Supplemental Material – Not for Publication

L Results from the Socio-Economic Panel

We use the German Socio-Economic Panel (SOEP) to assess how well our imputed LIAB home region captures individuals' true birth region. Reliable birth-region information is available for two SOEP samples. First, the 1984 SOEP wave covered only West Germans, while the 1990 wave covered only East Germans, identifying birth region with certainty. We refer to individuals from these waves who remain in the labor force in 2009–2014 as the "Old SOEP Sample". Second, for individuals who entered the survey as children, we infer home region from the location of their earliest observed non-tertiary schooling; we call this the "Young SOEP Sample". Although SOEP also asks some individuals about their residence in 1989, this variable is available for only about 0.5% of individual-year observations.

We construct an imputed home region in the same way and subject to the same restrictions as in the LIAB. Table S1 compares the imputed and actual home region for individuals that are in the labor force in 2009-2014. We find that in both samples the imputed and the actual home region match closely.

Table S1: Fraction of Individuals Where Imputed Home Region Matches Actual

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

Notes: We compute in the SOEP an imputed home region in the same way as in the LIAB. Specifically, we use only SOEP data from 1993 onward, exclude Berlin, and drop residence information prior to 1999. We then use the worker's region of residence at the first time he/she is observed in employment or unemployed, but not outside of the labor force, from 1999 onwards, or the worker's job region prior to 1999, to assign an imputed home region. The table reports the share of observations where the imputed birth region matches the actual birth region for individuals that are either employed or unemployed in 2009-2014.

As a more rigorous test, we compare the wage gap 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. Results in Table S2 show that the wage gap is similar under both the true and the imputed region definitions. Thus, we find no evidence that our misclassification of some workers quantitatively alters the wage gap.

Table S2: Individual-Level Wages by Imputed Home Region versus Birth Region

	Old SOEP Sample				New SOEP Sample			
$\log(w_{it})$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{I}_{i,East,imp}$	$-\mathbf{.346}^{***}$	$-\mathbf{.404}^{***}$			$-\mathbf{.160}^{***}$	$-\mathbf{.163}^{***}$		
	(.0212)	(.0196)			(.0325)	(.0309)		
$\mathbb{I}_{i,East,true}$			$-\mathbf{.338}^{***}$	$-\mathbf{.406}^{***}$			$-\mathbf{.133}^{***}$	$-\mathbf{.127}^{***}$
			(.0207)	(.0192)			(.0319)	(.0303)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Age/edu/male	–	Y	–	Y	–	Y	–	Y
Observations	15,240	15,210	15.240	15,210	2,894	2,540	2,894	2,540

Notes: $*$, $**$, and $***$ indicate significance at the 90th, 95th, and 99th percent level, respectively. Standard errors are clustered at the individual level. $\mathbb{I}_{i,East,imp}$ is a dummy for the worker's home region being East Germany, imputed using the same procedure as in the LIAB. $\mathbb{I}_{i,East,true}$ is a dummy for a worker's true birth region (Old SOEP sample) or region of earliest non-tertiary schooling (Young SOEP sample) being East.

M Comparison to the Burdett-Mortensen Model

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

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

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 $\hat{w} \equiv w\tau^i_j P^{-1}_j$ 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 (5) simplifies to

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

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

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

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

Define $\hat{p} \equiv p\tau_j P_j^{-1}$. The firm's maximization problem (11) becomes

$$\hat{\pi}_j\left(\hat{p}\right) = \frac{P_j}{\tau_j} \max_{\hat{w}} \left(\hat{p} - \hat{w}\right) \hat{l}\left(\hat{w}\right) \tag{61}$$

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

$$\hat{l}(\hat{w}) = \frac{\hat{\mathcal{P}}\left(\hat{w}\right)}{\hat{q}\left(\hat{w}\right)} \quad \text{if } \hat{w} \geq \hat{R}, \tag{62}$$

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

$$\hat{q}\left(\hat{w}\right) = \delta + \sum_{x\in\mathbb{J}} \bar{v}_x \left[1 - \hat{F}_x\left(\hat{w}\right)\right] \tag{63}$$

and

$$\hat{\mathcal{P}}(\hat{w}) = \sum_{x\in\mathbb{J}} \left[\hat{E}_x\left(\hat{w}\right) + u_x\right], \tag{64}$$

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

$$\frac{\left(\hat{p} - \hat{w}\right)\left(\frac{\partial \hat{l}(\hat{w})}{\partial \hat{w}}\right)}{\left(\hat{l}\left(\hat{w}\right)\right)} = 1, \tag{65}$$

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

$$\frac{\left(\hat{p}-\hat{w}\right)\left(\frac{\partial \hat{\mathcal{P}}(\hat{w})}{\partial \hat{w}}\hat{q}\left(\hat{w}\right) - \frac{\partial \hat{q}(\hat{w})}{\partial \hat{w}}\hat{\mathcal{P}}\left(\hat{w}\right)\right)}{\hat{\mathcal{P}}\left(\hat{w}\right)\hat{q}\left(\hat{w}\right)} = 1. \tag{66}$$

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

$$\hat{w}(\underline{\hat{p}}) = \hat{R}. \tag{67}$$

From (61), 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 (12) the mass of vacancies across regions, $\tilde{v}(\hat{p})$, is

