THE PINCHAS SAPIR CENTER

FOR DEVELOPMENT

Tel Aviv University

המרכז לפיתוח ע"ש פנחס ספיר

ליד אוניברסיטת תל אביב

"עמותה רשומה"

580010221

THE PINHAS SAPIR CENTER FOR DEVELOPMENT
TEL AVIV UNIVERSITY

Outside Options in the Labor Market

Sydnee Caldwell and Oren Danieli

Discussion Paper No. 9-2020

August 2020

אוניברסיטת תל-אביב, תל אביב יפו 6997801 | טל': 03-6409202 | פקס: 03-6405815

Tel Aviv University, Tel Aviv Yafo 6997801 Israel | Tel. 972-3-6409202 | Fax 972-3-6405815 | email: sapircen@tauex.tau.ac.il

Outside Options in the Labor Market

Sydnee Caldwell and Oren Danieli*

August 17, 2020

Abstract

This paper develops a method to estimate the outside employment opportunities available to each worker, and to assess the impact of these outside options on wage inequality. We outline a matching model with two-sided heterogeneity, from which we derive a sufficient statistic, the "outside options index" (OOI), that captures the effect of outside options on wages, holding productivity constant. This OOI uses the cross-sectional concentration of similar workers across job types to quantify the availability of outside options as a function of workers' commuting or moving costs, preferences, and skills. Higher concentration in a narrower range of job types implies lower OOI and higher dispersion across a wide variety of job types corresponds to higher OOI. We use administrative data to estimate the OOI for every worker in a representative sample of the German workforce. We estimate the elasticity between the OOI and wages using two sources of quasi-random variation in the OOI, holding workers' productivity constant: the introduction of high-speed commuter rail stations, and a shift-share ("Bartik") instrument. Using this elasticity and the observed distribution of options, we find that differences in options explain 30% of the gender wage gap, 88% of the citizen-non-citizen wage gap, and 25% of the premium for higher education. Differences in options between genders and education groups are driven mostly by differences in the implicit costs of commuting and moving.

*Danieli: Tel-Aviv University. Email: orendanieli@tauex.tau.ac.il. Caldwell: UC Berkeley. We thank Chiara Farronato, Roland Fryer and Lawrence Katz for advice and guidance throughout this project. We also thank Isaiah Andrews, David Autor, Gary Chamberlain, John Coglianese, Ashley Craig, Ellora Derenoncourt, Ed Glaeser, Nathan Hendren, Simon Jäger, Myrto Kalouptsidi, Robin Lee, Ro'ee Levy, Alan Olivi, Amanda Pallais, Will Rafey, Assaf Romm, Jonathan Roth, Neil Thakral and Dan Svirsky for helpful comments. Participants at Harvard and MIT seminars also provided useful feedback. This paper uses data from the German Institute for Employment Research (IAB) under the project "Market Power in the Labor Market" (project number 1241/1242). All results based on IAB micro-data have been cleared for disclosure to protect confidentiality. We thank Johannes Schmieder and Daniel Heuermann for generously sharing their data on German train expansions with us. Danieli thanks the Sapir Center at Tel-Aviv University for financial support

1 Introduction

In almost every model of the labor market, wages depend on a workers' outside options: the amount of compensation they could receive from different employers. In a perfectly competitive labor market, an equally attractive outside option always exists, and competition between identical employers sets compensation at the marginal product. However, in reality, a worker's next best option could require different skills, working hours or be located in a different city. The availability of outside options could be systematically worse for some workers due to the health of their local labor market, because they are unwilling or unable to commute, or because their skills are valuable only for a few employers or industries. Such differences could have significant implications for their incomes.

A key challenge for empirical research on this topic is that a worker's outside option set is not typically observed. Even within the same firm and occupation, workers may face different options due to their specific set of skills, their preferences or their constraints. As a result, little is known about which workers have better outside options and what role options play in generating wage inequality.

The first contribution of this paper is to develop an empirical procedure to uncover a key latent parameter in most wage-setting models: the value of an individual's option set. We show how this latent parameter can be derived from the cross-sectional concentration of similar workers across jobs. If similar workers are concentrated in a certain region, industry, occupation or other job characteristics, then the worker's options are more limited. We quantify this concentration in a single "outside options index" (OOI). We show that in a matching model of heterogeneous workers and jobs this OOI is a sufficient statistic for the effect of outside options on compensation, when holding productivity constant. We then estimate the OOI for every worker using administrative matched employer-employee data from a 1% representative sample of workers in Germany. Examining the distribution of the OOI, we find what workers' characteristics are associated with better outside options. Next, we quantify the impact on wages by estimating the elasticity between the OOI and wages using two quasi-random sources of variation in the OOI, that holds workers' productivity constant: the introduction of high-speed commuter-rails, and a shift-share ("Bartik") instrument.

Our second contribution is to show that differences in outside options explain substantial portions of several widely-discussed wage gaps between different segments of workers. Outside options explain 30% of the gender wage gap in Germany. This gender difference is driven entirely by differences in willingness to commute or move. We also find that differences in outside options account for 88% of the wage gap between German citizens and non-citizens, and about 25% of the high-education wage premium. The availability of more options also increases the wage premium for urban residents. In contrast, differences in outside options reduce inequality between occupations, since high-paying occupations tend to be more specialized and workers in them therefore have fewer options.

We start by outlining a static model of the labor market that illustrates how, with two-sided heterogeneity, differences in outside options lead to differences in compensation, even for equally productive workers. Our model is based on the classic Shapley and Shubik (1971) assignment game - a two-sided matching model with transfers. Compensation in this setting is set to prevent workers from moving to their outside options; because of heterogeneity, this will be below their full productivity in the first-best option. A direct implication is that workers' compensation is not only determined by what they produce, but also by their ability to produce in more places.

We derive a sufficient statistic from this model, the "outside options index" (OOI), that summarizes the impact of options on compensation. It measures the quantity of relevant jobs for a given worker. If a worker gets access to more similar jobs, their compensation would increase by exactly the increase in OOI times a constant elasticity, even though their productivity remains constant. The OOI depends on two factors: the supply of jobs, and worker flexibility (i.e. a worker's ability or willingness to take jobs in more places, more occupations, more industries, etc.). Workers with more relevant jobs, as captured by the OOI, will on average have both a better outside option, and will be able to sort into better matches, conditional on their productivity.

We show that the OOI is equal to a standard concentration index: workers with more options are those who, in equilibrium, are found in a greater variety of jobs. Under standard assumptions on the distribution of match quality (Choo and Siow, 2006; Dupuy and Galichon, 2014), the OOI is equal to the entropy index. This index, with a negative sign, is used in the industrial organization literature as a measure of market concentration (Tirole, 1988), similar to the Herfindhal-Hirschman Index (HHI), which has also been used to measure concentration in labor markets (Azar et al., 2017; Benmelech et al., 2018). In contrast to most concentration indices, our index is not measured on a specific dimension such as occupation, or industry. Instead, workers with more options are those that are less concentrated across jobs, on all dimensions included in our data set. Options here are estimated in equilibrium, based on matches we actually observe in cross-sectional data. Jobs that the worker will never take in practice because they are less attractive will not enter the OOI nor affect compensation even if the employer is willing to hire. To isolate the effect of more options from the effect of productivity, the OOI is calculated without using any information on wages or wage offers.

We develop a method that estimates the OOI for each worker in the labor market, which is computationally feasible even in large datasets. The OOI is a function of the joint probability of every worker to be in every job. Our method estimates this probability, using the cross-sectional distribution of similar workers. We show that this problem can be translated into a logistic regression framework. We then use the fast implementation of logistic regressions to estimate the probabilities for every worker-job combination. From those probabilities we can directly calculate the OOI for each worker.

We then use the OOI to analyze the impact of outside options on inequality, starting with identification of which workers have better outside options. Specifically, we estimate the OOI for every worker in a representative sample of German workers in 2014 using administrative linked employer-employee data. Looking across observed workers' characteristics, we find that the OOI is higher for men, German citizens, city residents, more educated and more experienced workers. We also find that higher skill workers such as medical doctors or pilots tend to be more specialized in their current industry, which narrows down their outside options. The OOI also predicts which workers will be less affected by a mass-layoff: workers with better outside options recover more quickly from a displacement. Because we do not use wages to calculate the OOI, there is not a mechanical link between the OOI and wages.

We use two sources of quasi-random variation in options, that do not affect productivity, in order to estimate the elasticity between the OOI and wages: the introduction of high-speed commuter rail stations (Heuermann and Schmieder, 2018), and a standard industry shift-share ("Bartik") instrument (Beaudry et al., 2012). These sources of variation allow us to verify that, even if our model is not perfectly specified, there is a link between our outside options index and wages in the data. Our first source of variation focuses on the introduction of new train stations that were constructed along existing routes. These stations effectively increased the labor market size for workers in small German cities that happened to live along the shortest route between two major cities. The second source of variation in outside options uses differences in exposure to industry growth trends between local labor markets. We compare workers who work in the same industry, but have outside options in different industries because they reside in different parts of the country. We instrument for the growth in outside options in other industries with the national industry trends to exclude the impact of local productivity shocks. Both quasi-random sources of variation yield a similar semi-elasticity of roughly .17-.32 between the OOI and wages.

Combining this elasticity with the estimated distribution of the OOI, we find that differences in outside options tend to increase wage inequality. Differences in options lower compensation for women by six percentage points, explaining roughly thirty percent of the overall gap in Germany. They also account for an eight percentage points difference in compensation between immigrants and natives, which is 88% of the overall gap. We also find large effects on the return to higher-secondary education.[1] Graduates from higher-secondary education have access to more options, which increases their compensation by seven percentage points. This is about a quarter of the total return to higher-secondary education.

Finally, we examine the reasons why workers face different options. We start by examining the parameters that determine workers' options, to understand which ones are most significant. We then use the underlying model to create counterfactual changes to the OOI under different scenarios. These exercises show that the heterogeneity in the ability to commute or move is a key factor in explaining variation in outside options. This factor can account for the full gender gap in outside options. We also find that without their higher willingness to work at more distant jobs, high-educated workers would actually have fewer options. Our analysis suggests that this is likely because their skills tend to be more industry specific.

Related Literature Our paper contributes to at least three distinct literatures. First, we contribute to a large literature on imperfect competition in the labor market by estimating the impact of outside options for every worker in the labor market. While outside options are a key parameter in many labor models, prior work has not focused on estimating the distribution of this parameter across different workers. Most empirical work on imperfect competition has used natural experiments in specific segments of the labor market to show that firms face upward sloping supply curves (see, e.g. Naidu, 2010; Naidu et al., 2016; Ransom and Sims, 2010; Staiger et al., 2010). Beaudry et al. (2012) and Caldwell and Harmon (2018) take a different approach, and provide direct evidence that outside options directly impact workers' earnings, but do not investigate which workers have better options, nor the consequences for between-group wage inequality.[2] Our paper adds to this literature by providing estimates of the distribution of options and by providing

[1] The level that grants a certificate allowing college admission.

[2] Beaudry et al. (2012) show that shocks to one industry "spill over" onto the wages of other industries. Caldwell and Harmon (2018) show that workers with better information about their outside options see greater wage growth. Jäger et al. (2018) focus on a specific outside option– unemployment insurance– and find that changes in UI generosity has little to no effect on workers' wages. Their result fits our finding that what matters for wage-setting is the value of a worker's best alternative to a match. For most workers, this is likely the value of working in another job, not the value of unemployment.

descriptive evidence on why some workers have more options than others. By combining this distribution with a causal estimate of the impact of options on wages, we are also able to present the first estimates of the impact of outside options on each individual's wages.

Our theoretical framework emphasizes market imperfections arising from worker and employer heterogeneity. This is similar to the approach taken by Card et al. (2018), and is a standard approach in the industrial organization literature for analyzing market imperfections (for instance, Berry et al., 1995). Recent work by Dube et al. (2018) shows that even small amounts of heterogeneity can generate substantial market imperfections. One difference between our approach and that in the search literature is that we focus on a static equilibrium. While work by Postel-Vinay and Robin (2002) shows that differences in options (as the result of on-the-job search) can impact wage growth during an employment spell, these dynamic considerations are beyond the scope of this paper.

Second, our paper contributes to a small literature on the impact of imperfect labor market competition on between-group wage inequality. Theoretical papers in this literature have argued that some groups such as women or minorities have systematically worse options, enabling their employers to pay them lower wages. These worse options may generate either higher search costs (Black, 1995) or less elastic supply to a particular firm (Robinson, 1933), and can lead to racial or gender wage gaps. Empirical papers in this literature have shown evidence that group differences do exist in both labor supply to a firm (Manning, 2003; Hirsch et al., 2010; Ransom and Oaxaca, 2010) and in rents (Card et al., 2016). A key advantage of our setting is that we are able to combine our estimates of group differences in outside options with a causally estimated elasticity between options and wages. This allows us to translate the estimated group differences in options into group differences in wages, and quantify the portion of between-group inequality that can be attributed to imperfections in the labor market.[3]

Finally, our paper contributes to a recent empirical literature on labor market size and concentration, by characterizing workers options using multiple worker and job characteristics at once. Manning and Petrongolo (2017) and Nimczik (2017) develop methods to uncover the size of a workers' labor market based on willingness to commute and on observed firm-firm transitions. Azar et al. (2017) and Benmelech et al. (2018) examine trends in labor market concentrations by calculating Herfindahl-Hirschman indices (HHI's) by occupation/industry, within a geographic area. Hsieh et al. (2013) estimate concentration

[3]Our setting expands the setting of Bidner and Sand (2016) who quantify the portion of the gender gap that can be attributed to differences in outside options driven solely by differences in access to industries. Our method includes several additional factors, such as differences in commuting costs, that we find to be generating the majority in differences in outside options between genders. We also analyze additional wage gaps beyond gender such as the education, city and citizenship premium.

trends by occupations and demographics such as gender and race using a model similar to ours.

