr} = m_r\left(\overline{\phi}(j)\right) - m_r\left(\underline{\phi}(j)\right).$$

Thus, Δ_{jr} measures how much moment r would change if we changed parameter j by 10% around the estimated value while keeping all the other parameters constant.

Overall, we have 305 moments and 21 parameters, which would create a matrix with 6,405 cells; hence, impossible to read. Therefore, for the exposition we reduce the dimensionality by taking averages by groups of moments and parameters that are similar. Specifically, for the moments, we follow Table 4, and compute the averages by the 16 blocks shown there. For the parameters, we bundle together the following: i. the two relative amenities τ_{SW} and τ_E (referred to as τ_j in Figure A5); ii. the two home biases τ_l and τ_r $\left(\tau_j^i\right)$; iii. the relative search efficiencies between regions z_0, z_1, $z_{l,2}$ and z_r $\left(z_{jx}^i\right)$; iv. the cost of moving κ_0 and κ_1 (κ); v. the two relative productivities A_{SW} and A_E (A); vi. the two costs of vacancy posting $\xi_{0,W}$ and $\xi_{0,E}$ (ξ_0). In this way, we reduce the number of parameters to be shown to 13.

To ease comparison, we normalize Δ_{jr} for each parameter j so that, when rounded, it sums to 32 across all moments: $\sum_r Round\left(\Delta_{jr}\right) = 32$, i.e., twice the number of moment blocks. The result of this procedure is the Jacobian matrix shown in Figure A5, which illustrates which parameter is most important for each moment. Our normalization helps to generate interpretable magnitudes: if all moments are impacted in the same way by a specific parameter, then we should see a value of 2 for each parameter in the corresponding row; if only four moments are impacted by a parameter, with equal relevance, then we should see a value 8 for those moments and 0 otherwise, and so on.

Figure A5: Normalized Partial Derivatives of Moments with Respect to Parameters

	τ_j	τ_j^i	z_{jx}^i	z_{jj}^j	κ	A	ξ_0	ξ_1	σ	ν	ϵ	Σ	ι
Wage gains, w/i locations	1	2	1	0	0	0	0	1	3	0	2	6	2
Wage gains, b/w locations	2	4	3	2	7	1	1	1	2	0	0	4	1
Job flows, w/i locations	2	0	1	2	0	0	5	1	1	1	6	1	0
Job flows, b/w locations	3	4	9	3	6	1	3	0	0	0	4	0	0
Employment shares, by (i,j)	5	5	4	6	6	2	2	0	1	0	1	0	0
Unemployment shares, by (i,j)	4	5	5	6	6	2	2	0	1	0	2	0	0
Firm wages, by (i,j)	4	3	3	2	3	9	1	1	2	0	2	1	1
Average firm wage, by j	4	3	2	4	1	10	1	1	1	0	2	1	1
Output per capita, by j	0	4	2	4	3	7	2	0	0	8	3	0	0
Unemployment rate, by j	1	0	0	1	0	0	4	2	1	22	3	1	0
Firm size distributions	1	0	0	0	0	0	1	9	3	0	1	3	0
Relationship firm wage & size	0	0	0	0	0	0	0	9	0	0	0	4	21
Wage gains, by firm wage	1	1	1	0	0	0	0	1	5	0	0	6	0
Workers separation, by firm wage	1	0	0	1	0	0	5	2	4	1	2	0	1
Standard deviation of wage gains	1	0	0	0	0	0	3	1	2	0	3	5	3
Profit shares, by j	2	1	1	1	0	0	2	3	6	0	1	0	2

Notes: The matrix includes the normalized values of Δ_{jr} computed as described in the text. Each row is a block of moments and each column represent one or more parameters.

H 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 here in figures, we explicitly present their (305) numerical values in tables in Supplemental Appendix U.

