DISCUSSION PAPER SERIES

DP17494

Labor Misallocation Across Firms and Regions

Sebastian Heise and Tommaso Porzio

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

LABOUR ECONOMICS

MACROECONOMICS AND GROWTH

ISSN 0265-8003

Labor Misallocation Across Firms and Regions

Sebastian Heise and Tommaso Porzio

Discussion Paper DP17494
Published 22 July 2022
Submitted 20 July 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Labor Misallocation Across Firms and Regions

Abstract

We develop a frictional labor market model with multiple regions and heterogeneous firms to study how frictions impeding labor mobility across space aaffect 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 Classification: J6, O1, R1

Keywords: N/A

Sebastian Heise - sebastian.heise@ny.frb.org
Federal Reserve Bank of New York

Tommaso Porzio - tommaso.porzio@columbia.edu
Columbia Business School and CEPR

Labor Misallocation Across Firms and Regions[*]

Sebastian Heise[†]
Federal Reserve Bank of New York

Tommaso Porzio[‡]
Columbia University, NBER and CEPR

June 14, 2022

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, Paolo Martellini, Mushfiq Mobarak, Ben Moll, Simon Mongey, Melanie Morten, Todd Schoellman, and Jonathan Vogel for very useful comments that improved the paper. We have also benefited from the reactions of several seminar and conference audience, including participants at the NBER SI EFMPL, NBER Growth, Berkeley, Columbia, LSE, UBC, UCLA, UPenn, University of Toronto. Rachel Williams provided excellent research assistance. The views and opinions expressed in this work do not necessarily represent the views of the Federal Reserve Bank of New York. This study uses the weakly anonymous Establishment History Panel (Years 1975 - 2014) and the Linked-Employer-Employee Data (LIAB) Longitudinal Model 1993-2014 (LIAB LM 9314). Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and remote data access. The study also uses data made available by the German Socio-Economic Panel Study at the German Institute for Economic Research (DIW), Berlin. Neither the original collectors of the data nor the archive bear any responsibility for the analyses or interpretations presented here.

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

[‡]665W 130th St, New York, NY 10027, email: tommaso.porzio@columbia.edu.

1 Introduction

In many countries, large differences in real wages and labor productivity across regions have persisted for long periods of time.[1] These persistent differences suggest the presence of spatial frictions, such as moving costs, that limit the ability of workers to arbitrage the gaps away by migrating towards more productive regions. A large literature has shown that such frictions to labor mobility could entail large aggregate losses by misallocating labor across space.[2]

In this paper, we study and quantify a new margin through which spatial frictions misallocate labor and reduce aggregate productivity. We develop a general equilibrium framework that embeds frictional labor markets as in Burdett and Mortensen (1998) within a multi-region economy subject to a variety of spatial barriers. Estimating the model with matched employer-employee data from Germany, we find that, beyond misallocating labor across space, spatial frictions have an additional impact on the worker allocation: they misallocate labor across firms *within* regions.

In a frictional labor market, spatial barriers deprive workers of job opportunities in other regions, which they could use to move up the job ladder, thus slowing the reallocation of labor towards more productive firms. Additionally, spatial barriers increase firms' local monopsony power, which allows in particular low productivity firms to stay in business and to attract workers. We estimate that the aggregate losses due to these mechanisms in Germany amount to about 5% of GDP. These losses depend crucially on the details of the local labor market: as we show, two economies could exhibit the same wage or productivity gap between regions, yet the aggregate gains from removing spatial frictions could vary dramatically between the two economies 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 justify the ingredients of our model.

First, we use the Establishment History Panel (BHP), a 50% sample of all establishments in Germany, to study the distribution of wages and employment within and between regions. We document a large real wage gap between East and West Germany, but also significant heterogeneity in wages across firms *within* the two regions. Overall, it would be possible to completely close the regional wage gap by just reallocating labor within East Germany towards high-wage firms.

Second, we use the Linked Employer-Employee Data (LIAB) to study workers' wage gains as they climb the job ladder. We show that East Germans get very large wage increases when moving West, suggesting substantial gains from regional integration. At the same time, workers experience sizable wage gains for any job-to-job move, even within region, thus implying that

[1] Examples are the Italian Mezzogiorno, Andalusia in Spain, and the East of Germany.

[2] We discuss the relevant literature in detail in a section below.

frictions hindering within-region mobility could be as costly as those limiting migration towards high productivity regions.

Third, we use again the LIAB data to study workers' flows. We show that the job ladder is distorted. Workers switch jobs mostly locally and exhibit home bias (i.e., workers have a preference for their home region), leading to a job ladder that is characterized by frequent return migration of workers that have left their home region.

Motivated by the evidence, in the second part of the paper we develop a framework to study the 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 structurally identify the different spatial frictions and to isolate them from general labor market frictions. Separating the different types of spatial frictions is important as they have distinct aggregate effects on the economy and are amenable to different policy interventions.[3] 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. Given a set of 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 observed wage gains for workers moving out of their home region relative to other worker types making the same move identify home preferences. The relative frequency of job switches, instead, disciplines the spatial search costs. Relatively lower worker flows across regions, compared to between firms within region, indicate in our model 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 being relatively parsimonious

[3]For example, tax vouchers to limit moving expenses may increase mobility if spatial frictions represent moving costs, but may be less effective if spatial frictions are due to worker preferences for their home region, which are difficult to affect with policy.

with 21 parameters being used to match 305 micro and aggregate moments.

Our model estimates imply non-negligible spatial barriers, especially 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 only 1/20th as many job applications when searching for jobs across locations as within. We estimate a direct cost of moving between any two locations of 3.1%-5.3% of life-time income (dependent on the distance of the move), and find that workers need to be paid, everything else equal, 7.4% of their yearly income to work away from their home location and maintain the same utility. These relatively small moving costs and home biases reflect our model's ability to separate their impact from spatial search frictions and from general labor market frictions.

In the third part of the paper, we compute a series of counterfactual equilibria to quantify the aggregate and distributional costs of spatial frictions. Removing all spatial frictions, including workers' home bias, would raise GDP per capita in Germany by almost 5%, and average real wages by 9%. These sizable 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. Removing spatial frictions reduces firms' local monopsony power by exposing them to more competition from other locations, which forces unproductive firms to shrink or to exit the market and reallocates labor towards high productivity firms. Moreover, workers now have more job opportunities as they climb an integrated Germany-wide job ladder. Our empirical estimates indicate that the aggregate gains are mainly due to the former channel, i.e., the endogenous response of firms to the changes in the competitive environment. If we hold fixed firms' wage posting and vacancies, removing spatial frictions generates only a modest increase of 0.5% in output per capita, significantly smaller than the 5% gain including firms' equilibrium response.

The gains are not equally distributed across locations and workers' types. When spatial frictions are removed, East Germany sees an increase in output per capita of 17%, while aggregate gains in the West are only 4%. Similarly, East Germans see their wage rise by almost 25%, while West Germans gain less than 9%. Both the reallocation of labor within locations and across locations are important for these distributional effects. Labor reallocation within East Germany is larger than in the West since there are more unproductive firms there, which are more affected by spatial frictions. Reallocation of workers across locations is important because eliminating spatial frictions reduces dramatically the sorting of East and West Germans towards their home region, allowing more East Germans to benefit from the higher wages paid in the West. Moreover, since we estimate that West Germans have higher unobserved skills, migration of some West Germans to the East raises East German output and wages by increasing the average skill-level of the East German labor force.

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 in place workers' preference for their home region. However, we find strong complementarities between technological spatial frictions and workers' home bias. Summing over the gains from removing the two types of spatial frictions separately, we obtain only about half of the gains from removing both sources of frictions at the same time. Thus, the gains from an integrated labor market are largest when workers have access to opportunities to move to more productive locations (technology), but are also willing to accept these opportunities (preferences).

Our main conclusions are robust to allowing for more than four locations. As we increase the number of geographic units, we continue to find large aggregate gains due to the within-location reallocation of workers, although the estimated gains from reallocation of workers across locations also rise. Additionally, increasing the number of firms and workers in a location has virtually no effect on the location's aggregate gains from removing spatial frictions.

Finally, 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 regions, 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); Restuccia and Rogerson (2013)). In particular, we fit within a growing quantitative macro literature studying the role of labor market frictions in misallocating labor (Bilal, Engbom, Mongey, and Violante (2019, 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 the role of spatial frictions in shaping the competitive environment in the local labor market, leading to misallocation of labor across firms and within regions. Our focus on spatial frictions in the labor market is motivated by recent work showing that workers direct most of their search effort locally (e.g., Manning and Petrongolo (2017)).[4]

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, Dvorkin, and Parro (2019); Bryan and Morten (2019); Hsieh, Hurst, Jones, and Klenow (2019)).[5] Our contribution to this

[4]See also Schmutz and Sidibé (2018) and Le Barbanchon, Rathelot, and Roulet (2020).

[5]See also Diamond (2016); Giannone (2017); Caliendo, Opromolla, Parro, and Sforza (2017); Desmet, Nagy, and Rossi-Hansberg (2018); Hsieh and Moretti (2019); Fajgelbaum and Gaubert (2020).

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.[6] The quantitative spatial frameworks typically allow for a rich spatial heterogeneity, and consider workers that draw a vector of comparative advantages towards regions and/or occupations. Barriers to labor mobility may thus trap workers into a mismatched job, leading to misallocation of talent and aggregate productivity losses.[7] This channel has a very limited role in our framework, which instead focuses on the misallocation of labor across firms with different productivities and across regions with different distributions of firms. In our model, misallocation is not driven by mismatch of talent, but rather by misallocation of inputs, closer to the wedge approach of Hsieh and Klenow (2009) and Restuccia and Rogerson (2013) but generated endogenously by the interaction between labor and spatial frictions.[8]

Methodologically, we extend a wage posting model à la Burdett and Mortensen (1998) to incorporate a spatial structure. Our framework is most closely related to job ladder models with labor mobility across sectors, such as Meghir, Narita, and Robin (2015), Hoffmann and Shi (2016), and Bradley, Postel-Vinay, and Turon (2017).[9] In contrast to our work, however, these models do not consider switching costs between sectors, and therefore two workers with the same current value of employment accept the same job offers regardless of their current sector. In our setup, instead, workers' acceptance decisions not only depend on their current value but also on their current location: two workers with the same value may decide differently if they are in different locations. The existing frameworks cannot be applied to our research question since to do so we would need to assume that there is no cost of moving between locations.

At a conceptual level, our work also contributes to the fast-growing literature on local monopsony power (e.g., Berger, Herkenhoff, and Mongey (2022), Yeh, Macaluso, and Hershbein (2022)), and in particular to work that links labor market power to spatial frictions such as commuting costs (Caldwell and Danieli (2021), Datta (2022)).[10] Relative to this work, our paper analyzes how changes to spatial frictions affect monopsony power and endogenously reallocate workers within local labor markets.[11] The reallocation of workers towards higher productivity

[6]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, Meghir, and Robin (2016); Bagger and Lentz (2019); Bonhomme, Lamadon, and Manresa (2019).

[7]See Nakamura, Sigurdsson, and Steinsson (2022) for causal evidence supporting this mechanism.

[8]Krueger and Pischke (1995), Hunt (2001, 2006), Fuchs-Schündeln, Krueger, and Sommer (2010), Uhlig (2006, 2008), Dauth, Lee, Findeisen, and Porzio (2021), and Boeri, Ichino, Moretti, and Posch (2021) 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.

[9]A large literature has estimated versions of Burdett and Mortensen (1998) models. Examples are Van Den Berg and Ridder (1998); Bontemps, Robin, and Van den Berg (2000); Burdett and Coles (2003); Manning (2013); Burdett, Carrillo-Tudela, and Coles (2020); Moser and Engbom (2021).

[10]See also Manning (2013); Manning and Petrongolo (2017); Hirsch, Jahn, Manning, and Oberfichtner (2022).

[11]Similar to us, Galenianos, Kircher, and Virag (2011) and Bachmann, Bayer, Stüber, and Wellschmied (2021)

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 a very different mechanism: 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 (Hicks, Kleemans, Li, and Miguel (2017), Lagakos, Marshall, Mobarak, Vernot, and Waugh (2020)).[12] 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, and estimate the size of spatial search frictions compared to relocation costs.[13] 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.

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

2 Data

We use two main datasets provided by the German Federal Employment Agency (BA) to document a set of stylized facts that motivate our analysis: 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 defined as a company's unit operating in a distinct county and industry.[14] For simplicity, we will refer to these units as "firms". For each such firm in each year, the dataset contains information on location, average wages, the number of employees, and employee characteristics (education, age, gender).

The LIAB data contain records for more than 1.9 million individuals drawn from the In-

emphasize how firms' monopsony power reduces employment at highly productive firms; however, these papers do not analyze the role of spatial frictions in generating local monopsony power.

[12] See also Combes, Duranton, and Gobillon (2008); Roca and Puga (2017); Alvarez (2020).

[13] See also Amior (2019).

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

tegrated 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.[15] 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 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 three additional datasets. First, we obtain information on cost of living differences across German counties from the Federal Institute for Building, Urban Affairs and Spatial Development (BBSR (2009)), which we will use to construct real wages.[16] Second, we supplement our main analysis with annual survey data from the German Socio-Economic Panel (SOEP) to examine additional demographic characteristics and to corroborate some of our main findings. Finally, we use information on firms' profit shares from the ORBIS database by Bureau van Dijk 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, Jacobebbinghaus, Ludsteck, and Witter (2011). Since wages are only reported to the IAB up to the upper limit for statutory pension insurance contributions, the

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

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

BHP contains an imputed average wage variable which estimates the censored wages based on Card, Heining, and Kline (2013). For the LIAB, no such variable is provided and we replicate the imputation steps ourselves. We use the corrected, 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 datasets and on data construction in Appendix A.

3 Motivating Facts

To motivate our model and the relevance of our setting, we document three sets of facts on: (i) the distribution of firm wages within and between regions; (ii) wage gains of job-to-job movers within and between regions; (iii) workers' job flows.

3.1 Significant Wage Heterogeneity Between and Within Regions

We first study firm wages and show that there is significant heterogeneity in wages both across regions and across firms within regions.

Figure 1a plots the average real daily wage in each county in the period 2009-2014 from the BHP. What stands out from the figure is the large real wage difference between East and West. To examine whether this East-West wage gap is due to observables, we run firm-level regressions of the form[17]

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

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 wage gap to $\gamma = -.2052$ (s.e. .0027), but about 80% of the real wage gap remains unexplained.

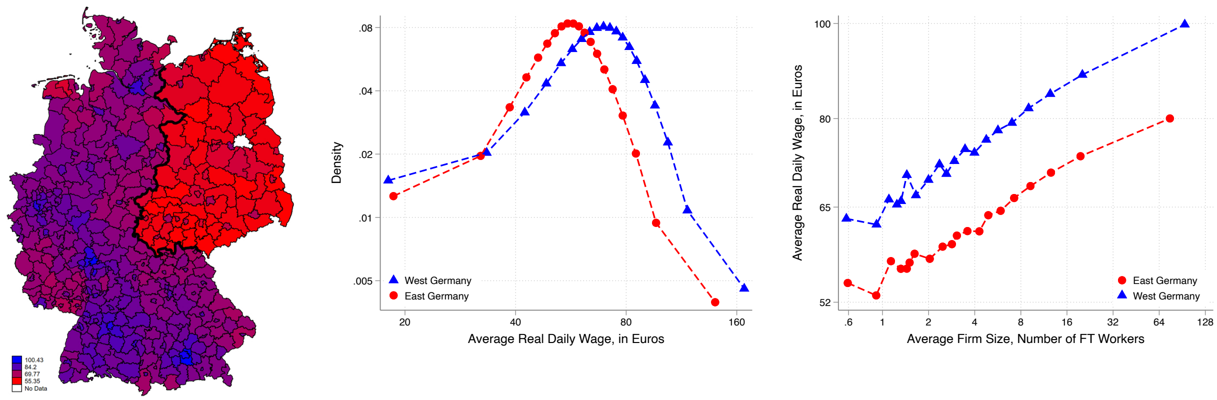
While the wage gap between East and West Germany is striking, we next show that there is even larger wage heterogeneity between firms within each region. Figure 1b plots PDFs of firms' average real wage from the BHP separately for both East and West Germany. We residualize log real wages by regressing them on year dummies and 3-digit industry dummies to remove across industry variation. The figure shows that the wage gap between the lowest- and highest-paying firms in each region exceeds the average wage gap between East and West.[18]

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

[18] In Supplemental Appendix K, available on the authors' websites, we drill more deeply into this pattern and 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 cross-county wage dispersion shown in Figure 1a.

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, 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 wages across firms separately for East and West Germany for the period 2009-2014. Wages are residualized by regressing the log real wage on 3-digit industry dummies and time dummies, for East and West Germany separately. We generate the cleaned wage as the residuals from this regression plus the mean of the log wage in the given region and transform these log wages back into levels. We then find the twentiles of the residualized wage distribution, compute the average wage within each twentile, and transform it into a density. While all firms are weighted equally, only a very small share of overall employment is at the lowest wage firms. 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, for both East and West Germany, where wages and size are residualized using the same procedure as before.

Figure 1c further plots the average firm size against the firms' average real wage for twentiles of the firm size distribution. Wage and size are residualized by year and industry dummies as before. 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 K[19], 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 German workers; (iii) there are no clearly delineated regional differences in tax rates.

3.2 Large Wage Gains of Movers Across and Within Regions

We next focus on wage gains of movers and show that workers obtain large gains when migrating, especially if moving away from their home region. However, each move across regions is also a move across firms, and we show that workers also experience sizable wage gains for any job-to-job move, even within region.

We analyze workers' wage dynamics around the time of a job-to-job move by running a

[19] This Supplemental Appendix is not meant for publication and includes additional material. It is available on the authors' websites.

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:[20]

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

where $w_{i\tau}$ is an individual's weighted average wage across all employment spells in year τ, where we use each spell's length as its weight. We define a job-to-job move as a job switch between two firms without an intermittent unemployment spell. The variable $\Delta \log(w_{i\tau})$ is the log change of this average wage between year τ and the previous year except for $t+1$, where it is the difference with respect to $t-1$. We drop wages from the year of the move to avoid contaminating our results by other types of payments in the year of the move.[21]

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 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.[22]

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 additional 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.[23]

The controls X_{it} include dummies for the current work region, home region, and their inter-

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

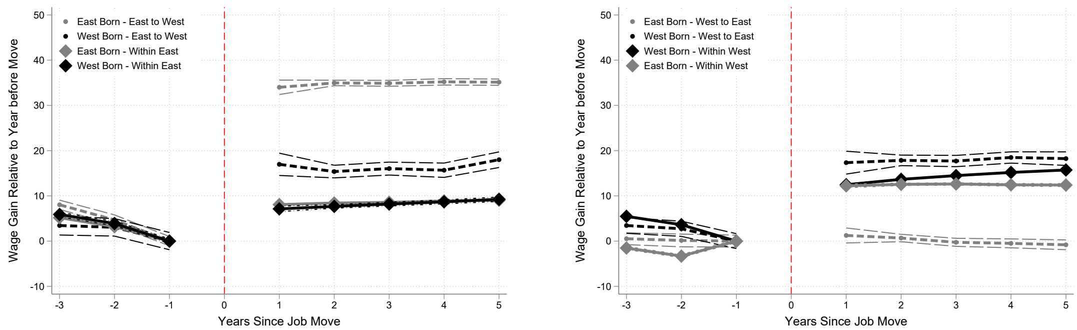
[21]The results are similar if we include year t, see Supplemental Appendix L.

[22]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.

[23]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 $i \in \{\text{West, East}\}$, and subtract from this sum the term $\sum_{u=-3}^{-1} \beta^{i}_{s,u}$ to normalize the coefficients with respect to period $\tau = -1$. The dotted lines represent the 95% confidence intervals. The 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 $\beta^{East}_{WW,\tau}$.

action, 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 $\beta^{West}_{s,\tau}$ and $\beta^{East}_{s,\tau}$ 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.

The dashed lines in Figure 2a plot the estimated wage gains for East-to-West migration by East Germans (gray) and West Germans (black) – i.e. the predicted wage from the relevant coefficients $\beta^{West}_{s,\tau}$ and $\beta^{East}_{s,\tau}$, translated into levels, and normalized around the wage level prior to the year of the migration. 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. The dashed lines in Figure 2b present the analogous wage gains for West-to-East migration. Moves to the East are associated with sizable wage gains for West-born 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 their home region (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.

The solid lines with diamonds plot instead the estimated wage gains for within-region job-to-job switches from regression (2) in East Germany (left panel) and in the West (right). Workers experience wage gains of around 10% for any job-to-job move, even within-region, 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 L, we list the full estimates from specification (2), and show that our results are robust to alternative definitions of job-to-job switches and migration.