In this paper we develop a method to estimate labor market size and concentration that incorporates five key features. First, when estimating workers options, we account for all job characteristics in our data. This combines all the dimensions that previous papers have used, such as geography, occupations and industry, together with job characteristics that were not used before such as working hours. Second, we account for outside options in different industries and occupations. Third, we allow each worker to have a different set of options depending on their demographics, locations, skills and preferences. Fourth, instead of assuming workers can be partitioned into distinct local labor markets, we allow options sets to overlap between workers. We also allow the distance workers are willing to travel to vary by their characteristics. Fifth, we introduce a more continuous notion of options, accounting for the fact that some options are more relevant than others. These five features allow us to estimate the value of an option set more precisely for every individual worker.

The remainder of the paper proceeds as follows: Section 2 outlines the theoretical matching model and derives the Outside Options Index (OOI). Section 3 describes the relevant features of the German labor market and the key features of the administrative linked employer-employee data that we use. Section 4 explains the empirical procedure of estimating the OOI. Section 5 describes the empirical estimates of worker outside options and presents descriptive statistics on their distribution. Section 6 estimates the elasticity between the outside options index and wages using two quasi-random sources of variation in options. Section 7 analyzes the overall effect on wage inequality. Section 8 concludes.

2 A Model of Outside Options and Wages

This section derives a model of a heterogeneous competitive labor market. We use this model to derive the outside options index (OOI), and show it is a sufficient statistic for the impact of outside options on wages. To provide additional intuition for the OOI and its effect in the model, we describe a simple parametric example. We summarize this section by discussing what is and what is not captured in the OOI using the model's assumptions.

2.1 Setup

There is a continuum set of workers $\mathcal{I}$ with measure I and a continuum set of one-job firms $\mathcal{J}$ with a measure J which we pin down to 1. If a worker $i \in \mathcal{I}$ works at job $j \in \mathcal{J}$, they produce a value of y_{ij} to the employer and a job-specific amenity valued a_{ij} to the worker. The value of y_{ij} is net of all costs, including capital and amenities. The value for a_{ij} includes all non-pecuniary impacts on worker i's utility including effort, interest, number of vacation days and more. The sum of these two values is the total value of a match, τ_{ij}. This is defined for every potential worker-job pair, even those that are not observed in equilibrium.[4] The value of τ_{ij} is taken as exogenous; all decisions by workers and employers that could affect this value such as investment in capital or human capital and location choices are pre-determined.[5]

Employers and workers decide how to split the total surplus τ_{ij} into worker compensation (ω_{ij}) and employer profits (π_{ij}).

$$\tau_{ij} = \pi_{ij} + \omega_{ij} = y_{ij} + a_{ij}$$

This division is accomplished via a set of transfers (wages) w_{ij}, which allow the worker and employer to divide the total value produced in any way between them:

$$\begin{aligned} \pi_{ij} &= y_{ij} - w_{ij} \\ \omega_{ij} &= a_{ij} + w_{ij} \end{aligned}$$

2.2 Equilibrium

We next derive the allocation of workers into jobs and equilibrium wages. We use an equilibrium notion based on cooperative game theory, which is identical to the assignment game, first analyzed by Shapley and Shubik (1971). We assume a static framework with perfect information. There are additional equilibrium concepts that lead to the same result.[6] We use a cooperative framework since it is somewhat more general as it does not make any assumption about how agents reach this equilibrium (e.g. who makes offers).

An allocation is defined as a set $M = \{(i,j)\,|i \in \mathcal{I}, j \in \mathcal{J}\}$ in which no i or j appears twice, so every worker can work only in one job, and every job can hire exactly one

[4]Formally the value is a function $\tau : \mathcal{I} \times \mathcal{J} \to \mathbb{R}$.

[5]This is similar to Kreps and Scheinkman (1983) who show how even with competition on prices, pre-determined quantities would deviate from a Bertrand competition.

[6]Pérez-Castrillo and Sotomayor (2002) show one specific mechanism that leads to the same equilibrium using sub-game perfect Nash equilibrium.

worker. For a given allocation M we can define an invertible function on the domain of matched workers $m(i)$ such that $(i, m(i)) \in M$. Note that we do not require all workers and jobs to be in M; some workers can be unemployed and some jobs could be vacant. If a worker is unmatched, she produces u_i, which could be thought of as a combination of unemployment insurance and home production. Similarly, a vacant job produces v_j.

Shapley and Shubik (1971) show that a stable equilibrium (core allocation) includes an allocation M, and a transfer w_{ij} for each $(i, j) \in M$ which satisfies

$$\begin{aligned} \forall i' \in \mathcal{I}, j' \in \mathcal{J} \quad &: \quad \omega_{i'} + \pi_{j'} \geq \tau_{i'j'} \\ \forall i' \in \mathcal{I} \quad &: \quad \omega_{i'} \geq u_{i'} \\ \forall j' \in \mathcal{J} \quad &: \quad \pi_{j'} \geq v_{j'} \end{aligned} \tag{1}$$

where $\omega_{i'} = \omega_{i',m(i')}$ if worker is matched and $\omega_{i'} = u_{i'}$ otherwise, and similarly $\pi_{j'} = \pi_{m^{-1}(j'),j'}$ or $v_{j'}$.

The first condition says that there is no single worker-employer combination that could deviate from their current allocation, produce together, and split the surplus in such a way that both the employer and the worker would be better off. Note that this condition includes all possible combinations, including those that are not matched in equilibrium. The second and third conditions are participation constraints which require that every worker and employer obtain no less than their unemployment or vacancy value.[7] Shapley and Shubik (1971) shows that a stable allocation M^* is also optimal in the sense that the maximum total value is produced.

Workers' compensation in this model depends not only on the value they produce in their workplace, but also on the value they produce in other jobs. Compensation is strictly bounded by the worker and the employer's marginal contributions to the entire market (Roth and Sotomayor, 1992). Because this marginal contribution to the market is weakly smaller than the productivity at the workplace, workers are paid below their full productivity. Workers who are able to produce a similar value in more places will get a larger portion of their productivity to keep the equilibrium stable.

2.3 Deriving an Index for Outside Options

We next examine the role of outside options in this equilibrium. In particular, we derive the outside options index (OOI), a sufficient statistic for the impact of outside options on

[7] Unemployment and vacancies can exist simultaneously, as long as $u_i + v_j \geq \tau_{ij}$ for every possible match of non-participants.

wages.

In any stable equilibrium, each worker must earn more in her current match than she could earn at a different employer.

$$\omega_{ij} \geq \max_{j' \neq j} \omega_{ij'} \tag{2}$$

This outside option $\omega_{ij'}$ is exactly what will make employer j' indifferent between their equilibrium match, and hiring i (formally, $\tau_{ij'} - \omega_{ij'} \geq \pi_{j'}$). Hence

$$\omega_{ij'} = \underbrace{\tau_{ij'}}_{\substack{\text{potential} \\ \text{value } i, j'}} - \underbrace{\pi_{j'}}_{\substack{j' \text{ equilibrium} \\ \text{compensation}}} \tag{3}$$

Combined we get a lower-bound for worker compensation[8]

$$\omega_{ij} \geq \max_{j' \neq j} \tau_{ij'} - \pi_{j'} \tag{4}$$

The employer decision can thus be written as the solution to a simple profit maximization problem.[9]

From these equations, we can derive an expression for worker's compensation that we can take to the data. First, define $\mathcal{X} \subseteq \mathbb{R}^{d_x}, \mathcal{Z} \subseteq \mathbb{R}^{d_z}$ to be the characteristic spaces of workers and jobs accordingly. Let X_i and Z_j denote the observed worker and job characteristics which are distributed with a density $f(X_i), f(Z_j)$ respectively.[10] We next add an assumption on the distribution of τ_{ij} based on these observables. We follow Dupuy and Galichon (2014) and assume that the value of τ_{ij} conditional on the observables is drawn from a sum of two continuous logit models, one for the workers and one for the employers. This is a generalization of the classic multinomial logit for a continuous case.

Assumption 1. *The match value τ_{ij} between a worker with observable characteristics x_i, and a job with observable characteristics z_j, can be written as*

$$\tau_{ij} = \tau(x_i, z_j) + \varepsilon_{ij}$$

[8]Equilibrium compensation ω_i must satisfy $\omega_i + \pi_{j'} \geq \tau_{ij'}$, yielding this equation. This bound will be tight as long as $\max_{j' \neq j} \tau_{ij'} - \pi_{j'} \geq u_i$. It holds with equality under an additional assumption (Assumption 1).

[9]Equation 2 defines the effective price that employer j needs to pay to hire worker i. In order to maximize profit, the employer needs to choose a worker that maximizes the value net of cost: $\max_{i'} \pi_{i'j} = \max_{i'} \tau_{i'j} - \omega_{i'j}$.

[10]Formally, there are measurable functions $X_i : \mathcal{I} \to \mathcal{X}$, $Z_j : \mathcal{J} \to \mathcal{Z}$.

where ε_{ij} has the following distribution

$$\begin{aligned} & \varepsilon_{ij} = \varepsilon_{i,z_j} + \varepsilon_{j,x_i} \\ s.t. \quad & \varepsilon_{i,z_j} \perp \varepsilon_{j,x_i} \\ & \varepsilon_{i,z_j}, \varepsilon_{j,x_i} \sim CL(\alpha) \end{aligned}$$

$CL(\alpha)$ is the continuous logit distribution, that closely resembles an extremum value type-1 distribution with scale α (Dagsvik, 1994). For details on this distribution, see Appendix B.1.[11]

This assumption simplifies the math considerably. However, it is strong; it implies that workers have an unobserved utility or productivity in jobs with specific observed characteristics, and those unobserved shocks are uncorrelated, even between jobs with similar characteristics. Employers also have similar unobserved independent shocks based on the workers observables. Moreover, the assumption that $\varepsilon_{i,z_j} \perp \varepsilon_{j,x_i}$ implies that there are no interactions between the worker and job unobserved characteristics.

We can rewrite the latent value of outside options from Equation 3 as

$$\omega_{ij'} = \tau(x_i, z_{j'}) - \pi_{j'} + \varepsilon_{ij'} \tag{5}$$

Using $*$ to denote the best alternative offer ($\omega^*_{ij'} = \tau^*(x_i, z_{j'}) - \pi^*_{j'} + \varepsilon^*_{ij'} = \max_{j'} \omega_{ij'}$), we get a simple expression for the expected value of the best alternative offer:

$$\underbrace{E\left[\omega^*_{ij'}\right]}_{\text{Best Alternative Offer}} = \underbrace{E\left[\tau^*(x_i, z_{j'})\right]}_{\text{Mean Value}} - \underbrace{E\left[\pi_{j'}\right]}_{\text{Employer Rents}} + \underbrace{E\left[\varepsilon^*_{ij'}\right]}_{2\alpha \cdot OOI} \tag{6}$$

This decomposition is the key result of our theoretical analysis.

The first component reflects the mean value the worker can produce where they typically work (without strategic sorting on ε[12]). Therefore, as in almost all labor models, workers that produce a higher value would earn a larger compensation. The second component reflects the mean employer profit, beyond costs. This could be zero, or constant if we think the employer market clears perfectly through entry, but we do not assume this is necessarily the case. This component is affected by several factors including the firm productivity, and the market price of their workers.

In this paper, we focus on the third component $E\left[\varepsilon^*_{ij'}\right]$. This expression depends on

[11] Formally, every worker i draws ε_{i,z_j} shocks from a Poisson process on $\mathcal{Z} \times \mathbb{R}$. As a result, for every subset $S \subseteq \mathcal{Z}$, $\max_{z \in S} \{\varepsilon_{i,z}\} \sim EV_1(\alpha \log P(S) + const, \alpha)$ where EV_1 is extremum-value type-1 distribution, and $P(S) = \int_S f(z)\, dz$. A similar process exists for ε_{j,x_i}. More details in Appendix B.1.

[12] $\tau(x_i, z_j)$ reflects the expected value a worker can produce in a random job with characteristics z_j.

the measure of relevant options the worker has. Even though for a random match, $E\left[\varepsilon_{ij}\right]$ is constant, the expectation of $\varepsilon^*_{ij'}$ is higher because j' is positively selected, since it is the second best option. The more similar options in expectation worker i will have, the larger this component will be.

We derive the Outside Option Index, OOI, directly from this expectation. Specifically, we define the OOI to be the standardized expectation $\frac{1}{2\alpha}E\left[\varepsilon^*_{ij'}\right]$, where $\alpha \geq 0$ is the scale parameter that depends on the distribution of ε_{ij}. This descaling guarantees that the value of OOI is independent of the standard deviation of ε, and therefore of the units in which we define $\tau\left(x,z\right)$. α also sets the link between the OOI and wages, which we'll estimate using two distinct quasi-random sources of variation in options in Section 6. We assume that α is constant across all workers, implying a constant elasticity between OOI and wages. Our results on the heterogeneous impact of options on wages in Section 6 are consistent with this assumption.

The standard result that workers earn what they produce is a particular case of this setting. This occurs when $\alpha = 0$ (ε_{ij} is constant at 0) and entry decision of employers are optimal, such that profit is zero. This emphasizes the key distinction of this more general setting from the perfectly competitive model: heterogeneity. When $\alpha > 0$, there is no identical employer to bid wages up to the worker's full product.

Under Assumption 1, workers and employers are indifferent between matches with the same characteristics. Formally, defining f^i_j the probability density of worker i to work at job j in equilibrium we get the following lemma:

Lemma 1. *Under Assumption 1, the probability density of worker i to work at job j satisfies*

$$f^i_j = \frac{f\left(X_i, Z_j\right)}{f\left(X_i\right)f\left(Z_j\right)}$$

where $f\left(X_i, Z_j\right)$ is the joint density of matched worker and job observables in equilibrium.

Intuitively, this lemma implies that f^i_j is equal for all jobs with the same characteristics. Appendix B.2 provides a full proof, as well as formal definitions for those densities.

This assumption also yields a closed-form expression for the outside options defined in Equation 5.