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

Define $\tilde{x}(\hat{p}) \equiv \hat{x}(\hat{w}(p))$ for any $\hat{x}$. We can then re-define (63) and (64) using these definitions to obtain

$$\tilde{q}\left(\hat{p}\right) = \delta + \bar{\tilde{v}}\left[1-\tilde{F}(\hat{p})\right] \tag{68}$$

and

$$\tilde{\mathcal{P}}(\hat{p}) = \tilde{E}(\hat{p}) + u \equiv (1-u)\tilde{G}(\hat{p}) + u, \tag{69}$$

where $\tilde{E}(\hat{p}) \equiv \sum_{x\in\mathbb{J}} \tilde{E}_x(\hat{p})$ 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 (66) as

$$\frac{\partial\hat{w}(p)}{\partial\hat{p}} = \frac{(\hat{p}-\hat{w}(\hat{p}))\left(\frac{\partial\tilde{\mathcal{P}}(\hat{p})}{\partial\hat{p}}\tilde{q}(\hat{p}) - \frac{\partial\tilde{q}(\hat{p})}{\partial\hat{p}}\tilde{\mathcal{P}}(\hat{p})\right)}{\tilde{\mathcal{P}}(\hat{p})\,\tilde{q}(\hat{p})}. \tag{70}$$

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

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

and hence

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

The mass of unemployed is given from (21) by

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

Substituting these expressions into (69) gives

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

Plugging this expression for the acceptance probability and its derivative into (70), we obtain

$$\frac{\partial\hat{w}(\hat{p})}{\partial\hat{p}} = \frac{-2\,(\hat{p}-\hat{w}(\hat{p}))\,\frac{\partial\tilde{q}(\hat{p})}{\partial\hat{p}}}{\tilde{q}(\hat{p})}. \tag{71}$$

Together, equations (63), (64), (67), and (71) 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. □

N Parameters and Empirical Moments

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

N.1 Calibrated Parameters

We first describe how we compute the calibrated parameters shown in Table 3.

(1) Worker Skills

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

$$\log(w_{it}) = \alpha_i + \psi_{J(i,t)} + \beta \mathbb{I}^{(h_i \neq R(J(i,t)))} + BX_{it} + \epsilon_{it}, \tag{72}$$

and describe details on the identification in Appendix G. 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.[47] This sample includes approximately 97% of West and East workers in the LIAB.

The estimation yields a comparative advantage estimate of $\beta = 0.027$, indicating a small *negative* comparative advantage towards the home region. Thus, a typical East-born worker is paid, controlling for firm characteristics, almost 1.4% more if she works in the West.[48] 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

$$\hat{\alpha}_i = \eta^h \mathbb{I}_i^h + CX_i + \varepsilon_i, \tag{73}$$

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

[48] 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.

where $\hat{\alpha}_i$ is the estimated worker fixed effect, $\mathbb{I}_i^h$ are dummies for the workers' home location, and X_i are dummies for worker age groups, gender, and college. We let NW be the omitted category, and obtain the η^h for the remaining three regions. We take their exponent since the AKM was estimated in logs, and present the exponentiated estimates in Table 3. We find that conditional on age, gender, and schooling, West-born workers earn, within the same firm, around 10% 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 average monthly probabilities, computed in the LIAB data, that workers separate into unemployment. Specifically, we compute in each month the share of employed workers that become unemployed. 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. 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) Discount Rate

Since individuals in our model are infinitely lived, the discount rate ρ accounts for both discounting and rates of retirement or death. We pick a monthly discount rate equal to 1.5%, which implies an annual rate of 18%, around the lower end of the range estimated by Andreoni and Sprenger (2012).

N.2 Moments for Estimation

Next, we turn to the 284 empirical moments targeted in the estimation and described in Table A9. We follow the order of the table in describing each set of moments in detail.

N.2.1 Log-Odds Ratio of Acceptance within Locations

We asked participants in the survey about their current salary and presented them with four hypothetical offers for jobs with similar characteristics in the same city, with salaries equal to their current salary, as well as 70–90%, 110–140%, and 150–170% of their current salary. Respondents reported their probability of accepting each offer. Let $P_{i,m}$ denote respondent i's reported acceptance probability for offer m in the current city, and let $D^k_{i,m}$ be a dummy indicating that the hypothetical salary of offer m equals k percent of the respondent's current salary. Across respondents and questions, offer ratios range from 70% to 170%; we group them into 11 bins, $K = 70\%, 80\%, \ldots, 170\%$. We then run the regression

$$\log\left(\frac{P_{i,m}}{1-P_{i,m}}\right) = \sum_{k\in K} \beta_k D^k_{i,m} + \upsilon_i + \varepsilon_{i,m}, \tag{74}$$

where υ_i are individual fixed effects, and $k = 11$ is the omitted category. The first two columns of Table S3 report the regression coefficients and standard errors. The final column reports rescaled coefficients: we add a common constant to the fitted coefficients so that the value at an offer ratio of 100% matches the average log-odds ratio of acceptance at 100% in the raw data. These are the 11 values we target, depicted in Figure 5(b).

