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

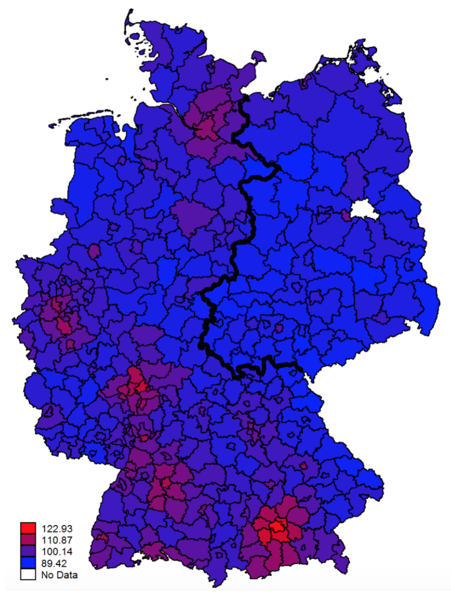
	Variable	Obs	Mean	Std. Dev.
(1)	Real wage of FT employed	7,963,537	111.890	76.967
(2)	Real wage of unemployed	1,254,063	27.580	12.469
(3)	Employed dummy	9,485,701	0.849	0.358
(4)	Age	9,485,701	40.172	11.538
(5)	Male dummy	9,485,701	0.696	0.460
(6)	College dummy	5,904,697	0.205	0.403
(7)	Work county East	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)			(2)		
		Home: West			Home: East		
(1)	Crossed border	3.8%			21.9%		
(2)	Returned movers	41.9%			32.3%		
(3)	Mean years away (returners)	1.86			2.34		
(4)	Mean years away (non-returners)	5.38			6.65		
		Stayers	Movers	Returners	Stayers	Movers	Returners
(5)	Age at first move	–	35.9	35.5	–	32.3	32.2
(6)	Share college	0.22	0.34	0.32	0.19	0.19	0.19
(7)	Share male	0.70	0.75	0.80	0.57	0.73	0.78

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

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

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

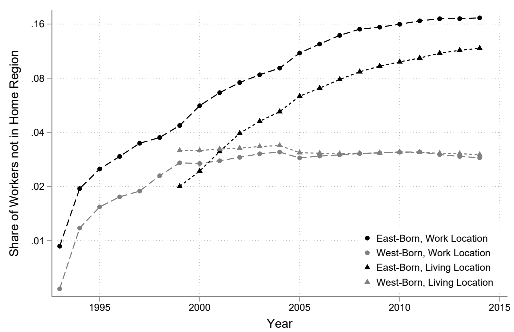
Source: LIAB. Notes: The table shows statistics for workers with at least one employment spell in our core sample period 2009-2014. For these workers, we compute the distribution of the number of cross-border moves throughout their lifetime, going back as many years as available. The first two columns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

Table A8: Mobility by Cohort

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

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

Figure A2: Stock of Individuals away from Home Region



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

C Results from the Socio-Economic Panel

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

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

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

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

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

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

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

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

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

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

	Old SOEP Sample				New SOEP Sample			
$\log(w_{it})$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{I}_{i,East,imp}$	$-$**.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

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

as claimed in the main text.

D.2 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

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

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

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

From equation (30), we find

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

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

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

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

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

which can itself be substituted into (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*

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

where

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

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

and

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

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

Proof. Define the real wage, adjusted for amenities, as $\hat{w} \equiv w\tau_j P_j^{-1}$, where we have used that $\tau_j^i = \tau_j$. By assumption, $\hat{b} \equiv b_j^i \tau_j P_j^{-1}$ is constant across regions. Define $\hat{F}_j(\hat{w}) \equiv F_j(w\tau_j P_j^{-1})$. 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

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

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

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

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

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

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

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

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

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

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

and

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

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

The first-order condition of the wage posting problem is

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

where

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

Plugging this latter expression into the first-order condition gives

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

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

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

Define $\tilde{x}(\hat{p}) \equiv \hat{x}(\hat{w}(p))$ for any $\hat{x}$. We can then re-define (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. □

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

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

Since $\bar{a}^{E,i}_{jx} = z^i_{jx}\bar{s}^{E,i}_x$, where $\bar{s}^{E,i}_x \equiv \int s^{E,i}_{jx}\left(w\right)\mathrm{d}E^i_j\left(w\right)$, a lower search efficiency z^i_{jx} leads to lower job-job flows from location j to x given the acceptance probability $\mu^{E,i}_{jx}\left(w,w'\right)$, 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