Lemma 2. *Under Assumptions 1, in equilibrium, worker i with characteristics x_i is facing a continuous logit choice between employers who are offering*

$$\max_{z_j} \omega\left(x_i, z_j\right) + \varepsilon_{i,z_j}$$

and

$$\omega\left(x_i, z_j\right) = \tau\left(x_i, z_j\right) - \pi\left(z_j\right) - \alpha \log f^i_{j'}$$

where $\pi\left(z_j\right) = E\left[\pi_{j'}|Z_j = z_j\right]$. *Similarly, employers choose between*

$$\max_{x_i} \pi\left(x_i, z_j\right) + \varepsilon_{j,x_i}$$

This lemma simplifies the matching procedure into two one-sided continuous logit choices. Because employers with the same characteristic z_j are willing to make the same offer, the best alternative offer ω^* equals the maximal offer the equilibrium employer is willing to make. Hence, the lower bound from Equation 4 can be replaced with an equality. The workers are facing a choice between the employers who are looking for workers with their observed characteristics x_i.

The market clears when the supply of workers with characteristics x_0 to jobs with characteristics z_0 equals demand. Demand is decreasing with quantity, because the marginal employer has a lower value of ε_{j,x_0}. This is why the compensation $\omega_{ij'}$ in lemma 2 depends negatively on $\alpha \log f^i_j$. However, supply increases with compensation, hence f^i_j will be increasing in $\tau\left(x_i, z_z\right) - \pi\left(z_j\right) - \alpha \log f^i_j$. Those two equalize exactly when

$$f^i_j \propto \exp \frac{1}{2\alpha}\left[\tau\left(x_i, z_j\right) - \pi\left(z_j\right)\right]$$

This result implies that we can learn about the quality of outside options based on how similar workers sort into different jobs. Workers tend to sort into jobs where their net productivity is highest, where net productivity is the difference between the mean value they can produce $\tau\left(x_i, z_j\right)$, minus the employer's expected profits $\pi\left(z_j\right)$. Therefore, the only valuable outside options for a worker are those that are taken, in equilibrium, by similar workers.

With this distributional assumption we can find an analytical expression for the OOI. The expected value of ε^*_{ij} simplifies with the following lemma

Lemma 3. *Under Assumption 1:*

$$E\left[\varepsilon^*_{ij}\right] = E\left[\varepsilon^*_{i,z_j} + \varepsilon^*_{j,x_i}\right] = -2\alpha \int f^i_j \log f^i_j \tag{7}$$

The last equality follows because both $\varepsilon_{i,z}, \varepsilon_{j,x}$ are drawn from a continuous logit with scale parameter α. To measure the OOI for worker i, we need to take the integral over all their potential matches. Using the definition in Equation 6 yields

$$OOI = \frac{1}{2\alpha} E\left[\varepsilon_{ij}^{*}\right] = -\int_{j} f_{j}^{i} \log f_{j}^{i} \tag{8}$$

This expression is the well-known entropy index. Entropy is frequently used to measure industry concentration. Analogously, the OOI can be thought as a concentration index across jobs. A worker with more options (a worker who is less concentrated) will have a higher OOI because their probability of being in a specific job is lower. This is concentration on all (observable) dimensions: location, occupation, industry etc. Empirically, we will estimate it based on the concentration of workers with similar observables. If similar workers tend to be concentrated in a specific region of the country, small number of occupations or industries we will estimate a lower OOI for them. We will describe this procedure in detail in Section 4.

The entropy index is also commonly used in measuring unpredictability. In our context, this would be the difficulty to predict the worker's job. Workers whose jobs are harder to predict, are those with more options.[13] The OOI takes values on $(-\infty, 0]$. As the measure of jobs a worker can take approaches zero, $OOI \to -\infty$; if a worker is equally likely to take any job, $OOI = 0$.

This OOI is driven by two factors. First is worker flexibility, the ability of the worker to take jobs at different locations, use their skills in different occupations, industries etc. All of which we will measure empirically. Second is the supply of relevant jobs. More relevant jobs that the worker can take will increase the OOI directly. The OOI is only driven by relevant outside options – jobs that similar workers are actually observed taking in equilibrium ($f_j^i > 0$). Empirically, this will be options that are actually sometimes executed by workers with similar observables. Therefore, jobs that a worker could do but never would do in practice will not enter the OOI, and won't affect the equilibrium outcome.

The key advantage of using the OOI is that it does not depend on any information on a worker's alternative wages. This is useful because information on potential wages at other jobs is typically unavailable. Moreover, a worker's alternative wages also depend directly on their productivity. The OOI captures the impact of options on wages, holding productivity constant. This also implies that any link that we find between the OOI and wages is not mechanical.

[13] It is possible that it's easier to predict the job of certain workers due to better data quality. This would imply that those workers will have a lower scaling parameter α, and therefore a lower elasticity between the OOI and wages. To test this, we estimate the OOI-wage elasticity in Section 6 separately by gender and education. Our results are consistent with a constant value of α for all workers.

2.4 Sufficient Statistic

The OOI is a sufficient statistic for the effect of access to more options on workers compensation, under our model assumptions. Access to options has two distinct effects on workers, both of which are captured in the OOI. First, it improves workers compensation at the same job by improving their outside options. Second, the improvement in options allows some workers to find better matches.

We first define an improvement in access to options. We define λ_x to be the measure of a random set of jobs that are accessible to workers with observables x. All jobs that are not accessible have $\tau_{ij} = -\infty$ and are therefore never chosen in equilibrium. We model an increase in access to more jobs would be an increase to this λ_x. In Appendix B.2 we show that other definition of λ such as a linear commuting cost would yield the same results.

Theorem 1 shows that workers who get access to more outside options get an increased wage offer from their employer that equals to α times the change in their OOI.

Theorem 1. *Let j be $i's$ equilibrium match. Access to outside options λ_{x_i} has the following effect on the maximum offer j is willing to make in the new equilibrium:*

$$\frac{d\omega_{i,j}}{d\lambda_{x_i}} = \alpha \frac{dOOI}{d\lambda_{x_i}}$$

The second effect of access to more options is an improvement in match quality. An improvement in outside options is only an improvement, if some workers would in practice match into those additional jobs in equilibrium. Therefore, the overall effect of access to more options is a combination of the better outside options, and the option to improve match quality. The following theorem shows that in this model, the overall effect is exactly twice the size of the effect only through outside options.

Theorem 2. *Access to options λ_{x_i} has the following overall effect on expected worker compensation in equilibrium*

$$\frac{dE\left[\omega_{i,j}\right]}{d\lambda_{x_i}} = 2\alpha \frac{dOOI}{d\lambda_{x_i}}$$

Different choices of counterfactuals could potentially lead to different results. The counterfactual we consider is giving a small group of workers access to more similar jobs. If the increase in λ affects a non-zero measure of workers, then there will be general equilibrium impacts on employer profits . For instance, mandating stable working hours in all jobs will give all women access to more jobs. This counterfactual may decrease the profits of employers that were already hiring mostly women. In such cases, the OOI would only be a sufficient statistic if the employers' market is perfectly competitive such that profit are kept constant through entry and exit. Access to better jobs (as opposed to similar jobs), in which the worker can produce greater value will also affect workers productivity, and therefore will affect compensation beyond the effect on the OOI.

2.5 Parametric Example

To give further intuition for the OOI, and the additional components in our key decomposition (Equation 6) we go over a simple parametric example.

In this simple setting, workers are characterized only by their productivity and their amount of options. Assume workers and jobs are equally dispersed across the real line $\mathbb{R}$.[14] Each worker can be described as a 3-dimensional tuple (l_i, y_i, d_i) which is her location on the real line, her productivity and the maximal distance she is able to commute. Jobs are identical other than their location l_j. The value of a match is then

$$\tau_{ij} = \begin{cases} y_i + \varepsilon_{ij} & |l_i - l_j| < d_i \\ -\infty & else \end{cases}$$

where ε_{ij} are the sum of two continuous logit distribution as before.

In this simple setting, the OOI corresponds to the log measure of options. The PDF of a worker distribution across jobs is constant at $\frac{1}{2d_i}$ for all jobs within feasible range. Therefore, the OOI is $-\log \frac{1}{2d_i} = \log 2d_i$ which is the log of the measure of jobs a worker can take. Differences in OOI are therefore the log ratio in the measure of relevant options. This result will generally hold for every pair of workers with similar distribution of jobs and different sizes of support, not only in this example. In this setting, λ is exactly $2d$, hence from Theorems 1 and 2, an infinitesimal increase in d_i leads to an increase of $\frac{\alpha}{2d_i}$ if they stay at the same job, and $\frac{\alpha}{d_i}$ overall.

The first component of Equation 6 (mean value) captures a worker's baseline productivity; in this case this is equal to y_i. This component represents the expected productivity in a random job that a worker could take. Equivalently, it captures productivity differences, conditional on having the same amount of options (OOI). The final component, employer rents, will be equal for all workers, as all jobs are equivalent.[15]

This example shows clearly how two workers who are on average equally productive, could still earn different wages due to differences in outside options. Assume $l_1 = l_2$, $y_1 = y_2$, and $d_1 < d_2$. Worker 2 earns a higher wage because her OOI is greater. In

[14] Formally, assume each interval $[a, b]$ has a measure of $b - a$ workers and jobs. This implies an infinite measure of both workers and jobs.

[15] Its exact value would be pinned down depending on the value of unemployment, and vacant jobs.

expectation, workers 1 and 2 are equally productive at every job in $[l - d_1, l + d_1]$. Since worker 2 has a higher price, most jobs in this range would prefer to hire worker 1. Still, as a result of heterogeneity, some employers would be willing to pay the higher price. Because worker 2 has more options than worker 1, there are enough employers who are willing to pay the higher price, so that the market clears.

2.6 Discussion

We summarize this section by re-examining the model assumptions and their implication on what is and what is not captured with the OOI. The primary advantage of the OOI is that it more precisely captures the size of a worker's relevant option set. It allows workers to use their skills in different occupations and industries. By contrast, measures such as the HHI assume that workers belong to only one industry or occupation. Similarly, the OOI accounts for heterogeneity in commuting and moving costs. Instead of assuming each worker is assigned to a specific local labor market, the OOI empirically assess the distance over which each type of worker searches for a job. Finally, the OOI accounts for variation in employer characteristics even within the same industry. For instance, if some workers are unable to work on weekends, their OOI will only be affected by employers who do not require that.

The main limitation of the OOI is that it does not account for any dynamic considerations. This is because it was derived from a static model. Dynamic considerations such as switching costs, firm-specific human capital that is acquired over time, and learning tend to limit a worker's ability to move to their outside options, but are beyond the scope of this analysis.

A second limitation is that the OOI calculates the measure of relevant jobs, not relevant employers. We assumed that employers are 1-job firms and do not account for the fact that many jobs are under the same employer. While the model will, with minor adjustments, accommodate firms, we focused on jobs due to limitations of our data (see Section 3.1). Therefore, the OOI will over-estimate options for workers who are more likely to work in large firms.

In contrast, some aspects of the labor market that are not explicitly modeled above could still be captured in the OOI. The most prominent one is information frictions that would generate search costs. Black (1995) has analyzed a search model where some workers have more options, and showed that in this setting as well, more options would lead to higher wages in equilibrium. Hence, it is possible that some of the effect of the OOI on wages is operating through this channel as well.

3 Empirical Setting and Data

We use administrative data from Germany to generate measures of individual workers' outside options. The data includes detailed information on establishment and worker characteristics, including information on a variety of amenities provided by different establishments, which allow us to estimate workers' options more accurately. Excluding some idiosyncratic features which we will now discuss, the German labor market is comparable to other low-regulated labor markets, making wages more directly affected by the market forces we want to study.

3.1 Data

Administrative German Employer-Employee Data

Our primary source of data is a panel of German worker employment histories known as the "LIAB Longitudinal" dataset. It is a matched employer-employee administrative data, based on a sample from the universe of German Social Security records from 1993-2014. There are four key features of the data which make it ideal for our setting. First, it is a large dataset, including about 1% representative sample of the entire German labor force. Second, there is detailed establishment-level survey information with information such as hours requirements, profitability, leave/maternity policies etc. This allows us to account for differences in outside options that may be due to differences between establishments, even within industries. Third, the panel structure of the data, allows us to track workers over long periods of time. This gives us valuable information about the workers such as their specific experience in the market, and their location before taking their job. Fourth, this data provides 4-digit occupational classification which highly improves our precision in measuring relevant options. To our knowledge, this combination of data is not available in the United States.

The data come from the Integrated Employment Biographies (IEB) dataset, which is collected by the German Institute for Employment Research (IAB). Employers are required to report daily earnings (subject to a censoring limit at the maximum taxable earning level)[16], education, occupation, and demographics for each of their employees at least once per year, and at the beginning of any new employment spell. New spells can arise due to changes in job status (e.g. part-time to full-time), establishment, or occupation.

Each year the IAB selects a stratified random sample of establishments from the pool of all German establishments with at least one employee liable to Social Security. These

[16] 11% of the sample is censored. As we do not use wages to calculate OOI, it is not affected by censoring.

establishments are required to complete a series of surveys on organizational structure, personnel policies, financing, and research activities. In particular, the establishments are asked for information on their annual sales, profits, establishment size and leave policies. The survey data are then merged with the complete employment histories of all individuals who worked at least one day in any of these firms between 1993 and 2014.

There are several limitations for the data, that may affect our calculations of outside options. The data do not cover civil servants or the self-employed, which comprise 18% of the German workforce. They also do not cover labor force non-participants. Therefore, we do not account for any of those options when calculating the OOI. Since the sample is done at the establishment level, we usually observe only few establishments in each industry-region combination. This is why we construct the OOI at the job level, and not the employer level.

Because our model is static, we rely on repeated cross-sections of data. For each year, we use data on employment relations on June 30th of each year. Our descriptive analysis is done for our last year in the sample, that is June 30th 2014. We use data from 1999, 2004 and 2012 in Section 6 to examine how quasi-random variations in the OOI effect wages.