Table S3: Log-Odds Ratio of Acceptance within Location

Wage offer ratio	Regression		Moments
	Reg coef.	SE	Standardized coef.
70%	−2.960	0.078	−1.363
80%	−3.015	0.072	−1.417
90%	−2.826	0.069	−1.228
100%	−2.038	0.056	−0.441
110%	−1.147	0.069	0.451
120%	−0.928	0.069	0.669
130%	−0.750	0.071	0.847
140%	−0.623	0.069	0.974
150%	−0.010	0.072	1.587
160%	0.072	0.079	1.670
170%	0	.	1.597

Notes: Columns 1 and 2 report raw regression coefficients and their standard errors. The third column adds to these coefficients a constant so that the value at an offer ratio of 100% matches the average log-odds ratio of acceptance at 100% in the raw data.

N.2.2 Log-Odds Ratio of Acceptance across Locations

We asked participants in the survey about their probability of accepting job offers from different cities. Each respondent was shown one random city from each of the four "locations" (NW, SW, NE, SE). To give some context to respondents, we also provided the size of the population, the state, and the location of the city. We also asked respondents whether they live or grew up in counties different from their current workplace and, if so, added those residence and birth counties to the list. We then asked respondents about their likelihood of accepting the same salary as the salary of the "benchmark likelihood" in each of these four to six cities or counties. By holding the salary fixed, we focus on differences in acceptance that are due to the fact that the offer is from a different location. The "benchmark likelihood" is the lowest acceptance probability higher than 70% that the respondents gave to a hypothetical job offer for jobs within their current city. If respondents only assigned probabilities between 50% and 70%, we used the highest probability from this interval. If none of the within-city offers was accepted with probability above 50%, we asked respondents to state the salary required to accept with 70% probability.

Let $P_{i,c}$ be individual i's reported acceptance probability of an offer with the salary of the "benchmark likelihood" from city c, where c is one of the four to six cities described above or the current city (where the acceptance probability is the "benchmark probability"). We run the regression

$$\log\left(\frac{P_{i,c}}{1-P_{i,c}}\right) = \sum_{x\in\mathbb{X}} \beta_x D_{x,o,c} + \phi \mathbb{I}_{c\in h} + \eta \mathbb{I}_{R=h,c\notin h} + \upsilon_i + \varepsilon_i, \tag{75}$$

where $D_{x,o,c}$ are six distance dummies between current city o and destination city c (10-25km, 25-50km, 50-100km, 100-200km, 200-400km, and $\geq$ 400km), $\mathbb{I}_{c\in h}$ is a dummy that is equal to one if the destination is in the home "location" (NW, SW, NE, SE), and $\mathbb{I}_{R=h,c\notin h}$ is a dummy indicating that the city is in the home region (East or West Germany) but not in the home location. The current city is the omitted category.

The first two columns of Table S4 show the regression coefficients and standard errors from the regression. The remaining four columns then show the 19 moments we obtain from this exercise, which are also depicted in Figure 5(b). We construct these by rescaling the regression coefficients: we add a common constant to all so that the estimated coefficient on the regression constant matches the average log-odds ratio in the current city of work in the raw data. Column "Within" shows this rescaled coefficient on the regression constant. Column "Not home" shows the rescaled coefficients on the distance dummies $\hat{\beta}_x$ for moves not to the home location or region. Column "Home location" shows the rescaled coefficients

$\hat{\beta}_x + \hat{\phi}$ for moves to the home location. Finally, the last column shows the rescaled coefficients $\hat{\beta}_x + \hat{\eta}$ for moves to the home region.

Table S4: Log-Odds Ratio of Acceptance across Locations

	Regression		Moments			
	Reg coef.	SE	Within	Not home	Home location	Home region
Constant	1.265	0.024	1.219			
10–25km	−0.771	0.108		0.448	1.474	0.649
25–50km	−0.971	0.062		0.248	1.273	0.448
50–100km	−1.410	0.052		−0.191	0.835	0.010
100–200km	−1.830	0.037		−0.611	0.415	−0.410
200–400km	−1.971	0.032		−0.752	0.274	−0.551
≥400km	−2.011	0.040		−0.792	0.233	−0.592
$\mathbb{I}_{c\in h}$	1.026	0.053				
$\mathbb{I}_{R=h,c\notin h}$	0.201	0.072				

Notes: Columns 1 and 2 report raw regression coefficients and their standard errors. The remaining columns construct standardized moments by adding to these coefficients a constant so that the value of the regression constant matches the average log-odds ratio of acceptance for offers within the same city or county. Column "Within" shows the acceptance probability within the current location (the rescaled constant from the regression). Column "Not home" shows the rescaled coefficients on the distance dummies $\hat{\beta}_x$ for moves not to the home location or region. Column "Home location" shows the rescaled coefficients $\hat{\beta}_x + \hat{\phi}$ for moves to the home location. Column "Home region" shows the rescaled coefficients $\hat{\beta}_x + \hat{\eta}$ for moves to the home region.

N.2.3 Search Intensity

We asked participants in the survey about their current salary and presented them with four hypothetical offers for jobs with similar characteristics in the same city, with salaries equal to their current salary, as well as 70–90%, 110–140%, and 150–170% of their current salary. Respondents reported their probability of accepting each offer. We take an average of the acceptance probabilities across the four offers. Let d_i^a be dummies that are equal to one if the average acceptance probability is 0-20%, 20-40%, 40-60%, 60-80%, and 80-100%, respectively. We then asked respondents if they spent any time searching for a job in the last three months. For those who did, we asked them in how many weeks they searched and how many hours they spent searching, separately for their current city and for the rest of the country.