3.3 Distorted Job Ladder

Finally, we study job flows and show that workers climb a country-wide job ladder, which is, however, 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 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 as

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

where $\mathbb{D}$ is the set of all the 402 counties in East and West Germany.[24] We use these shares to fit the gravity equation

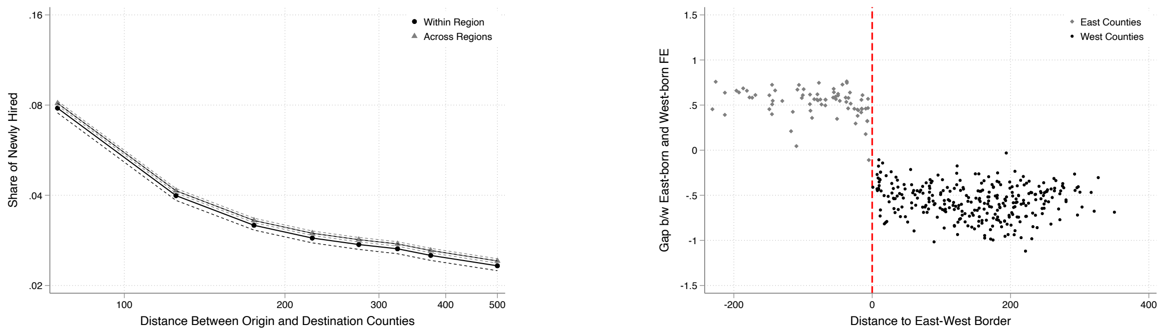
$$\log s^h_{o,d} = \delta^h_o + \gamma^h_d + \sum_{x \in \mathbb{X}} \phi_x D_{x,o,d} + \rho \mathbb{I}_{R(o) \neq R(d)} + \epsilon^h_{o,d}, \tag{3}$$

where δ^h_o and γ^h_d are county of origin and destination fixed effects, respectively, which differ by workers' home region, $D_{x,o,d}$ are dummies for buckets of distance traveled between origin and destination, and $\mathbb{I}_{R(o) \neq R(d)}$ is a dummy that is equal to one if the job switch is between East and West Germany. The set of buckets $\mathbb{X}$ contains seven 50km intervals from 50km-99km onward to 350km-399km and an eighth group for counties that are further than 399 km apart. The 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^h_o and γ^h_d capture the fact that some counties may be more attractive to workers of home region h, for example due to preferences, comparative advantage, or possibly due to a social network that allows them to find job opportunities.

[24] 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 M – we have 95,275.

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.

Figure 3a plots the estimated distance coefficients ϕ_x (black line), which we re-normalize into interpretable shares of switchers.[25] Workers move mostly locally, and job switches become less likely for counties that are further apart. The gray line plots the same results for flows between East and West Germany (the coefficients $\hat{\phi}_x + \hat{\rho}$), taking the origin and destination effects as constant. The lines are almost on top of each other. Thus, conditional on distance and fixed effects, we do not find a role for geographical barriers at the East-West border.

Figure 3b shows that there is strong home bias. For each county, we compute the difference between the destination fixed effect for East- and West-born workers, $\gamma_d^{East} - \gamma_d^{West}$. We then plot these differences against each county's distance to the East-West border, defined so that East counties have negative distance.[26] The figure shows that East individuals have significantly higher destination fixed effects for the East, indicating that they are relatively more likely to move to counties in the East than West workers regardless of their current location. Conversely, East-born workers are less likely to move to counties in the West. Supplemental Appendix M provides additional robustness checks for different sub-groups of the population and for different definitions of cross-border mobility.

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 that on average 1% of all employed West and East Germans switch jobs within-region in an average month (row

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

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

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.

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).[27] Overall, workers climb a country-wide job ladder, but this ladder is distorted by spatial frictions which lead workers to change jobs mostly locally and to frequently return home. The substantial return migration implies that the gains obtained 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 over the recent period.[28]

4 Model

We now develop a model to quantify how spatial barriers and labor market frictions jointly affect worker mobility across space and firms. 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, spatial search frictions, and regional comparative advantages. The model's ingredients are tied to the empirical facts shown in the previous section: first, the

[27] 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.

[28] This fact, together with the stable wage gap, motivates our analysis in steady state below.

wage dispersion and wage gains within-region call for a model with heterogeneous firms and labor market frictions. Second, the spatial wage gaps and the asymmetries in wage gains and job flows necessitate a model with mobility costs and home bias. Third, the presence of repeated moves across East and West suggests a framework in which individuals draw (infrequently) jobs from different regions.

We solve the model in general equilibrium, which will allow us 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. There are $\mathbb{J} = \{1, ..., J\}$ sites, which we refer to as *locations*.[29] The economy is inhabited by a continuum of workers of types $i \in \mathbb{I}\{1, .., I\}$ with mass $\bar{D}^i$, where $\sum_{i\in\mathbb{I}} \bar{D}^i = 1$. Throughout the text, we will use superscripts for worker types and subscripts for locations. Workers of type i have a preference parameter τ_j^i for being at location j, and consume both a tradable and a local good, such as housing. Their utility is $\mathcal{U}_j^i = \tau_j^i c^\eta h^{1-\eta}$, where c and h are the amounts of tradable good and local good, respectively. A worker of type i produces θ_j^i units of output per time unit in location j. Hence if this worker is employed at wage rate w per efficiency unit, she earns an income of $w\theta_j^i$. Worker i's indirect utility from receiving wage rate w in location j is then $\mathcal{V}_j^i = w\theta_j^i\tau_j^i/P_j$, where $P_j = (P_c)^\eta (P_{h,j})^{1-\eta}$ is the location's price level, P_c is the price of the tradable good, and $P_{h,j}$ the price level of the local good in location j.[30] We normalize $P_c = 1$.

Workers and firms operate in a frictional labor market. We define by e_j^i and u_j^i the mass of employed and unemployed workers of type i in location j, respectively. We introduce spatial search frictions that make it easier for workers to find job opportunities locally, building on a recent literature which uses job application data to show that workers' number of applications declines sharply with the distance of the vacancy (Manning and Petrongolo (2017); Le Barbanchon, Rathelot, and Roulet (2020)).[31] Specifically, workers of type i currently in location j must spend search effort s_x to send $a_{jx}^i(s_x) = z_{jx}^i s_x$ job applications towards location x, where z_{jx}^i is the worker's relative search efficiency, which depends on the worker's current and destination locations (j, x). Search efficiency also depends on the worker's type i. For example, it may be easier for workers to find open positions in their home location due to reliance on social networks

[29] 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 that are sub-units of the East and West German regions.

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

[31] Schmutz and Sidibé (2018) also incorporate spatial search frictions into their model to capture the lack of migration between areas with very different unemployment rates. In Bilal (2021), unemployed workers search for jobs only in the local labor market.

or referrals (as in, e.g., Galenianos (2013)). Search effort is subject to a cost, to be paid 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. Unemployed workers face a cost $\psi_u(s_x) = \nu^{-\epsilon}\frac{s_x^{1+\epsilon}}{1+\epsilon}$, where $\nu \geq 1$ modulates a potential difference in search intensity between employed and unemployed workers along the lines of Moscarini and Postel-Vinay (2016).

On the firm side, there is a mass M_j of firms exogenously assigned to locations $j \in \mathbb{J}$, with $\sum_{j\in\mathbb{J}} M_j = 1$. Within each location, firms are distributed over labor productivity p according to density function $\frac{\gamma_j(p)}{M_j}$ with support in a location-specific closed set $[\underline{p}_j, \bar{p}_j] \subseteq \mathbb{R}^+$.[32] Each firm p in location j decides how many vacancies $v_j(p)$ to post, subject to a vacancy cost $\xi_j(v)$, and what wage rate $w_j(p)$ to offer, determining the endogenous distributions of wage offers $\{F_j\}_{j\in\mathbb{J}}$. Firms cannot discriminate between worker types, hence they must offer identical wages per efficiency unit to all their workers. Also, firms cannot change their locations.[33]

Matches in location j are created as a function of the total mass of applications filed by workers, $\bar{a}_j$, and vacancies posted by firms, $\bar{v}_j$, according to a matching function $M(\bar{a}_j, \bar{v}_j) = \bar{a}_j^{\chi}\bar{v}_j^{1-\chi}$ as in Diamond-Mortensen-Pissarides models (e.g., Pissarides (2000)). We define market tightness in location j as $\vartheta_j \equiv \frac{\bar{v}_j}{\bar{a}_j}$. Thus, the rate at which a vacancy is filled is $\vartheta_j^{-\chi}$, and the rate at which an application is accepted and becomes a job is $\vartheta_j^{1-\chi}$. Offers are randomly drawn from the endogenous wage offer distributions $\{F_j\}_{j\in\mathbb{J}}$.

Upon receiving an offer from location x, workers draw idiosyncratic preference shocks for locations x and j and decide whether to accept or decline the offer. Movers between j and x incur a utility cost κ_{jx}^i that captures any monetary and non-monetary one-time cost associated with the move across locations, similar to Caliendo, Dvorkin, and Parro (2019). Workers separate exogenously into unemployment at location-type-specific rate δ_j^i and receive an unemployment benefit rate equal to b_j^i per efficiency unit when unemployed. They also always have the possibility to quit and become unemployed, keeping their draw of the preference shocks.

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

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

[33]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 by going from zero to positive vacancies or vice versa.

goods, taking prices in the output market as given.

In our model, firms compete for all worker types in one unified labor market. That seems an adequate description of the German labor market since we will define worker types based on their home region below, and firms cannot explicitly hire only West Germans, for example. Previous work with heterogeneous types (e.g. Moser and Engbom (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.

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

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

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

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

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

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

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

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

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

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

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

where $W_j^i(w)$ is the value of employment at wage w in location j, $W_x^i(w')$ is the value of employment in location x at wage w', and $\kappa_{jx}^i = 0$ if $j = x$. The terms ε_j and ε_x are idiosyncratic shocks drawn from a type-I extreme value distribution with zero mean and standard deviation σ, which capture shocks to workers' preferences for being in a specific firm and location.[34] These shocks simplify the model characterization and computation. 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).

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

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

and the expected value of an offer is

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

The problem for an unemployed worker is identical, but replacing $W_j^i(w)$ with the value of unemployment U_j^i.[35] We use notation $\mu_{jx}^{U,i}\left(b_j^i, w'\right)$ to refer to the probability that an unemployed worker accepts an offer w', and use $V_{jx}^{U,i}\left(b_j^i, w'\right)$ for the expected value of the offer.

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

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

Similarly, the unemployment value is:[36]

$$rU_j^i = \frac{b_j^i\theta_j^i\tau_j^i}{P_j} + \max_{\{s_x\}_{x\in\mathbb{J}}} \sum_{x\in\mathbb{J}} \left(a_{jx}^i(s_x)\vartheta_x^{1-\chi}\left[\int V_{jx}^{U,i}\left(b_j^i, w'\right) dF_x(w') - U_j^i\right] - \psi_u(s_x)\right). \tag{8}$$

[34] This 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.

[35] We show these equations in Appendix D.2.

[36] Appendix D.2 shows expressions for (7) and (8) once we solve out for optimal search effort.

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

$$\bar{a}_j^i \equiv \sum_{x \in \mathbb{J}} \left[\int a_{xj}^{E,i}(w)\, dE_x^i(w) + a_{xj}^{U,i}(b)\, u_x^i \right], \tag{9}$$

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

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

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

where $p \sum_{i \in \mathbb{I}} \theta_j^i l_j^i(w)$ are the net revenues from the goods market from (6).

Firms choose the number of vacancies to maximize total profits

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

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

Firms' vacancy posting and wage policies give the total mass of offers posted in each location and the endogenous offer distribution

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

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

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

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

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

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

which consists of the exogenous separation rate into unemployment plus the rate at which workers receive and accept offers from other firms – i.e. poaching within and across locations. As usual, we can use the law of motion (14) to solve for the steady state mass of workers per vacancy

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

and zero otherwise.

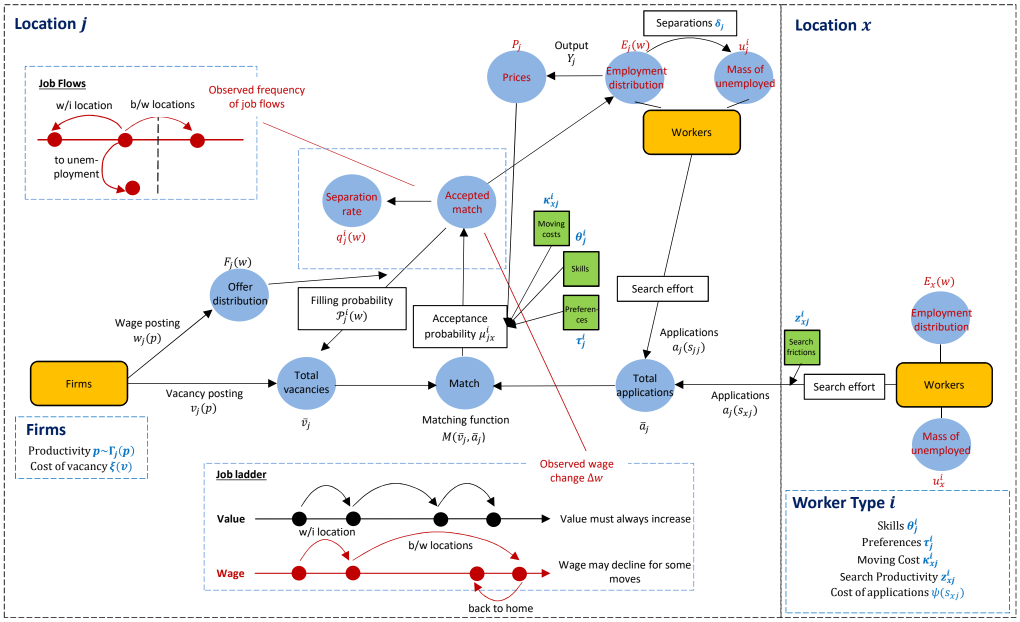
The mass of employed workers i in location j at firms paying at most w satisfies

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

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

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

Figure 4: Illustration of the Model



where $\bar{D}_j^i = e_j^i + u_j^i$ is the total mass of workers i in region j.

Figure 4 illustrates the main building blocks of our model and how they fit together. Yellow boxes denote the model's agents, blue circles endogenous objects, and green squares spatial frictions. We use red text for observable objects and black text for unobservables. The right-hand side of the diagram shows employed and unemployed workers in some location x. These workers exert search effort to post applications to some location j, subject to spatial search frictions. Workers already in location j also exert search effort but do not face the same spatial frictions since they search within-location. The left-hand side of the diagram shows the firm side. Heterogeneous firms post vacancies as well as wages, summarized by the wage offer distribution. Vacancies and applications meet in a frictional labor market, where the meeting probability depends on the ratio of total vacancies and applications, i.e., tightness. Given a match, workers' acceptance probability depends on the wage offered as well as the worker's moving costs, preferences, skills, and the price level. We illustrate the worker's acceptance decision in the dashed box at the bottom of the diagram. Workers' accept any offer that offers a higher value than their current one. However, workers' wage does not necessarily have to increase, since a wage loss can be compensated for example by a higher location preference. Workers that accept an offer separate from their previous job if they were employed, generating an endogenous separation rate. We illustrate worker flows in the box at the top left. Matches and separations determine the employment distribution and unemployment in location j, which in turn determine output and hence the price level.

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 in the labor market consists of a set of wage and vacancy posting policies* $\{w_j(p), v_j(p)\}_{j\in\mathbb{J}}$*, search efforts* $\left\{s^{E,i}_{jx}(w), s^{U,i}_{jx}(b)\right\}_{j\in\mathbb{J},x\in\mathbb{J},i\in\mathbb{I}}$*, wage offer distributions* $\{F_j(w)\}_{j\in\mathbb{J}}$*, acceptance probabilities* $\left\{\mu^{E,i}_{jx}(w,w'), \mu^{U,i}_{jx}(b,w')\right\}_{j\in\mathbb{J},x\in\mathbb{J},i\in\mathbb{I}}$*, labor per vacancy for each worker type* $\left\{l^i_j(w)\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$*, unemployment* $\left\{u^i_j\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$*, and market tightness* $\{\vartheta_j\}_{j\in\mathbb{J}}$ *such that*

1. *workers file applications and accept offers to maximize their expected present discounted values taking as given tightness* $\{\vartheta_j\}_{j\in\mathbb{J}}$ *and the wage offer distributions,* $\{F_j(w)\}_{j\in\mathbb{J}}$*;*
2. *firms set wages to maximize per vacancy profits, and choose vacancies to maximize overall firm profits, taking as given the function mapping wage to firm size,* $\left\{l^i_j(w)\right\}_{j\in\mathbb{J},i\in\mathbb{I}}$*;*
3. *the arrival rates of offers and wage offer distributions are consistent with aggregate applications, wage policies, and vacancy posting, according to equations (9), (10), and (12);*
4. *firm sizes and worker distributions satisfy the stationary equations (17), (18), and (19).*

The model does not admit an analytical solution. Yet, 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 a system of differential equations*

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

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

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

and

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

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

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

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

Proof. See Appendix D.3. □

Our framework is a generalization of Mortensen (2005). In Appendix D.4, we show that our model collapses to the standard framework if we shut down the spatial heterogeneity and the preference shocks.

5 Estimation

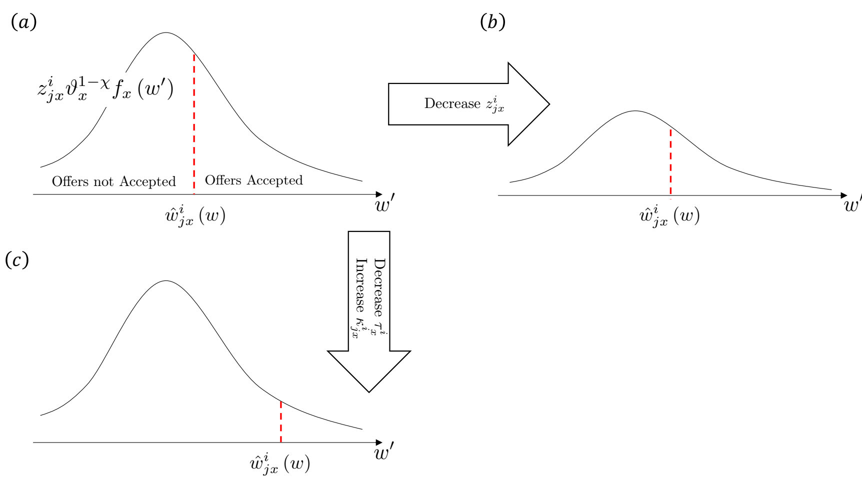
We next estimate the model using the German data described in Section 2.

5.1 Identifying the Spatial Frictions

Estimating the model requires us to separately identify the spatial frictions (κ_{jx}^i, τ_j^i, and z_{jx}^i) from the labor market frictions. Our strategy relies on the insight that labor market frictions mostly affect the allocation of labor within each location, and can therefore be identified from within-location moments, using similar moments as in the standard estimation of Burdett-Mortensen models (see, e.g., Bontemps, Robin, and Van den Berg (2000)). Given the labor market frictions, the spatial frictions can then be inferred from cross-location moments. While all model parameters are jointly identified, we illustrate our identification argument in Figure 5. 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, $z_{jx}^i \vartheta_x^{1-\chi} f_x(w')$.

To simplify the exposition, we assume that $\sigma \to 0$ so that an offer w' from location x is accepted by a worker of 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 the ones to the right of $\hat{w}_{jx}^i(w)$, and the mass of job flows per unit of search effort is the integral under the wage offer density to the right of $\hat{w}_{jx}^i(w)$. Going from panel (a) to (b), a decrease in the search efficiency z_{jx}^i reduces the mass of offers received, and hence the

Figure 5: Identifying Spatial Frictions



worker flows from location j to x. However, it does not affect the cutoff $\hat{w}^i_{jx}(w)$, and hence has no effect on average wage gain of workers that accept the offer and move from j to x.

Panel (c) shows the effect of a decline in the worker's preference for location x, τ^i_x, which shifts the acceptance location to the right (a similar argument applies for the moving cost, κ^i_{jx}). A decrease in τ^i_x (or an increase in κ^i_{jx}) raises the cutoff wage for any level of w. As the worker accepts only relatively better offers, the expected wage gain of a move increases in κ^i_{jx} and decreases in τ^i_x.

The identification argument illustrates that we need both worker flows and wage gains to separate the effect of the search efficiency from location preferences and moving costs. However, without further restrictions, we cannot separate the moving costs from location preference. We therefore assume that moving costs are identical for all worker types, reflecting for example relocation costs and transaction costs on the housing market. Under that assumption, we can separately identify the location preferences using the 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]

Discussion of Identifying Assumptions. Our argument is based on two core assumptions of the Burdett and Mortensen (1998) framework: wage posting and random search.

The wage posting protocol implies that firms cannot discriminate based on workers' type or current location. This assumption is supported by recent evidence that shows that the outside option has a limited effect on workers' wages (Jäger, Schoefer, Young, and Zweimüller (2020))

[37] Appendix E.1 provides further details, and Appendix H.3 uses simulations of the model to verify that the identification argument holds in practice.

and that, conditional on the current firm, a worker's previous firm has almost no effect on current wages (Kline, Saggio, and Sølvsten (2019)). Nonetheless, we note that under a different wage setting method 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, for example, by workers i employed in location j being more likely to sample from the left tail of the distribution in location x. While our assumption is strong, it does not affect the overall 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.