BIBB Task Data

We supplement these data with survey information on the characteristics of occupations and industries. It includes information on the tasks completed, hours requirements and typical working conditions in these occupations/industries. These data are similar to the O*NET series, but allow us to account for possible differences in the task content of occupations between the United States and Germany, as well as differences in coding.[17] The survey is conducted by the IAB and includes information on respondents' occupation, industry, in addition to responses on questions related to organizational information, job tasks, job skill requirements, health and working conditions.

3.2 Empirical Setting: German Labor Market

There are several distinctive features of the German labor market which are relevant for our analysis. First, there are different levels of secondary-school leaving certificates, which depend on the number of years and type of education. Our data allows us to distinguish between three categories: lower-secondary, which typically requires nine years of schooling, intermediate-secondary, which typically requires ten years of schooling, and

[17] These data have been used in prior publications on the German task structure including Gathmann and Schönberg (2010).

high-secondary, which requires twelve to thirteen years of schooling, and allows the student to pursue a university degree. In our analysis we use indicators for the type of secondary education to account for years of schooling, and school quality.

Second, in addition to (or sometimes instead of) formal education, many German workers receive on-the-job training through formal apprenticeships. Individuals in apprenticeship programs complete a prescribed curriculum and obtain occupation-specific certifications (e.g. piano maker). We use this information to precisely identify the types of jobs a worker could perform.

Third, eleven percent of workers in Germany work under "fixed-term contracts" (as of 2014). These contracts expire automatically without dismissal at the end of the agreed term, at no cost to the employer. The maximal period for employment under such contract varies between 6 to 18 months over the period for which we have data. At the end of a contract, the worker and employer may choose to continue the employment relationship, but cannot use another fixed term contract to do so (Hagen, 2003).

Fourth, two percent of workers are hired through temporary work agencies. This is a triangular employment relationship, which involves the temporary work agency, a client company and a temporary worker. Historically these working relations were limited to 24 months; their duration is no longer regulated. There are additional regulations on the pay received by workers hired through temporary agencies (in particular relating to how these workers are paid relative to other workers at the same firms) but the rules vary significantly over time. In our analysis, we distinguish between employment found via temporary work agency and work found via more traditional means (Mitlacher, 2008).

While wage setting in Germany was historically governed by strong collective bargaining agreements, employers today have considerable latitude in setting pay (Dustmann et al., 2009). While employers could always raise wages above the agreed-upon levels, it only became common for contracts to include "opening clauses" allowing employers to negotiate directly with workers to pay below-CBA wages in the 1990s. Today these clauses are very common.

3.3 Summary Statistics

Table 1 describe the characteristics of workers and jobs in our sample, for the full sample, as well as by gender.

Our sample is roughly evenly split between male and female workers. The mean age for a worker in our sample is forty-five years old and the vast majority (97%) are citizens. The workers are divided about equally between the three types of secondary education. In nineteen percent of the sample the lower- and intermediate- secondary education categories are aggregated.[18] Men and women have similar age, education and citizenship status.

On the job side, thirty one percent of the jobs in our sample are part-time. Eleven percent of jobs are on fixed contracts and only two percent are from temporary agencies. The distribution of establishment size is very skewed, with mean of 1,552 workers and a standard deviation of five times that size. The mean annual sales per worker are 163,000 Euros. Twenty-six percent of the establishments report to have females in managerial positions.

It can already be observed that men and women sort into different types of jobs. Females are much more likely to work in part-time jobs (53% compared to 13%), which is relatively high compared to other countries.[19] They also work at smaller establishments with 827 employees on average and mean annual sales of 130,000 Euros, compared to 2,166 and 191,000 accordingly for males. Females are also more concentrated at establishments with higher share of female-management (36% compared to 17%).

4 Estimating Outside Options

In this section we describe how we estimate the outside options index. Our method uses the cross-sectional allocation of observably similar workers to estimate the relevant options of each worker . This allocation teaches us about the worker's ability or willingness to commute, about the set of industries or occupations that are suitable to the worker's skills, and about the worker's demand for certain workplace amenities. Section 4.1 states the key assumption, Section 4.2 describes the estimation procedure, and Section 4.3 describes the worker and job characteristic we use as inputs.

The OOI of a given worker requires an estimate of their probability to work in each one of the jobs observed in the data. We calculate the OOI using Equation 8 which shows that the OOI is only a function of the different f_j^i .This requires us to estimate N^2 distinct probabilities.

Earlier methods that were developed to estimate such densities do not work on data sets of our size. Non-parametric approaches cannot be used due to the large number of worker and firm characteristics. Parametric methods that were designed specifically for this model, work well when the number of possible combinations is around a few mil-

[18] More details in data Appendix C.1

[19] Germany is ranked 6th out of 35 OECD countries in female part-time employment (OECD, 2018).

lions.[20] However, given the size of our data, we need to calculate the probability density of approximately 250 billion possible combinations, making these methods computationally not feasible.

To overcome this challenge, we develop a new method that is computationally feasible for large data sets, and uses a similar set of assumptions to those used in the prior literature (Dupuy and Galichon, 2014). Our method relies on an equivalent representation of the probability densities as the ratio between the likelihood of a matched pair to appear in the equilibrium allocation compared to a random one. These ratios can be estimated quickly using logistic regressions. We discuss the links and differences between our method and prior methods in more detail in Appendix B.3.

4.1 Assumptions

In this section we state the parametric assumptions we make to link the f_j^i densities to the data. Our data are comprised of pairs of matches between workers, and jobs (x_k, z_k), where the x_k/z_k are observed worker/job characteristics we discuss in the Section 4.3.

We first use the result of Lemma 1

$$f_j^i = \frac{f\left(X_i, Z_j\right)}{f\left(X_i\right) f\left(Z_j\right)}$$

$f\left(X_i, Z_j\right)$ is the probability of observing a match between a worker with characteristics X_i and a job with characteristics Z_j. $f(X_i)f(Z_j)$ is the product of two the marginal distributions for workers and job characteristics. This is the probability of observing a match with such observables, under a random assignment. The basic intuition for this result is that the probability of observing i matched with j depends on the frequency that workers and jobs with such observables are matched, accounting for the total measure of workers and jobs with these observables (if there are more jobs with a particular set of observables, the probability to match to a specific one is smaller). This result can be derived from weaker assumptions as well.[21]

Our second assumption parametrizes f_j^i as a function of the observables. We follow Dupuy and Galichon (2014) in assuming that the log density is linear in the interaction of worker and job characteristics.

[20] Choo and Siow (2006) develop a non-parametric method where the number of possible combinations is finite and small. Dupuy and Galichon (2014) use Iterative-Proportional-Fitting algorithm to estimate a continuous density. Their data set had about $N^2 = 10^6$ possible combinations.

[21] It is sufficient to assume If $X_i = X_{i'}$ and $Z_j = Z_{j'}$ then $f_j^i = f_{j'}^{i'}$, instead of Assumtion 1.

Assumption 2. *The log of the probability density is linear in the interaction of every worker and job characteristic:*

$$\log f_j^i = X_i A Z_j + a\left(X_i\right) + b\left(Z_j\right)$$

The matrix A includes all the coefficients on each of the interactions between worker and job characteristics. The marginal distributions $f\left(x\right), f\left(z\right)$ are fully determined by $a\left(x\right)$ and $b\left(z\right)$.[22]

This assumption reduces the dimension of the problem significantly, while allowing the relationship between each pair of covariates to remain unrestricted. Dupuy and Galichon (2014) show that A is proportional to the cross-derivative of τ

$$2\alpha A = \frac{\partial^2 \tau}{\partial x \partial z} \tag{9}$$

where α is the scale parameter of ε we defined in Assumption 1. Intuitively, this means that if a worker characteristic and a job characteristic are complements, they will be observed more frequently in the data.

4.2 Empirical Procedure

Under these two assumptions, we can estimate the OOI using a simple procedure that we will now describe. The key idea of our method is to use the result of Lemma 1, that the probability density f_j^i can be written as the ratio between the probability of observing a match in the real distribution to its probability under a random assignment.

We start by expanding our data set of worker and job matches. We simulate data from a distribution $\widetilde{f}\left(x,z\right) = f\left(x\right) \cdot f\left(z\right)$, where x and z are independent. This is done by randomly sampling an observed worker and an observed job independently. We simulate a total number of random matches equal to our original data size, such that the share of real and simulated data is exactly one half. We define a binary variable Y that equals to one whenever the match is 'real' (taken from the data) and zero whenever it is simulated.

We then estimate all our parameters using a logistic regression. We regress the binary variable we constructed Y_k on the matched worker and job characteristics $\left(X_k, Z_k\right)$. Note that, as a result of Lemma 1, and a simple Bayes rule, the match probability density f_j^i is proportional to the ratio of observing this match in the real or simulated data, conditional

[22]While Dupuy Galichon are able to fit the marginal distribution precisely to their observed value in the data, we won't be able to this with our data size. Therefore, we take linear functions of all X variables and Z variables. We also include indicators for district. As we discuss in Appendix B.3 this specification fits the first moments of the marginal distributions.

on the observed worker and job characteristics.

$$\frac{P\left(Y_k=1|x_k,z_k\right)}{P\left(Y_k=0|x_k,z_k\right)}=\frac{f\left(x_k,z_k\right)}{f\left(x_k\right)f\left(z_k\right)}\frac{P\left(Y_k=1\right)}{P\left(Y_k=0\right)}=f_j^i\cdot const$$

Combining this result with Assumption 2 yields

$$\log\frac{P\left(Y_k=1|x_k,z_k\right)}{P\left(Y_k=0|x_k,z_k\right)}=x_kAz_k+a\left(x_k\right)+b\left(z_k\right) \tag{10}$$

We can estimate this equation using a logistic regression where we approximate $a\left(x\right),b\left(z\right)$ with linear functions. Under the assumptions this produces consistent estimates for $\widehat{A},\widehat{a}\left(x_k\right),\widehat{b}\left(z_k\right)$. We discuss the intermediate results from this estimation procedure, in Section 7.2 where we analyze the underlying reasons for differences in the OOI.

We use the estimates from the logistic regression to estimate the probability density of every potential match. Specifically, we estimate the probability density of worker i to work in job j to be

$$\widehat{f_j^i}=\exp\left[x_i\widehat{A}z_j+\widehat{a}\left(x_i\right)+\widehat{b}\left(z_j\right)\right]$$

We calculate this value for all possible worker-job combination in our data set.[23] This simple functional form allows us to make this calculation directly and with minimum computational burden.

With these result in hand we can calculate the outside options index for every worker in our sample using Equation 8:

$$\widehat{OOI}_i=-\sum_j\widehat{f_j^i}\log\widehat{f_j^i} \tag{11}$$

This yields a consistent estimate of the OOI, if both assumptions are correctly specified.

We verify that the OOI is robust to different choices of functional form. Instead of estimating it using the entropy index, we use the same probabilities we estimated in an HHI formula: $-\sum_j\widehat{f_j^i}^2$. We find that the results are very similar. The correlation between the two indices is .62.

In Appendix B.3 we discuss the properties of this method, in the case where these assumptions do not hold. We show this method can be written as a GMM estimator, and discuss the moments that are being matched. We also show that if we increase the size of the simulated data, and fully saturate the functions a and b, our method becomes equivalent to Dupuy and Galichon (2014).

[23]We normalize those estimated densities, such that $\sum_j\widehat{f_j^i}=1$. So effectively, we don't use $\widehat{a}\left(x_i\right)$.

4.3 OOI Input: Job and Worker Characteristics

To estimate f_j^i using Equation 10 we include three groups of variables: worker characteristics (x), job characteristics (z), and the geographical distance between workers and jobs. Information on wages is intentionally not used in any of these groups, to avoid a mechanical link between the OOI and wages.

Worker Characteristics x We use x to denote the variables that describe worker demographics and worker training. The demographic variables include workers' gender, worker's level of secondary education, an indicator for whether the worker is a citizen, and a quadratic in age. For training we use the occupation in which they undertook their apprenticeship. If we do not have information on a worker's apprenticeship (e.g. if it occurred before our data begin in 1993), or if a worker did not complete an apprenticeship, we use their first occupation observed in the data, as long as this is at least ten years old.

Job Characteristics z The job characteristics z variables fall into three categories: (1) characteristics of establishments, (2) characteristics of employment contracts, and (3) characteristics of jobs. First, we take several establishment-specific variables directly from the establishment survey: size, sales and the share of females in management . We also use the first two principal components of each of the six categories of the establishment survey: business performance, investments, working hours, firm training, vocational training, and a general category. Appendix Table A1 shows the most weighted questions in each category.

Second, we use several variables which relate to the structure of the employment contract: whether the job is part-time, whether the contract is fixed term, and whether the position was filled by a temporary agency.

Finally, to describe the characteristic of the job we use information on the occupation and industry. Because it would not be feasible to include interactions between all of our industry and occupation codes, we use data from the BIBB to identify the characteristics associated with different industries and occupations. The BIBB survey contains modules on working hours, task type, requirements, physical conditions and mental conditions. For each 3-digit occupation and 2-digit industry, we include the first two principal components for each module. We use these to code both the occupation and industry that describe the job, and the training occupation that describes the worker. Appendix Table A2 shows the most weighted questions in each module. We also include occupation complexity, which codes occupations into four categories based on the type of activity they require: (1) simple , (2) technical (3) specialist and (4) complex.[24] We use a total of 18 worker characteristics and 39 job characteristics.

Geographical Distance We include the geographical distance between workers and employers. For workers we use their last place of residence before taking the job.[25] This distance could capture both the commuting, as well as the moving costs between places; empirically we cannot directly distinguish the two. Both locations are given at the district (kreis) level.