We then run a regression of a dummy that is equal to one if workers searched within their current city on the average within-location acceptance probability dummies:

$$s_i = \sum_{a\in A} \beta_a d_i^a + \varepsilon_i,$$

where s_i is a dummy equal to one if worker i searched for jobs in her county in the last three months, and d_i^a are the dummies constructed above.

The first two columns of Table S5 show the regression coefficients and standard errors from this regression. The final column shows standardized coefficients, normalized so that the value of the 40-60% group is 1. These moments are depicted in Figure 5(c).

Table S5: Search Effort by Acceptance Probability

Acceptance probability	Regression		Moments
	Reg coef.	SE	Standardized coef.
0-20%	0.092	0.031	0.472
20-40%	0.138	0.019	0.707
40-60%	0.195	0.014	1.000
60-80%	0.284	0.019	1.458
80-100%	0.371	0.032	1.904

Notes: Columns 1 and 2 show the regression coefficient and standard errors. The last column shows standardized coefficients, normalized so that the value of the 40-60% group is equal to one.

N.2.4 Directed Search Shares

We showed each respondent one random city from each of the four "locations" (NW, SW, NE, SE). To give some context to respondents, we also provided the size of the population, the state, and the location of the city. We also asked respondents whether they live or grew up in counties different from their current workplace and, if so, added those residence and birth counties to the list. We finally added the worker's current city or county to the list. We then introduced a hypothetical scenario where workers were asked to imagine that they suddenly lost their job, and asked respondents what share of their search effort for a new job (out of 100%) they would devote to finding a job in each of the cities or counties.

Let s_{ioc} be the share of the search effort of worker i in city o directed to destination city c, relative to the share of search effort directed to the current city. We run the regression

$$s_{ioc} = \sum_{x \in \mathbb{X}} \beta_x D_{x,o,c} + \phi \mathbb{I}_{c \in h} + \eta \mathbb{I}_{R=h, c \notin h} + \psi \mathbb{I}_{c=o, c \neq h_c} + \upsilon_i + \varepsilon_i,$$

where $D_{x,o,c}$ are six distance dummies between current city o and destination city c (10-25km, 25-50km, 50-100km, 100-200km, 200-400km, and $\geq$ 400km), $\mathbb{I}_{c \in h}$ is a dummy that is equal to one if the destination is in the home "location" (NW, SW, NE, SE), $\mathbb{I}_{R=h, c \notin h}$ is a dummy indicating that the city is in the home region (East or West Germany) but not in the home location, and $\mathbb{I}_{c=o, c \neq h_c}$ is a dummy that is equal to one if the destination is the current work city but not the home city. The omitted category is the current work city when it is the home city.

The first two columns of Table S6 show the regression coefficients and standard errors from the regression. The remaining four columns then show the 19 moments we obtain from this exercise, which are also depicted in Figure 5(d). We construct these by adding the value of the constant to all coefficients and then normalize them so that search effort in the current city when it is not the home city is one. Column "Within" shows this rescaled coefficient. Column "Not home" shows the rescaled coefficients on the distance dummies $\hat{\beta}_x$ for moves not to the home location or region. Column "Home location" shows the rescaled coefficients $\hat{\beta}_x + \hat{\phi}$ for moves to the home location. Finally, the last column shows the rescaled coefficients $\hat{\beta}_x + \hat{\eta}$ for moves to the home region.

Table S6: Relative Search Effort across Locations

	Regression		Moments			
	Reg coef.	SE	Within	Not home	Home location	Home region
Constant	1.063	0.076				
$\mathbb{I}_{c=o,c\neq h_c}$	1.724	0.111	1.000			
10–25km	0.711	0.130		0.636	1.247	0.775
25–50km	0.783	0.081		0.662	1.273	0.801
50–100km	0.477	0.084		0.553	1.164	0.691
100–200km	−0.159	0.081		0.324	0.935	0.463
200–400km	−0.464	0.078		0.215	0.826	0.353
≥400km	−0.593	0.082		0.169	0.780	0.307
$\mathbb{I}_{c\in h}$	1.703	0.063				
$\mathbb{I}_{R=h,c\notin h}$	0.385	0.085				

Notes: Columns 1 and 2 report raw regression coefficients and their standard errors. The remaining columns construct standardized moments by adding to each coefficient the estimated coefficient on the regression constant. We normalize them so that search effort in the current city when it is not the home city is one. Column "Within" shows this rescaled coefficient. Column "Not home" shows the rescaled coefficients on the distance dummies $\hat{\beta}_x$ for search not in the home location or region. Column "Home location" shows the rescaled coefficients $\hat{\beta}_x + \hat{\phi}$ for search in the home location. Column "Home region" shows the rescaled coefficients $\hat{\beta}_x + \hat{\eta}$ for search in the home region.