5.2 Parametrization and Calibrated Parameters

While the model can be solved for arbitrarily many locations, adding locations reduces the computation speed significantly and increases the number of targeted moments exponentially since we need to match worker flows and wage gains between every pair of locations for every worker type. To keep the estimation time feasible and to limit the number of moments and parameters, 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.[38] This number of locations and types allows us to distinguish the role of the former East-West border from other spatial frictions between locations. We will continue to refer to East and West Germany overall as "regions".

Functional Forms. We set a unit interval of time to be one month.[39] Firms' log productivity is drawn from a log-normal distribution with equal variance in all locations, Σ, and mean A_j.

[38] This parametrization implies that we need to match $4\times4\times4 = 64$ wage gains and 64 worker flows. Appendix F provides further details on the locations. In robustness checks below, we analyze the effect of increasing the number of locations on our main findings.

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

We normalize $A_{NW} = 1$ and refer to A_j as the relative aggregate productivity in location j.

We parametrize the vacancy cost function as $\xi_j(v) = \frac{\xi_{0,j}^{-\xi_1}}{1+\xi_1} v^{1+\xi_1} \bar{\pi}_j(p)$, where $\xi_{0,j}$ and ξ_1 are parameters to be estimated, and $\bar{\pi}_j(p)$ is the average firm profit in location j. This parametrization implies that the equilibrium mass of vacancies posted by a firm with productivity p is $v_j(p) = \xi_{0,j} \left(\frac{\pi_j(p)}{\bar{\pi}_j(p)}\right)^{\frac{1}{\xi_1}}$.[40] We assume that the curvature ξ_1 is constant across locations but allow $\xi_{0,j}$ to be specific to the overall region – i.e. we estimate $\xi_{0,W}$ and $\xi_{0,E}$.

For the unemployment benefit rates b_j^i, we set the reservation wages directly and assume that all worker types have identical reservation wage R_j within each location, where R_j is equal to a fraction ι of the productivity of the lowest productivity firm, $R_j = \iota \underline{p}_j$.[41] Note that due to the spatial frictions, workers' utility differs across types despite equal reservation wages.

Parametrizing Spatial Frictions. We interpret the moving cost κ_{jx}^i as the opportunity cost of foregone wages (Sjaastad (1962)), and assume that it is symmetric and proportional to the average value for each worker: $\kappa_{jx}^i = \hat{\kappa}_{jx} \bar{W}^i$, where $\bar{W}^i = \frac{1}{e^i} \sum_{j \in \mathbb{J}} \int W_j^i(w)\, dE_j^i(w)$ and $e^i \equiv \sum_{j \in \mathbb{J}} e_j^i$.[42] We assume that $\hat{\kappa}_{jx}$ is a function of distance between locations j and x, and, given our identification argument above, identical for all workers,

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

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

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

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

[40] The normalization can be interpreted as capturing the fact that vacancy creation requires work by individuals whose outside option would be to start a firm (getting a productivity draw from the existing distribution). The normalization is without loss of generality in the estimation as it simply changes the estimated values of $\xi_{0,j}$. Instead, in the counterfactuals it limits the relationship between average firm profits and vacancy posting. Our assumption implies that in the counterfactuals the overall amount of posted vacancies is relatively constant, and the results are mostly driven by a shift in the relative vacancy posting across the firm productivity distribution rather than by an aggregate shift in the demand for labor. Since our firm data do not allow us to discipline the total amount of vacancies posted, we prefer to be conservative and to limit the quantitative relevance of this channel.

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

[42] We impose this scaling because if κ_{jx}^i were a constant for all i, then the moving cost would be more binding for East-born workers since these have on average lower wages at any firm, as we show below.

both their home location and their overall home region, i.e., East or West Germany.

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

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

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

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

Calibrated Parameters. We calibrate eight sets of parameters listed in Table 2. Most of them are straightforward and we describe how we set their values in Appendix G. We focus here on how we set workers' relative productivity, θ^i_j (row 1). We use the fact that due to wage posting and the assumption that firms post the same wage to all workers' types, the model yields a log additive wage equation

$$\log w^i_j(p) \quad = \quad \log \theta^i_j + \log w_j(p).$$

This equation is similar to the specification by Abowd, Kramarz, and Margolis (1999), with the main difference that in our specification θ^i_j is not an individual fixed effect but both individual- and location-specific. This allows for *comparative advantage* when a worker is employed in her home location, i.e., she could have higher productivity there.[43] To estimate this parameter, we run a modified AKM regression

$$\log(w_{it}) = \alpha_i + \psi_{J(i,t)} + \beta \mathbb{I}^{(h_i \neq R(J(i,t)))} + BX_{it} + \epsilon_{it}, \tag{20}$$

where α_i is the worker component of worker i, $\psi_{J(i,t)}$ 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

[43] In practice, we only allow for *regional* comparative advantage. Our main hypothesis is that East Germans might have obsolete skills that cannot easily be transferred to the West German firms which have more advanced technologies.

Table 2: Calibrated Parameters

	Parameters	Source		Values: *West*	Values: *East*
(1)	θ^i: Workers' skills	AKM in LIAB, see Appendix E.2	*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.

home region h_i is currently employed at a firm in the other region.[44] We show in Appendix E.2 that β identifies the comparative advantage of workers in their home region, $\theta_i^{r(i)}$, and implement the estimation in Appendix G.1. We estimate $\beta = 0.019$, indicating a small *negative* comparative advantage towards the home region. 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 compute workers' average skills, θ^i, from the average worker fixed effects α_i. Based on our estimation, the average East German worker's unobserved skills are about 9 percentage points below those of a West German worker.

5.3 Estimation Targets

We are left with 21 parameters that we jointly estimate through simulated method of moments, and target the 305 moments summarized in Table 3. Appendix G.2 includes details on how each moment is computed. We provide a brief rationale on how the targets are chosen.

As described in Section 5.1, targeting the wage gains and job flows (rows 1-4 of Table 3) is key to identify the spatial frictions. In a steady state equilibrium, the job flows are directly tied

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

to the allocation of labor across locations, hence we target those moments as well (rows 5-6). We ask our model to match aggregate measures of economic performance (wages, output per worker, unemployment; rows 8-10) to be consistent with the stark regional gaps.

For the within-location allocation of labor to firms, we target several moments disciplining the way in which the job ladder works: the joint distribution of firm wages and size (rows 11-12); the relationship between firm wages and the separation rates and wage gains of their workers (rows 13-14); and how directed the typical job-to-job moves are, as measured by the standard deviation of the wage gains across workers (rows 15). Finally, we target the firm profitability (row 16) as a natural way to discipline the overall extent of monopsony power in each location.

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. One small complication is that 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. We thus need to decide how to define job flows across locations. As our baseline, we count as cross-location migrants all individuals that change their work location and satisfy either one of these two conditions: i) they update their residence; ii) their new job is further 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).[45] In Supplemental Appendix N, 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.

5.4 Identification and Model Fit

We estimate the model using a standard indirect inference approach and provide more details on our estimation algorithm in Appendix H.[46] While all parameters are jointly identified in equilibrium, we next analyze the connection between all parameters and moments via model simulations, and verify that the heuristic identification argument for the spatial frictions holds

[45]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 the Supplemental Appendix we consider the more extreme definitions.

[46]Figure A5 shows that the model likelihoods are locally single-peaked around each parameter estimate.

Table 3: Targeted Moments

	Moments	N	Source	Model Fit	Key Parameters
(1)	Wage gains w/i locations, by (i, j)	16	Sect G.2.1	Fig 6	Σ, σ
(2)	Wage gains b/w locations, by (i, j, x)	48	Sect G.2.1	Fig 6	κ, τ_j^i, Σ
(3)	Job flows w/i locations, by(i, j)	16	Sect G.2.2	Fig 6	ϵ, ξ_0
(4)	Job flows b/w locations, by (i, j, x)	48	Sect G.2.2	Fig 6	z_{jx}^i, κ
(5)	Employment shares, by (i, j)	16	Sect G.2.3	Fig A7	κ, z_{jj}^i, τ_j, τ_j^i, z_{jx}^i
(6)	Unemployment shares, by (i, j)	16	Sect G.2.4	Fig A7	κ, z_{jj}^i, τ_j, τ_j^i, z_{jx}^i
(7)	Firm component of wages, by (i, j)	15	Sect G.2.5	Fig A7	A_j, τ_j
(8)	Average firm component of wages, by j	3	Sect G.2.6	Fig A7	A_j, τ_j ,z_{jj}^i, τ_j^i
(9)	Relative output per worker, by j	3	Sect G.2.7	Fig A7	A_j, ν
(10)	Unemployment rates, by j	4	Sect G.2.8	Fig A7	ν
(11)	Deciles of firm-size distributions, by j	40	Sect G.2.9	Fig A8	ξ_1
(12)	Slope of wage vs firm size relationship, by j	4	Sect G.2.10	Table A28 and Fig A9	ξ_1 , ι
(13)	Slope of J2J wage gain vs firm wage, by j	4	Sect G.2.11	Table A28 and Fig A9	Σ, σ
(14)	Slope of separation rate vs firm wage, by j	4	Sect G.2.12	Table A28 and Fig A9	ξ_0 , σ, ϵ
(15)	Std of job-job wage gains, by (i, j, x)	64	Sect G.2.13	Table A28 and Fig A10	Σ, ξ_0 , ϵ, ι, σ
(16)	Profit to labor cost ratio, by j	4	Sect G.2.14	Table A28	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 Section H.3.

in practice. We then evaluate the model fit.

Identification. We compute the elasticity of each (model generated) moment to each parameter, and provide the Jacobian matrix in Appendix H.3. The last column of Table 3 reports the most important parameters for each moment based on this exercise.[47]

A few comments are in order. First, and importantly, the Jacobian matrix verifies that the heuristic argument made in Section 5.1 holds. The wage gains between regions (row 2 of Table 3) are crucial for the moving costs κ and the preference τ_j^i, while the job flows between regions (row 4) are especially important for the relative search efficiencies z_{jx}^i. The spatial frictions are also crucial for the steady state allocation of employed and unemployed workers (rows 5-6). As expected, the within region wage gains and flows (rows 1 and 3) are, instead, not relevant for the spatial frictions. Instead, large within region wage gains are driven by either a large variance

[47]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 3 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 $\xi_{0,E}$ ($\xi_0 \equiv \{\xi_{0,W}, \xi_{0,E}\}$).

of the productivity distribution (Σ), or a low variance of the taste shock (σ) so that workers only accept job offers with an associated wage increase. The job flows are mainly related to the parameters modulating the cost of filing applications (ϵ) and of posting vacancies (ξ_0).

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

Third, the moments that matter most for the efficiency of the job ladder (rows 11-15) are mostly linked to 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). This is expected: as already noted, Σ is crucial to determine 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 on average directed towards higher wage offers.

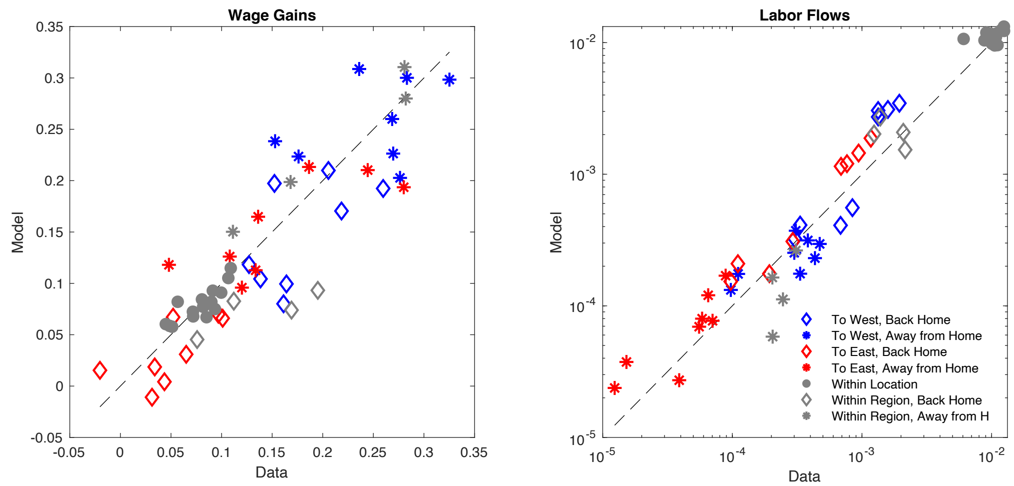
Fourth and last, the firm profitability (row 16) is a function of the labor market friction parameters (σ, ξ_0, ξ_1), as expected since they determine the extent of monopsony power, and the reservation wage ι, which mechanically decreases profitability. The home preference τ_j^i also plays a relevant role: when workers are more attached to a location, firms face effectively less competition from other locations.

Model Fit. The model matches well the key moments that help to identify the spatial frictions. The left panel of Figure 6 plots the wage gains of job-to-job movers in the data against those in the model (from rows 1 and 2 of Table 3).[48] Each dot is for one of the 64 different types of moves by origin-destination-home location, which we color code by direction and type of worker. As in the data, the model generates larger wage gains for moves towards the West (blue symbols), for within-region moves away from the home location (gray stars), and for moves away from the home region, in particular to the West (blue stars). The right panel presents a similar plot for the monthly share of movers in all employed workers (from rows 3 and 4). As in the data, in our model individuals are more likely to move within-location (gray circles) and to move back to their home location and region (diamonds) than away from home (stars).

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

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

Figure 6: Wage Gains and Frequency of Job Flows



Notes: The left panel shows the average wage gains of different types of job-to-job moves in the data (x-axis) against the average wage gains in the model (y-axis). The right panel shows the frequency of each type of job-to-job move in the data (x-axis) against the frequency in the model (y-axis). Different types of moves are identified by a mix of colors and symbols, listed in the right panel. In total, there are 64 possible types of moves by origin location, destination location, and home location. The data moments are listed in Appendix G.2.1 and G.2.2.

is in a 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.[49]

Overall, the model displays a very good fit considering that we estimate 21 parameters to target 305 moments.[50] Several structural restrictions imposed by the model on the joint distributions of firm wages, employment, wage gains, and labor flows are satisfied in the data, building confidence in our estimated frictions.

[49]In Figure A9 we show the non-parametric relationships for the moments in rows 12, 13, and 14 of Table 3. In Figure A10, 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.

[50]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.

5.5 Parameter Estimates

We present the estimated spatial frictions in Table 4, and show the remaining parameters in Appendix H.2. Row (1) reports the one-time moving costs, $\hat{\kappa}_{jx}$, as a fraction of the present discounted value of income. Since these costs vary with distance, we present a range of costs for moves between the closest two locations and moves between the farthest two locations. Our estimates indicate moving costs in the range of $3 - 5\%$ of the PDV of income, implying that an individual earning a yearly salary of 36,000€ for a work life of 45 years faces a moving cost of between 17,453 € and 29,704 €.[51]

Rows (2) and (3) show that a worker employed not in her home location but still in her home region would need to be paid, in real terms, about 7.4% more than in her home location to obtain the same flow utility. Moving away from both home location and region requires a yearly compensation almost 10% higher than at home.[52]

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.

Rows (4) and (5) report the estimated search efficiencies, relative to the within-home location level, which is normalized to 100%. Individuals that are in a location away from home and search within that location are slightly less effective than at home, filing only about 90% as many applications per unit of search effort as at home (row 4). More importantly, however, all individuals have a much lower search efficiency for cross-location searches, consistent with evidence that workers search for jobs primarily locally.[53] As before, we provide a range for searches between the two closest locations and between the two farthest locations. Row (5.i) shows that one unit of search effort expended across locations in the non-home region translates into filing only about 1/20th as many applications as in the home location. Cross-location searches directed towards the home region, but not to the home location, are only slightly more effective (5.ii). Row (5.iii) shows that one unit of search effort by workers currently away from their home location that is directed towards the home location generates 24.11% to 17.22% as many applications as searches within the home location. Hence, workers searching across

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

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

[53] See Manning and Petrongolo (2017), Le Barbanchon, Rathelot, and Roulet (2020), Datta (2022).

Table 4: Estimated Spatial Frictions

Moving Costs $\{\kappa\}$			
(1)	Moving cost as share of PDV of income: $\kappa_0 e^{\kappa_1 dist_{jx}}$ (b/w closest to b/w furthest locations)		3.12% to 5.31%
Preferences $\{\tau\}$			
(2)	Cost of not living in the home location but in the home region, as share of income: τ_l		7.41%
(3)	Cost of not living in the home region, as share of income: τ_r		9.88%
Relative Search Efficiency $\{z\}$			
(4)	w/i location, away from home location: $1 - z_{l,1}$		90.52%
		5.i) not to home region: $z_0 e^{-z_1 dist_{jx}}$	6.10% to 4.95%
(5)	b/w locations (closest to furthest locations)	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 A27. Row 1 provides a range of the estimated moving costs, ranging from costs for moves between the closest two locations to moves between the furthest two locations. Rows 2-3 present the values of the estimated preference parameters. Search efficiencies in rows 4 and 5 are expressed as a percentage of the efficiency within the home location, z^j_{jj}, which is normalized to 1. Rows 5i-5iii show the efficiencies for searching across locations outside of the home region, in the home region but not the home location, and in the home location, respectively. The efficiencies are again reported as a range for searching between the two closest locations to searching between the two furthest locations.

locations are about four times as efficient in searching in their home location than in their non-home region. The lower efficiency away from home could reflect social connections, which facilitate finding jobs at home (Burchardi and Hassan (2013), Bailey, Farrell, Kuchler, and Stroebel (2020)).

While we leave the discussion of the remaining parameters to Appendix H, we note here 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, Seidel, and Suedekum (2021)) and it could additionally reflect remaining cost of living differences that are not picked up by our price indices.

6 Labor Misallocation across Firms and Regions

We use the estimated model to study the role of spatial frictions in the allocation of labor across firms and regions. First, we analyze the aggregate effects of spatial frictions and the mechanisms through which they unfold. Then, we turn to the distributional effects across regions and workers' types, and we explore whether the results are robust to changing the number of locations or their relative size. Finally, we study the extent to which the aggregate results are affected by the estimated frictions in the local labor markets. Throughout this section, we present the results by region rather than for individual locations to facilitate the exposition and since the heterogeneity across locations within regions is minimal.

6.1 Aggregate Effects of Spatial Frictions

We recompute the equilibrium keeping all the parameters at their estimated values, but remove all spatial frictions: the moving cost ($\kappa_0 = 0$), the preferences for the home location or region ($\tau_l = \tau_r = 0$), and the differences in search efficiency for searches across and within locations ($z_{l,1} = z_{l,2} = z_r = z_1 = 0$ and $z_0 = 1$). We then compute five core statistics for the long-run steady state equilibrium in the baseline and the counterfactual: (i.) output per capita (and hence labor productivity, p); (ii.) the average of workers' value functions across all employed and unemployed workers; (iii.) average wage, $w_j(p)\,\theta_j^i$; (iv.) average real wage, $w_j(p)\,\theta_j^i/P_j$; and (v.) the share of the overall employment in West Germany.

The results for Germany overall are shown in the first column of Panel (a) of Table 5. Removing all spatial frictions leads to an increase in output per capita, hence in labor productivity, of slightly less than 5%.[54] Despite these relatively modest output gains, the increase in the average worker's value is much larger.[55] The reason is twofold. First, without spatial frictions workers no longer incur the moving cost $\hat{\kappa}_{jx}$ or the utility cost (τ_l, τ_r) when they cross locations. Moreover, workers' search efficiency across locations rises, which increases their continuation value and allows workers to focus on matches with a high taste shock ε. Second, eliminating spatial frictions exposes firms to stronger competition for workers from firms in other locations, which raises wages more than the increase in labor productivity due to a reduction in firms' monopsony rents. We show below that the reduction in monopsony power is responsible for the majority of the aggregate output gains and a sizable share of the increase in workers' value.

Row 5 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 due to home bias and search frictions. 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 our findings. First, we analyze the impor-

[54] 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 the fact that each individual draws a vector of location-specific comparative advantages.