Figure 1 presents a map of the 402 districts in Germany. The size of the districts varies across the country and, importantly, it tends to be smaller in highly populated areas. In many cases, the major city is its own district, allowing us to separately identify the city center and the suburbs. Though not perfect, this coding allows us to get a reasonable approximation of commuting and moving patterns by workers. Appendix Table A3 shows the mean of the distance variable by gender and education groups. We find that the mean is 15.5 miles, but there is significant variation across groups.

We allow distance to have a non-linear effect on match probability that is different for each worker type. When we estimate Equation 10 we use a 4th degree polynomial of the distance between a worker's lagged home district and their location of work to account for the non-linear impact of distance.[26] To account for heterogeneity in willingness to commute or move, we interact the polynomials in distance with all worker characteristics x. This allows workers to be affected differently by distance, depending on their gender, education, age, citizenship and training. As we discuss in Section 7.2 this turns out to be the main driver of differences in outside options.

5 The Empirical Distribution of Outside Options

We next turn to describing the distribution of the OOI, and the characteristics of workers with better and worse options, as measured by it. We find that the OOI is higher for men, German citizens, city residents, more educated and more experienced workers. We also

[24]These four categories usually reflect the type of qualification needed to perform the job, which ranges between none, vocational training, some tertiary degree and higher education. For instance, different occupations in nursing that fall under the same occupational coding (813) will be coded with different complexity, ranging between a nursing assistant, nurse, specialist nurse and general practitioner.

[25]Workers current place of residence is affected by their match. Their place of residence before taking the job better reflects the actual radius over which people are searching for jobs.

[26]We find that for more than 100 miles, the effect of distance is constant. This is consistent with the idea that individuals do not commute more than 100 miles; to switch to a job that is much further away, they have to move. This moving cost may not vary significantly with distance.

find that higher skill workers tend to be more specialized in their current industry, which narrows down their outside options.

Figure 2 plots the raw distribution of the OOI for every worker in our data. The mean of the distribution is -4.85. We can interpret the mean by considering the share p of options a worker with this OOI would have if the probability density they worked at any given job was either $\frac{1}{p}$ or 0. A worker with an OOI of -4.85 would be found in a share $p = 0.8\%$ of jobs. The distribution is skewed, with a long left tail, indicating that there are many workers who are extremely concentrated. The standard deviation of the distribution is also quite sizable: .93. For comparison, duplicating the worker's option set by generating an additional identical job for every job option they have would increase the OOI by only .69.

We estimate the following regression to decompose the average OOI by worker characteristics

$$OOI = \beta_0 Female + \beta_1 Education + \beta_2 Citizen + \beta_3 Age + \beta_4 Age^2 + \epsilon \tag{12}$$

Figure 3 plots the results. With controls, the average OOI for women is .237 units below that of men. The average OOI for German citizen is higher by .217 units. Assuming similar distributions across jobs, this would imply that male (German citizens) have 27% (24%) more options than women (non-citizens).

Options are also better for higher-educated workers. Lower-secondary (intermediate secondary) school workers' options are on average .70 (.32) units lower than higher-secondary workers. This implies 101% (38%) more options assuming similar distribution across jobs.

Figure 4 shows that men have more options than women, not just on average, but across the entire distribution. The figure shows that the cumulative distribution function for men is shifted to the right. We cannot reject that the distribution for men stochastically dominates that of women.

We find an inverse U-shape relationship between the OOI and age. Figure 5 plots the mean OOI by age and shows that workers' options tend to improve with age before flattening off at age thirty. Older workers (over 50) see declining values of options. As we discuss in Section 2.6, the OOI does not capture any dynamic considerations that are particularly likely to have a differential effect across ages. Accounting for this could potentially change these results.

A large portion of the variation in options is driven by geographical variation in labor market size and density. The last category in Figure 3 shows a positive correlation between the district density and its OOI, controlling for other demographics. Figure 6 graphs the mean value of the OOI by German district. As the figure illustrates, workers in cities tend to have better options, as measured by the OOI. Workers near these cities, also appear to have better options. This result is robust for adding controls for worker demographics.

While most of our results indicate that high earning workers, such as workers with higher education or city residents, tend to have more options, this relationship is reversed at the occupation level. This is because high-skilled workers tend to have more specialized skills, which are valued by a smaller number of employers. Controlling for all other observables, workers who completed their training (apprenticeship) in higher earning occupations tend to have lower options, as measured by the OOI (raw correlation equals -.022). Figure 7 plots each training occupation by their (residualized) log wages and OOI.

We next look at which occupations drive the negative correlation. The upper-left corner of this figure, comprises high paying occupations with few relevant options: medical doctors, pilots, dentists. These are textbook examples of high-wage occupations with skills that cannot be easily transferred. While wages in these occupations are still high, our model predicts that, at least in partial equilibrium, if these workers were able to use their skills in more industries, their wages would have been even higher. The bottom right corner shows occupations like meter reader or car sales that have lower wages and more options. These are examples of low-wage occupations with highly general skills.

Table 2 presents coefficients from Equation 12, including additional controls for training occupation, district of residence and establishment. Controlling for training occupations (column 2) does not change the results significantly. However, adding controls for worker's district of residence as well (column 3) reduces some of the education gap, and increases the gap between German citizens to non-citizens. This suggests that higher educated, and non-citizen workers are more concentrated in large cities where there are more job options, and therefore their OOI is lower once controlling for that. Controlling for establishments (column 4) yields results that are similar to the results with controls for districts and occupation, as workers in the same establishment tend to live closely. However, we find smaller gender differences in options within establishments.

5.1 Mass Layoffs

We next show that the OOI is able to predict the ease with which workers recover from a job separation. These separations force workers to move to their outside option. To identify exogenous separations we follow the prior literature in focusing on mass-layoffs.

We start by constructing a sample of workers who were involved in mass-layoff between 1993 and 2014, following the approach in Jacobson et al. (1993). We define a plant in our sample as undergoing a mass layoff if it has a decline in its workforce of at least thirty percent over the year. We consider only mass layoffs that occur in establishments with at least fifty workers. We restrict our analysis to workers who had been employed at the establishment for at least three years prior to the mass layoff and who are below the age of 55. This leaves us with a final sample of 13,681 workers from 583 distinct mass-layoffs.

The outcome variable we use is relative income: the ratio between current daily income and the last daily income before the layoff. Formally we define relative income as $\widetilde{w}_t = \frac{w_t}{w_0}$, where t is months after the layoff, ranging from one to thirty-six. In case a worker is unemployed during this period their relative wage is set to zero. This choice of outcome variable takes out all productivity differences that can be captured with worker fixed effects and might be correlated with the OOI.

Appendix Figure A1 replicates the main result in Jacobson et al. (1993). Specifically, we look at

$$\widetilde{w}_{i,t} = \beta_t + \psi_{j(i)} \tag{13}$$

where $\psi_{j(i)}$ is establishment fixed effect. We plot β_t, the mean relative income of workers each month, for three years after the layoff. We find that on average workers lose 80% of their income in the month following the layoff. Their income gradually returns to its previous value over the next three years.

We next look at the differences in recovery for workers with different value of OOI. Within each establishment in our sample, we divide the laid-off workers into two groups, based on whether they are above or below the establishment median of the outside options index. Appendix Table A4 shows summary statistics for the two groups.

We then calculate

$$\widetilde{w}_{i,t} = \rho_t High_i + \delta_{j(i),t} \tag{14}$$

where $High_i$ indicates above the establishment median OOI, and $\delta_{j(i),t}$ is establishment by month fixed effects. Figure 8 plots ρ_t, the difference in relative income between those two groups for each month in the three years after the layoff. Our point estimates show that workers with better options, as captured by the OOI, gain an additional 8 percent of their previous income during the first year after the layoff. The groups seem to converge throughout time and after three years there are no differences.

The lower relative income is driven by both longer search time, and lower wage after search. To show this we recalculate Equation 14, replacing the outcome variable with $e_{i,t}$

an indicator for being employed, which we define as having a positive wage

$$e_{i,t} = \rho_t^e High_i + \delta_{j(i),t}^e \tag{15}$$

Figure 9 plots ρ_t^e, the differences in the share of employed workers on both groups. Our point estimate show that about 2% more people with higher OOI are working compared to the lower OOI. Therefore, the relative income in the new job must also be lower to explain an 8% difference between the groups.

We then repeat this analysis with a continuous measure of OOI, and a varying set of controls. We regress relative income at month t on the OOI, with fixed-effects for establishment by month $\delta_{j(i),t}$.

$$\widetilde{w}_{i,t} = \lambda_t OOI_i + X_{it} + \delta_{j(i),t} \tag{16}$$

We also repeat this analysis with additional worker controls X_{it} including tenure, gender, age and education. The results are reported in Table 3. We find virtually the same patterns we found when divided by the median, for all choices of controls.

Our findings suggest that the OOI may have an impact on additional dynamic aspects of the labor market that are not captured in our static-model. The OOI quantifies the number of options workers have. It is therefore not surprising that workers with more options (higher OOI) are able to find a new job more quickly. However, the effect on relative wage is less obvious, as the OOI is likely to already affect income before the layoff. One interpretation of our findings is that workers with higher OOI face a labor market that is closer to perfectly competitive, such that the impact of a single employer on equilibrium wages is negligible, and not affected by their exit.[27]

6 Effect on Wages

In this section we use two different sources of quasi-random variation in options to estimate the elasticity between the OOI and wages. The first (Section 6.1) focuses on the introduction of high-speed commuter rail stations in small German towns, and the second (Section 6.2) uses a shift-share ("Bartik") instrument. These methods yield semi-elasticities between .17-.32 .

Estimating this elasticity using instruments allows us to translate differences in OOI

[27] This intuition can be captured with the predicted effect of a plant closure on the OOI. Removing a job with low probability have only a small effect on the OOI. Since workers with higher OOI, have on average a lower probability to be in every job, their OOI is less affected from the destruction of only a few jobs. Therefore their wage will be less affected as well.

to differences in wages, even if the model is misspecified. In our model, the effect of the OOI on wages is determined by the parameter α (Theorems 1, 2). Our model only implies that this parameter is non-negative and that in a perfectly competitive labor market, this parameter is zero. Since we do not use any of the model assumptions to estimate this parameter, the results capture the elasticity between the OOI and wages, even if the model is misspecified. This exercise can also be seen as a test for whether the OOI has an impact on wages.

Identifying α, the relationship between options and wages is challenging for two reasons. First, the OOI estimator is a function of the observed characteristics X_i. These observables are likely to also capture differences in productivity, thus creating a problem of omitted variable bias. Second, the OOI is estimated with a potentially large amount of noise, which would create an attenuation bias, especially when adding several controls. Because of those reasons, the OLS estimates of this parameter (Appendix Table A5), strongly depend on the set of controls we use. Adding more controls shrinks the results towards zero, and can even affect the sign of the coefficient.[28] In order to cope with both issues we use two sources of quasi-random variation in workers' options that do not affect their productivity.

6.1 High-Speed Commuter Rail

We leverage the expansion of the high-speed rail network in Germany as an exogenous shock to workers' outside options, following prior work by Heuermann and Schmieder (2018).[29] High-speed trains were first introduced in Germany in 1991-1998. During that period, stations were placed in major cities. We focus on the second wave of expansion, which began in 1999. During this wave, new stations were added in cities along existing routes. Cities were chosen for new stations after the routes already existed, and mostly on the basis of political considerations, and not labor market factors. As a result, towns as small as 12,000 residents were connected to the train network. Heuermann and Schmieder (2018) show that this increase in infrastructure led to an increase in commuting probability. Figure 10 shows the map of districts that got stations in the two waves of installation.

There are several threats to identification that should be considered. One potential

[28]Our findings in Section 5 suggest that the bias could go both ways, therefore the sign of the OLS coefficient could be both positive and negative. The OOI is correlated with high experience, or high education, but also with lower wage occupations, potentially because high-skill occupations tend to be more specific.

[29]Daniel Heuermann and Johannes Schmieder generously provided the train data for our use in this project.

concern is that the cities which received new train stations were selected on the basis of expected increase in productivity. Institutional details and prior research suggest that this is unlikely.[30] Another concern is that the new infrastructure can also be used for the transportation of goods. This would impact the workers' wages through their productivity. However, Heuermann and Schmieder showed that the introduction of these stations had no effect on the product market, as the trains were only used to transport passengers. One remaining concern is that the trains also had a similar effect on employers in those small towns, by allowing them to recruit workers from major German cities. We do not have detailed enough data on the employers to reject this possibility.

We supplement our data with train schedules in the years 1999 and 2012, before and after the installment of the second-wave stations. From this data we construct an indicator variable for every match for whether the worker can use a direct high-speed line to get to this job. We add this variable for our estimation of the match probabilities f_j^i. Therefore, we allow the match probability to depend on whether there's a high-speed line connecting the worker and the employer. To allow for heterogeneity between workers in their demand for trains, we interact this variable with all worker characteristic X_i. We then use these probabilities to estimate the OOI, as explained in Section 4.2.

The treatment group consists of all workers who, in 1999, lived in districts that got high-speed connections during the second-wave expansion (1999-2012). The control group consists of all workers who, in 1999, lived in districts that never received stations. Since the major cities were connected in the first wave, they are effectively excluded in this analysis. We then follow the same workers to the year 2012, regardless of where they live. We match workers from the treatment and the control group based on their gender, age, citizenship, education level, training occupation, state (Bundesländer) and lagged income using nearest-neighbor matching with replacement.[31] Appendix Table A6 presents a balance table for this match.

We estimate the following system of equations:

$$\begin{aligned}\Delta_{1999}^{2012} \log w_{im} &= \alpha \Delta_{1999}^{2012} OOI_i + \mu_m + \upsilon_{im} \\ \Delta_{1999}^{2012} OOI_{im} &= \delta Treat_i + \lambda_m + \epsilon_{im}\end{aligned}$$

where $Treat_i$ is an indicator for living in a treated district in 1999 and μ_m, λ_m are match

[30] See Heuermann and Schmieder (2018) for the full list of criteria that were used for location choices.