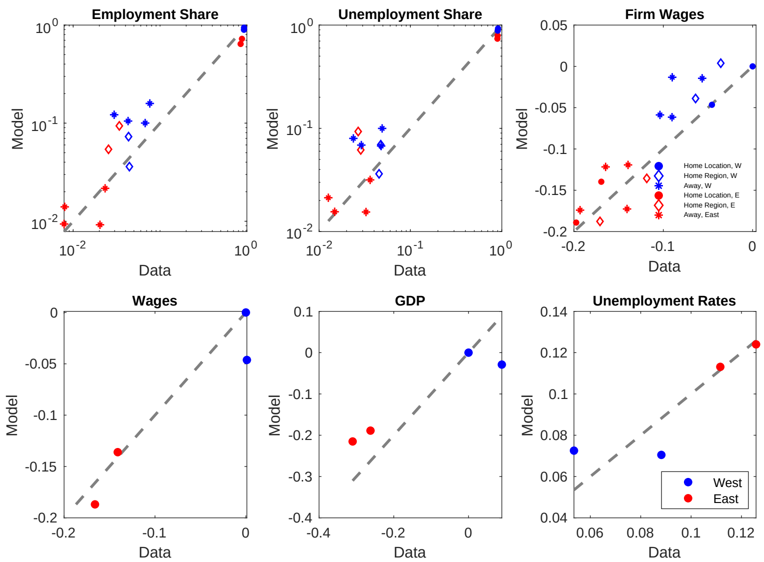
Figure A6 shows that the model fits well the empirical moments on distribution of employment, output and wages across locations and by workers types. Each panel plots a set of moments in the data (x-axis) against their values in the model (y-axis), with the 45-degree line indicating a perfect fit.

Figure A7 plots the firm size distributions in each location in the model and in the data. 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 A8 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 core mechanism of the model generates a positive relationship between firm size and firm wage (row 1 of Table A8), 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 non parametrically, for each of the four locations. 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 A8). 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

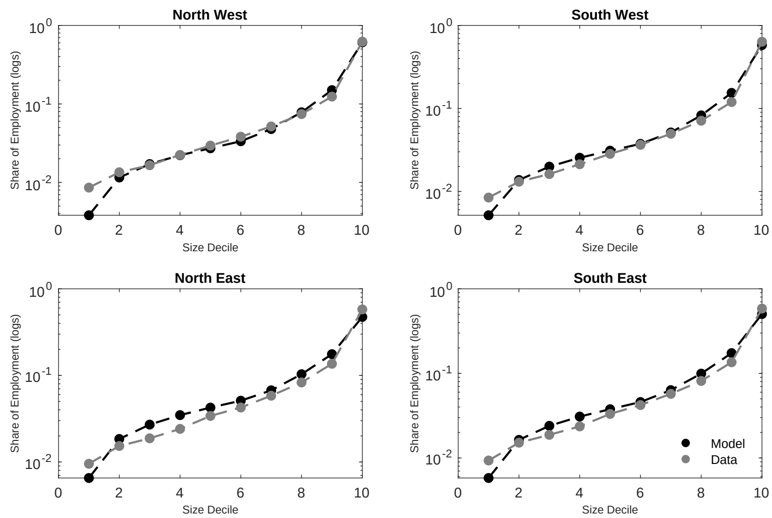
Figure A6: Employment, Wages, and GDP by Location and Worker-Type



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 Supplemental Appendices Q.2.3 to Q.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.

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 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 Supplemental Appendix Q.2.9.

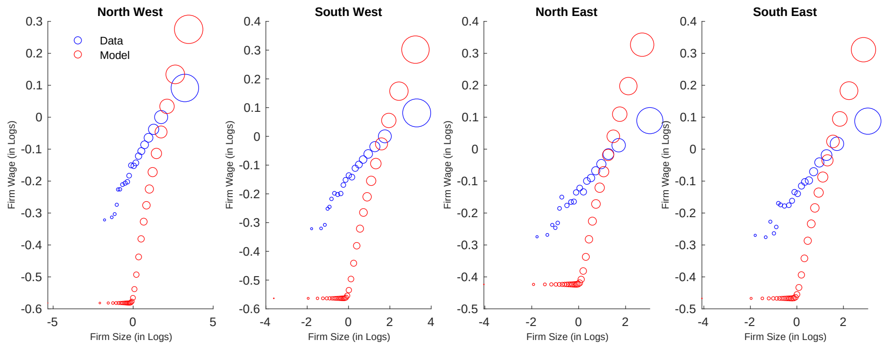
Table A8: Model Fit for Additional Moments

	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

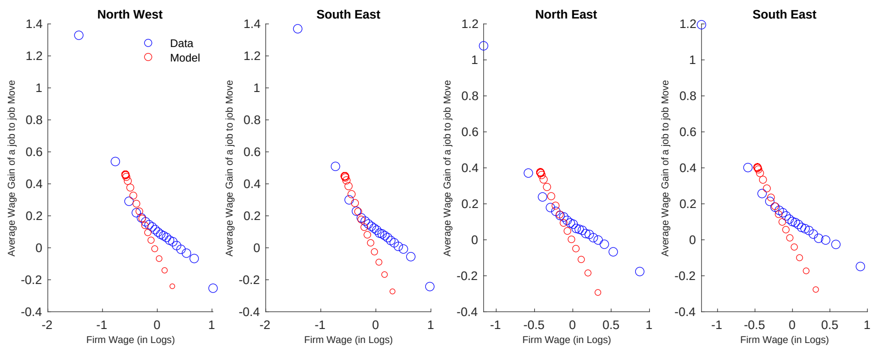
Notes: The table compares several moments in the model to their data analogues, by location of the firm. The construction of these moments is described in Supplemental Appendices Q.2.10 to Q.2.14. The first row shows the slope of the wage function with respect to firm size. The second row presents the 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 U. The last row shows the average ratio of profits to labor costs in each location.