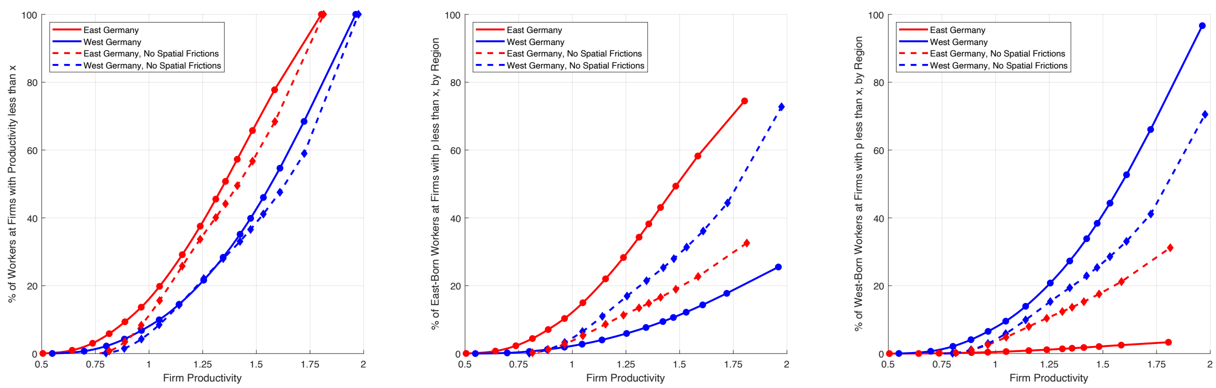
[55] We use the term workers' value rather than welfare since we are, in the counterfactual, effectively changing preferences through the taste spatial friction τ_j^i.

Table 5: Model Counterfactuals with Reduced Spatial Frictions

			All Frictions	w/i Locations	Partial Eq.	Technology	Preferences
			(1)	(2)	(3)	(4)	(5)
			Panel (a): Aggregate				
Overall	(1)	Output pc	+ 4.7 %	+ 6.6 %	+ 0.5 %	+ 2.7 %	+ 0.7 %
	(2)	Value Function	+ 37.0 %	+ 37.1 %	+ 22.0 %	+ 25.1 %	+ 2.9 %
	(3)	Wage	+ 9.1 %	+ 11.3 %	- 2.1 %	+ 3.8 %	+ 1.7 %
	(4)	Real Wage	+ 9.6 %	+ 11.3 %	- 1.6%	+ 4.2 %	+ 1.8 %
	(5)	% Workers in West	- 10.9 %	/	- 8.7 %	- 8.2 %	- 0.6 %
			Panel (b): By region				
West	(6)	Output pc	+ 4.2 %	+ 6.0 %	+ 0.4 %	+ 2.5 %	+ 0.1 %
	(7)	Value Function	+ 33.3 %	+ 35.0 %	+ 18.8 %	+ 22.1 %	+ 1.8 %
	(8)	Wage	+ 8.6 %	+ 10.5 %	- 1.5 %	+ 4.1 %	+ 0.8 %
	(9)	Real Wage	+ 9.2 %	+ 11.1 %	- 0.9 %	+ 4.6 %	+ 0.9 %
	(10)	Wage per eff. unit	+ 10.2 %	+ 10.5 %	+ 0.4 %	+ 5.6 %	+ 1.4 %
East	(11)	Output pc	+ 17.0 %	+ 9.6 %	+ 10.0 %	+ 12 %	+ 4.5 %
	(12)	Value Function	+ 53.7 %	+ 46.2 %	+ 36.6 %	+ 39.1 %	+ 8.1 %
	(13)	Wage	+ 24.6 %	+ 16.6 %	+ 6.2 %	+ 13.3 %	+ 7.6 %
	(14)	Real Wage	+ 21.1 %	+ 13.3 %	+ 3.8 %	+ 10.8 %	+ 7.2 %
	(15)	Wage per eff. unit	+ 17.4 %	+ 16.6 %	+ 0.4 %	+ 7.1 %	+ 5 %
			Panel (c): By worker type				
Born West	(16)	Output pc	+ 1.9 %	+ 6.0 %	- 2.1 %	+ 0.3 %	- 0.4 %
	(17)	Value Function	+ 34.3 %	+ 34.5 %	+ 19.8 %	+ 23.2 %	+ 1.9 %
	(18)	Wage	+ 6.0 %	+ 10.6 %	- 5.0 %	+ 1.3 %	+ 0.3 %
	(19)	Real Wage	+ 7.5 %	+ 11.1 %	- 3.6 %	+ 2.6 %	+ 0.8 %
	(20)	% Workers in West	- 27.3 %	/	- 25.1 %	- 23.2 %	- 6.8 %
Born East	(21)	Output pc	+ 15.9 %	+ 8.7 %	+ 11.3	+ 12.1 %	+ 5.1 %
	(22)	Value Function	+ 47.2 %	+ 47.0 %	+ 30.5	+ 32.1 %	+ 6.6 %
	(23)	Wage	+ 23.1 %	+ 14.8 %	+ 10.4	+ 15 %	+ 8 %
	(24)	Real Wage	+ 18.9 %	+ 12.7 %	+ 6.7	+ 11.2 %	+ 6.2 %
	(25)	% Workers in West	+ 43.5 %	/	+ 45.6	+ 41.4 %	+ 20.6 %

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 same semi-CDF for West Germans.

tance of within-location reallocation of labor compared to worker reallocation across locations. Second, we discuss the role of the equilibrium response of firms. Finally, we separately analyze the different types of frictions.

The Importance of the Within-Location Allocation of Labor. To analyze the importance of a better allocation of workers to firms within locations, in the second column of Table 5 we recompute the aggregate gains holding fixed the share of workers in each location at the baseline level, thus shutting down migration. We continue to change the within-location distribution of workers and firms' policy functions as in the full counterfactual. We find that shutting down the migration across locations actually *raises* output and wages. This outcome arises because, as already noted, workers migrate towards the lower productivity East in the full counterfactual, decreasing aggregate output and wages.

Panel (a) of Figure 7a analyzes how the within-region reallocation of workers generates the aggregate gains. The figure shows the CDFs of employment to firms of different productivity within East and West Germany for the baseline (solid) and the counterfactual without spatial frictions (dashed).[56]

Removing spatial frictions shifts both 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

[56] Since the baseline was estimated from the data moments, it is consistent with the within-region wage distributions shown in Figure 1b if wages are increasing in productivity, as in our model.

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. While removing spatial frictions also makes it easier for unproductive firms to hire from the other region, on net the negative effect dominates. As a result, the lowest productivity firms stop posting vacancies as they are not able to offer a higher value than unemployment. This effect is stronger in the East because it has the lowest productivity firms overall. The reallocation towards higher productivity firms is similar in spirit to the within-industry reallocation observed in international trade models such as Melitz (2003) after an economy opens up to trade. Yet, it comes from a very different mechanism: competition for workers in the labor market, rather than for customers in the output market.

We can further unpack the drivers behind the within-region reallocation of labor by decomposing the total labor employed at a firm of productivity p in region j as

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

The first term captures the local market tightness and thus only affects the allocation of labor between, but not within, regions. The other three terms could, in principle, explain the 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 $\mathcal{P}_j^i(w)$), or facilitate worker retention (low $q_j^i(w)$).

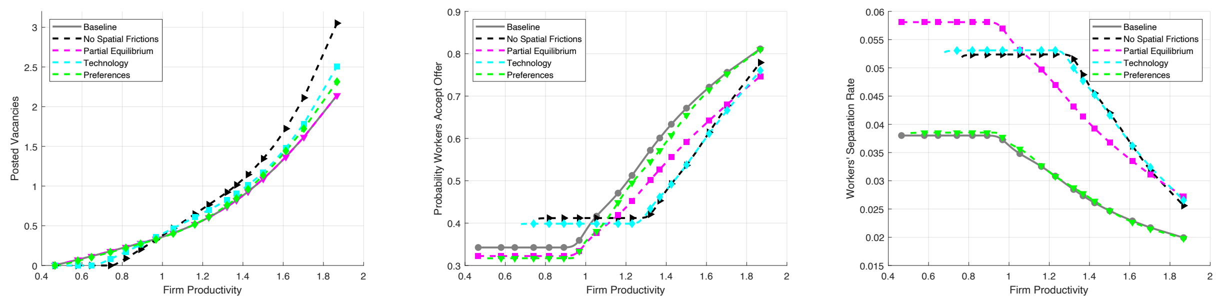
In Figure 8, we plot these three objects as a function of firm productivity, for both the baseline economy (solid gray line) and the economy without spatial frictions (black dashed).[57] Panel (a) shows that the number of posted vacancies contributes positively to the reallocation of labor from low- 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 positively 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 particularly large 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.[58]

[57]The plots are for East Germany. The ones for West Germany are similar and are included in Appendix J.

[58]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.

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.

Large Equilibrium Effects due to Lower Monopsony Power. How important is the change in labor market competition, and the resulting decline in most firms' local monopsony power, for the aggregate gains? In the counterfactual equilibrium, both workers and firms change their behavior. 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 results in column 3 of Table 5 show that when firms' wage and vacancy policies are held fixed, the output per capita increases by only 0.5% (row 1). 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.[59] While the separation rate still rises more for low-productivity firms than for high-productivity ones (Panel (c)), contributing to a better allocation of workers, this channel alone has only a modest effect.

Lack of Opportunities or Unwillingness to Take Them? There are significant complementarities between the different types of spatial frictions. Our model contains both *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, a faster railway system or rental subsidies to facilitate the housing search could decrease κ, while an integrated online job portal could reduce z. Instead, preference frictions are plausibly harder to affect, as they are typically a slow moving object that changes across generations (Alesina and Fuchs-Schündeln (2007)).

[59] In Appendix J we replicate Figure 7 for this alternative counterfactual.

To analyze their effects separately, we recompute the equilibrium of the economy when we remove either only the technological spatial frictions or the home bias. We find substantial aggregate gains from removing technological barriers that are about a third to one half as large as the baseline (column 4 of Table 5). In contrast, removing home preferences generates only small gains (column 5 of Table 5). The blue and the green line in Figure 8 illustrate the drivers behind these effects.[60] When only home preferences are eliminated, workers' separation decisions remain nearly unchanged and vacancies shift by less than when the technological spatial frictions are removed, leading to smaller gains. However, there are important complementarities: summing over the aggregate gains from both exercises in Table 5 yields only about half the effect of removing both sources of frictions at the same time.

Our results show that both a lack of opportunities and an unwillingness to take them contribute to the aggregate effects in Germany. The technological spatial frictions mainly generate a lack of opportunities: even in the absence of home preferences, East-born individuals have a hard time accessing jobs in the West because they do not obtain job offers, and many of the ones they find are not good enough to compensate for the cost of moving there. Instead, the home preferences affect individuals' willingness to take the available opportunities: many East Germans remain in the East because, everything else equal, they are more likely to accept offers received from their birth-region. To generate large effects from integrating the labor market, both sources of spatial frictions must be addressed simultaneously.

6.2 Distributional Effects of Spatial Frictions

We next show that spatial frictions also have large effects on the distribution of resources across regions and worker types.

Differences by Region. We first examine the effects of removing spatial frictions separately for individuals in the West and in the East of Germany, for all five exercises run before (Panel (b) of Table 5).[61] We also report the wage per efficiency unit, $w_j(p)$, to highlight the difference with the average wage $w_j(p)\,\theta_j^i$, which depends on the composition of workers (θ_j^i) in each region.

Column 1 highlights that the baseline gains from removing spatial frictions are much larger in the East than in the West. There are two main reasons for this result. The first one is mechanical: despite similar observable characteristics, we estimated a large gap between East and West workers in unobservable skills from the AKM (see row 1 of Table 2). Therefore, as West workers move East and East workers move West, average human capital improves in

[60] In Appendix J we replicate Figure 7 for these alternative counterfactuals.

[61] 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 computed statistics will, of course, take into consideration the possibility that individuals move across locations and regions.

the East and declines in the West, which can be seen by comparing the changes in wages and wages per efficiency unit in the two regions. 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.

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 that case we would observe an absolute *decline* of the wage and labor productivity in the West. In our model, 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.

Column 2 highlights that worker migration across regions is important for the distributional effects. When worker mobility across regions is shut down, the output and wage gains in East Germany fall by nearly half. A large part of the East German gains is due to the increase in average human capital resulting from the inflow of West German workers. Instead, in West Germany, migration has a negative effect on output and wages. The take-aways from the counterfactual exercises in columns 3-5 are qualitatively similar to before.

Differences by Worker Type. Panel (c) shows the effect of removing spatial frictions for East versus West Germans. While everyone benefits, East Germans see a larger increase in their output per capita, wages, and values than West Germans 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.[62] The share of East Germans in the West rises significantly.

For West Germans, as shown in Panel (c) Figure 7, the net migration towards the East leads them, on average, to work for lower productivity firms than in the West. 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.

Implications for the West-East Gaps. Table 6 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. For reference, Column (1) presents the East-West gaps in the baseline economy. Column (2) shows that eliminating spatial frictions shrinks the gaps in output, value, and wages considerably, but does not eliminate them. The remaining gaps are due to the average higher productivity of firms in the West, the higher

[62]The semi-CDF means that each line does not end at one but at the share of East German workers in each region. Adding up the last point on the two solid lines or on the two dashed lines gives one.

Table 6: 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 %

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.[63]

The gap between East and West Germans is purely due to the estimated differences in workers' skills θ. Even in the absence of spatial frictions, West Germans earn a higher wage, produce more GDP per capita, and have higher value due to their higher estimated skills.[64]

6.3 Increasing the Number of Locations and Market Size Effects

We next 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.

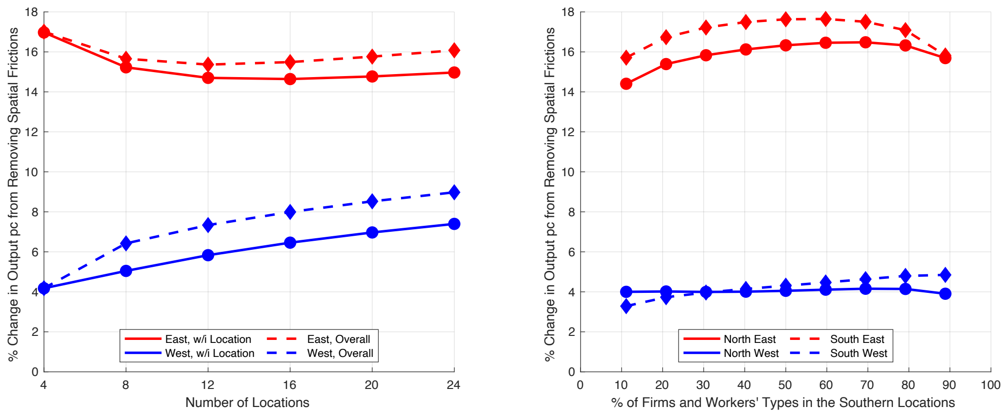
First, we vary the number of locations. In principle, the large role we find for the within-location reallocation of labor could be due to the fact that geographic units in our model are large, thus limiting the possibility of between-location reallocation. To assess this concern, we solve a version of our model in which we allow for more locations, splitting each of the four

[63]Since the residual real wage gap is 13.6% and we estimated an amenity difference between East and West of 11%, there is, even once we account for amenity differences, a roughly 2.6% higher real wage in the West.

[64]A small difference remains between East and West Germans even in wage per efficiency unit since West Germans, due to their higher skills, search more intensively for jobs.

Figure 9: 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.

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.[65]

Figure 9a presents the output per capita gains from removing spatial frictions as a function of the total number of locations in the economy. The solid lines show the gains due to the reallocation of labor *within* the locations, averaged across all locations in East and in West Germany, respectively. The dashed lines provide the total gains, i.e., including those from worker reallocation across locations. Overall, we continue to find large gains from the within-location reallocation of labor even as we increase the number of locations. While reallocation across locations becomes more important as the number of locations grows, most gains are within-location in all cases. 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. In principle, the larger effects of removing spatial frictions for East Germany could be driven not by the lower productivity of East firms, but by

[65]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 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^i_{jx} so that $\sum_{x\in\mathbb{J}} z^i_{jx}$ is constant across all scenarios.

the smaller size of the East German labor market in terms of the number of firms and workers. As a result, spatial frictions could be more binding in the East. To assess whether the size of a location affects its aggregate gains, 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 9b 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. This result is driven by the interplay of two counteracting forces. On the one hand, without spatial frictions firms in smaller locations have a bigger relative increase in the mass of workers that can now apply to their vacancies. On the other hand, they face a relative bigger increase in the competition for labor from firms in other locations. These two effects roughly balance each other out so that on net the market size is not an important driver of the regional heterogeneity.[66]

6.4 The Role of the Local Labor Market for Aggregate Gains

Since the aggregate impact of spatial frictions is mediated by their impact on the allocation of labor across firms, our results naturally depend on the micro-level details of the labor market, as we next show.

We vary, one at a time, three core parameters which modulate the strength of the 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 by searching for better jobs while at the low rungs. We show that varying these parameters has a significant impact on within-location labor market moments and on the aggregate gains, but does not affect the East-West wage gap.

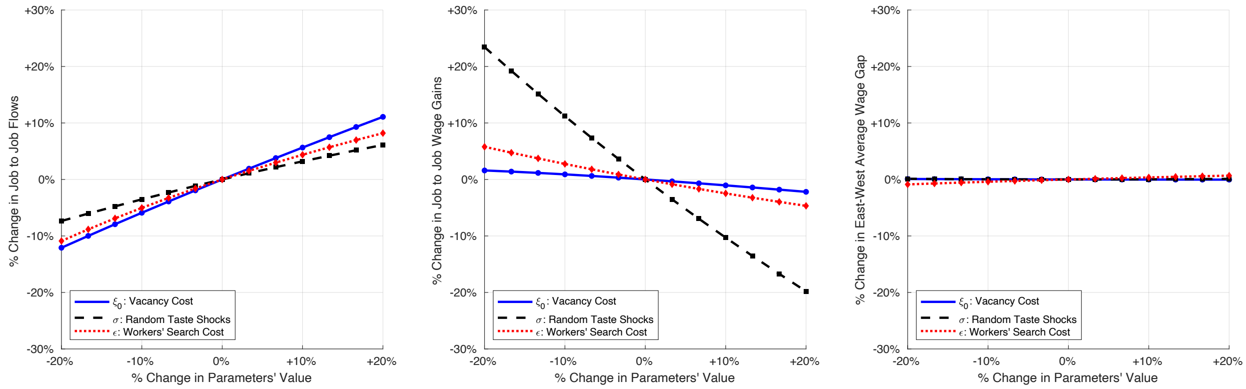
We first show the effect of these parameters on within-location labor market moments and on the wage gap in Figure 10. Panel (a) illustrates that within-location job-to-job flows increase relative to the baseline with each parameter. Panel (b) shows that the within-location wage gains for movers decline sharply with the variance of preference shocks σ, but are relatively unaffected by the other two parameters.[67] When σ is large, workers' moves are more frequently due to preferences rather than wage differences, reducing the average wage gain. Finally, Panel (c) highlights that changing the labor market frictions has no significant effect on the aggregate wage gap between East and West Germany since, as argued in Section 5.1, these parameters mainly affect the distribution of labor within, rather than between, regions.

[66] This result is possibly not very surprising as there are no real scale economies in our model. The matching function has constant returns to scale in vacancies and applications, and each location and each firm produce an identical good.

[67] While we report in the figure only the within-location moments, we note that the changes in cross-location flows and wage gains are very similar (in percentage terms).

Figure 10: Sensitivity of Micro and Macro Moments to Labor Market Parameters

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



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). The middle panel shows the wage gains obtained from move within region (marked with a cross) and between regions. The right panel shows the gap in average wage between West and East Germany.

We next compute, just as in Section 6.1, the gains from removing spatial frictions in these alternative economies with different labor market frictions. As Panel (a) of Figure 11 shows, the aggregate gains in output per capita decline substantially for higher values of the labor market parameters, i.e., as 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%. This result is intuitive: higher labor mobility implies smaller potential gains from improving the within-region allocation of labor. This result is also important: two economies could look identical in terms of their wage gap between regions (as shown in Figure 10), yet removing spatial frictions could lead to very different aggregate outcomes dependent on the economies' labor market frictions.

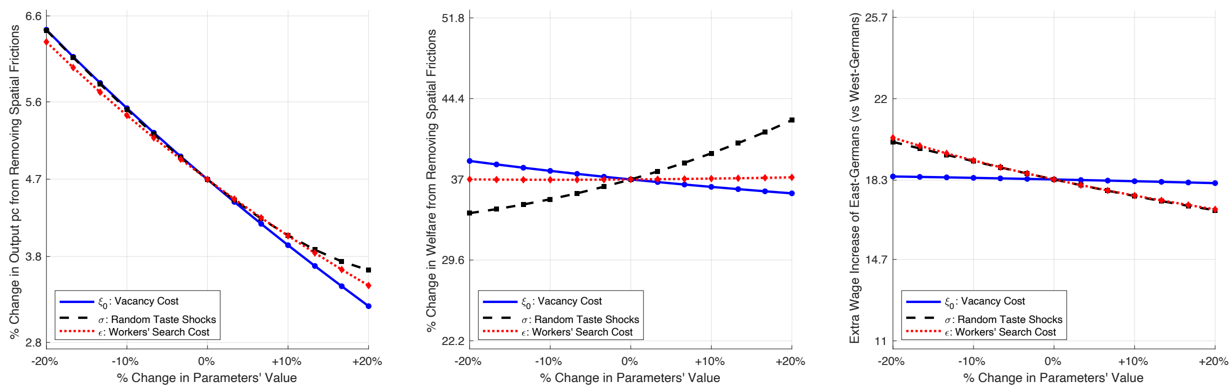
The impact of the spatial frictions on either the workers' value or the relative wage of East Germans is much less sensitive to the value of the labor market parameters (Panels (b) and (c)). For these two statistics, the allocation of labor within location is less relevant: removing spatial frictions mostly changes the value functions because workers receive more job opportunities and no longer pay the moving or utility cost, rather than because of within-location frictions. Similarly, East Germans' wages rise relative to West Germans' mainly because they move to the higher productivity West.