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

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},

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

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

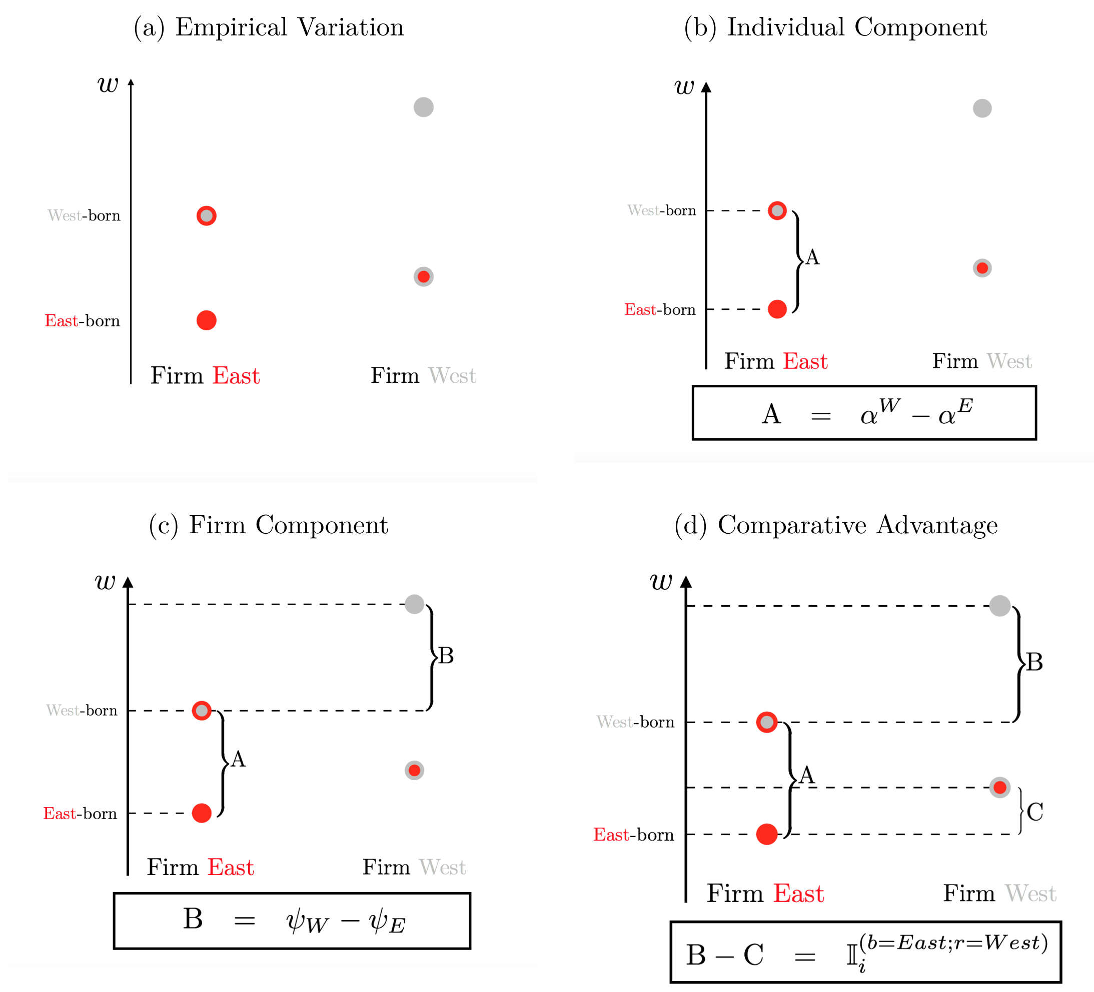
Consider four wage observations associated with two workers: an East-born and a West-born individual working in one firm in the East, and the same two individuals working in one firm in the West. Figure A3a plots an example of these two workers' wages, where the x-axis is the identity of the firm, the 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