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

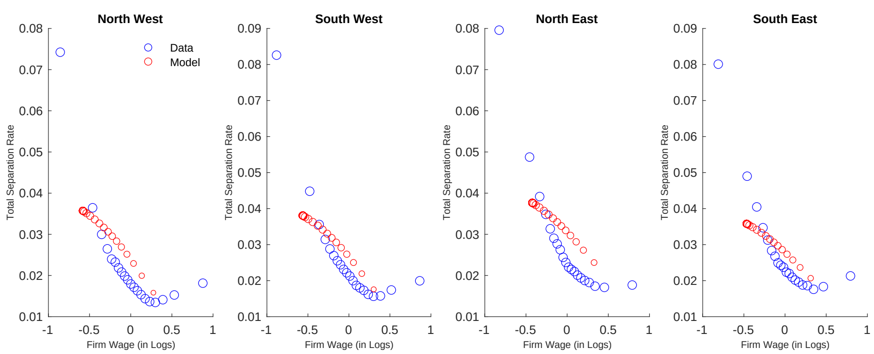
(a) Relationships between Firm Sizes and Average Wages



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

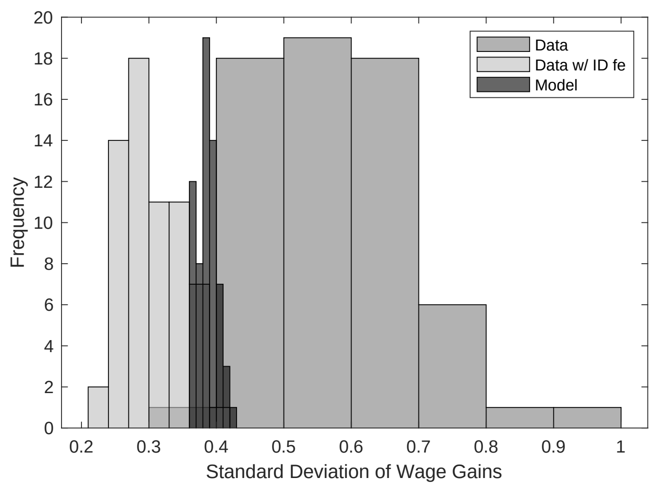


(c) Relationships between Firm Wages and Separation Rates



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 Supplemental Appendices Q.2.10 to Q.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 Supplemental Appendix Q.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 Supplemental Appendix Q.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.

I Additional Quantitative Results

Robustness. We explore the quantitative role of two key assumptions of our model: (i.) there are only two locations in each region; (ii.) the locations in East Germany are smaller, hence have fewer firms and workers.

First, we vary the number of locations. We solve a version of our model in which we split each of the four locations in the benchmark model into either 2, 3, 4, 5, or 6 sub-locations. We randomly draw each sub-location k's average firm productivity, $A_j(k)$, from a normal distribution with mean equal to the overall location's estimated productivity, A_j, and standard deviation equal to the East-West productivity gap to allow for possibly large gains from the reallocation of labor across the sub-locations. We keep the spatial frictions exactly as estimated in the baseline, and we split the workers' types to match the new locations.[59]

[59]Two complications arise. First, we need to recompute the distance between the new sub-locations. Given the scope of this exercise, we keep the average distance between locations as in

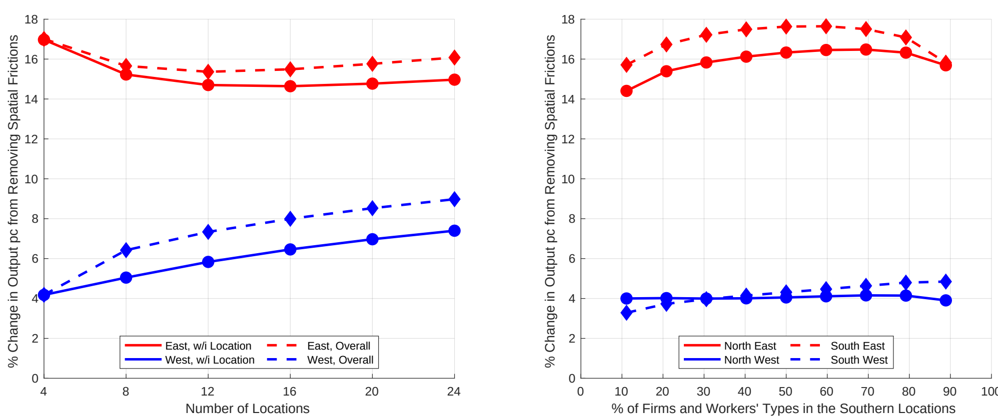
Figure A10a shows that we continue to find large gains from the within-location reallocation of labor even as we increase the number of locations. Intuitively, there is significant scope for within-location reallocation due to substantial heterogeneity across firms, and congestion forces due to prices and labor market tightness limit the gains from labor reallocation across space.