7 Conclusion

This paper's main finding is that taking into account how local labor markets function, and how they are hampered by search and matching frictions, is important to quantify the effects of

Figure 11: Sensitivity of the Aggregate Effects to Labor Market Parameters

(a) Output per Capita Gains (b) Welfare Gains (c) Wage Gains of East Germans



Notes: We vary three labor market parameters and recompute the effect of removing spatial frictions under these alternative calibrations. The three panels show the effect on GDP per capita (left), workers' value function (middle) and relative wage increase of East-born (right), plotted as a function of the change in the primitive parameters relative to the baseline. To ease comparability across the different panels, we standardized the y-axis to cover changes of + 40 % to - 40 % relative to the baseline value of the statistic.

spatial frictions on the aggregate economy. Barriers impeding labor mobility across space affect firms' local monopsony power, which, in turn, shapes the allocation of labor to firms and thus aggregate productivity.

To reach this overall 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 new insights that make our core takeaway concrete and that are relevant beyond the context of Germany.

First, removing spatial frictions can improve the allocation of workers to firms within locations, generating aggregate gains in addition to the typically emphasized reallocation of workers across locations. 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 leads to the reallocation of 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.

References

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

Alesina, A., and N. Fuchs-Schündeln (2007): "Good-bye Lenin (or not?): The effect of communism on people's preferences," *American Economic Review*, 97(4), 1507–1528.

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

Alvarez, J. A. (2020): "The agricultural wage gap: Evidence from Brazilian micro-data," *American Economic Journal: Macroeconomics*, 12(1), 153–73.

Amior, M. (2019): "Education and geographical mobility: the role of the job surplus," .

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

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

Bachmann, R., C. Bayer, H. Stüber, and F. Wellschmied (2021): "Monopsony Makes Firms not only Small but also Unproductive: Why East Germany has not Converged," mimeo.

Bagger, J., and R. Lentz (2019): "An empirical model of wage dispersion with sorting," *The Review of Economic Studies*, 86(1), 153–190.

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

Bassi, V., R. Muoio, T. Porzio, R. Sen, and E. Tugume (2021): "Achieving scale collectively," Discussion paper, National Bureau of Economic Research.

Baum-Snow, N., and R. Pavan (2012): "Understanding the city size wage gap," *The Review of economic studies*, 79(1), 88–127.

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

Berger, D., K. Herkenhoff, and S. Mongey (2022): "Labor market power," *American Economic Review*, 112(4), 1147–93.

Bilal, A. (2021): "The geography of unemployment," Discussion paper, National Bureau of Economic Research.

Bilal, A., N. Engbom, S. Mongey, and G. L. Violante (2021): "Labor Market Dynamics When Ideas are Harder to Find," Discussion paper, National Bureau of Economic Research.

Bilal, A. G., N. Engbom, S. Mongey, and G. L. Violante (2019): "Firm and worker dynamics in a frictional labor market," Discussion paper, National Bureau of Economic Research.

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

Boeri, T., A. Ichino, E. Moretti, and J. Posch (2021): "Wage equalization and regional misallocation: evidence from Italian and German provinces," *Journal of the European Economic Association*, 19(6), 3249–3292.

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

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

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

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

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

Burdett, K., C. Carrillo-Tudela, and M. Coles (2020): "The cost of job loss," *The Review of Economic Studies*, 87(4), 1757–1798.

Burdett, K., and M. Coles (2003): "Equilibrium wage-tenure contracts," *Econometrica*, 71(5), 1377–1404.

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

Caldwell, S., and O. Danieli (2021): "Outside options in the labor market," mimeo.

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

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

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

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

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

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

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

Dauth, W., S. Y. T. Lee, S. Findeisen, and T. Porzio (2021): "Transforming Institutions: Labor Reallocation and Wage Growth in a Reunified Germany," mimeo.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2018): "The geography of development," *Journal of Political Economy*, 126(3), 903–983.

Diamond, R. (2016): "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000," *American Economic Review*, 106(3), 479–524.

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

Elsby, M. W., and A. Gottfries (2022): "Firm dynamics, on-the-job search, and labor market fluctuations," *The Review of Economic Studies*, 89(3), 1370–1419.

Engbom, N. (2020): "Misallocative Growth," mimeo.

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

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

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

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

Galenianos, M., P. Kircher, and G. Virag (2011): "Market power and efficiency in a search model," *International economic review*, 52(1), 85–103.

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

Giannone, E. (2017): "Skill-biased technical change and regional convergence," mimeo.

Henkel, M., T. Seidel, and J. Suedekum (2021): "Fiscal transfers in the spatial economy," *American Economic Journal: Economic Policy*, 13(4), 433–68.

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

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

Hirsch, B., E. J. Jahn, A. Manning, and M. Oberfichtner (2022): "The urban wage premium in imperfect labor markets," *Journal of Human Resources*, 57(S), S111–S136.

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

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

Hsieh, C.-T., and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124(4), 1403–1448.

Hsieh, C.-T., and E. Moretti (2019): "Housing constraints and spatial misallocation," *American Economic Journal: Macroeconomics*, 11(2), 1–39.

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

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

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

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

Kaboski, J. P., and R. M. Townsend (2011): "A structural evaluation of a large-scale quasi-experimental microfinance initiative," *Econometrica*, 79(5), 1357–1406.

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

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

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

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

Le Barbanchon, T., R. Rathelot, and A. Roulet (2020): "Gender Differences in Job Search: Trading off Commute against Wage," *Quarterly Journal of Economics*, 136(1), 381–426.

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

Manning, A. (2013): *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.

Manning, A., and B. Petrongolo (2017): "How local are labor markets? Evidence from a spatial job search model," *American Economic Review*, 107(10), 2877–2907.

Martellini, P. (2022): "Local labor markets and aggregate productivity," Discussion paper, Working Paper.

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

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *econometrica*, 71(6), 1695–1725.

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

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

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

Nakamura, E., J. Sigurdsson, and J. Steinsson (2022): "The gift of moving: Intergenerational consequences of a mobility shock," *The Review of Economic Studies*, 89(3), 1557–1592.

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

Pavcnik, N. (2002): "Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants," *The Review of economic studies*, 69(1), 245–276.

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

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

Restuccia, D., and R. Rogerson (2013): "Misallocation and Productivity," *Review of Economic Dynamics*, 16(1), 1–10.

Roca, J. D. L., and D. Puga (2017): "Learning by working in big cities," *The Review of Economic Studies*, 84(1), 106–142.

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

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

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

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

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

Van Den Berg, G. J., and G. Ridder (1998): "An Empirical Equilibrium Search Model of the Labor Market," *Econometrica*, 66(5), 1183–1221.

Yeh, C., C. Macaluso, and B. Hershbein (2022): "Monopsony in the US Labor Market," SSRN Working Paper No. 4049993.

Online Appendix

A Further Details on Data and Data Construction

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

BHP Data

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

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

We obtain the mean gross daily wage paid to full-time employees by each firm in each year. Since the social security notifications contain earnings only reported up to the upper limit for earnings for statutory pension insurance contributions, approximately 10% of full-time employees' earnings are censored. To remedy this issue, the BHP provides a corrected mean gross daily wage for each firm (te_imp_mw), which we use for all our analyses. This variable imputes the missing wages for each worker before the mean firm wage is calculated. The imputation procedure follows Card, Heining, and Kline (2015).

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

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

Table A1: Summary Statistics of the BHP Dataset

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

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

LIAB Data

The LIAB data provide matched employer-employee data that link more than 1.9 million individuals to about 400,000 firms for which these individuals work, for 1993-2014. The data contain information for the unemployment spells during which workers receive unemployment insurance benefits. Workers do not appear in the data if they are self-employed, in the public sector, or unemployed without receiving UI benefits.

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

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

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

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

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

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

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

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

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

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

Table A3: Summary Statistics of the LIAB Dataset

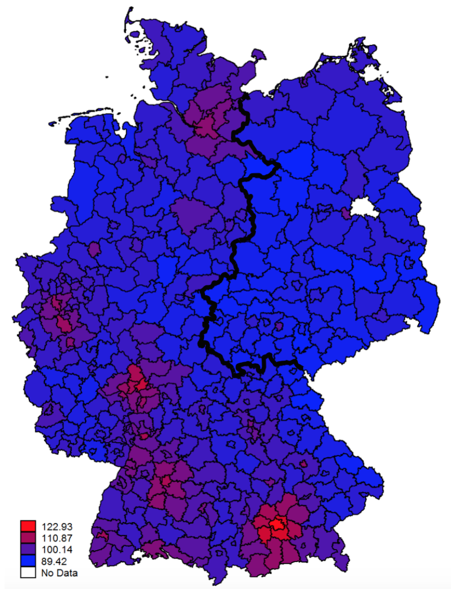
	Variable	Obs	Mean	Std. Dev.
(1)	Real wage of FT employed	$7,963,537$	111.890	76.967
(2)	Real wage of unemployed	$1,254,063$	27.580	12.469
(3)	Employed dummy	$9,485,701$	0.849	0.358
(4)	Age	$9,485,701$	40.172	11.538
(5)	Male dummy	$9,485,701$	0.696	0.460
(6)	College dummy	$5,904,697$	0.205	0.403
(7)	Work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

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

Price Deflators

We obtain data on regional prices from a study of the Federal Institute for Building, Urban Affairs and Spatial Development (BBSR (2009)). The study computed prices in 2007 for 393 micro regions covering all of Germany that correspond to cities, counties, or slightly larger unions of counties. The data cover about two thirds of the consumption basket, including housing rents, food, durables, holidays, and utilities. Of the 402 counties in the IAB data, 311 are directly represented in the BBSR data. A further 81 counties in the IAB data can be mapped to 41 regions in the BBSR data that are slightly larger than a county or combine multiple counties. The remaining 10 counties in the IAB data are represented in the BBSR data by the main town within them. Using this mapping, we obtain a price level in 2007 for each of the 402 counties in the IAB data. We then obtain for each federal state the GDP deflator from the growth accounting of the states (Volkswirtschaftliche Gesamtrechnungen der Länder, VGRdL). We apply each state's deflator to all counties in that state to obtain a county-level price index for each year in 2009-2014. Figure A1 shows a map of the price levels in 2007.

Figure A1: Price Level, 2007



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

B Additional Statistics on Worker Mobility

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

Column 1 in the top panel of Table A4 presents the number of cross-region migrants in our core sample. Migrants are defined, as in the main text, as all workers moving job-to-job between East and West Germany that change their residence in the year of the move compared to one year earlier. Job-to-job moves are defined as job switches between two firms without an intermittent unemployment spell (but possibly non-employment), as in the main text. Our sample contains about $14,000$ job-to-job migrants between East and West Germany (row 1), with slightly more switches from East-to-West than from West-to-East (rows 2 and 3). Column 2 of the top panel presents the same statistics using all job-to-job switchers across regions, including those that do not change their residence. Comparing the total number of job-to-job movers in column 2 to the number of migrants in column 1, we find that about 80% of cross-region job moves are done without a reported change in residence. We will refer to such moves as "commuting". However, as discussed in the main text, social security reporting regulations do not prescribe which residence to report for individuals with multiple residences, and therefore some individuals may not list the residence closest to their job. It is therefore not possible to know with certainty whether individuals that do not report a change in residence are in fact commuting or whether their residence location is misclassified. As we discuss in Section 5.3, in our estimation we therefore consider a third, "intermediate" version of cross-region migration. This variable is defined as all migration moves plus all cross-region job switches without a change in residence where the distance between residence and work is less than 200km at both the origin and the destination, provided that the move takes the worker further away from her current residence. We impose the upper bound on the distance between work and residence to remove workers with implausibly long commutes. Moreover, we require the distance to the residence to increase to remove job changes that take the worker closer to her current residence, since such moves do not really impose a moving cost on the worker. Column 3 presents statistics for moves based on this definition. We explore the sensitivity of our structural estimates to different alternatives in a dedicated Supplemental Appendix N.

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

200km of their location of residence. Adding these workers lowers the average distance between jobs for cross-region movers slightly, to 234km.

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

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

The bottom panels of Table A5 and A6 show some characteristics of stayers, movers, and movers that return home. We find that the share of college-educated migrants is significantly higher than the share of college-educated movers overall. West German migrants and movers are significantly more likely to be college-educated than East German migrants and movers. Considering the gender of migrants, we find that the male share among migrants is comparable to the male share among non-migrants for both East and West Germans. However, East German movers overall are significantly more likely to be male than stayers.

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

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

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

Table A4: Number of Movers Between East and West Germany

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

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

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

Table A5: Summary Statistics for Migrants

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

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

Table A6: Summary Statistics for Job Moves since 1999

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

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

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

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

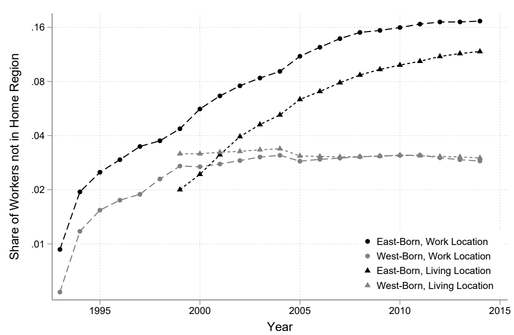
Source: LIAB. Notes: The table shows statistics for workers with at least one employment spell in our core sample period 2009-2014. For these workers, we compute the distribution of the number of cross-border moves throughout their lifetime, going back as many years as available. The first two columns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

Table A8: Mobility by Cohort

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

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

Figure A2: Stock of Individuals away from Home Region



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

C Results from the Socio-Economic Panel

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

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

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

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

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

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

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

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

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

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

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

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

D Proofs

D.1 Equilibrium in the Goods Market

The firm's problem in the goods market is

$$\hat{\pi}_j(w) = \max_{n_h, n_c, k} pn_c + P_{h,j} \left(pn_h\right)^{1-\alpha} k^{\alpha} - \rho_j k \tag{21}$$

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

$$n_h = \frac{\rho_j}{p} \frac{1-\alpha}{\alpha} k \tag{22}$$

and assuming that both goods are supplied in equilibrium

$$P_{h,j} = \rho_j^{\alpha} \alpha^{-\alpha} \left(1-\alpha\right)^{-(1-\alpha)} . \tag{23}$$

We can plug (22) and (23) into (21) to obtain

$$\hat{\pi}_j(w) = pn_j(w) = p \sum_{i \in \mathbb{I}} \theta_j^i l_j^i(w), \tag{24}$$

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

$$P_{h,j} H_j = \left(1-\eta\right) P_j Y_j, \tag{25}$$

where, assuming that consumers own the firms and using (24), their total income is

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

and Y_j is real GDP.

Using the production function $h = (pn_h)^{1-\alpha} k^{\alpha}$, and plugging in (22), 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 (23), implies

$$P_{h,j} H_j = \frac{1}{\alpha} \rho_j K_j. \tag{26}$$

Combining demand and supply yields

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

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

We can express the equilibrium condition in terms of ratios as follows. Starting from $P_j = (P_{h,j})^{1-\eta}$, we can substitute in with (23) and use the supply equation (26) 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 (25) 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

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

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

The probability of an unemployed worker accepting this offer is

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

and the expected value of an offer is

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

We next provide expressions for (7) and (8) after solving out for the optimal search effort.

Defining the expected value gain from location x as $\bar{V}_{jx}^{E,i}(w) \equiv \int V_{jx}^{E,i}(w,w')\,dF_x(w') - W_j^i(w)$ and replacing the functional forms for $a_{jx}^i(\cdot)$ and $\psi(\cdot)$, we get

$$rW_j^i(w) \quad = \quad \frac{w\theta_j^i\tau_j^i}{P_j} + \frac{\epsilon}{1+\epsilon}\sum_{x\in\mathbb{J}}\left[z_{jx}^i\vartheta_x^{1-\chi}\bar{V}_{jx}^{E,i}(w)\right]^{\frac{1+\epsilon}{\epsilon}} + \delta_j^i\left[U_j^i - W_j^i(w)\right]. \tag{27}$$

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

$$rU_j^i = \frac{b_j^i\theta_j^i\tau_j^i}{P_j} + \nu\frac{\epsilon}{1+\epsilon}\sum_{x\in\mathbb{J}}\left[z_{jx}^i\vartheta^{1-\chi}\bar{V}_{jx}^{U,i}(b)\right]^{\frac{1+\epsilon}{\epsilon}}. \tag{28}$$

These expressions highlight the relationship between the values of employment and unemployment and the primitive parameters ν and ϵ.

D.3 Proof of Proposition 1

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

$$\pi_j(p) = \max_w (p-w)\sum_{i\in\mathbb{I}}\theta_j^i l_j^i(w) \tag{29}$$

and, as shown,

$$l_j^i(w) = \frac{\mathcal{P}_j^i(w)\,\vartheta_j^{-\chi\frac{\bar{a}_j^i}{\bar{a}_j}}}{q_j^i(w)} \quad \text{if } w \geq R_j \tag{30}$$

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 (29) and obtain

$$\frac{(p-w_j(p))\left(\sum_{i\in\mathbb{I}}\theta_j^i\frac{\partial l_j^i(w_j(p))}{\partial w}\right)}{\left(\sum_{i\in\mathbb{I}}\theta_j^i l_j^i(w_j(p))\right)} \quad = \quad 1. \tag{31}$$

From equation (30), we find

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

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 (31) to find a differential equation for $w_j(p)$

$$\frac{\partial w_j(p)}{\partial p} = \frac{(p - w_j(p))\left(\sum_{i\in\mathbb{I}}\theta_j^i\frac{\frac{\partial \tilde{\mathcal{P}}_j^i(p)}{\partial p}\tilde{q}_j^i(p) - \tilde{\mathcal{P}}_j^i(p)\frac{\partial \tilde{q}_j^i(p)}{\partial p}}{\tilde{q}_j^i(p)^2}\vartheta_j^{-\chi}\frac{\bar{a}_j^i}{\bar{a}_j}\right)}{\left(\sum_{i\in\mathbb{I}}\theta_j^i\frac{\tilde{\mathcal{P}}_j^i(p)}{\tilde{q}_j^i(p)}\vartheta_j^{-\chi}\frac{\bar{a}_j^i}{\bar{a}_j}\right)}. \tag{32}$$

Since $w_j(p)$ is continuous at p by assumption, the differential equation (32), 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

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

We have thus proved that

$$w_j(p) = w_j\left(\underline{p}_j\right) + \int_{\underline{p}_j}^{p}\frac{\partial w_j(z)}{\partial z}\gamma_j(z)\,dz \tag{33}$$

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

D.4 Comparison to the Burdett-Mortensen Model

Lemma 1. *If $a_{jx}^i(s_x) = 1$ and $\kappa_{jx}^i = 0$ for all i, j, and x, $\theta_j^i = 1$, $\tau_j^i = \tau_j$, $\delta_j^i = \delta$, $b_j^i\tau_j^i P_j^{-1} = \hat{b}$, and $R_j^i\tau_j^i P_j^{-1} = \hat{R}$ for all i and j, $\nu = 1$, $\chi = 0$, and $\sigma \to 0$, then the ODEs for the wage functions simplify to*

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

where

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

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

and

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

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

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

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

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

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

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

Define $\hat{p}\equiv p\tau_jP_j^{-1}$. The firm's maximization problem (10) becomes

$$\hat{\pi}_j\left(\hat{p}\right)=\frac{P_j}{\tau_j}\max_{\hat{w}}\left(\hat{p}-\hat{w}\right)\hat{l}\left(\hat{w}\right) \tag{34}$$

for all j, where $\hat{l}\left(\hat{w}\right)\equiv l_j(w\tau_jP_j^{-1})$. From $a_{jx}^i(s_x)=1$ and $\chi=0$ it follows that

$$\hat{l}(\hat{w})=\frac{\hat{\mathcal{P}}\left(\hat{w}\right)}{\hat{q}\left(\hat{w}\right)}\quad\text{if }\hat{w}\geq\hat{R}, \tag{35}$$

where $\hat{R}\equiv R_j^i\tau_jP_j^{-1}$ is constant across regions by assumption. Since $\delta_j^i=\delta$, we have

$$\hat{q}\left(\hat{w}\right)=\delta+\sum_{x\in\mathbb{J}}\bar{v}_x\left[1-\hat{F}_x\left(\hat{w}\right)\right] \tag{36}$$

and

$$\hat{\mathcal{P}}(\hat{w}) = \sum_{x\in\mathbb{J}} \left[\hat{E}_x(\hat{w}) + u_x\right], \tag{37}$$

where $\hat{E}_x(\hat{w}) \equiv E_x(w\tau_j P_j^{-1})$.