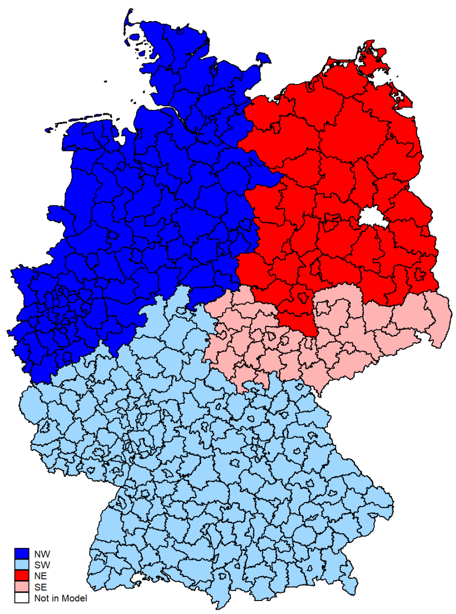
In this section, we provide more details on the four locations in our estimated model and the mobility between them. Figure A4 visualizes the four locations.

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_j^i depend only on the work location j and set them equal to the monthly probabilities, computed in the LIAB data, that workers separate into unemployment or permanent non-employment (i.e. either retired or dropping out of the labor force). 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, $\hat{\epsilon}_{jlt}^{wage}$, and for the log number of workers, $\hat{\epsilon}_{jlt}^{size}$. We add back the mean of each variable in each location, $\overline{\ln(y_{jlt}^{wage})}$ and $\overline{\ln(y_{jlt}^{size})}$, to obtain a cleaned log real wage, $\ln(\hat{y}_{jlt}^{wage}) = \overline{\ln(y_{jlt}^{wage})} + \hat{\epsilon}_{jlt}^{wage}$ and a cleaned log number of workers, $\ln(\hat{y}_{jlt}^{size}) = \overline{\ln(y_{jlt}^{size})} + \hat{\epsilon}_{jlt}^{size}$ for each firm. We then regress the residualized log real wage on the residualized log number of workers in each location

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

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

Table A22: Log Wage on Log Firm Size by Location

Dep. var.: $\ln(\hat{y}^{wage}_{jlt})$	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, $\hat{\epsilon}^{gain}_{ilt}$. 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, $\Delta \ln(\hat{w}_{ilt}) = \overline{\Delta \ln(w_{ilt})} + \hat{\epsilon}^{gain}_{ilt}$. 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_{ilt}^{sep}. 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, $\hat{\epsilon}_{ilt}^{wage}$, and for the separation dummy, $\hat{\epsilon}_{ilt}^{sep}$, and add back the mean of each variable in each location, $\overline{\ln(w_{ilt})}$ and $\overline{d_{ilt}^{sep}}$, to obtain a cleaned log wage, $\ln(\hat{w}_{ilt}) = \overline{\ln(w_{ilt})} + \hat{\epsilon}_{ilt}^{wage}$ and a cleaned separation dummy $\hat{d}_{ilt}^{sep} = \overline{\ln(d_{ilt}^{sep})} + \hat{\epsilon}_{ilt}^{sep}$. We then regress the residualized separation dummy on the residualized log wages for each location

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

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:

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

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.

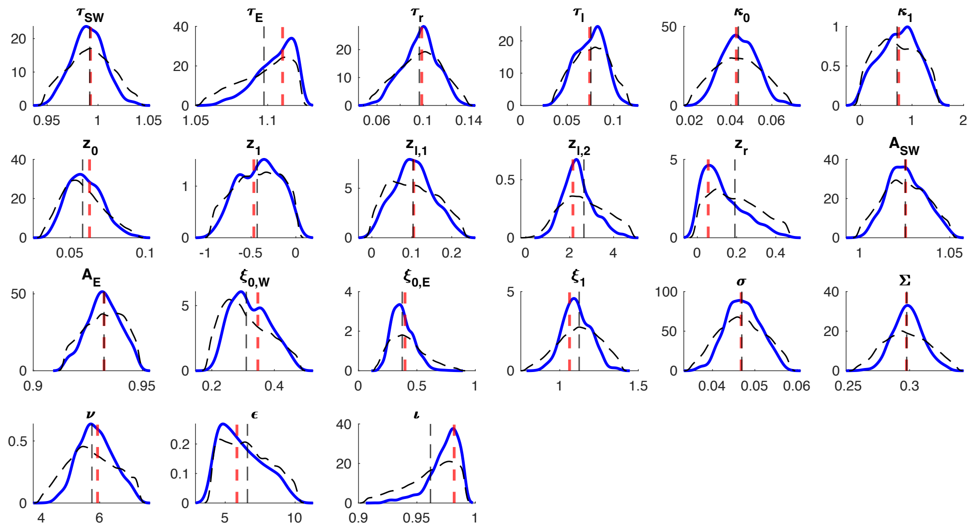
We simulate, using Markov Chain Monte Carlo for classical estimators as introduced in Chernozhukov and Hong (2003), 200 strings of length 10,000 (+ 1,000 initial scratch periods used only to calculate posterior variances) starting from 200 different guesses for the vector of parameters ϕ_0. In the first 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