Second, we vary the labor market size. We proportionally vary the mass of firms (M_j) and workers $\left(\bar{D}_j^i\right)$ that are in the South versus in the North in both East and West Germany, keeping the total mass of workers and firms in the overall region and the other structural parameters constant. Figure A10b shows that increasing the mass of workers and firms in the South relative to the North has only small effects on the aggregate gains in both locations. While removing spatial frictions gives firms in smaller locations a bigger relative increase in the mass of workers that can now apply to their vacancies, they also face a relative bigger increase in competition. These two effects roughly balance each other out.

Figure A10: Aggregate Cost of Spatial Frictions as a Function of Size and Number of Locations

(a) Number of Locations

(b) Location Size



Notes: The left panel shows the change in output per capita from removing spatial frictions computed for East Germany (in red) and West Germany (in blue) as we vary the overall number of locations. The solid lines show the average of the gains from within-location reallocation across all locations in the region. The dashed lines show the total gains, including from reallocation across locations. The right panel shows the change in the output per capita for the two locations in the East (in red) and the two in the West (in blue) plotted as a function of the share of the population in the Southern locations.

the baseline, and we assign the sub-locations to be equally distanced on a line. Second, we need to re-normalize the search productivity z as we vary the number of sub-locations, since otherwise the overall ability of workers to search would scale up. We proportionally scale all z_{jx}^i so that $\sum_{x \in \mathbb{J}} z_{jx}^i$ is constant across all scenarios.

Regional Gaps. Table A9 provides another perspective on the results by showing the percentage differences in our variables of interest between East and West Germany and between East and West German workers. Eliminating spatial frictions shrinks the gaps in output, value, and wages considerably, but does not eliminate them. The remaining East-West gap is due to the average higher productivity of firms in the West, the higher estimated amenity in the East, and the presence of labor market frictions. The gap between East and West Germans is purely due to the estimated differences in workers' skills θ.

Table A9: West-East Gaps with Reduced Spatial Frictions

			Baseline (1)	All Frictions (2)	w/i Locations (3)	Partial Eq. (4)	Technology (5)	Preferences (6)
By Region	(1)	Output pc	30.3 %	16 %	26 %	18.9 %	19.2 %	24.8 %
	(2)	Value Function	15.8 %	0.4 %	6.9 %	0.8 %	1.7 %	9.1 %
	(3)	Wage	35.4 %	17.9 %	28.3 %	25.6 %	24.4 %	26.9 %
	(4)	Real Wage	26 %	13.6 %	23.5 %	20.3 %	18.9 %	18.6 %
	(5)	Wage (per eff. unit)	25.6 %	17.9 %	19 %	25.6 %	23.7 %	21.3 %
By Birth	(6)	Output pc	26.4 %	11.2 %	23.4 %	11.2 %	13.1 %	19.7 %
	(7)	Value Function	18.7 %	8.3 %	8.5 %	9 %	10.7 %	13.4 %
	(8)	Wage	29.8 %	11.7 %	25.1 %	11.7 %	14.3 %	20.6 %
	(9)	Real Wage	23.5 %	11.7 %	21.8 %	11.7 %	14 %	17.2 %
	(10)	Wage per eff. unit	18.1 %	1.7 %	13.8 %	1.8 %	4 %	9.7 %
	(11)	% of West-born in the West	96.7 %	69.3 %	96.7 %	71.6 %	73.5 %	89.9 %
	(12)	% of East-born in the West	25.5 %	69.1 %	25.5 %	71.1 %	66.9 %	46.1 %

J Evidence Supporting the Model Mechanism

We provide further evidence supporting the model's mechanism from the NY Fed's Survey of Consumer Expenditures (SCE) job search supplement for the years 2013-2020. The data is a series of repeated cross-sections with roughly 1,200 individuals each year. We use a confidential version of the data which identifies respondents' ZIP codes and individual demographic information. We merge each worker to the total employment in the worker's industry and commuting zone (CZ) using the Census Bureau's County Business Patterns (CBP), and obtain the wage distribution of the worker's industry and CZ from the 5-year American Community Survey (ACS) for 2015-2019. We provide more information on data preparation in Supplemental Appendix X.