N.2.5 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

$$\Delta \log(w_{it}) = \sum_{h\in\mathbb{H}} \sum_{s\in\mathbb{S}} \beta_{hs} d_{it}^{s} \mathbb{I}_i^h + BX_{it} + \gamma_t + \epsilon_{it}, \tag{76}$$

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 moves (NW-SW, NW-NE, NW-SE, SW-NW, and so on) and the 4 possible within-location moves. As in the main text, we define 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 is within 200km of her job. 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, a dummy for the worker's gender, and dummies for the number of previous job switches. The controls also include 12 dummies for the remaining worker job moves that we exclude from our baseline definition (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 movers relative to stayers, the omitted category. Table S7 shows the estimated coefficients on the move 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 S7: 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	0.118	(0.002)	0.288	(0.012)	0.154	(0.023)	0.249	(0.041)
	SW	0.232	(0.013)	0.111	(0.007)	0.027	(0.060)	0.132	(0.053)
	NE	0.170	(0.022)	0.262	(0.062)	0.079	(0.009)	0.029	(0.046)
	SE	0.154	(0.036)	0.209	(0.038)	0.157	(0.061)	0.086	(0.013)
SW	NW	0.116	(0.006)	0.222	(0.013)	0.114	(0.069)	0.121	(0.069)
	SW	0.288	(0.011)	0.109	(0.002)	0.275	(0.059)	0.180	(0.022)
	NE	0.274	(0.069)	0.120	(0.048)	0.090	(0.014)	0.073	(0.051)
	SE	0.190	(0.049)	0.131	(0.020)	0.202	(0.044)	0.092	(0.007)
NE	NW	0.112	(0.004)	0.169	(0.030)	0.037	(0.017)	0.124	(0.050)
	SW	0.181	(0.029)	0.108	(0.006)	-0.008	(0.024)	0.097	(0.045)
	NE	0.231	(0.011)	0.301	(0.028)	0.070	(0.002)	0.179	(0.014)
	SE	0.308	(0.059)	0.316	(0.034)	0.112	(0.025)	0.115	(0.009)
SE	NW	0.113	(0.009)	0.260	(0.032)	0.031	(0.051)	0.043	(0.028)
	SW	0.224	(0.031)	0.098	(0.006)	0.020	(0.059)	0.035	(0.016)
	NE	0.176	(0.056)	0.206	(0.049)	0.067	(0.011)	0.114	(0.023)
	SE	0.314	(0.024)	0.277	(0.013)	0.113	(0.014)	0.111	(0.002)

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

N.2.6 Within-Location Wage Gain Distribution

We obtain in the cleaned LIAB data all job-to-job moves that are within the same location (NW, SW, NE, SE). For each of these job changes, we then calculate the associated wage change, computed 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 compute the share of these wage changes associated with a within-location move that are at least -10%, at least -5%, at least 0, at least +5%, and at least +10%. We present these shares in Figure 6(b) and in Table S8.

Table S8: Distribution of Wage Gains

Wage gain at least...	Share
-10%	0.800
-5%	0.749
0%	0.673
+5%	0.584
+10%	0.499

Notes: The table shows the share of wage gains associated with within-location job-to-job moves that are at least equal to the stated percentage change.

N.2.7 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 cross-location moves (NW-SW, NW-NE, NW-SE, SW-NW, and so on) and 4 possible within-location job moves. As in the main text, we define 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 is within 200km of her job. We compute these movers by worker home location (i.e., their type). In total, there are thus $16\times4=64$ 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 S9 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 S9: Job-to-Job Migration Flows Between Locations by Birth Location

	Move to Location:	NW	SW	NE	SE
Birth Location	Current Work Location				
NW	NW	1.012%	0.023%	0.004%	0.002%
	SW	0.243%	1.163%	0.008%	0.010%
	NE	0.228%	0.039%	0.957%	0.031%
	SE	0.155%	0.081%	0.048%	1.071%
SW	NW	1.045%	0.250%	0.008%	0.008%
	SW	0.027%	1.283%	0.001%	0.006%
	NE	0.101%	0.184%	0.915%	0.076%
	SE	0.038%	0.185%	0.029%	1.137%
NE	NW	1.087%	0.036%	0.093%	0.013%
	SW	0.081%	1.279%	0.084%	0.033%
	NE	0.047%	0.011%	0.936%	0.034%
	SE	0.041%	0.052%	0.152%	1.051%
SE	NW	1.070%	0.099%	0.028%	0.116%
	SW	0.047%	1.206%	0.011%	0.137%
	NE	0.036%	0.034%	0.627%	0.164%
	SE	0.013%	0.037%	0.022%	1.107%

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.

N.2.8 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. As in the main text, we keep only workers that report a residence location within 200km of their job; otherwise we suspect the residence location might be misreported. We then average across months. We also compute the share of employed workers working in each location. Table S10 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.

Table S10: Share of Employed Workers by Location of Residence or Work Location

Home Location	Location of...	...Residence	...Work
NW	NW	93.2%	92.0%
	SW	4.2%	5.6%
	NE	1.9%	1.6%
	SE	0.7%	0.8%
SW	NW	4.0%	6.1%
	SW	93.1%	90.9%
	NE	0.7%	0.8%
	SE	2.2%	2.2%
NE	NW	7.5%	12.8%
	SW	4.2%	5.8%
	NE	85.0%	77.1%
	SE	3.3%	4.4%
SE	NW	2.9%	4.4%
	SW	6.6%	9.8%
	NE	2.5%	3.9%
	SE	88.1%	81.9%

Notes: The table shows the fraction of employed workers of the home location indicated in column 1 that live or work in the location indicated in column 2. Column 3 shows the share of workers in each location based on residence, and column 4 shows the share of workers based on work location.

N.2.9 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 S11 presents these worker shares. In our baseline estimation, we use the residence location as target for the distribution of labor.