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

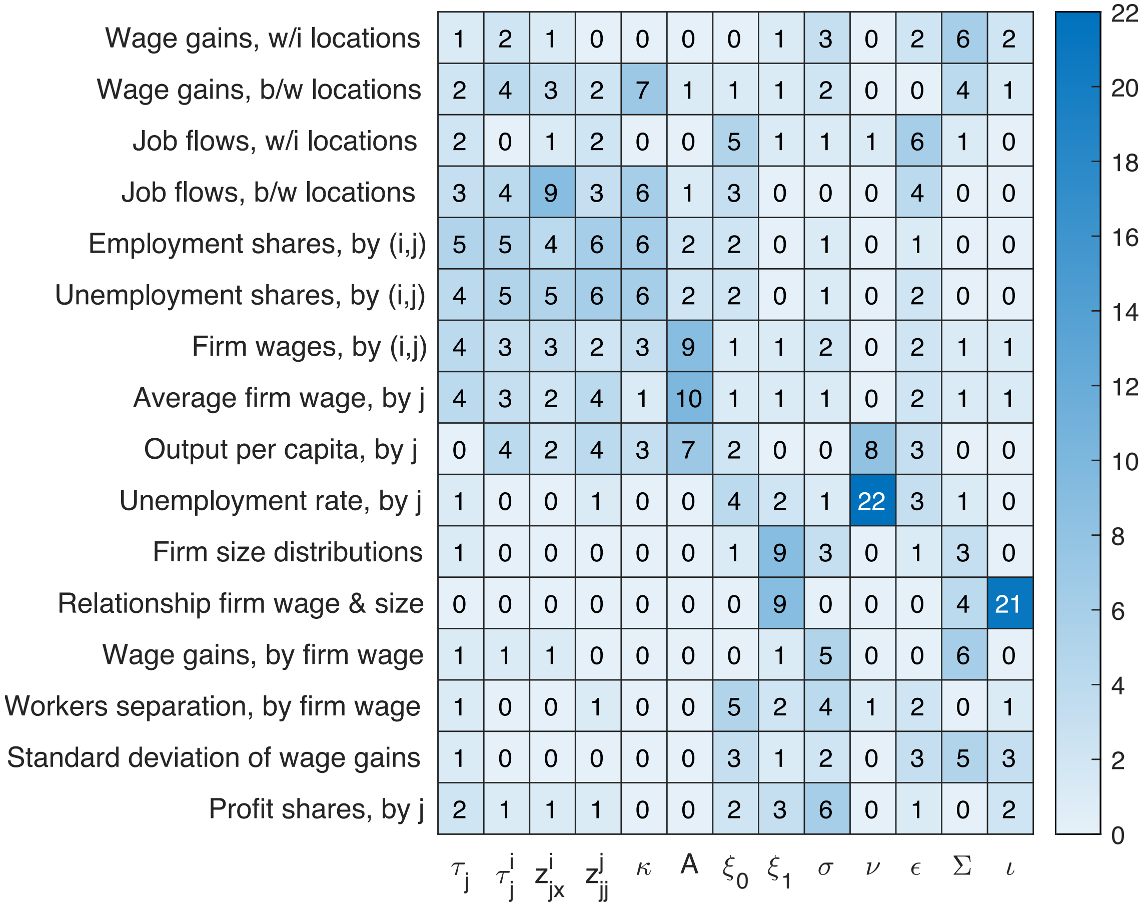
Thus, Δ_{jr} measures how much moment r would change if we changed parameter j by 10% around the estimated value while keeping all the other parameters constant.

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).