The first-order condition of the wage posting problem is

$$\frac{(\hat{p}-\hat{w})\left(\frac{\partial \hat{l}(\hat{w})}{\partial \hat{w}}\right)}{\left(\hat{l}(\hat{w})\right)} \quad = \quad 1, \tag{38}$$

where

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

Plugging this latter expression into the first-order condition gives

$$\frac{(\hat{p}-\hat{w})\left(\frac{\partial \hat{\mathcal{P}}(\hat{w})}{\partial \hat{w}}\hat{q}(\hat{w}) - \frac{\partial \hat{q}(\hat{w})}{\partial \hat{w}}\hat{\mathcal{P}}(\hat{w})\right)}{\hat{\mathcal{P}}(\hat{w})\,\hat{q}(\hat{w})} = 1. \tag{39}$$

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

$$\hat{w}(\underline{\hat{p}}) = \hat{R}. \tag{40}$$

From (34), firms with the same $\hat{p}$ post the same wage $\hat{w}$ and therefore attract the same number of workers. Moreover, from the usual complementarity between firm size and productivity, more productive firms post higher real wages $\hat{w}$. Define a job offer distribution across regions as a function of productivity

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

where

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

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

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

Define $\tilde{x}(\hat{p}) \equiv \hat{x}(\hat{w}(p))$ for any $\hat{x}$. We can then re-define (36) and (37) using these definitions to obtain

$$\tilde{q}\left(\hat{p}\right) = \delta + \bar{\bar{v}}\left[1 - \tilde{F}(\hat{p})\right] \tag{41}$$

and

$$\tilde{\mathcal{P}}(\hat{p}) = \tilde{E}(\hat{p}) + u \equiv (1-u)\tilde{G}(\hat{p}) + u, \tag{42}$$

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

Using

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

we re-write the first-order condition (39) as

$$\frac{\partial\hat{w}(p)}{\partial\hat{p}} = \frac{\left(\hat{p} - \hat{w}(\hat{p})\right)\left(\frac{\partial\tilde{\mathcal{P}}(\hat{p})}{\partial\hat{p}}\tilde{q}\left(\hat{p}\right) - \frac{\partial\tilde{q}(\hat{p})}{\partial\hat{p}}\tilde{\mathcal{P}}\left(\hat{p}\right)\right)}{\tilde{\mathcal{P}}\left(\hat{p}\right)\tilde{q}\left(\hat{p}\right)}. \tag{43}$$

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

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

and hence

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

The mass of unemployed is given from (19) by

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

Substituting these expressions into (42) gives

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

Plugging this expression for the acceptance probability and its derivative into (43), we obtain

$$\frac{\partial\hat{w}(\hat{p})}{\partial\hat{p}} = \frac{-2\left(\hat{p} - \hat{w}(\hat{p})\right)\frac{\partial\tilde{q}(\hat{p})}{\partial\hat{p}}}{\tilde{q}\left(\hat{p}\right)}. \tag{44}$$

Together, equations (36), (37), (40), and (44) are the functions stated in the proposition, redefined on $\hat{p}$ instead of on p, and are the same as in the standard Burdett-Mortensen model. $\square$

E Identification

E.1 Identification of Moving Costs, Preferences, and Search Efficiency

In this section, we provide further details on how various spatial frictions are identified.

Moving Costs and Location Preferences: τ and κ. We can pin down these moments using the average wage gain conditional on a move for an individual of type i, employed in location j, and taking a job in location x[68]

$$\underbrace{\mathbb{E}\left[\log(w_x^i\theta_x^i) - \log(w_j^i\theta_j^i)\right]}_{\text{Average Observed Wage Gain}} = \underbrace{\log\left(\theta_x^i\right) - \log\left(\theta_j^i\right)}_{\text{Comparative Advantage}} + \int\left(\int \underbrace{(\log w' - \log w)}_{\text{Wage Gain}} \underbrace{\frac{\mu_{jx}^{E,i}(w,w')}{\bar{\mu}_{jx}^{E,i}(w)}}_{\text{Rel. Prob. Accept}} \underbrace{\mathrm{d}F_x(w')}_{\text{Offers CDF}}\right) \underbrace{\frac{a_{jx}^{E,i}(w)}{\bar{a}_{jx}^{E,i}}\mathrm{d}E_j^i(w)}_{\text{Weighted Employment CDF}}, \tag{45}$$

where $\bar{a}_{jx}^{E,i} \equiv \int a_{jx}^{E,i}(w)\,\mathrm{d}E_j^i(w)$ and $\bar{\mu}_{jx}^{i,E}(w) \equiv \int \mu_{jx}^{i,E}(w,w')\,\mathrm{d}F_x(w')$.

Given offer distributions $F_x(\cdot)$, employment distributions $E_j^i(w)$, and the share of applications coming from each firm $\frac{a_{jx}^{E,i}(w)}{\bar{a}_{jx}^{E,i}}$, which are all mostly shaped by labor market frictions and therefore identified from within-location moments, as well as an estimate of skills θ, the equation directly relates the moving costs κ and local preferences τ to the relative wage gains of cross-location movers. Consider the limiting case when $\sigma \to 0$. In that case, workers accept an offer if and only if $W_x^i(w') - \kappa_{jx}^i \geq W_j^i(w)$. Since the value functions are increasing, the cutoff wage level $\hat{w}_{jx}^i(w)$ at which an individual of type i employed in location j would accept an offer from location x is an increasing function of w. An increase in κ_{jx}^i, or a decrease in τ_x^i, would raise this cutoff wage for any level of w. As the worker accepts only relatively better offers, the expected wage gain of a move increases in κ_{jx}^i and decreases in τ_x^i. As discussed in the main text, we separately identify moving costs and preferences by assuming that moving costs are identical for all worker types. Under that assumption, the location preferences are identified from the differences in wage gains for individuals of different types that make the same migration move.

Search Efficiency: z. Given an estimate of the labor market frictions, as well as estimates of skills, moving costs, and preferences (θ, κ, τ), we can recover the relative search efficiencies

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

from the relative job-to-job flows within and between locations. The rate at which workers of type i currently employed in location j move towards a job in location x is given by

$$\underbrace{\psi^{i}_{jx}}_{\text{Quit Rate}} = \left[\underbrace{\vartheta_x^{1-\chi}}_{\text{Tightness}} \underbrace{\bar{a}^{E,i}_{jx}}_{\text{Applications}}\right] \times \left[\int \left(\int \underbrace{\mu^{E,i}_{jx}(w,w')}_{\text{Prob. Accept}} \underbrace{\mathrm{d}F_x(w')}_{\text{Offer CDF}} \right) \underbrace{\frac{a^{E,i}_{jx}(w)}{\bar{a}^{E,i}_{jx}} \mathrm{d}E^{i}_{j}(w)}_{\text{Weighted Employment CDF}}\right] \tag{46}$$

Since $\bar{a}^{E,i}_{jx} = z^{i}_{jx} \bar{s}^{E,i}_{x}$, where $\bar{s}^{E,i}_{x} \equiv \int s^{E,i}_{jx}(w)\,\mathrm{d}E^{i}_{j}(w)$, a lower search efficiency z^{i}_{jx} leads to lower job-job flows from location j to x given the acceptance probability $\mu^{E,i}_{jx}(w,w')$, which is not directly affected by z^{i}_{jx} itself.

E.2 Identification of Worker Skills

In this section, we describe how an augmented AKM specification can recover the comparative advantage of individuals across locations. We discuss here a specification for East and West Germany. We estimate the model in Section G.1 and show that the data do not show any evidence of comparative advantages between these two regions. Given the lack of an East-West comparative advantage, we do not extend the analysis to the level of the four finer locations we use in the estimation in Section 5. However, the same insights and identification strategy would apply and could be performed.

Specification of the Baseline Model

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

$$\log(w_{it}) = \alpha_i + \psi_{J(i,t)} + \beta \mathbb{I}^{(h_i \neq R(J(i,t)))} + BX_{it} + \epsilon_{it}, \tag{47}$$

where i indexes full-time workers, t indexes time, and $J(i,t)$ indexes worker i's firm at time t.[69] In this specification, α_i is the worker component, $\psi_{J(i,t)}$ is the component of the firm j for which worker i works at time t, and $\mathbb{I}^{(h_i \neq R(J(i,t)))}$ is a dummy that is equal to one if worker i with home region h_i (either East or West Germany) is currently employed at a firm in the other region. This term picks up the comparative advantage of workers in their home region. Finally, X_{it} is a centered cubic in age and an interaction of age and college degree, as in Card, Heining, and Kline (2013).

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

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

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

In this specification, the mean-zero match effect $\eta_{iJ(i,t)}$ represents an idiosyncratic wage premium or discount that is specific to the match, ζ_{it} reflects the drift in the persistent component of the individual's earnings power, which has mean zero for each individual, and ϵ_{it} is a mean-zero noise term capturing transitory factors. We estimate the model on the largest connected set of workers in our data.[70]

Identification of the Model with Comparative Advantage

We now discuss how the specification (47) allows us to identify, through β, the comparative advantage effect by region. The same idea immediately extends to more locations.

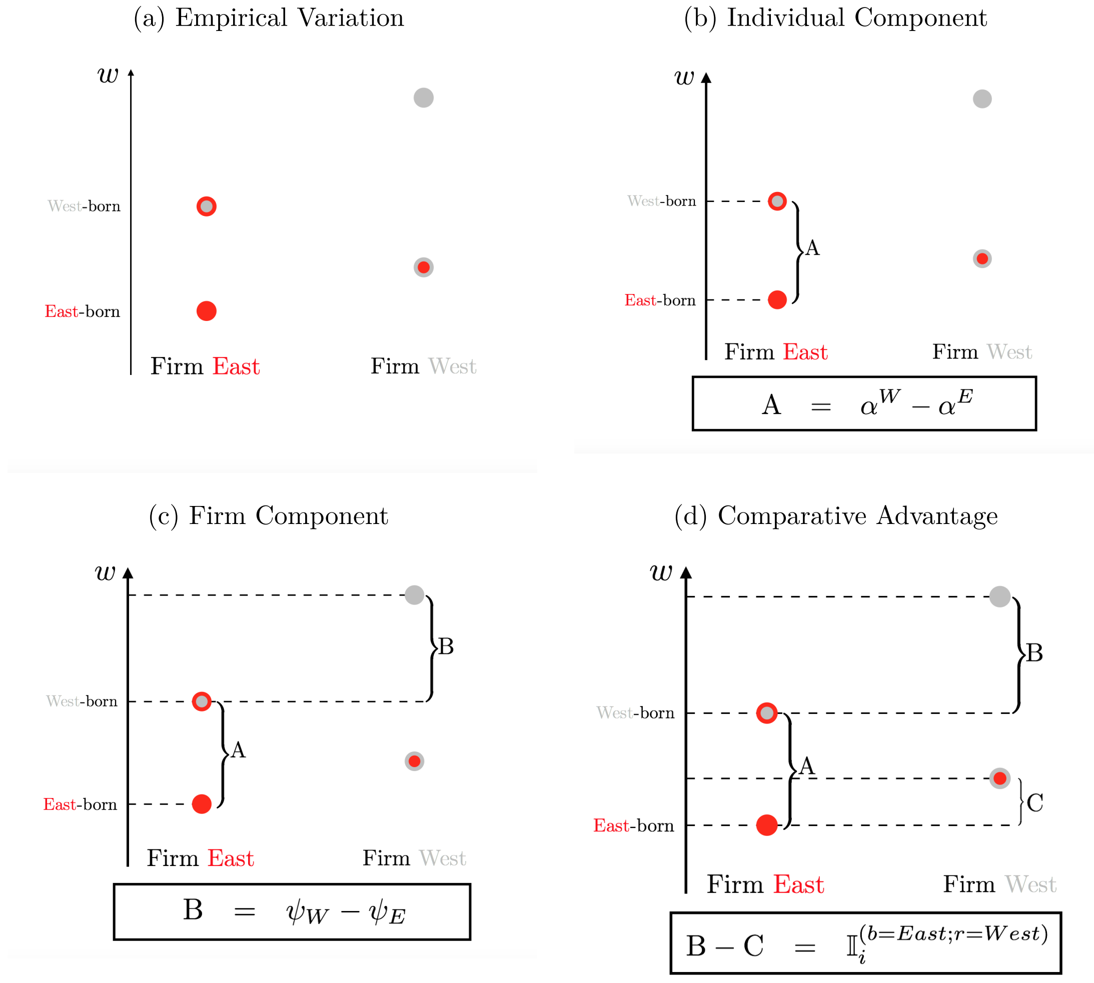
Consider four wage observations associated with two workers: an East-born and a West-born individual working in one firm in the East, and the same two individuals working in one firm in the West. Figure A3a plots an example of these two workers' wages, where the x-axis is the identity of the firm, the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

Note that the methodology cannot separately identify whether it is the East or the West-born worker that has a comparative (dis)advantage since all that is observed is their relative wage gap. For example, if the East German worker's wage is relatively lower than the West

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

German's wage at a firm in the West than at a firm in the East, then this difference could either arise because the East-born worker has a relative disadvantage in the West or because the West-born worker has a relative disadvantage in the East. As a result, the estimated β captures the sum of the two comparative advantages (East-born for East-Germany and West-born for West-Germany) and we need to make an arbitrary assumption in order to separately identify the two. In practice, we side-step this issue since we do not find evidence of comparative advantages. We show the estimation results Section G.1.

Figure A3: Identification of the AKM Components



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

F Additional Information on the Location

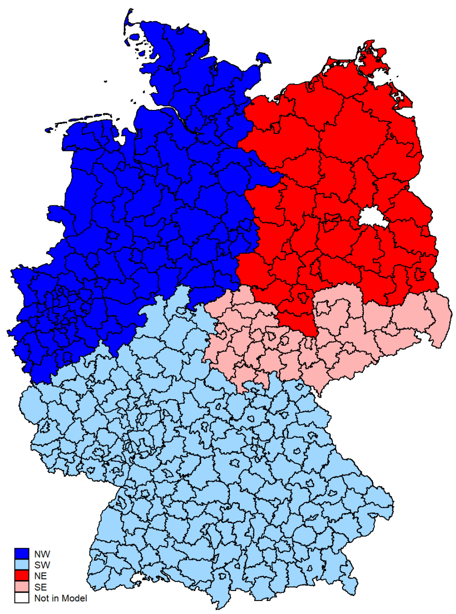
In this section, we provide more details on the four locations in our estimated model and the mobility between them. Figure A4 visualizes the four locations.

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

Table A12 presents statistics on worker mobility across locations, analogous to the discussion of mobility between the East and West German region in Appendix B. Column 1 in the top panel of Table A12 presents the number of cross-location migrants in our core sample. Migrants are defined, as in the main text, as all workers moving job-to-job between any two locations that change their residence in the year of the move compared to one year earlier. Our sample contains about $32,000$ job-to-job migrants between locations (row 1). Column 2 of the top panel presents the same statistics using all job-to-job switchers across locations, including those that do not change their residence. Similar to the cross-region job movers in Appendix B, about 80% of cross-location job moves are done without a reported change in residence. As discussed in the main text, social security reporting regulations do not prescribe which residence to report for individuals with multiple residences, and therefore some individuals may not list the residence closest to their job. Column 3 shows our third, "intermediate" version of cross-region migration, as discussed in Section 5.3. This variable is defined as all migration moves across locations plus all cross-location job switches without a change in residence where the distance between residence and work is less than 200km at both the origin and the destination, provided that the move takes the worker further away from her current residence. We impose the upper bound on the distance between work and residence to remove workers with implausibly long commutes. Moreover, we require the distance to the residence to increase to remove job changes that take the worker closer to her current residence, since such moves do not really impose a moving cost on the worker.

The bottom panel of Table A12 shows some additional statistics for cross-location job-to-job movers, analogous to Table A4. The columns titled "Work" show moments for the distance of the cross-location job-to-job move. The columns titled "To Live" present analogous statistics for the distance between the worker's new job after the cross-location job switch and the worker's residence. The same comments as in Appendix B apply. We note that the distances for cross-location moves are actually slightly larger than the distances for cross-region movers in Appendix B, reflecting the possibility to move large distances even within-region.

Figure A4: Locations in the Estimation



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

Table A11: Descriptive Statistics of the Locations

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

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

Table A12: Number of Movers Between Locations

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

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

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

G Parameters and Empirical Moments

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

G.1 Calibrated Parameters

We first describe how we compute the calibrated parameters shown in Table 2.

(1) Worker Skills

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

$$\log(w_{it}) = \alpha_i + \psi_{J(i,t)} + \beta \mathbb{I}^{(h_i \neq R(J(i,t)))} + BX_{it} + \epsilon_{it}, \tag{48}$$

and describe details on the identification in Section E.2. As is standard, we estimate the model on the largest connected set of workers in our data, since identification of workers and firm fixed effects requires firms to be connected through worker flows.[71] This sample includes approximately 97% of West and East workers in the LIAB.

The estimation yields a comparative advantage estimate of $\beta = 0.019$, indicating a small *negative* comparative advantage towards the home region. Thus, a typical East-born worker is paid, controlling for firm characteristics, almost 1% more if she works in the West.[72] One possible explanation for this finding could be selection, since the workers that move to the West could be those whose skills are particularly valuable there. Since the presence of the premium would require the remaining frictions to be larger to rationalize the lack of East-to-West mobility, we conservatively set the comparative advantage to zero in our estimation.

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

$$\hat{\alpha}_i = \eta^h \mathbb{I}_i^h + CX_i + \varepsilon_i, \tag{49}$$

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

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

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

estimated in logs, and present the exponentiated estimates in Table 2. We find that conditional on age, gender, and schooling, West-born workers earn, within the same firm, around 9% higher wages. The differences between locations within the East and within the West are small.

(2) Number of Firms by Region

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

(3) Workers by Birth Region

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

(4) Separation Rate

We assume that the separation rates δ^i_j depend only on the work location j and set them equal to the monthly probabilities, computed in the LIAB data, that workers separate into unemployment or permanent non-employment (i.e. either retired or dropping out of the labor force). Specifically, we compute in each month the share of employed workers that become unemployed or permanently move out of the sample. We do not include workers that are temporarily out of the sample between employment spells since such workers are included in our definition of job-to-job movers. Notice that workers move out of the sample if they are either self-employed, not employed, or employed in a public sector job. We drop 2014, the last year of our sample, to avoid misclassifying workers. We then take a simple average across months for each location.

(5) Price Level

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

(6) Payments to Fixed Factors

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

(7) Elasticity of the Matching Function

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

(8) Interest Rate

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

G.2 Moments for Estimation

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

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

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

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

$$\Delta \log(w_{it}) = \sum_{h \in \mathbb{H}} \sum_{s \in \mathbb{S}} \beta_{hs} d_{it}^{s} \mathbb{I}_{i}^{h} + BX_{it} + \gamma_t + \epsilon_{it}, \tag{50}$$

where $\Delta \log(w_{it})$ is the difference between a worker's log average real wage in the year after the job-to-job move and her log real wage in the job before the switch, d_{it}^{s} are dummies that are equal to one if worker i makes a job-to-job switch of type s at time t, and γ_t are year fixed effects. Here, $\mathbb{S}$ is the set of the 12 possible cross-location migration moves (NW-SW, NW-NE, NW-SE, SW-NW, and so on) and the 4 possible within-location moves. We define migration moves as all job switches across locations that entail the worker updating her residence county, plus all job moves that take the worker further away from her current residence as long as the worker's residence remains within 200km of her job, as discussed in more detail in Section 5.3. We interact the move dummies with four indicator variables $\mathbb{I}_{i}^{h}$ for worker i's home location (NW, SW, NE, or SE) to identify average wage gains separately for different types of workers. Thus, in total we have 16*4=64 move-by-birth dummies of interest. The controls X_{it} contain dummies for eight age groups (26-30 years, 31-35 years, ... 56-60 years, older than 60 years, where the group of under 26 year olds is the omitted category), a dummy for whether the worker has a college degree, and a dummy for the worker's gender. The controls also include 12 dummies for non-migration cross-location job moves (for example because the worker did not change residence location and moved closer to her residence), interacted with birth location dummies. We include these latter controls so that the variables of interest, d_{it}^{s}, pick up wage gains of migrants relative to stayers, the omitted category. Table A13 shows the estimated coefficients on the migration dummies, and their standard errors. All coefficients are tightly estimated given the very large sample size. For each coefficient, the first column indicates the worker's home location, the second column shows the location of the worker's initial job, and the top row shows the location of the worker's new job.

Table A13: Average Log Wage Gains for Job-Job Movers by Birth and Migration Locations

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

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

G.2.2 Flows of Job-to-Job Movers

We compute in our cleaned LIAB data in each month the number of workers making a job-to-job move between any combination of locations. There are 12 possible migration moves (NW-SW, NW-NE, NW-SE, SW-NW, and so on) and 4 possible within-location job moves. We define migration moves as all job switches across locations that entail the worker updating her residence county, plus all job moves that take the worker further away from her current residence as long as the worker's residence remains within 200km of her job, as discussed in more detail in Section 5.3. We compute these movers by worker home location (i.e., their type). In total, there are thus 16*4=64 worker flows. We translate these raw flows into shares by dividing them in each month by the total number of employed workers of the given type in the location of the origin job. We exclude workers that leave the sample in the next month from this calculation, since we do not have information on whether they move or stay within the location. We also exclude the last month in our data, December 2014, for the same reason. We then take the average of these shares across months.