[31] We require the match to be exact on gender, education, state and 2-digit occupation.

fixed-effect. Because this is a binary instrument, $\widehat{\alpha}$ collapses to a Wald estimator

$$\widehat{\alpha} = \frac{\overline{\Delta_{1999}^{2012} \log w_{treat} - \Delta_{1999}^{2012} \log w_{control}}}{\overline{\Delta_{1999}^{2012} OOI_{treat} - \Delta_{1999}^{2012} OOI_{control}}} \tag{17}$$

where the average is taken over matched pairs. We develop a procedure that calculates standard errors, building on the approach of Abadie and Imbens (2006). More details in Appendix B.4.

Table 4 shows the main result: an elasticity of .32 between options and wages. Column 1 shows that the OOI increased by .07 in treated districts following the introduction of the new stations. The reduced form results in column 2 suggest an increase of about 2.5% increase in income in the treated districts. Combining both estimates into a 2SLS estimator in column 3 yields a semi-elasticity of approximately .32 between our measure of outside options and wages. Column 4 shows that our matching process worked: there are no pre-trends. Our OLS results in column 5 show a precise zero. This is likely to be driven by an attenuation bias that is amplified substantially when using first differences of noisy variables in estimation.[32]

We next verify that our effect is driven by workers who are more likely to use the train. The high-speed commuter rail is a fairly expensive commuting option.[33] As a result, the introduction of train stations should primarily affect high-income workers. We break our sample into three education groups, which we use as a proxy for potential income. Figure 11 plots the first stage and the reduced form results, together with our point 2SLS estimates for each group. We find a higher first stage for workers with higher education. These are the workers we would expect to use the train the most. The reduced form is also higher for the more educated workers, though the estimate is imprecise. We cannot rule out a zero effect on the low education group. The two-stage least squares estimates are similar for all three groups, and we cannot rule out homogeneous effects by education group.

[32] Duncan and Holmlund (1983) show that this depends on the level of autocorrelation between the measurement errors, and true signal. Since worker observables tend to be constant while the measurement error might change between years, we expect attenuation bias to be much stronger in first difference.

[33] For example, a round-trip between Montabaur and Frankfurt takes 45 minutes each way and costs 60 Euros.

6.2 Shift-Share ("Bartik")

We next use a standard shift-share ("Bartik") instrument to estimate the elasticity between wages and options.[34] Though this exercise gives a lower point estimate of .17, we cannot reject that the elasticity we estimate is identical to the one estimated in the prior section.

The idea behind this strategy is to compare workers who work in the same industry, but who have different outside options, because they reside in different parts of the country with different industry mixes. Some workers happen to live near industries that are growing, while others happen to live near industries that are contracting. Because local growth of certain industries may be due to the impact of local productivity shocks, we use national industry trends as an instrument.

The instrument is a weighted average of national industry growth, weighted by the initial share of each industry in the region. Formally, we define

$$B_r = \sum_j s_{jr}^{04} \times \widehat{g}_j$$

where s_{jr}^{04} is the share of employed workers in region r, working at industry j in the base year (2004) and $\widehat{g}_j$ is the national employment growth of industry j. Regions are defined by the administrative regions ("Regierungsbezirke") in Germany, the statistical unit which is closest to a commuting zone.[35] Industries are defined at the 3-digit level.

To estimate the national growth of different industries, controlling for region-wide shocks, we regress the change in employment in industry j in region r between 2004 and 2014 on industry and region fixed effects:[36]

$$\Delta_{04}^{14} \log E_{jr} = g_j + g_r + \varepsilon_{jr}$$

By construction, the estimator of $\widehat{g}_j$ is not driven by regional trends captured in $\widehat{g}_r$. We use the weighted average of the industry fixed effects $\widehat{g}_j$ by initial industry shares s_{jr}^{04} to calculate B_r. This construction verifies that B_r is not driven by local employment shocks in this region, or even in nearby regions.[37]

[34] These shocks were used in several papers including Bartik (1991); Blanchard and Katz (1992); Card (2001); Autor et al. (2013). It was used specifically for the context of a shock to outside options by Beaudry et al. (2012).

[35] We take all 39 regions based on NUTS2 level coding of the European Union. This includes historical administrative regions that have been disbanded. Results are robust to the definition of a region and hold for the NUTS3 level (district) as well.

[36] We make a Bayesian correction of uniform prior by adding one observation in each industry, region and year combination.

[37] This is different from a leave-one-out estimate, that might still be driven by local shocks in nearby

We estimate the following system of equations

$$
\begin{array}{rcll}
\Delta_{04}^{14} \log w_{ijr} & = & \alpha \Delta_{04}^{14} OOI_{ijr} & + \beta \Delta_{04}^{14} X_{ijr} + Ind_j^{04} + \upsilon_{ijr} \\
\Delta_{04}^{14} OOI_{ijr} & = & \gamma B_r & + \delta \Delta_{04}^{14} X_{ijr} + Ind_j^{04} + \epsilon_{ijr}
\end{array}
\tag{18}
$$

where we control for the industry of the worker in the beginning of the period (2004). We cluster standard errors at the level of the treatment, which is the region. The parameter of interest is α, the elasticity of wages with respect to options.

Table 5 presents the main results. Columns 1 and 2 show the first-stage and reduced form results. A 10% higher employment in other industries, which is about .1 increase in the instrument, translates to approximately 6% more relevant options, and 1% increase in wages. Combining both estimates yields a semi-elasticity of .17: a 10% increase in relevant options leads to a 1.7% increase in wages.

The identifying assumption is that growing industries are not systematically located in regions where wages are growing for other reasons (Borusyak et al., 2018). One way the assumption could be violated is if there are productivity spillovers. Workers that live near industries that are growing, may enjoy a local demand shock for their production due to the positive income effect on workers in that region. This could generate a wage increase, that is not driven by the improvement in their outside options. This is particularly a concern for workers who are producing non-tradable goods, whose productivity is set by local demand.

We address this concern by showing the results hold for workers in exporting industries, which are less likely to be affected by local demand shocks. We use information from the establishment survey to calculate the export share of each industry.[38] We divide our data into three groups based on the export share of the industry where the worker worked in 2004. Table 6 shows the results for each of the groups. We find a large and statistically significant elasticity between options and wages even among workers in industries with the highest exporting share. Column 1 indicates that in response to a 10% increase in OOI, workers in these industries see their wages rise by 1%. This elasticity is somewhat lower than that in our baseline results (.10 versus .17). However, we cannot reject that they are equal.[39]

regions.

[38] This is a lower bound for demand from outside the region, as it does not include sales to other regions in Germany, which we cannot see in our data. We calculate the mean industry level since we don't have the share of sales from export for all employers, but only a representative sample.

[39] Beaudry et al. (2012) find similar results when dividing the data into tradable and non-tradable industries, based on their geographical spread. They argue that non-tradable industries are geographically spread across different regions, while tradable goods could be concentrated in specific regions. They also address additional potential threats to the identification assumption and show they don't seem to have a

We next examine heterogeneity across gender and the three education groups. We estimate Equations 18 separately for each group. Figure 12 plots the results for all groups, as well the full population. While splitting the sample increases the size of the confidence intervals, the point estimates are quite close. This suggests that using the same value for α for all groups is a reasonable approximation.

We next use this setting to decompose the different effect of access to more options into impacts for job stayers and movers. Because the choice of whether to move is endogenous, we view this as a decomposition exercise. We interact the changes in OOI with an indicator variable for whether a worker stayed at their establishments during this period. The results are shown in Table 7. As our model predicts, we find that the effect on stayers is smaller. This is possibly because they only benefit through an improvement in their outside options. The larger effect on movers is consistent with an additional improvement in match quality.

While the elasticity we estimate in this exercise is lower from the one we estimated using the fast commuter rails, their difference is not statistically different from zero. Figure 13 compares our results in this section to the elasticity we estimated using the introduction of high-speed commuter rails. The fact that we found elasticities of a similar magnitude by using two distinct sources of variation suggests that this range of estimates is a reasonable benchmark for the value of α.

7 Effect on Wage Inequality

In this section, we combine our estimates on the distribution of the OOI, with our estimates of the OOI-wage elasticity, to assess the overall effect of options on the wage distribution. We then examine which covariates drive differences in options. We find that equalizing workers' ability to commute or move would eliminate the gender gap in OOI, and would reverse the sign of the OOI gap by education.

7.1 Overall Impact on the Wage Distribution

We examine what portion of between-group wage inequality can be attributed to difference in OOI. We first estimate a Mincer equation

$$\log w_i = \beta_0 X_i + \epsilon_i \tag{19}$$

significant effect on the result.

where X_i includes indicators for each education group, a quadratic in age, gender, citizenship status, log district density and an indicator for part-time job. Since wages are top-coded we use a Tobit model to estimate $\widehat{\beta}_0$. We then add the OOI to the set of dependent variables, with a fixed coefficient.

$$\log w_i = \widehat{\alpha} OOI_i + \beta_1 X_i + \epsilon_i$$

We use $\widehat{\alpha} = .26$ which is the average of the two point estimates we derived in Section 6 from the two quasi-random sources of variation. $\widehat{\beta}_0$ captures the overall gaps in wages between these demographic groups, $\widehat{\beta}_1$ is the remaining gaps that are driven by factors other than the OOI, and $\widehat{\beta}_0 - \widehat{\beta}_1$ is the part that can be attributed to the differences in OOI.

Figure 14 shows the main results. The full bars display the gaps that we estimated ($\widehat{\beta}_0$), where every bar is the wage premium for this group members. For instance, the premium for being a male (the gender gap) is .19 log units in Germany. The portion that can be attributed to the OOI $\left(\widehat{\beta}_0 - \widehat{\beta}_1\right)$ is colored in red, while the remaining gap ($\widehat{\beta}_1$) is left in blue.

The OOI explains significant portions of several German wage gaps. When we add OOI to the regression, the gender gap is cut by .06 log units (30% of the overall gap). This is driven by the .23 gender gap in OOI we found in Table 2, multiplied by $\widehat{\alpha}$. Our results also indicate that 88% of the gap between German citizens to non-citizens (.08 log units) can be attributed to differences in options. The wage difference between high-level and intermediate-level secondary schooling is cut by .07 log units, which is about 25% of the initial gap. Our results also attribute 39% of the return to experience at age 18 in access to options. Table 8 shows these results numerically in columns (3) and (4), as well as results from a winsorized OLS in columns (1) and (2).

7.2 Explaining Differences in Outside Options - The Role of Commuting Costs

We next examine which factors impact workers' options. We find that differences in commuting costs seems to be particularly important, especially in its effect on the gender gap and return to higher education.

We start by examining the impact of different variables on the probabilities of observing a match. We analyze our results from the estimation of matrix A defined in Assumption 2, which we estimated using a logistic regression. This matrix is also the cross-derivative of match quality τ (Equation 9). Appendix Table A7 shows the top absolute values of $\widehat{A}$, when variables are standardized so the results are not affected by specific units.[40]

These results indicate that the most important factor in determining match quality is commuting and moving costs. Distance has the largest standardized coefficient in absolute terms (-4.15). While distance to a job is an important factor for all workers, it is particularly important for female workers, for less-educated workers, and for non-German citizens. Appendix Table A8 presents the raw coefficient on distance for different worker characteristics. This coefficient is the effect of an additional mile on the log probability of a match, at mile zero. For our baseline group, forty year old males citizens from higher secondary schools, the coefficient is -.141. The interaction with female is -.024, so women are 17% more sensitive to distance than the baseline group. Lower educated workers are significantly more sensitive to distance (coefficient -.037). Non-German citizens seem to be more sensitive than citizens (coefficient -.019). Finally, workers at first become less sensitive to distance with age, but this is a concave function that reaches its maximum at age 42.

By simulating counterfactuals from the underlying model, we can quantify the overall effect of differences in commuting and moving costs on wages through their effect on the OOI. We estimate the wage gain for every worker, if they had the minimal commuting/moving costs. Based on our estimation, these are the costs of a 40 year old, high-educated, male citizen. We generate a matrix $\widetilde{A}$ where the coefficients on distance is set to this minimum level for all workers. We then simulate the probabilities $\widetilde{f}_j^i$ using this matrix, calculate the $\widetilde{OOI}_i$ and translate it to $\widetilde{\log w_i}$ using $\widehat{\alpha}$. This counterfactual should be thought of as changing only a zero measure number of workers each time, and keeping all other workers and employers unchanged, so that there are no general equilibrium effects. We compare the differential gains from this exercise to assess the importance of commute in generating wage gaps.

We run a regression of the counterfactual gains in wage over basic demographics

$$\Delta \log w_i = \beta_2 X_i + \epsilon_i$$

where $\Delta \log w_i = \widetilde{\log w_i} - \log w_i$ and X_i same as in Equation 19. The coefficients β_2 from this regression are the part of the wage gap that would be closed, if commuting and moving costs were equalized at the lowest level, for these workers. The results of this exercise are presented at Figure 15. The figure plots the full gap ($\widehat{\beta}_0$, blue), the portion that

[40]Online appendix table A1 shows the full standardized results for A ; online table A2 shows the raw results.

can be attributed to the OOI ($\widehat{\beta_0} - \widehat{\beta_1}$, red), and the part that will be closed by equalizing commuting costs at the minimal level ($\widehat{\beta_2}$, yellow).

Differences in commuting costs seem to explain all of the gender gap that is driven by differences in options. Equalizing commuting costs would increase wages for women by about .07 log units, relative to men. This is one third of the overall gender gap. Even though we find that men and women sort into different jobs, there seems to be a similar number of jobs for males and females, therefore the only difference is the distance in which workers are searching for jobs. This does not mean that there aren't other ways to increase the OOI for women, such as increasing the supply of jobs that women typically sort into.

In contrast, equalizing commuting costs increases the wage gap between German citizens and non-citizens. These results are surprising at first glance because we found that non-citizens are more sensitive to distance. However, they can be explained by the fact that non-citizens are more concentrated in large German cities. As a result, their commuting costs are already low. German citizens are more dispersed across rural areas, and are more dependent on their ability to commute to jobs in major cities.