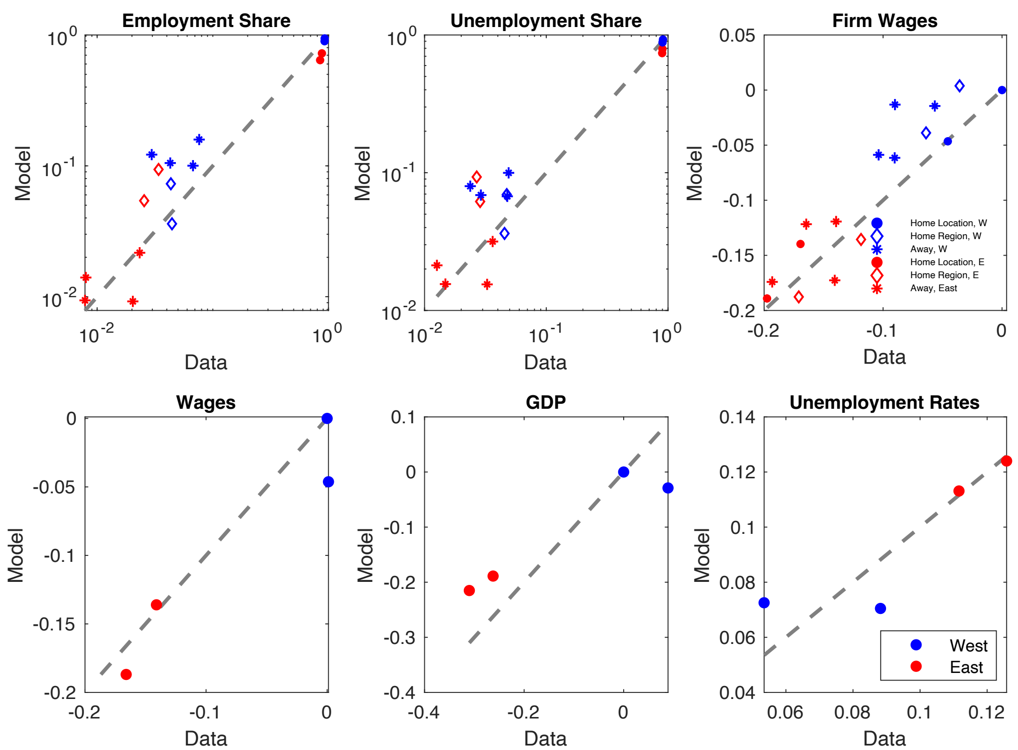
The bottom three panels of Figure A7 present the average firm component of wages (the empirical moments were computed in Section G.2.6), GDP per worker (Section G.2.7), and the unemployment rate (Section G.2.8) in each of the four locations. The model matches the data well, generating lower wages, lower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

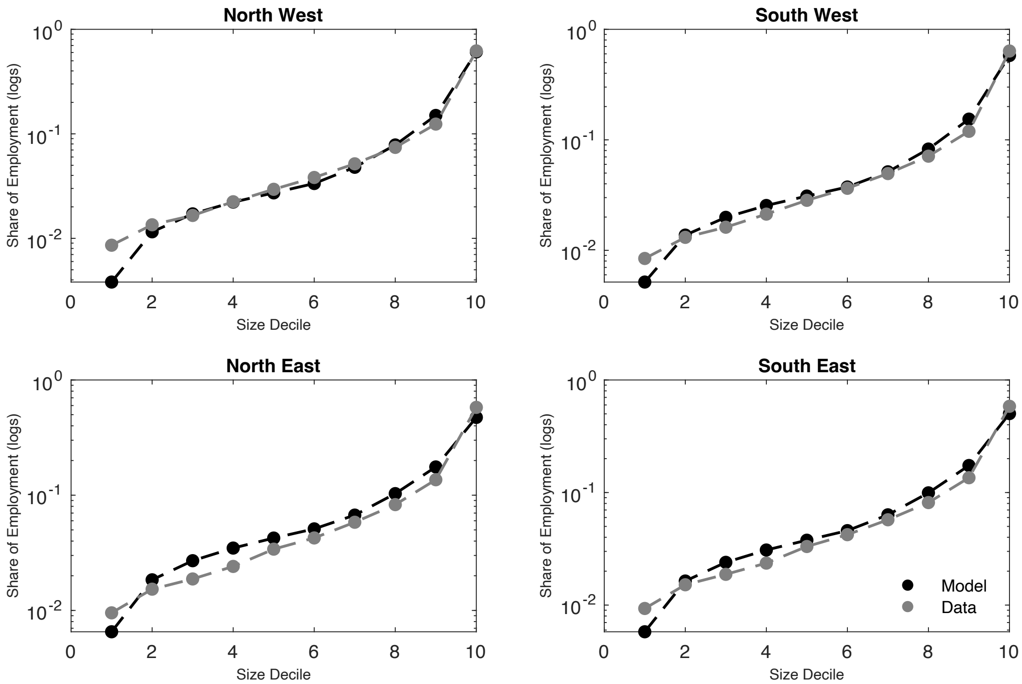
As noted in the main text, the model overestimates the relationship between job movers' expected wage gains and their current firm's average wage. Moreover, the model underestimates the standard deviation of wage gains of movers (row 4 of Table A28). This result is somewhat expected since in the model 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.