Table S11: Share of Unemployed Workers by Location of Residence or Location of Last Job

Home Location	Location of...	Residence	Last Job
NW	NW	91.0%	89.1%
	SW	4.5%	6.5%
	NE	3.3%	3.1%
	SE	1.3%	1.4%
SW	NW	4.6%	7.4%
	SW	90.2%	87.5%
	NE	1.5%	1.5%
	SE	3.7%	3.7%
NE	NW	4.9%	10.6%
	SW	2.9%	5.6%
	NE	89.6%	78.7%
	SE	2.7%	5.2%
SE	NW	2.3%	4.2%
	SW	4.7%	9.3%
	NE	2.8%	4.1%
	SE	90.2%	82.5%

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 or whose last job was in that location. Column 3 shows the share of unemployed workers of each home location type by residence location. Column 4 shows the share of unemployed workers of each type by location of the last job.

N.2.10 AKM Firm Fixed Effect 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:

$$fe_j = \sum_{l \in \mathbb{L}} \beta_l \mathbb{I}_j^l + \epsilon_j, \tag{77}$$

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 Appendix B. For similar reasons, we do not include demographic controls. Table S12 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 seven age groups (Less than 30 years, 31-35 years, ... 56-60 years, where the group of under 30 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.[49]

Table S12: Firm Fixed Effect by Location

Dep. var.: fe_j	Coef on Firm FE	SE
Location		
SW	.001	.002
NE	-.165	.002
SE	-.142	.002

Notes: The table presents the estimated coefficients β_l from specification (77) for firm location l indicated in column 1, where NW is the omitted category.

N.2.11 Wage Size Premia

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 et al. (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 et al.

[49] Specifically, the three coefficients for SW, NE, and SE become: -0.003, -0.157, -0.148.

(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

$$\ln(\hat{y}_{jlt}^{wage}) = \beta_{0,l} + \beta_{1,l} \ln(\hat{y}_{jlt}^{size}) + \varepsilon_{jlt}, \tag{78}$$

and report the slope coefficients $\beta_{1,l}$ in Table S13. These coefficients are also shown in Figure 8.

Table S13: Log Wage on Log Firm Size by Location

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

Notes: The table presents the coefficients $\beta_{1,l}$ of regression (78), by location of the firm, indicated in the first column. The residualization procedure is described in the text.

N.2.12 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 S14 shows the estimates. These rates are also shown in Figure 8.

Table S14: 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.

N.2.13 Slope of 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

$$y_{ilt} = B_l X_{ilt} + \gamma_t + \epsilon_{ilt}, \tag{79}$$

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 et al. (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

$$\hat{d}_{ilt}^{sep} = \beta_{0,l} + \beta_{1,l} \ln(\hat{w}_{ilt}) + \varepsilon_{ilt} \tag{80}$$

and report the slope coefficients $\beta_{1,l}$ in Table S15. 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. These coefficients are also shown in Figure 8.

Table S15: Avg. Separation Rates of Workers by Initial Wage

Dep. var.: $\hat{d}_{irt}^{sep}$	Coefficient	SE
Location		
NW	-.029	.000
SW	-.033	.000
NE	-.036	.000
SE	-.035	.000

Notes: The table presents the coefficients $\beta_{1,l}$ of regression (80), by location of the firm. The residualization procedure is described in the text.

N.2.14 Slope of Labor Share on Log TFP

We compute the relationship between firms' labor shares and total factor productivity (TFP) using the IAB Establishment Panel. The IAB Establishment Panel is an annual, representative survey of roughly 15,000 German establishments. We use three questions which capture establishment output, intermediate inputs, and investment. To capture sales, we use the question "What was your business volume in the last fiscal year?". We keep only observations in which the reported business volume represents sales, rather than balance-sheet totals, interest or commission income, insurance premia, or public budgets, which removes about 15% of observations. We use the question "What share of sales was attributed to intermediate inputs and external costs?" to construct intermediate input costs. Finally, we use the question "What was the approximate sum of all investments?" to capture investment. We compute material costs as sales multiplied by the reported intermediate-input share, and value added as sales net of material costs.

We then merge the data from the IAB Establishment Panel with variables from the BHP. For each establishment we observe the county, 3-digit time-consistent industry based on Eberle et al. (2011), the number of full-time employees, and the average daily wage of full-time workers. We compute annual labor costs as the number of full-time employees times the average daily wage times 365. We deflate sales, material costs, value added, and labor costs using the county-level price deflator described in the main text.

We next construct firm-level capital. To deflate investment, we use federal state-by-sector investment series from the investment tables of the Volkswirtschaftliche Gesamtrechnungen der Länder. For each federal state by sector, we combine nominal investment from the "Neue Anlagen in jeweiligen Preisen" tables with the real investment growth from the "Neue Anlagen (preisbereinigt, verkettet)" tables. The latter are reported as growth rates rather than levels and we therefore construct a real investment index. We normalize nominal investment to the same base, and then define the implicit investment deflator as the ratio of the nominal and real investment indices. We use this deflator to convert the establishment-level nominal investment into real investment.