We first analyze the effect of commuting time on employed workers' search behavior by running:

$$y_i = \beta_1 \ln(wage_i) + \beta_2 \ln(comm_i) + \alpha X_i + \epsilon_{ins}, \tag{33}$$

where y_i is the inverse hyperbolic sine (IHS) transformation of either i) employed worker i's number of applications sent to employers in the last four weeks; or ii) the number of hours spent searching for jobs in the last seven days. We use the IHS since many workers report zeros. The variable $wage_i$ is the worker's weekly wage at the current job, $comm_i$ is the commuting time in minutes, and X_i contains controls for gender, age dummies, a dummy for a college degree, industry fixed effects, and state fixed effects. The first two columns of Table A10 show that a greater commuting time for a given wage is positively associated with search effort.[60]

Second, we add to regression (33) dummies for whether the worker's current wage is in the second, third, or fourth quartile of the industry-CZ wage distribution. We focus on applications as our outcome variable; the results with search effort are similar and in Supplemental Appendix X. Columns 3 and 4 show that conditional on commuting time and wage, workers at the bottom of the wage distribution send more applications, consistent with our model.

Third, we add to the regression the total number of workers employed in the worker's industry and CZ. This variable is a measure of the density of the job market in the worker's location. In column 5, we use as LHS variable the worker's reservation wage for accepting a new job, and find that it rises with the density of the local job market, conditional on the current wage. In columns 6 and 7 we find that workers' search effort conditional on current wage is higher when the local job market is denser, even controlling for commuting time. Overall, these results highlight that the local labor

[60] In Supplemental Appendix X, we show that greater job dissatisfaction is also positively related to search effort.

market matters for workers' search decisions as highlighted by our model.

Table A10: Effect of Local Labor Market on Search

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$Apps_i$	$Search_i$	$Apps_i$	$Apps_i$	$\ln(ResWage_i)$	$Apps_i$	$Apps_i$
$\ln(wage_i)$	$-\mathbf{.1294}^{***}$	$-\mathbf{.1077}^{***}$		$-.0376$	$\mathbf{.5145}^{***}$	$-\mathbf{.1300}^{***}$	$-\mathbf{.1329}^{*}$
	(.0192)	(.0127)		(.0277)	(.0352)	(.0214)	(.0217)
$\ln(comm_i)$	$\mathbf{.0333}^{**}$	$\mathbf{.0359}^{**}$	$\mathbf{.0289}^{*}$	$\mathbf{.0301}^{**}$	.0152		$\mathbf{.0249}^{*}$
	(.0131)	(.0131)	(.0149)	(.0149)	(.0150)		(.0150)
$wage_i(Q2)$			$-\mathbf{.1928}^{***}$	$-\mathbf{.1605}^{***}$			
			(.0414)	(.0471)			
$wage_i(Q3)$			$-\mathbf{.2657}^{***}$	$-\mathbf{.2187}^{***}$			
			(.0400)	(.0523)			
$wage_i(Q4)$			$-\mathbf{.3623}^{***}$	$-\mathbf{.2914}^{***}$			
			(.0397)	(.0646)			
$\ln(emp_i)$					$\mathbf{.0253}^{**}$	$\mathbf{.0198}^{**}$	$\mathbf{.0179}^{**}$
					(.0109)	(.0081)	(.0082)
Industry FE	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y
Age, Sex, Ed	Y	Y	Y	Y	Y	Y	Y
Obs	4,960	4,960	4,150	4,150	3,344	4,151	4,151

Source: SCE and authors' calculations. Notes: Regressions are run on individual-level data for 2013-2020. $Apps_i$ is the IHS of the number of job applications sent by worker i in the last four weeks. $Search_i$ is the IHS of the number of hours spent searching for jobs in the last seven days. $ResWage_i$ is the reservation wage demanded for accepting a new job for workers looking for other employment. $wage_i$ are the weekly earnings at the main job. $comm_i$ is the average time spent commuting to the main job each day. $wage_i(Qx)$ is a dummy for whether the worker's weekly earnings are in the x percentile of worker i's commuting zone by industry wage distribution from the ACS. emp_i is the total employment in worker i's industry in her commuting zone from the CBP. Industries are 2-digit NAICS industries. Age controls are dummies for < 25, $25 - 54$, and $55+$ years. Sex is a dummy for males. Ed is a dummy for a bachelor's degree.