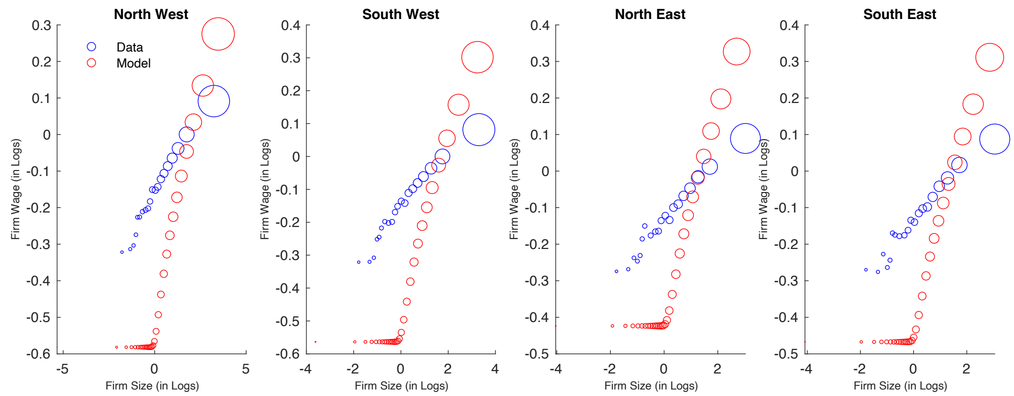
Table A28: Model Fit for Additional Moments

	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

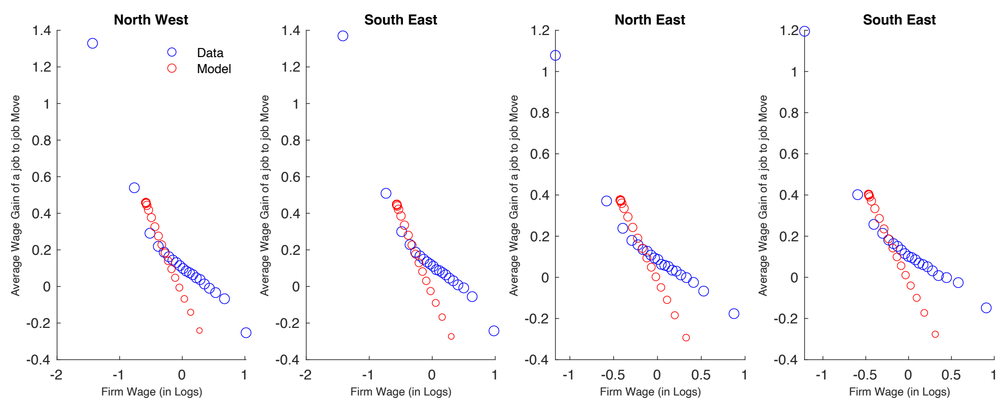
Notes: The table compares several moments in the model to their data analogues, by location of the firm. The construction of these moments is described in Sections G.2.10 to G.2.14. The first row shows the slope of the wage function with respect to firm size. The second row presents the slope of the 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