We also construct depreciation rates by state and sector, combining information from the capital and investment series of the Volkswirtschaftliche Gesamtrechnungen der Länder. We obtain nominal net capital stocks for equipment and buildings from the capital tables of the Volkswirtschaftliche Gesamtrechnungen der Länder. We also obtain the corresponding real net-capital-stock indices. We combine these data with the corresponding nominal investment and the real investment indices in equipment and buildings from the investment tables. For each state-sector cell, we then use the nominal series and the real indices to construct real capital and real investment series with a common base year. We then compute annual depreciation rates separately for equipment and buildings as the residual in the capital accumulation equation $\delta_t = (K_{t-1} + I_t - K_t)/K_t$. We average these rates over 2004–2014 and combine the equipment and building rates using their average capital-stock shares to obtain a single depreciation rate for each state-sector cell.

We use the perpetual inventory method to construct capital stocks. For each continuous firm spell in the IAB Establishment Panel of at least five years, we initialize real capital using the firm's average real investment in the first two years divided by the corresponding state-by-sector investment-to-capital ratio constructed above. This initialization assumes that the firm's capital stock is proportional to its investment in the same way as the state-sector average. We then update capital according to

$$K_{j,t} = (1 - \delta_{s(j)})K_{j,t-1} + I_{j,t}, \tag{81}$$

where $\delta_{s(j)}$ is the depreciation rate for firm j's state-sector cell. This method gives us a firm-level capital stock in each year.

We estimate TFP using the value-added production function approach of Wooldridge (2009). For each one-digit industry, we estimate

$$\ln(va_{jt}) = \beta_l \ln(pay_{jt}) + \beta_k \ln(K_{jt}) + \gamma_t + g\left(\ln(K_{j,t-1}), \ln(m_{j,t-1})\right) + \epsilon_{jt}, \tag{82}$$

where va_{jt} is real value added, pay_{jt} is real labor costs, K_{jt} is real capital stock, γ_t are year fixed effects, and $g(\cdot)$ is a third-order polynomial in lagged capital and lagged material costs. We instrument current labor costs with lagged labor costs. We cluster standard errors at the firm level, and compute log firm TFP as

$$\ln(TFP_{jt}) = \ln(va_{jt}) - \hat{\beta}_l \ln(pay_{jt}) - \hat{\beta}_k \ln(K_{jt}). \tag{83}$$

We normalize TFP by its one-digit industry mean. We then compute the labor share as $LS_{jt} = pay_{jt}/va_{jt}$. For each location, we trim observations below the 1st and above the 99th percentile of the TFP distribution.

To estimate the relationship the labor share and log TFP, we run the following regression

$$LS_{jkt} = \alpha + \beta \ln(\widehat{TFP}_{jkt}) + \rho_k + \gamma_t + \varepsilon_{jkt}, \tag{84}$$

where $\ln(\widehat{TFP}_{jkt})$ is normalized log TFP, ρ_k are one-digit industry fixed effects, and γ_t are year fixed effects. We estimate this regression separately for each of the four locations and report the slope coefficients β in Table S16. These coefficients are also shown in Figure 8.

Table S16: Regression of Labor Share on Log TFP

Dep. var.: LS_{jkt}	Coefficient	SE
Location		
NW	$-.320$	.048
SW	$-.322$	.053
NE	$-.314$	.038
SE	$-.356$	.037

Notes: The table presents the coefficients β from a regression of firms' labor share on their normalized log TFP, by location of the firm. The construction of labor share and TFP is described in the text.

N.2.15 Slope of 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

$$\Delta \ln(w_{ilt}) = B_l X_{ilt} + \gamma_t + \epsilon_{ilt}, \tag{85}$$

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 et al. (2011) (WZ93 classification). From these four regressions (one for each location l), we construct residuals for the log real wage gain, $\hat{\epsilon}_{ilt}^{gain}$. We add back the mean of the log real wage gain in each location, $\overline{\Delta \ln(w_{ilt})}$, to obtain a cleaned log real wage, $\Delta \ln(\hat{w}_{ilt}) = \overline{\Delta \ln(w_{ilt})} + \hat{\epsilon}_{ilt}^{gain}$. We similarly residualize the log real wage of the worker at the origin firm, $\ln(w_{ilt-1})$, to obtain the residualized initial log real wage, $\ln(\hat{w}_{ilt-1})$. We then regress the residualized log real wage gains on the residualized log initial real wages in each location

$$\Delta \ln(\hat{w}_{ilt}) = \beta_{0,l} + \beta_{1,l} \ln(\hat{w}_{ilt-1}) + \varepsilon_{ilt} \tag{86}$$

and report the slope coefficients $\beta_{1,l}$ in Table S17. 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. These coefficients are also shown in Figure 8.

Table S17: Log Wage Gain of Movers by Initial Wage

Dep. var.: $\Delta \ln(\hat{w}_{irt})$	Coefficient	SE
Location		
NW	-.564	.001
SW	-.586	.001
NE	-.570	.003
SE	-.569	.002

Note: The table presents the coefficients $\beta_{1,l}$ of regression (86), by location of the origin firm. The residualization procedure is described in the text.

N.2.16 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 GDP per capita with the resulting prices in each year to obtain state-level real GDP per capita 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 S18 presents the results. These coefficients are also shown in Figure 8.

Table S18: 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).

N.2.17 Deciles 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

$$\ln(y_{jlt}^{size}) = B_l X_{jlt} + \gamma_t + \epsilon_{jlt}, \tag{87}$$

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 et al. (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 S19 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 S19: 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 is residualized as described in the text.