Table A14 shows the resulting shares. For each worker home location (first column) and location of the current job (second column), we show the share of workers changing jobs to a given destination location (indicated in the top row) in an average month, as a fraction of all employed workers of the given home location and current location.

Table A14: Job-to-Job Migration Flows Between Locations by Birth Location

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

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

G.2.3 Employment Share

We count in our cleaned LIAB data in each month the number of employed workers of a given type (home location) living in each location, and we divide by the total number of employed workers of that type in our LIAB data to obtain shares. We then average across months. We similarly compute the share of employed workers working in each location. Table A15 presents these worker shares. The first column indicates the home location of the worker, and the second column indicates the residence/work location. Columns 3 and 4 show the shares of employed workers of the given home location that live in a given location (column 3) and work in a given location (column 4). In our baseline estimation, we use the residence location as target for the distribution of labor since it more closely reflects the way in which we define a cross-location move. We use the work location in some of the robustness checks in Supplementary Appendix N.

Table A15: Share of Employed Workers by Location of Residence or Work Location

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

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

G.2.4 Unemployment Share

We count in our cleaned LIAB data in each month the number of unemployed workers of a given type (home location) living in each location, and we divide by the total number of unemployed workers of that type to obtain shares. We then average across months. We similarly compute the share of unemployed workers by last work location of the worker. We obtain the last work location as the location of the most recent job before the unemployment spell, and we exclude unemployed workers whose last job was in Berlin and workers that do not have a prior employment spell. Table A16 presents these worker shares. In our baseline estimation, we use the residence location as target for the distribution of labor since it more closely reflects the way in which we define a cross-location move. We use the work location in some of the robustness checks in Supplementary Appendix N.

Table A16: Share of Unemployed Workers by Location of Residence or Location of Last Job

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

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

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

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

$$fe_{it} = \sum_{h \in \mathbb{H}} \sum_{l \in \mathbb{L}} \beta_{hl} \mathbb{I}_{it}^{l} \mathbb{I}_{i}^{h} + BX_{it} + \epsilon_{it}, \tag{51}$$

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

Table A17 shows the estimated coefficients. The first two columns with data show the estimated coefficients β_{hl} for workers with home location h indicated in column 1 and residence location l indicated in column 2, together with their standard errors. Each of the coefficients is relative to the coefficient of workers born in the Northwest and living in the Northwest. The last two data columns show the analogous estimates for workers with home location h indicated in column 1 and work location l indicated in column 2. In our baseline estimation, we use the moments related to the residence location as target since they more closely reflect the way in which we define a cross-location move. We use the moments related to the work location in some of the robustness checks in Supplementary Appendix N.

Table A17: Firm Fixed Effects by the Birth and Current Location of Workers

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

Notes: The table shows the estimated coefficients β_{hl} in specification (51). The first two columns with data show the coefficients for workers with home location h indicated in column 1 and residence location l indicated in column 2, together with their standard errors. Each of the coefficients is relative to the coefficient of workers born in the Northwest and living in the Northwest. The last two data columns show the analogous estimates for workers with home location h indicated in column 1 and work location l indicated in column 2.

G.2.6 Firm Component of Wages by Firm Location

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

$$fe_j = \sum_{l \in \mathbb{L}} \beta_l \mathbb{I}_j^l + \epsilon_j, \tag{52}$$

where fe_j is the estimated firm fixed effect of firm j, and $\mathbb{I}_j^l$ are dummies that are equal to one if firm j is in location l. Using the firm fixed effects instead of actual real wages isolates the firm component of wages and removes differences in wages due to worker composition. We do not include industry controls since we want our model to be consistent with the aggregate wage gaps between locations, which could partially be due to differences in industry composition. Our estimated productivity shifters therefore also reflect industry differences across locations, although they are not quantitatively important, as shown in Supplemental Appendix K. For similar reasons, we do not include demographic controls. Table A18 presents the estimated coefficients β_l for firm location l indicated in column 1, where NW is the omitted category.

While in our baseline specification we do not include controls since we simply want to capture the differences in average firm productivity across locations, we also computed an alternative specification with a vector of controls X_j. We control for firm-level averages, averaged across all workers at the firm, of dummies for eight age groups (26-30 years, 31-35 years, ... 56-60 years, older than 60 years, where the group of under 26 year olds is the omitted category), a dummy for whether a worker has a college degree, and a dummy for workers' gender. The results barely change.[73]

Table A18: Firm Fixed Effect by Location

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

Notes: The table presents the estimated coefficients β_l from specification (52) for firm location l indicated in column 1, where NW is the omitted category.

G.2.7 GDP per Capita

We obtain nominal GDP per capita for each federal state from the National Accounts of the States (Volkswirtschaftliche Gesamtrechnungen der Länder, VGRdL) for each year. To translate the nominal figures into real ones, we compute the price level in each state in 2007 as a population-weighted average across the county-level prices reported by the BBSR. We then extend the resulting state-level prices in 2007 forward to 2014 using the state-level deflators available in the VGRdL. We deflate each state's nominal GDPpc with the resulting prices in each year to obtain state-level real GDPpc in each year, and we aggregate to the location level using each state's population in each year, also reported in the VGRdL. We take a simple average over the years in our core sample period and normalize real GDP per capita in NW to 1. Table A19 presents the results.

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

Table A19: Average GDP per capita by Location

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

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

G.2.8 Unemployment Rate

We obtain the unemployment rate (Arbeitslosenquote bezogen auf abhängige, zivile Erwerbspersonen) of each federal state in each month from the official unemployment statistics of the German Federal Employment Agency. We compute this moment from the official statistics rather than from the smaller LIAB sample since the latter is not representative and includes unemployed individuals only for as long as they are receiving unemployment benefits. We aggregate across states to locations using each state's labor force as weight, and take a simple average across the months in our core sample period. Table A20 shows the estimates.

Table A20: Unemployment Rate by Location

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

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

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

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

$$\ln(y_{jlt}^{size}) = B_l X_{jlt} + \gamma_t + \epsilon_{jlt}, \tag{53}$$

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

Based on the four regressions (one for each work location l) we obtain residuals for the log number of workers at each firm j, $\hat{\epsilon}_{jlt}^{size}$. We add back the mean log number of workers in each location, $\overline{\ln(y_{jlt}^{size})}$, to obtain a cleaned number of workers, $\hat{y}_{jlt}^{size} = \exp[\overline{\ln(y_{jlt}^{size})} + \hat{\epsilon}_{jlt}^{size}]$. We then construct deciles of the distribution of residualized firm size in each location and compute the share of residualized workers employed in each decile. Table A21 presents the resulting shares. Each column of the table shows the share of the location's workers employed at firms in the decile of the location's residualized firm size distribution indicated in column 1.

Table A21: Share of Workers by Firm Size Decile and Location

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

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

G.2.10 Relationship between Firm Wage and Firm Size

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

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

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

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

$$\ln(\hat{y}^{wage}_{jlt}) = \beta_{0,l} + \beta_{1,l} \ln(\hat{y}^{size}_{jlt}) + \varepsilon_{jlt}, \tag{54}$$

and report the slope coefficients $\beta_{1,l}$ in Table A22. We also plot the non-parametric relationships between $\ln(\hat{y}^{wage}_{jlt})$ and $\ln(\hat{y}^{size}_{jlt})$ in Figure A9, panel (a).

Table A22: Log Wage on Log Firm Size by Location

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

Notes: The table presents the coefficients $\beta_{1,l}$ of regression (54), by location of the firm, indicated in the first column. The residualization procedure is described in the text.

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

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

$$\Delta \ln(w_{ilt}) = B_l X_{ilt} + \gamma_t + \epsilon_{ilt}, \tag{55}$$

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

$$\Delta \ln(\hat{w}_{ilt}) = \beta_{0,l} + \beta_{1,l} \ln(\hat{w}_{ilt-1}) + \varepsilon_{ilt} \tag{56}$$

and report the slope coefficients $\beta_{1,l}$ in Table A23. In this table, each row shows the estimated regression coefficient on the residualized log initial wage for job-to-job moves originating in the location indicated in the first column. We also plot the non-parametric relationships between $\Delta\ln(\hat{w}_{ilt})$ and $\ln(\hat{w}_{ilt-1})$ in Figure A9, panel (b).

Table A23: Log Wage Gain of Movers by Initial Wage

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

Note: The table presents the coefficients $\beta_{1,l}$ of regression (56), by location of the origin firm. The residualization procedure is described in the text.

G.2.12 Separation/Quit Rate by Initial Wage

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

$$y_{ilt} = B_l X_{ilt} + \gamma_t + \epsilon_{ilt}, \tag{57}$$

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

$$\hat{d}^{sep}_{ilt} = \beta_{0,l} + \beta_{1,l}\ln(\hat{w}_{ilt}) + \varepsilon_{ilt} \tag{58}$$

and report the slope coefficients $\beta_{1,l}$ in Table A24. In this table, each row shows the estimated regression coefficient on the residualized log initial real wage for separations from jobs in the location indicated in the first column. We also plot the non-parametric relationships between $\hat{d}_{ilt}^{sep}$ and $\ln(\hat{w}_{ilt})$ in Figure A9, panel (c).

Table A24: Avg. Separation Rates of Workers by Initial Wage

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

Notes: The table presents the coefficients $\beta_{1,l}$ of regression (58), by location of the firm. The residualization procedure is described in the text.

G.2.13 Standard Deviation of Wage Gains

We identify in our cleaned LIAB data all migration moves between any combination of locations (NW-SW, NW-NE, NW-SE, SW-NW, and so on). We define migration moves as all job switches across locations that entail the worker updating her residence county, plus all job moves that take the worker further away from her current residence as long as the worker's residence remains within 200km of her job, as discussed in more detail in Section 5.3. We also identify job-to-job moves within-location, for each of the four locations. We indicate for each move the home location of the worker making the move. We restrict the dataset to these job-to-job moves and compute the log real wage gain associated with each move, defined as the difference between a worker's log average real wage in the year after the job-to-job move and her log real wage in the job before the switch. We then residualize these wage gains to take out observable heterogeneity not present in our model by running, separately for each location of the initial job, the regression

$$\Delta \ln(w_{ilt}) = B_l X_{ilt} + \gamma_t + \epsilon_{ilt}, \tag{59}$$

where $\Delta \ln(w_{ilt})$ is the log real wage gain associated with the move of worker i with initial job in location l at time t. The controls X_{ilt} contain dummies for eight age groups (26-30 years, 31-35 years, ... 56-60 years, older than 60 years, where the group of under 26 year olds is the omitted category), a dummy for whether the worker has a college degree, and a dummy for the worker's gender. From these four regressions (one for each location of the initial job l), we construct residuals for the log real wage gain, $\hat{\epsilon}_{ilt}^{gain}$. We then compute the standard deviation of these residualized wage gains for each home location-origin-destination combination. These coefficients are in Table A25. For each worker home location (first column) and location of the current job (second column), we show the standard deviation of wage gains for workers changing jobs to a given destination location (indicated in the top row).

Table A25: Standard Deviation of the Residual Wage Gains for Job Movers

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

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

G.2.14 Ratio of Profits to Labor Costs

We obtain the pre-tax profits of all firms in Germany from the ORBIS database provided by the company Bureau van Dijk. We allocate firms to our four locations based on the ZIP code of their address, and drop firms with fewer than 5 employees since their profits are very noisy. We then construct the ratio of profits to labor costs by dividing pre-tax profits by total labor costs reported in ORBIS, and average across firms and years to compute the average ratio in each location. We drop outlier profit ratios below the 5th and above the 95th percentile of the distribution of profit ratios in each location, and compute the average over the remaining ratios. Table A26 presents the estimates.

Table A26: Average Ratio of Firm Profits to Labor Costs by Location

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

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

H Model's Computation and Estimation

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

H.1 Solution Algorithm

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

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

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

   where k indexes the external iteration loop.

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

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

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

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

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

and can go back to point 2.

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

In order to compute the general equilibrium counterfactuals, we follow the same algorithm, but with two differences. First, as mentioned in the main text, during the estimation of the model, we solve - within each loop - for the unemployment benefits that yield a reservation wage equal to $R_j = \iota \underline{p}_j$. In the counterfactuals, instead, we keep the unemployment benefits fixed at their estimated value, and solve for the implied reservation wage. Second, while during the estimation we can keep each location's prices fixed at their observed values, in the counterfactual we must solve for the new equilibrium prices. Therefore, within each loop, we calculate each location's GDP and then we use it to calculate the new aggregate equilibrium prices.

H.2 Estimation Algorithm and Outcomes

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

$$\phi^* = \arg\min_{\phi \in \mathbb{F}} \mathcal{L}(\phi) \tag{60}$$

where

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

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

The minimization algorithm that we use to solve the problem (60) combines the approaches of Jarosch (2016) and Lise, Meghir, and Robin (2016), and Moser and Engbom (2021), both adapted to our needs.

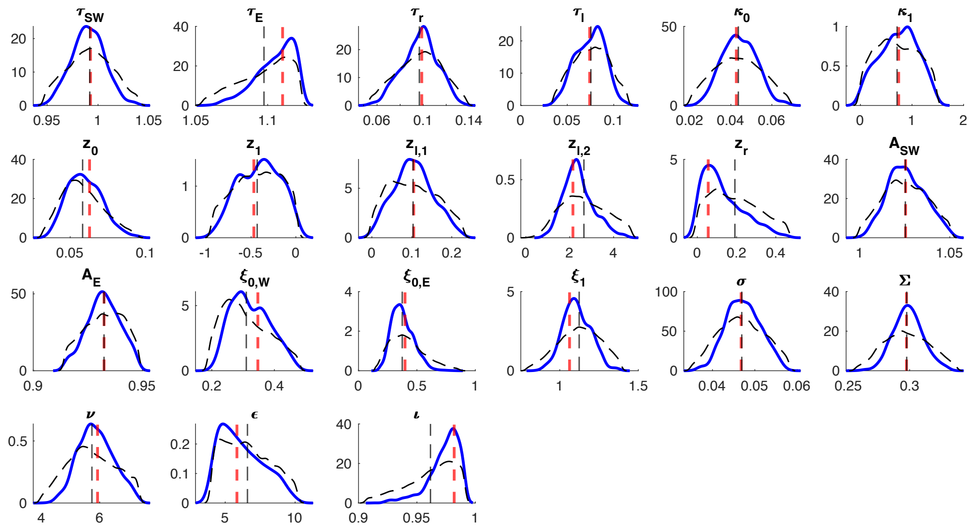
We simulate, using Markov Chain Monte Carlo for classical estimators as introduced in Chernozhukov and Hong (2003), 200 strings of length 10,000 (+ 1,000 initial scratch periods used only to calculate posterior variances) starting from 200 different guesses for the vector of parameters ϕ_0. In the first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

Figure A5 illustrates our approach and how it slightly differs from Jarosch (2016) and Lise, Meghir, and Robin (2016). The black dotted line shows the density function of the last 1,000 iterations across all strings. The usual approach is to pick the average across all these draws, which we highlight in the picture with a vertical black dotted line. However, this approach could be problematic if the parameter space is bounded, hence the estimated densities are not symmetric, as in our case for some parameters. Therefore, given our vector of parameters and likelihoods, we pick the optimal parameter following Moser and Engbom (2021), and simply select the vector of parameters that minimizes the objective function among all our draws.[74] Our estimates are shown with red dotted lines in the figure. For most parameters, they are almost identical to the alternative approach. Finally, the blue density functions shows the density, across all strings, of the 10 best outcomes within each string. This density provides a visual representation of the tightness of our estimates, which are, in general, quite good – especially for the key parameters that determine the spatial frictions. It is also relevant to notice that all the densities are single-peaked, which suggests that the model is, at least locally, tightly identified.

All the estimated parameters, corresponding to the vertical dotted red lines, are included in Table A27 below.

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

Figure A5: Estimation Outcomes



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

Table A27: 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.

H.3 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 (as commonly done in the literature, e.g., Kaboski and Townsend (2011)).

Specifically, we start from the estimated vector of parameters φ^*, and we create 28 alternative vectors, two for each parameter j, as follows: $\underline{\varphi}(j) = \left\{\varphi^*_{-j}, 0.95\varphi^*_j\right\}$ and $\overline{\varphi}(j) = \left\{\varphi^*_{-j}, 1.05\varphi^*_j\right\}$, where $\underline{\varphi}(j)$ keeps all parameters except for j constant and decreases j by 5%, while $\overline{\varphi}(j)$ 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

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

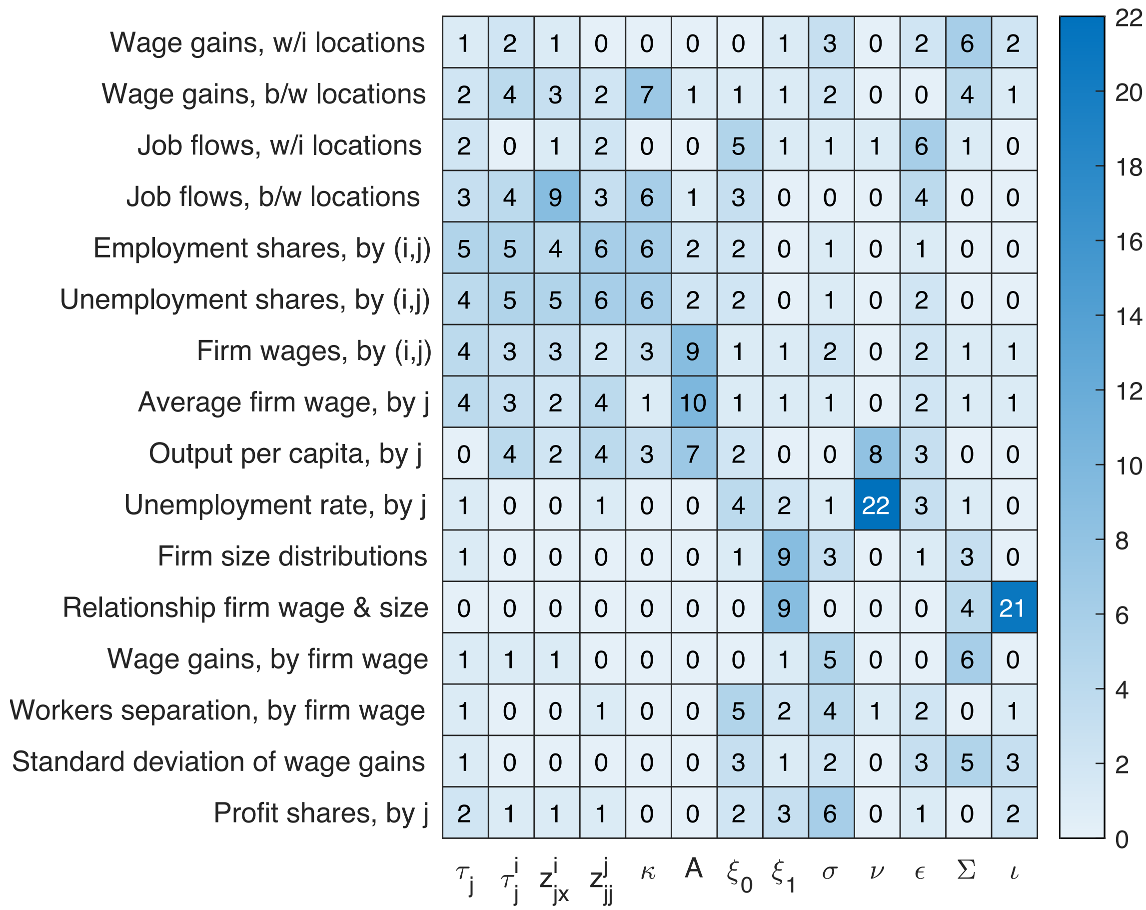
Thus, Δ_{jr} measures how much moment r would change if we changed parameter j by 10% around the estimated value while keeping all the other parameters constant.

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 3, 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 the Figure A6); ii. the two home biases τ_l and τ_r $\left(\tau^i_j\right)$; iii. the relative search efficiencies between regions z_0, z_1, $z_{l,2}$ and z_r $\left(z^i_{jx}\right)$; 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.

Following Bassi, Muoio, Porzio, Sen, and Tugume (2021), to ease comparison across the different parameters, 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 A6, 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.

The results from the Jacobian are quite intuitive and they connect different blocks of parameters to the moments that we would expect, as we discuss in details in Section 5.4. In Table 3 in the main text, we report for each moment the most relevant parameters.

Figure A6: Normalized Partial Derivatives of Moments with Respect to Parameters



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.