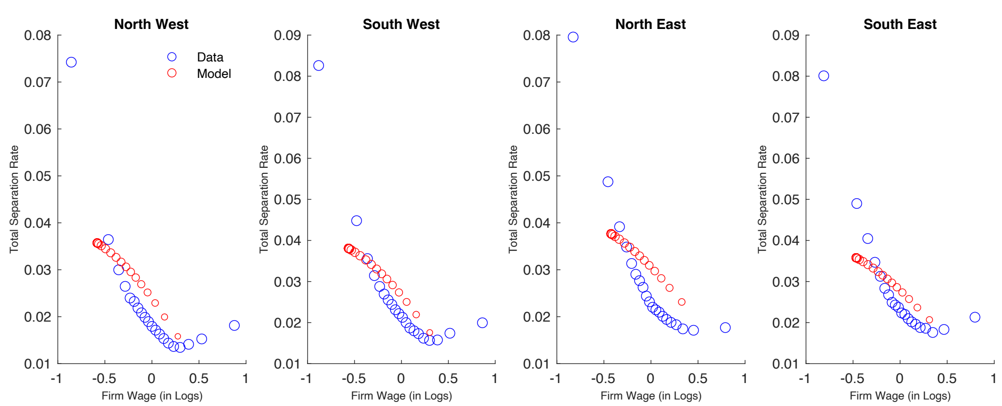
(a) Relationships between Firm Sizes and Average Wages



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

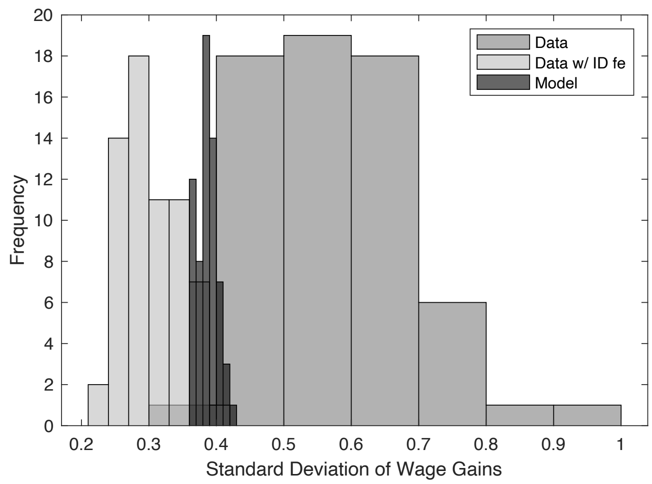


(c) Relationships between Firm Wages and Separation Rates



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.

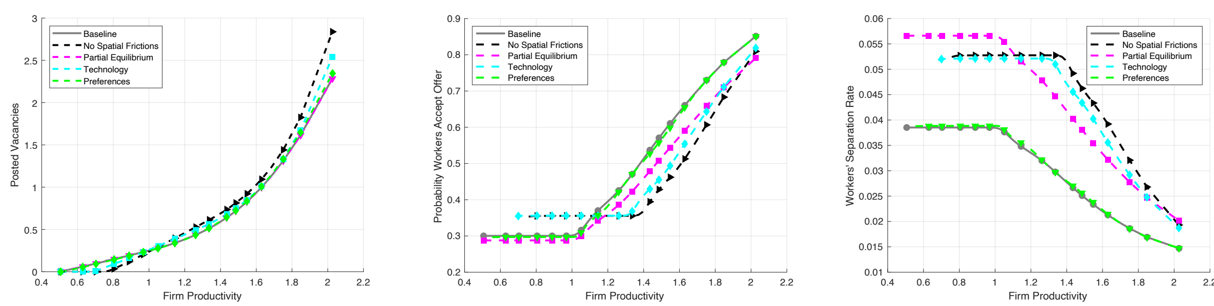
Figure A11 presents posted vacancies, workers' acceptance probability, and the separation rate as a function of firm productivity as in Figure 8, but for West Germany. The findings are similar to the figure shown in the main text: the number of vacancies and the separation rate contribute positively 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