O Details on Computation and Estimation

O.1 Solution Algorithm

We solve the model on a grid for firm productivity. Proposition 1 lets us work directly in productivity space: once the wage schedule $w_j(p)$ is known, the relevant worker and firm objects can be evaluated as functions of p rather than as functions of both wages and productivity. The numerical algorithm is a nested fixed point:

1. Start with guesses for the wage schedule $w_j(p)$, vacancies $v_j(p)$, market tightness ϑ_j, and firm size per vacancy $\tilde{l}^i_j(p)$ in each location and for each worker type.
2. Given wages, vacancies, and tightness, solve workers' value functions. This step delivers the probability of accepting each offer, $\tilde{\mu}^{E,i}_{jx}(p, p')$ and $\tilde{\mu}^{U,i}_{jx}(b, p')$, and directed search effort across destinations.
3. Given worker policies, compute separation rates using equation (18). We then iterate on equations (17), (19), and (20) to obtain employment distributions and firm sizes per vacancy that are consistent with steady state flows.
4. Given firm sizes and separation rates, update wages using the differential equation in Proposition 1, with the boundary condition for the lowest-productivity firm that posts vacancies. We then update vacancy postings and market tightness from firms' vacancy posting condition and aggregate market clearing.
5. Repeat the previous steps until wages, vacancies, tightness, worker values, and firm sizes per vacancy converge.

In the estimation, location prices are fixed at their empirical values. We also choose unemployment benefits so that the reservation wage in each location is consistent with the estimated reservation-wage parameter. In the counterfactuals, unemployment benefits are held fixed at their estimated values, so reservation wages are equilibrium objects. Counterfactual prices are also solved in equilibrium using the model-implied changes in local GDP.

O.2 Estimation Algorithm and Outcomes

The estimated parameter vector ϕ^* solves

$$\phi^* = \arg\min_{\phi \in \mathbb{F}} \mathcal{L}(\phi), \tag{88}$$

where

$$\mathcal{L}(\phi) = \sum_{x \in \mathcal{T}} \omega_x \left(\frac{m_x(\phi) - \hat{m}_x}{d_x} \right)^2 .$$

Here $\mathcal{T}$ is the set of 284 independent targeted moments, $\hat{m}_x$ is the empirical target, $m_x(\phi)$ is the same object computed in the model, d_x is a target-specific deflator, and ω_x is the moment weight. The estimation code stores one additional entry for the normalized 40–60% search-intensity bin, which is equal to one in both data and model and is used only to keep the plotted search-intensity profile aligned. The deflators put moments with different units on comparable scales. For most blocks of parameters, d_x is the mean absolute value of the corresponding empirical moments; for the survey acceptance and directed-search blocks, we use block-specific normalizations so that slopes, levels, and ratios enter in the same units as in the estimation. The weights give each block of moments comparable total mass, so that large blocks such as origin-destination flows do not mechanically dominate smaller blocks such as regional unemployment rates or survey acceptance moments.

We estimate 20 parameters and target the 284 moments listed in Table A9. The admissible set $\mathbb{F}$ imposes the sign and support restrictions required by the model. Some parameters are sampled in logs and then transformed back inside the model. In particular, the scale parameters of search efficiency and moving costs, the within-location switching cost, and the vacancy-cost intercept are sampled in log space to keep them positive.

The minimization uses Markov Chain Monte Carlo for classical estimators, following Chernozhukov and Hong (2003). The implementation is organized in two rounds. In the first round, we start 200 independent chains from dispersed initial values, using Sobol draws to cover the admissible support. In each chain, the first 1,000 iterations are used to tune the proposal variance and the next 10,000 iterations are used for the estimation. The proposal scale is adjusted during the run to keep the chain moving while avoiding excessive boundary hits.

The second round starts from the best first-round draws and uses tighter bounds around the promising region of the parameter space. Again, each of the 200 independent chains has a tuning block followed by 10,000 estimation iterations. The final estimates are selected from the second-round draws: we take the average of the 100 parameter vectors with the lowest objective value. This is useful because the objective is sometimes asymmetric near the boundary of the admissible set; averaging all late-chain draws can then move the estimate away from the best-fitting region. This final-selection rule follows the same practical logic as in Moser and Engbom (2022).

After estimation, we recompute all model moments at the selected parameter vector and compare them to the targets in the figures and tables in the main appendix. We also use the full set of evaluated second-round parameter vectors to build the local and global distance profiles and the Jacobian-style identification map reported in Appendix H.

References

Andreoni, J. and C. Sprenger (2012). Estimating time preferences from convex budgets. *American Economic Review 102*(7), 3333–3356.

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of west german wage inequality. *Quarterly Journal of Economics 128*(3), 967–1015.

Chernozhukov, V. and H. Hong (2003). An mcmc approach to classical estimation. *Journal of Econometrics 115*(2), 293–346.

Eberle, J., P. Jacobebbinghaus, J. Ludsteck, and J. Witter (2011). Generation of time-consistent industry codes in the face of classification changes - simple heuristic based on the establishment history panel (bhp). FDZ Methodenreport, 05/2011.

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

Moser, C. and N. Engbom (2022). Earnings inequality and the minimum wage: Evidence from brazil. *American Economic Review 112*(12), 3803–3847.

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

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

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics letters 104*(3), 112–114.