The education gap in OOI actually reverses once we equalize commuting costs: workers with intermediate-secondary education have more options than those with higher-secondary. Therefore, the higher-education premium drops by .15 log units (51% of the overall premium), which is more than the full effect of the OOI difference between these groups (.07 log units). This implies that, in a given area, workers with intermediate-secondary education have more relevant job options than workers with high-secondary education. It is only because higher-secondary education workers are willing to take jobs in more distant areas, that they end up with more options. This result can be explained by the fact that more educated workers tend to be more concentrated in occupations that have more industry specific skills, as shown in Figure 7. Additionally, intermediate-secondary workers can take both higher-skill, and lower-skill jobs in addition to staying at the same level. While high-secondary workers usually have fewer options to climb to jobs requiring even more skills.

Other than geographical distance, the most significant factor in determining a worker's match (Appendix Table A7) is their training occupation. Workers tend to stay in occupations similar to the ones in which they were trained. Our results in Section 5 show that those who undertook training in occupations with more transferable skills have more options than those who received more narrow training. Since transferable skills are more common in low-paying occupations, the OOI is reducing inequality between occupations.

8 Conclusion

In this paper we provide a distinctive and micro-founded approach to empirically estimate workers' outside options, and to measure the impact of outside options on the wage distribution. The starting point for our analysis was a two-sided matching model, which produced a sufficient statistic for the impact of outside options on wages, the OOI. We took the OOI to the data to identify the workers with better outside options. We then combined this result with a causal estimate of the elasticity between the OOI and wages, to assess the overall impact of options on wages. Our results suggest that differences in outside options generate lower income for females by six percent, non-citizens by eight percent and intermediate educated workers by seven percent (compared to high educated).

Our results indicate that policies that improve workers' options, including investments in transportation infrastructure or regulation of working hours, are likely to have significant general equilibrium effects. While such policies are usually analyzed only through their impact on workers that directly benefit from them, our results indicate that these policies will likely have important spillovers onto other workers through their outside options. These general equilibrium channels can be studied through their effect on the OOI.

One interesting direction for future work would be to use this framework to analyze specific industries in which outside options play a key role, and good micro-data is available. A similar analysis could also be done on the employer's side of the market, analyzing heterogeneity in the availability of options for firms, and the impact of outside options on profits. Finally, the ability to identify workers with better outside options could be useful in studying heterogeneous effects of various policies, or labor market shocks. Our analysis of the heterogeneous response to mass-layoff is one example for how this can be done.

References

Abadie, Alberto, and Guido W Imbens. 2006. "Large sample properties of matching estimators for average treatment effects." *Econometrica*, 74(1): 235–267.

Autor, David, David Dorn, and Gordon H Hanson. 2013. "The China syndrome: Local labor market effects of import competition in the United States." *American Economic Review*, 103(6): 2121–68.

Azar, José, Ioana Marinescu, and Marshall I Steinbaum. 2017. "Labor market concentration." National Bureau of Economic Research Working Paper 24147.

Bartik, Timothy J. 1991. *Who benefits from state and local economic development policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Beaudry, Paul, David A Green, and Benjamin Sand. 2012. "Does Industrial Composition Matter for Wages? a Test of Search and Bargaining Theory." *Econometrica*, 80(3): 1063–1104.

Benmelech, Efraim, Nittai Bergman, and Hyunseob Kim. 2018. "Strong employers and weak employees: How does employer concentration affect wages?" National Bureau of Economic Research Working Paper 24307.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–890.

Bidner, Chris, and Ben Sand. 2016. "Job Prospects and Pay Gaps: Theory and Evidence on the Gender Gap from US Cities."

Black, Dan A. 1995. "Discrimination in an Equilibrium Search Model." *Journal of Labor Economics*, 13(2): 309–334.

Blanchard, Olivier, and Lawrence F Katz. 1992. "Regional Evolutions." *Economic Studies Program, The Brookings Institution*, 23(1): 76.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2018. "Quasi-experimental Shift-share Research Designs." *arXiv preprint arXiv:1806.01221*.

Caldwell, Sydnee, and Nikolaj Harmon. 2018. "Outside Options, Bargaining, and Wages: Evidence from Coworker Networks." Massachusetts Institute of Technology Working Paper.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19(1): 22–64.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *Quarterly Journal of Economics*, 131(2): 633–686.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics*, 36(S1): S13–S70.

Choo, Eugene, and Aloysius Siow. 2006. "Who Marries Whom and Why." *Journal of Political Economy*, 114(1): 175–201.

Dagsvik, John K. 1994. "Discrete and continuous choice, max-stable processes, and independence from irrelevant attributes." *Econometrica*, 1179–1205.

Dube, Arindrajit, Jeff Jacobs, Suresh Naidu, and Siddharth Suri. 2018. "Monopsony in Online Labor Markets." National Bureau of Economic Research Working Paper 24416.

Duncan, Greg J, and Bertil Holmlund. 1983. "Was Adam Smith right after all? Another test of the theory of compensating wage differentials." *Journal of Labor Economics*, 1(4): 366–379.

Dupuy, Arnaud, and Alfred Galichon. 2014. "Personality Traits and the Marriage Market." *Journal of Political Economy*, 122(6): 1271–1319.

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg. 2009. "Revisiting the German Wage Structure." *The Quarterly Journal of Economics*, 124(2): 843–881.

Gathmann, Christina, and Uta Schönberg. 2010. "How General Is Human Capital? a Task-Based Approach." *Journal of Labor Economics*, 28(1): 1–49.

Hagen, Von Tobias. 2003. "Does Fixed-term Contract Employment Raise Firms' Adjustment Speed." *Jahrbücher f. Nationalökonomie u. Statistik*, 223(4).

Heuermann, Daniel F, and Johannes F Schmieder. 2018. "The Effect of Infrastructure on Worker Mobility: Evidence From High-speed Rail Expansion In Germany." *Journal of Economic Geography*, lby019.

Hirsch, B, T Schank, and C Schnabel. 2010. "Differences in Labor Supply to Monopsonistic Firms and the Gender Pay Gap: An Empirical Analysis Using Linked Employer - Employee Data from Germany." *Journal of Labor Economics*, 28(2): 291–330.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. 2013. "The Allocation of Talent and U.S. Economic Growth." National Bureau of Economic Research Working Paper 18693.

Jacobson, Louis S, Robert J. Lalonde, and Daniel G Sullivan. 1993. "Earnings Losses of Displaced Workers." *American Economic Review*, 83(4): 685–709.

Jäger, Simon, Benjamin Schoefer, Sammy Young, and Josef Zweimuller. 2018. "Wages, Bargaining, and the Nonemployment Option: Evidence from Unemployment Insurance."

Kreps, David M, and Jose A Scheinkman. 1983. "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes." *The Bell Journal of Economics*, 14(2): 326–337.

Manning, Alan. 2003. *Monopsony in motion: Imperfect competition in labor markets.* Princeton University Press.

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

Mitlacher, Lars W. 2008. "Job quality and temporary agency work: Challenges for human resource management in triangular employment relations in Germany." *The International Journal of Human Resource Management*, 19(3): 446–460.

Naidu, Suresh. 2010. "Recruitment Restrictions and Labor Markets: Evidence from the Postbellum US South." *Journal of Labor Economics*, 28(2): 413–445.

Naidu, Suresh, Yaw Nyarko, and Shing-Yi Wang. 2016. "Monopsony Power in Migrant Labor Markets: Evidence from the United Arab Emirates." *Journal of Political Economy*, 124(6): 1735–1792.

Nimczik, Jan Sebastian. 2017. "Job Mobility Networks and Endogenous Labor Markets."

OECD. 2018. "Part-time employment rate (indicator)."

Pérez-Castrillo, David, and Marilda Sotomayor. 2002. "A Simple Selling and Buying Procedure." *Journal of Economic Theory*, 2(103): 461–474.

Postel-Vinay, Fabien, and Jean-Marc Robin. 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica*, 70(6): 2295–2350.

Ransom, Michael R, and David P Sims. 2010. "Estimating the Firm's Labor Supply Curve in a "new Monopsony" Framework: Schoolteachers in Missouri." *Journal of Labor Economics*, 28(2): 331–355.

Ransom, Michael R, and Ronald L Oaxaca. 2010. "New Market Power Models and Sex Differences in Pay." *Journal of Labor Economics*, 28(2): 267–289.

Robinson, Joan. 1933. *The Economics of Imperfect Competition.* Palgrave Macmillan.

Roth, Alvin E, and Marilda Sotomayor. 1992. "Two-sided matching." *Handbook of Game Theory with Economic Applications*, 1: 485–541.

Shapley, Lloyd S, and M Shubik. 1971. "the Assignment Game I: The Core." *International Journal of Game Theory*, 1(1): 111–130.

Staiger, D, Joanne Spetz, and C Phibbs. 2010. "Is There Monopsony in the Labor Market? Evidence from a Natural Experiment." *Journal of Labor Economics*, 28(2).

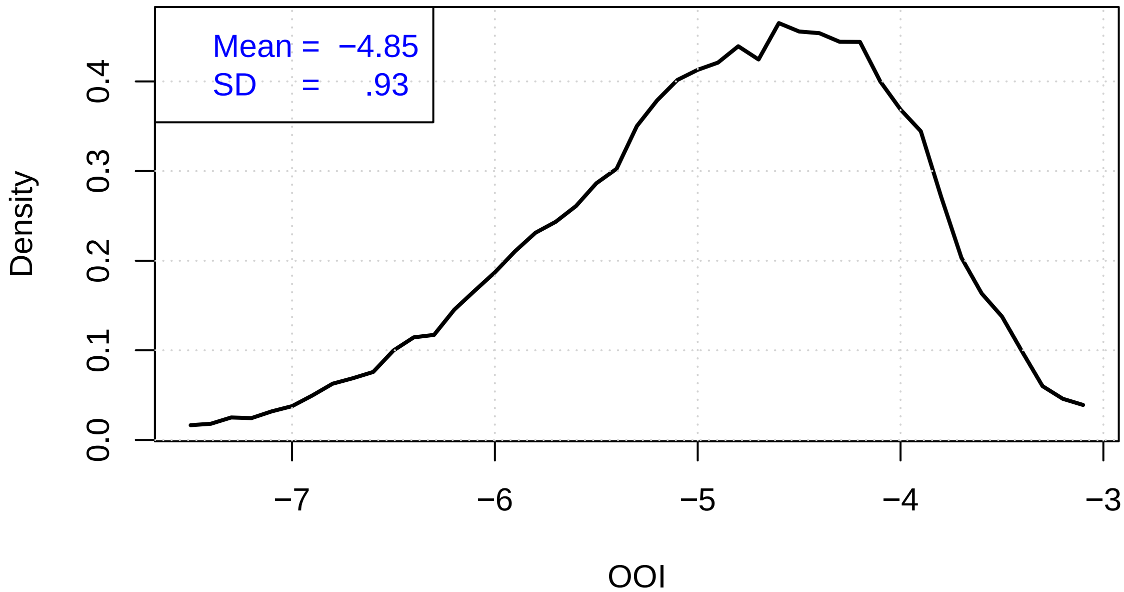
Tirole, Jean. 1988. *The Theory of Industrial Organization.* MIT press.

Figures and Tables

Figure 1: German Districts

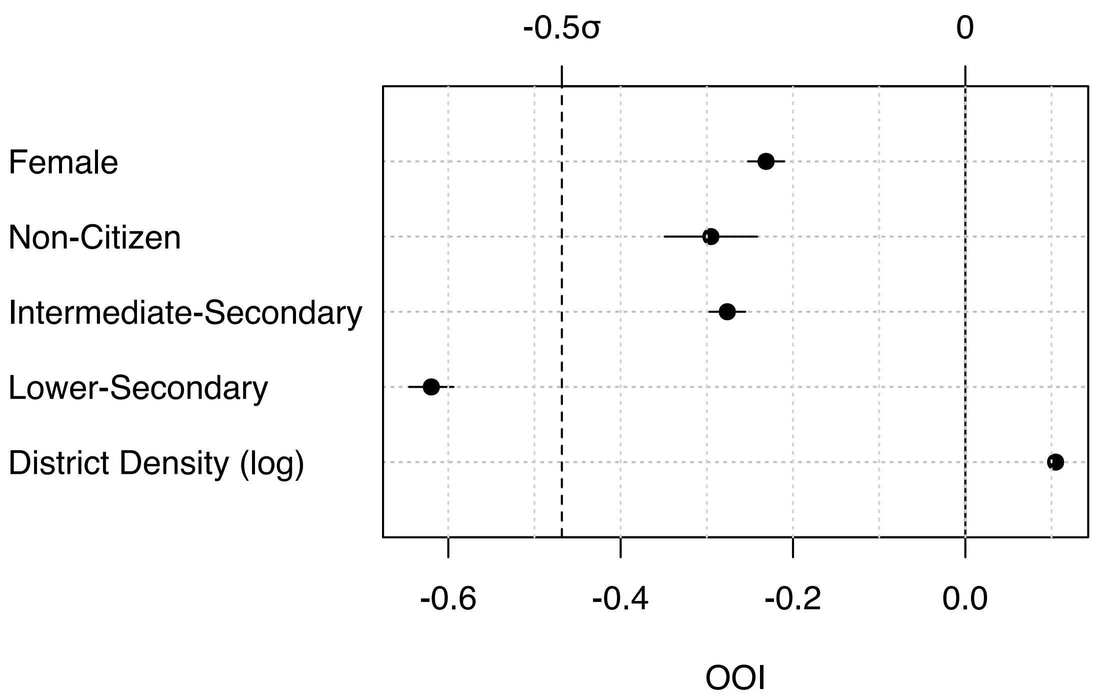
Note: This map illustrates the 402 districts (kreis) in Germany.