I Further Details on Model Fit

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

Figure A7 shows that the model fits the empirical moments well in several dimensions. Each panel plots a set of moments in the data (x-axis) against their values in the model (y-axis), with the 45-degree line indicating a perfect fit. In each of the top three panels, moments relating to West German workers are in blue and moments for East German workers are in red. The top left panel presents the share of employed workers of each type in a given location. We use each worker's residence to determine her location since our definition of a cross-location move is based on the worker's residence. The empirical values of these moments were computed in Section G.2.3. The panel shows that in our model, as in the data, most workers are in their home location (circles). Moreover, East-born workers are more likely to be in the West than West-born workers in the East (stars), consistent with the fact that the West has higher productivity and a larger ratio of firms to workers. The top middle panel shows the share of unemployed workers in each region, which is similar to the distribution of employed workers. The top right panel presents the average firm component of wages paid to workers of a given type in each region (as computed in Section G.2.5). Consistent with the data, workers in the East earn lower wages than workers in the West. Furthermore, the relative wage gaps differ by workers' location, in a similar way in the model and in the data. In particular, the average wage gap between East and West German workers is smaller for the group of workers that are away from their home region (red versus blue stars) than for the group of workers that are in their home location (red versus blue circles).

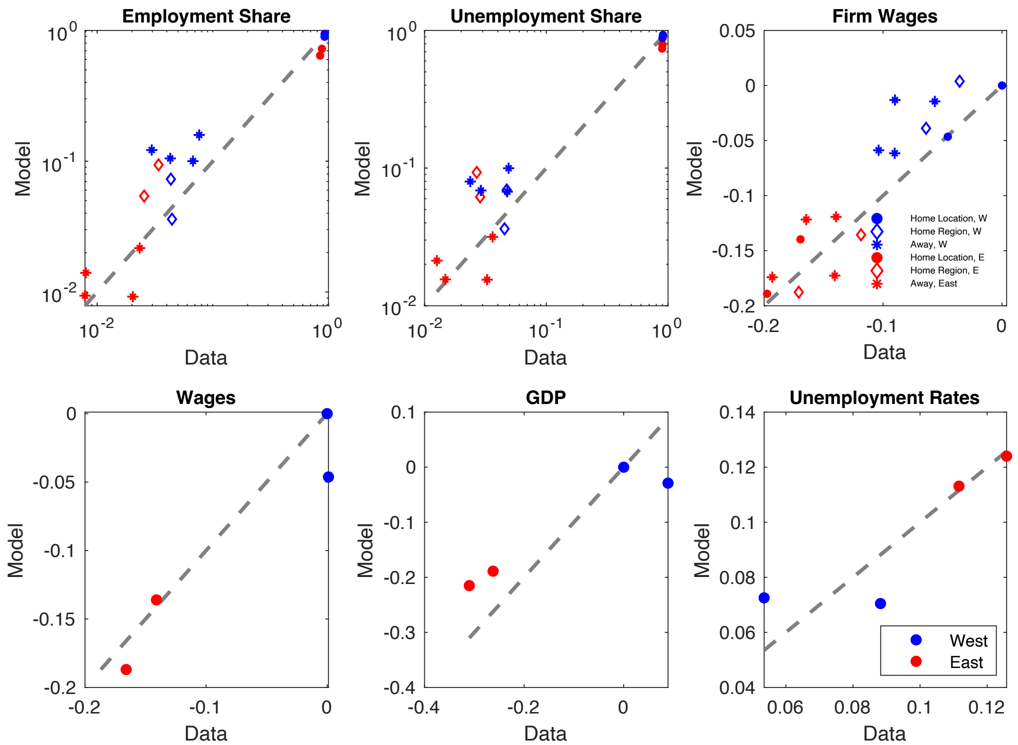
The bottom three panels of Figure A7 present the average firm component of wages (the empirical moments were computed in Section G.2.6), GDP per worker (Section G.2.7), and the unemployment rate (Section G.2.8) in each of the four locations. The model matches the data well, generating lower wages, lower GDP per worker, and higher unemployment in the East.

Figure A8 plots the firm size distributions in each location in the model and in the data, computed as described in Section G.2.9. The model matches almost perfectly the share of employment in the middle of the size distribution, and only slightly underestimates the mass of employment at the bottom and top deciles. In each location, approximately half of the overall employment is accounted for by the largest decile of firms.

Table A28 shows that the model also does a reasonable job in matching the joint distributions of firm wages, sizes, and separation rates, the standard deviation of wage gains, and the profit shares (the empirical moments were computed in Sections G.2.10 to G.2.14). The core mechanism of the model generates a positive relationship between firm size and the firm wage (row 1 of Table A28), 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 A9, where we plot these variables in the model and in the data nonparametrically, for each of the four locations. The top panels show the relationship between firm log size and log average wage in each location, the middle panels present the expected wage gains as a function of a worker's current firm's log average wage, and the bottom panels show the relationship between the separation rate and a worker's current firm's log average wage. In both the model and data, these relationships are roughly linear.

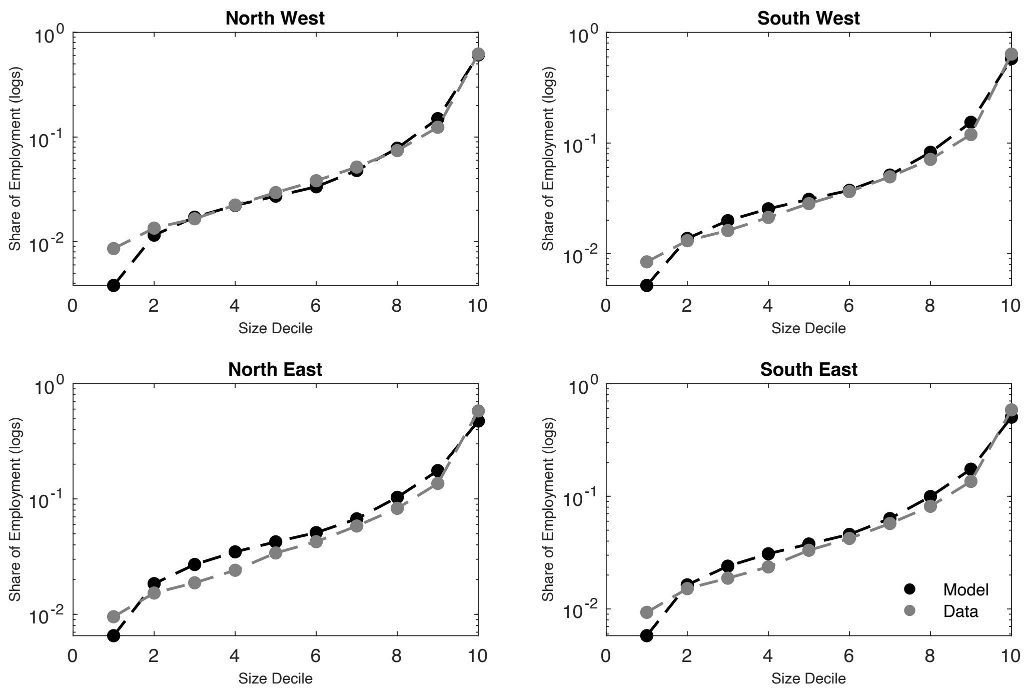
As noted in the main text, the model overestimates the relationship between job movers' expected wage gains and their current firm's average wage. Moreover, the model underestimates the standard deviation of wage gains of movers (row 4 of Table A28). This result is somewhat expected since in the model 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 A10 plots the distribution of the standard deviation of wage gains in the model and data for all 64 origin-destination-home location tuples. The standard deviations in the data are higher than in the model for nearly all combinations of moves. For comparison, we also plot in the figure an alternative empirical moment: the standard deviation of wage gains controlling for individual fixed effects (light gray). As expected, controlling for individual fixed effects reduces significantly the empirical variance (some individuals have persistently higher wage gains than others, as shown in the literature). Relative to this alternative target, our model slightly overestimates the standard deviation of wage gains.

Figure A7: Employment, Wages, and GDP by Location and Worker-Type



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

Figure A8: Within-Location Firm-Size Distributions



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

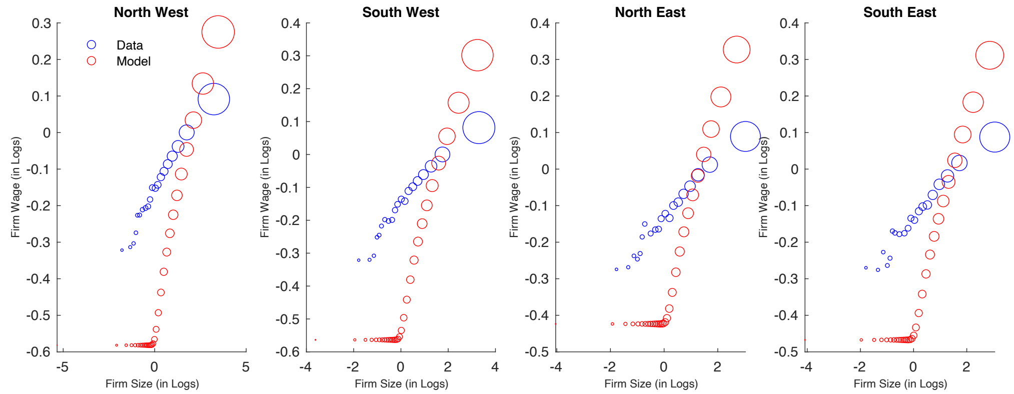
Table A28: Model Fit for Additional Moments

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

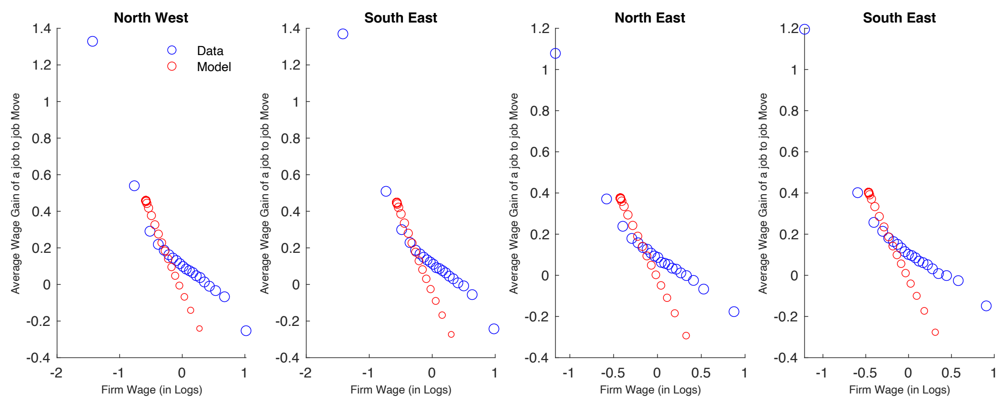
Notes: The table compares several moments in the model to their data analogues, by location of the firm. The construction of these moments is described in Sections G.2.10 to G.2.14. The first row shows the slope of the wage function with respect to firm size. The second row presents the slope of the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

Figure A9: Model Fit for Joint Distribution of Firm Wages, Sizes, and Separation Rates

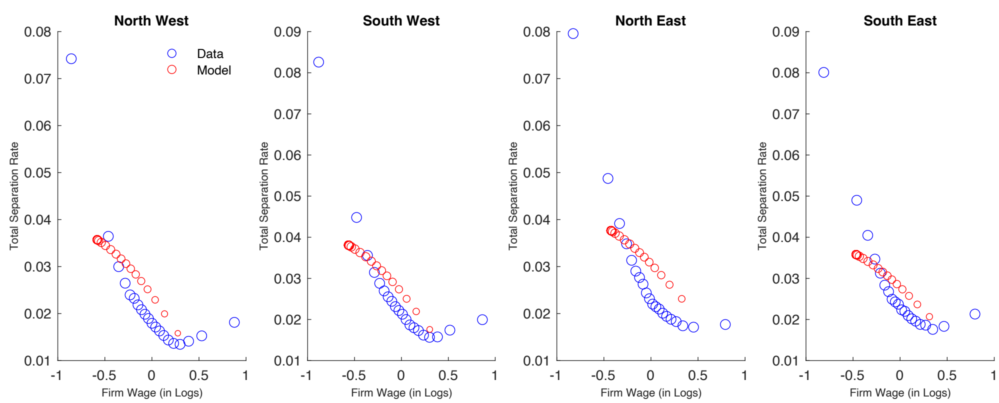
(a) Relationships between Firm Sizes and Average Wages



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

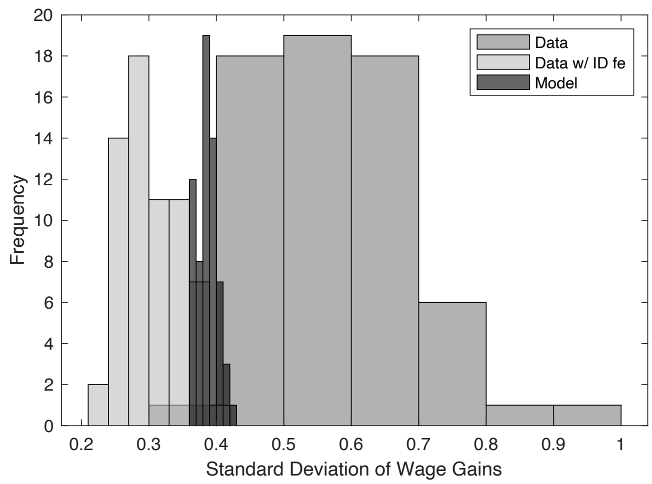


(c) Relationships between Firm Wages and Separation Rates



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

Figure A10: Standard Deviation of Wage Gains



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

J Additional Quantitative Results

We here include a few additional figures generated from the quantitative exercises of Section 6.

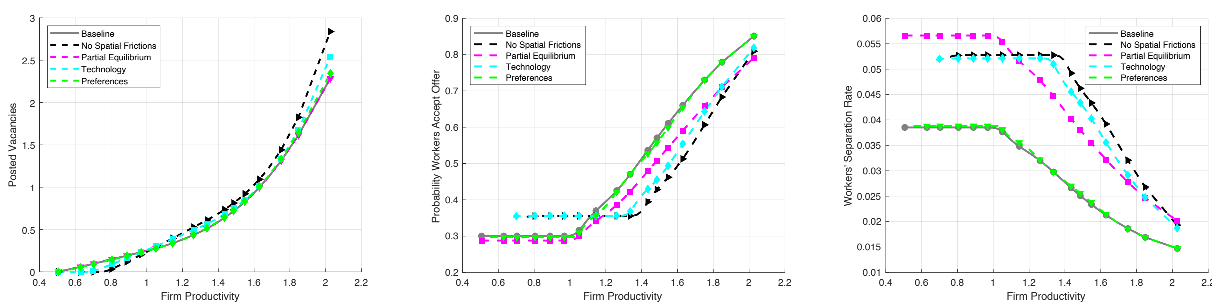
Figure A11 presents posted vacancies, workers' acceptance probability, and the separation rate as a function of firm productivity as in Figure 8, but for West Germany. The findings are similar to the figure shown in the main text: the number of vacancies and the separation rate contribute positively to the reallocation of labor from low- to high-productivity firms. In contrast, the acceptance probability mitigates the reallocation gains.

Figure A12 shows the distribution of workers to firms, analogously to Figure 7, for the partial equilibrium counterfactual where we hold fixed firms' wage and vacancy posting (column 3 of Table 5). Consistent with the relatively small aggregate effects, we see little change in the overall worker distribution (Panel (a)). However, there is reallocation across regions as East Germans move West and West Germans move East, as illustrated in Panels (b) and (c).

Figures A13 and A14 show the distribution of workers to firms, analogously to Figure 7, for the counterfactuals where only technological spatial frictions are removed or where only preference frictions are removed (columns 4 and 5 of Table 5). Removing only technological spatial frictions generates some improvement in the worker allocation both within and across regions. In contrast, removing preference frictions mostly changes only the allocation of workers across regions.

Figure A11: Margins of Employment, West Germany

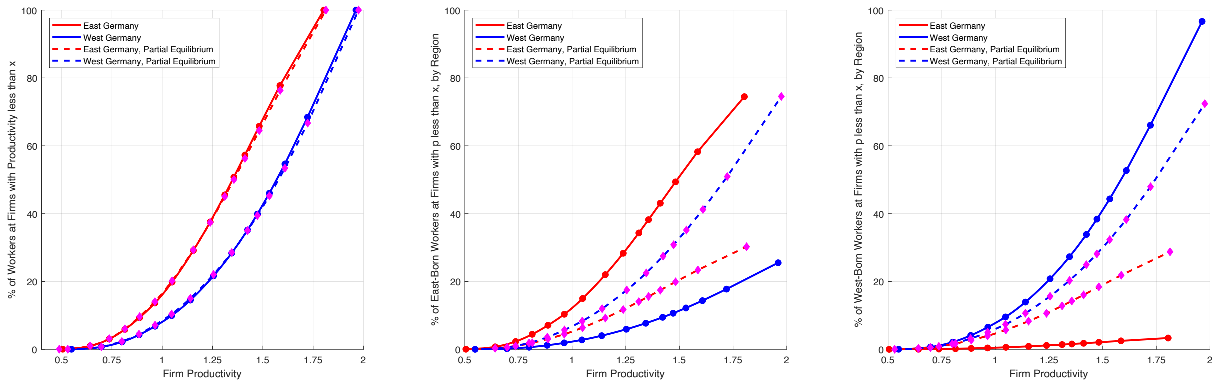
(a) Posted Vacancies (b) Acceptance Probability (c) Separation Rate



Notes: All panels are for firms in West 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.

Figure A12: Labor Allocation Across Firms and Regions, Partial Equilibrium

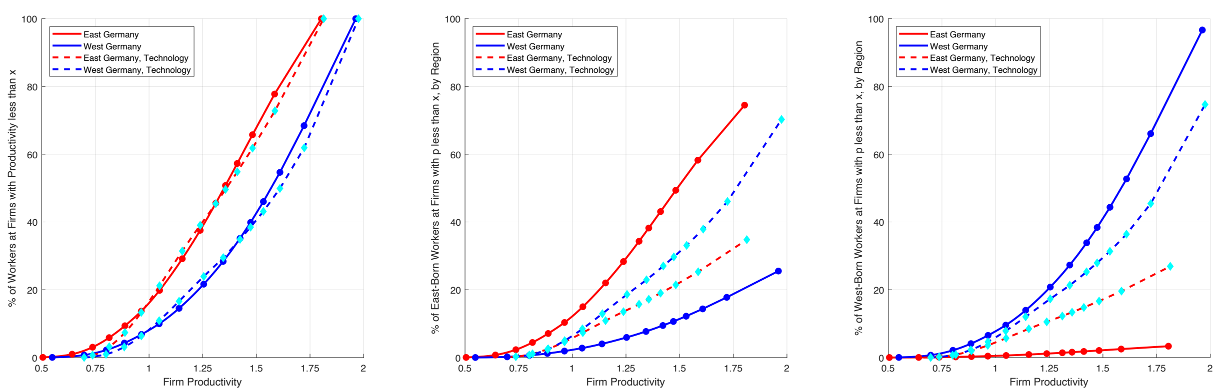
(a) All Workers (b) East Germans (c) West Germans



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 when we keep constant the firm equilibrium response. 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 same semi-CDF for West Germans.

Figure A13: Labor Allocation Across Firms and Regions, Technology

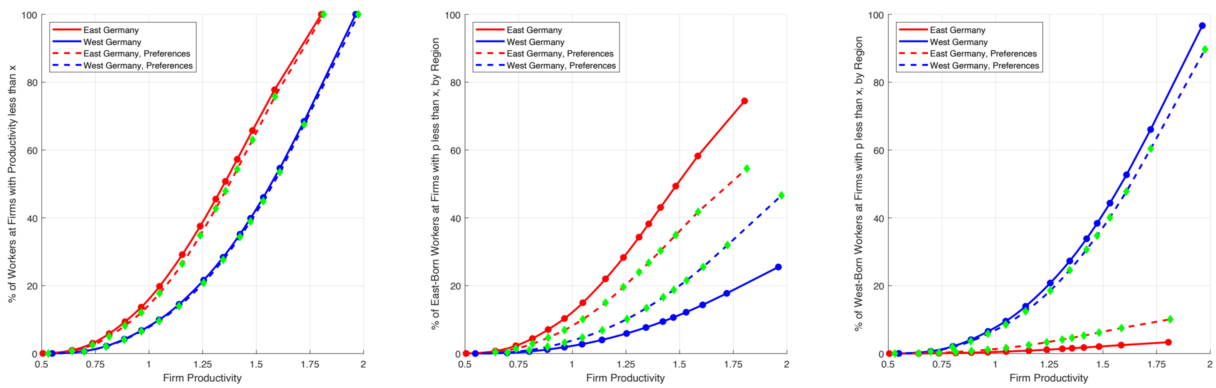
(a) All Workers (b) East Germans (c) West Germans



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 in which we eliminate spatial frictions due to technology (i.e. z and κ). 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 same semi-CDF for West Germans.

Figure A14: Labor Allocation Across Firms and Regions, Preferences

(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 in which we eliminate spatial frictions due to preferences (i.e. τ). 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 same semi-CDF for West Germans.