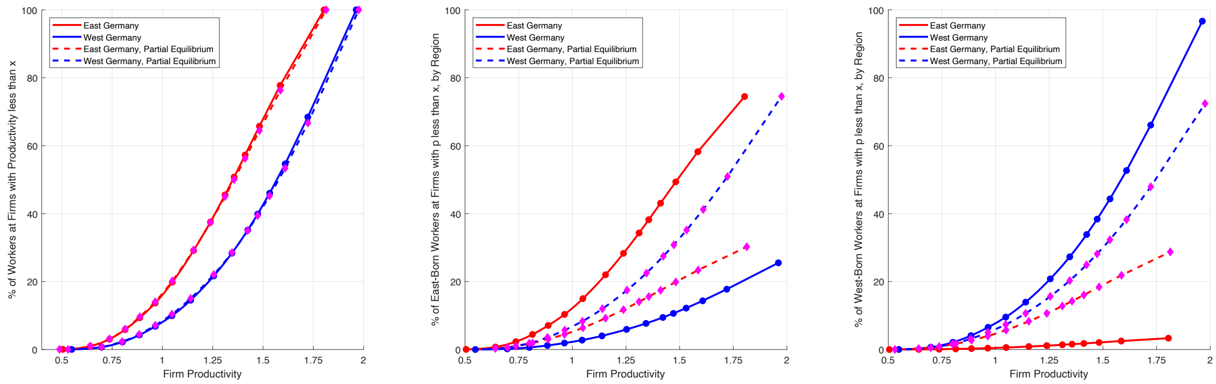
(a) Posted Vacancies (b) Acceptance Probability (c) Separation Rate



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

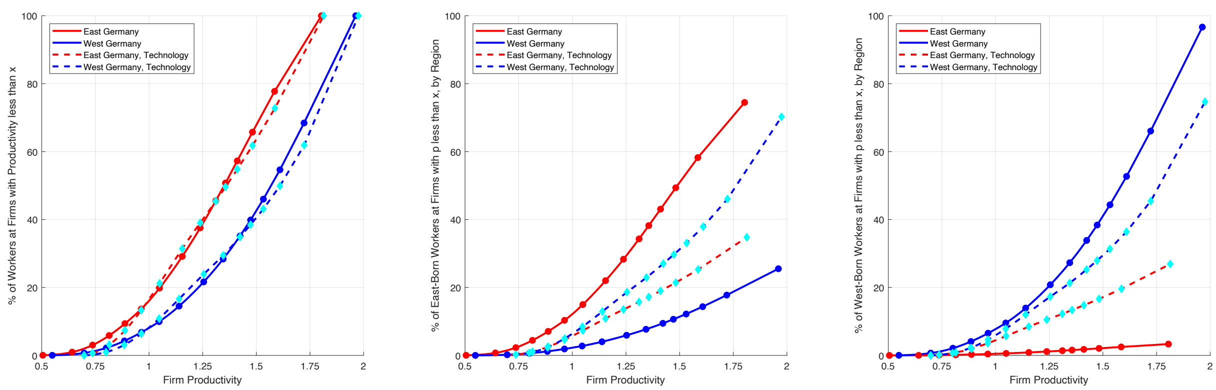
(a) All Workers (b) East Germans (c) West Germans



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

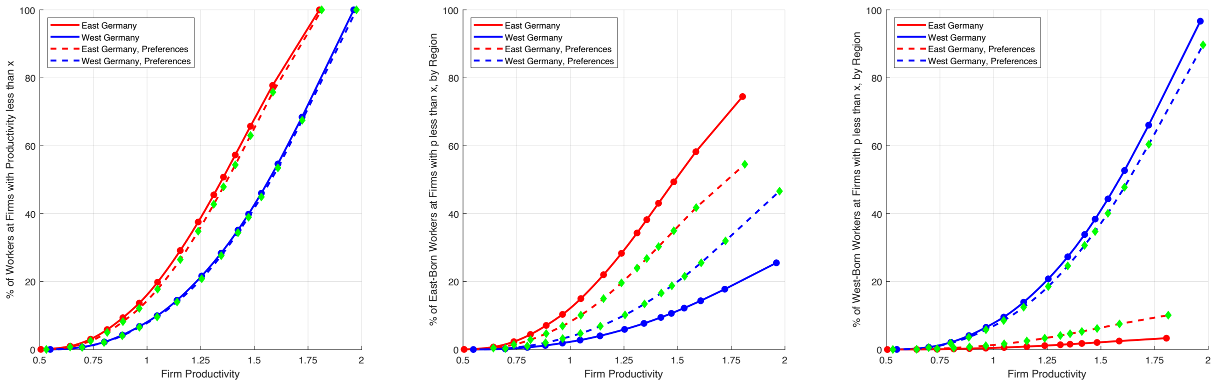
(a) All Workers (b) East Germans (c) West Germans



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.