Figure 2: Distribution of Outside Option Index



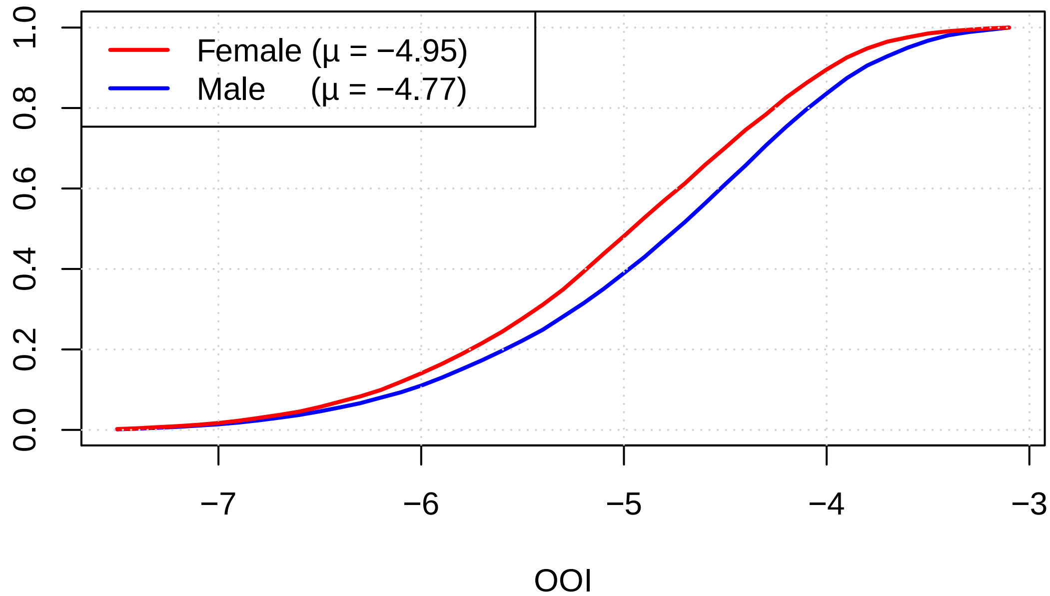
Note: This figure plots the distribution of the outside options index as calculated for the population of German workers as of June 30th, 2014. The OOI was calculated using the procedure described in Section 4.2. LIAB sample weights are used to make the distribution representative of the German population.

Figure 3: OOI by Characteristics



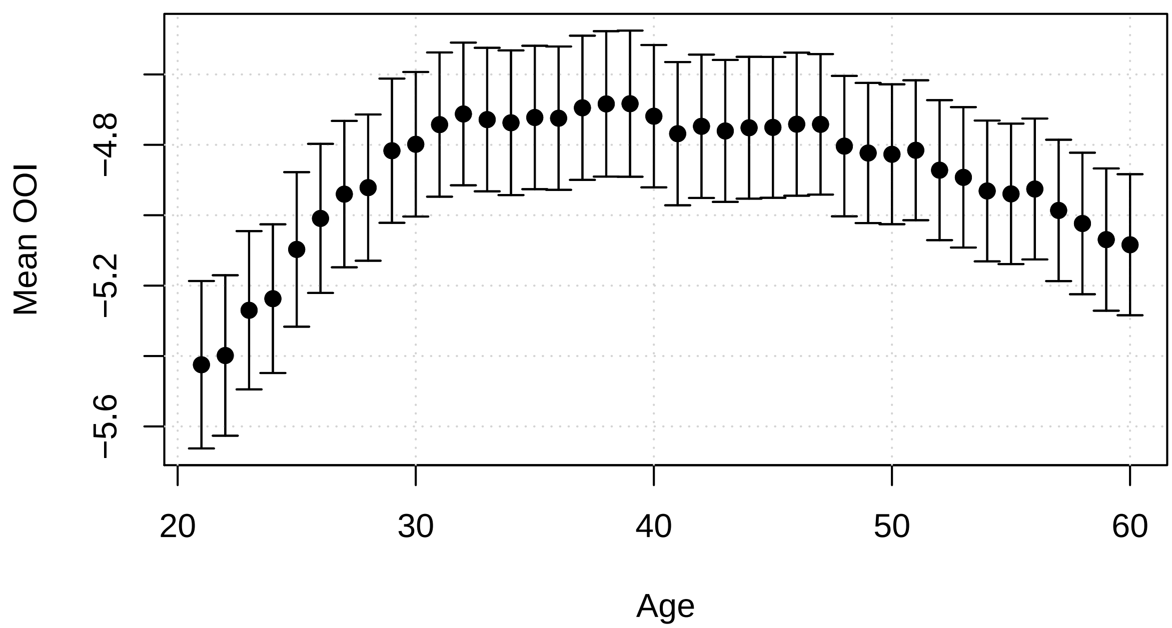
Note: This figure plots the coefficients from a regression of OOI on education, gender, citizenship and a quadratic in age. The results are also presented on column 1 of Table 2. Confidence intervals are plotted at the 95% level. The lower axis shows raw OOI units, while the upper axis uses standard deviation units.

Figure 4: Cumulative Distribution of Outside Option Index by Gender



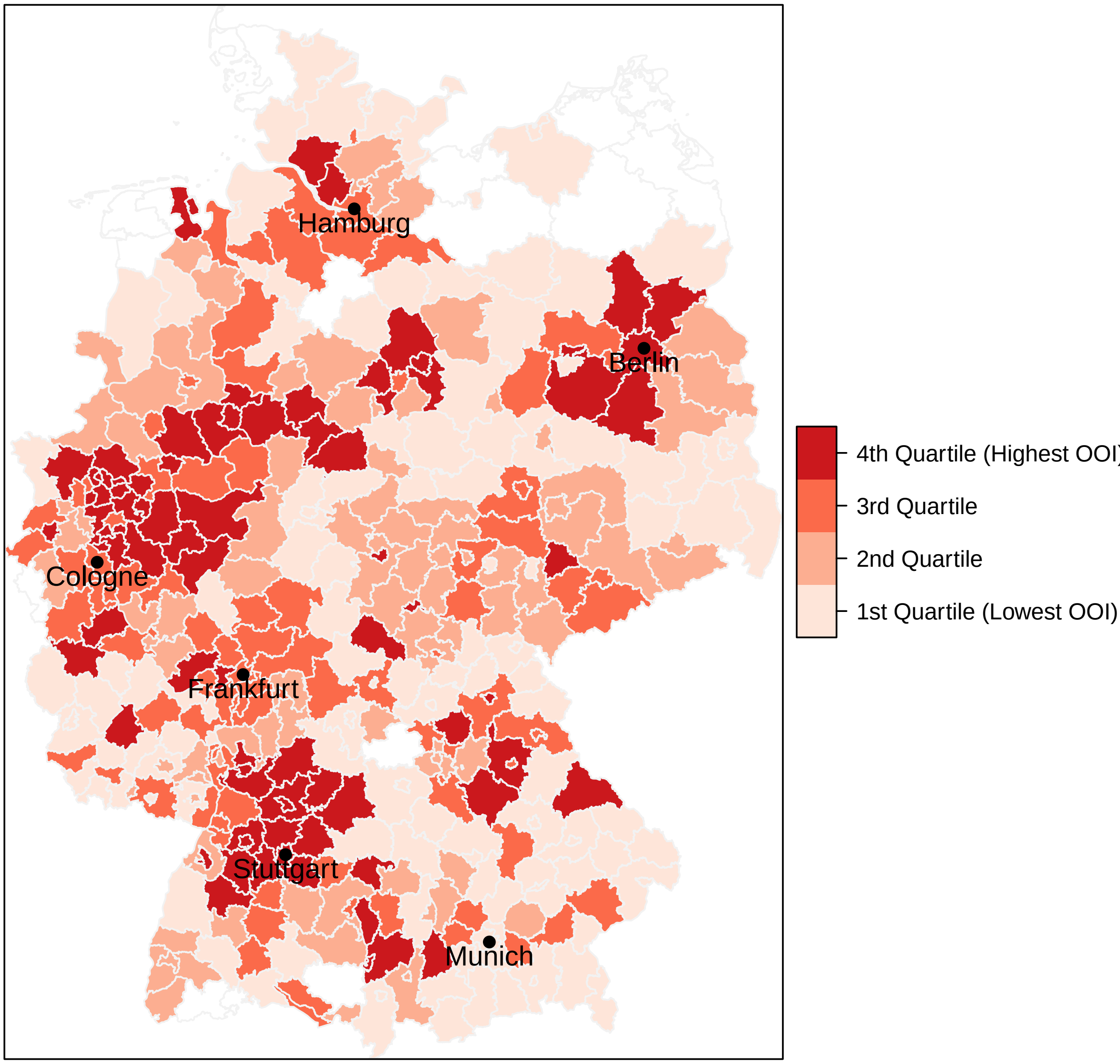
Note: This figure plots the cumulative distribution function of the outside options index by gender, as calculated for the population of German workers as of June 30th, 2014. The OOI was calculated using the procedure described in Section 4.2. LIAB sample weights are used to make the distribution representative of the German population.

Figure 5: OOI by Age



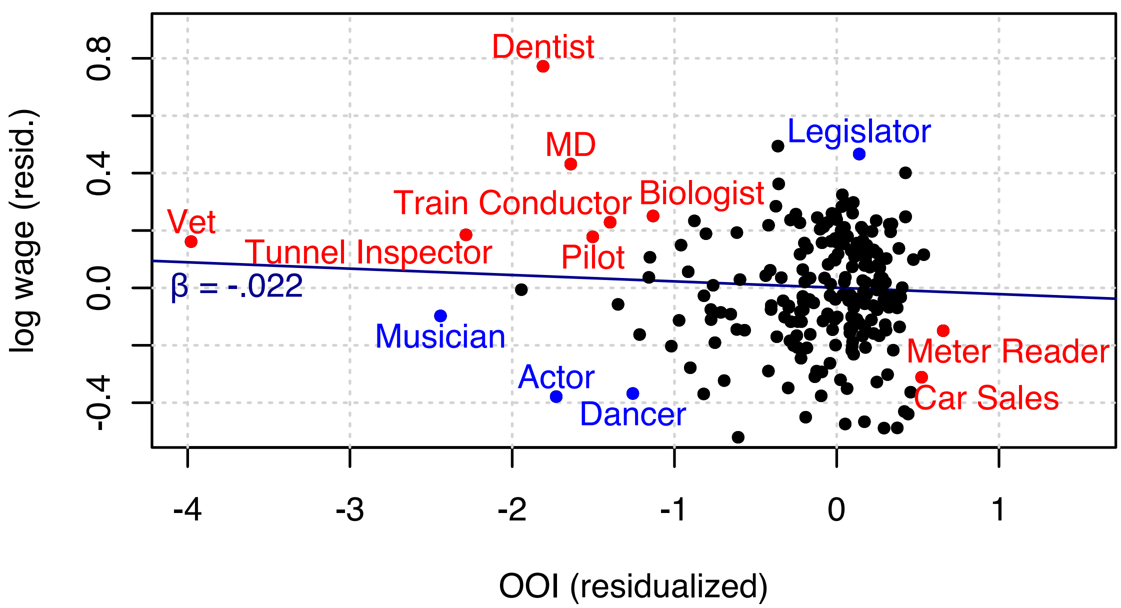
Note: This figure plots the mean OOI by age in the German population. LIAB sample weights are used to make the sample representative of the German population. Confidence intervals are plotted at the 95% level.

Figure 6: OOI Distribution by Region



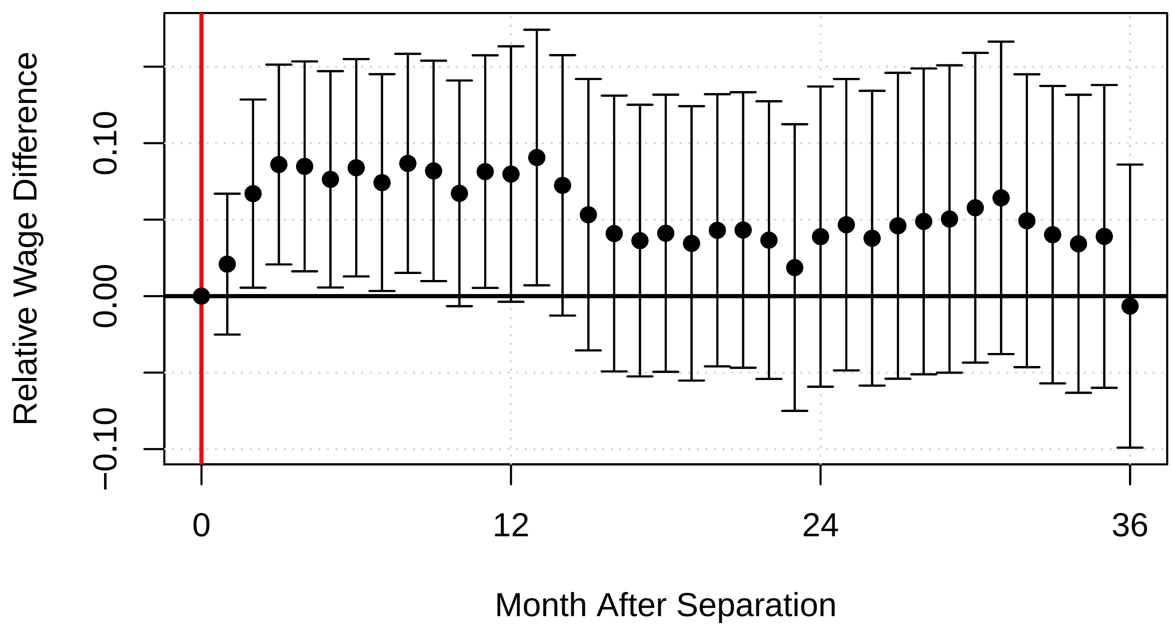
Note: This figure plots the distribution of the outside options index by district (kreis) as calculated for the population of German workers as of June 30th, 2014. The OOI was calculated using the procedure described in Section 4.2. The value for each district is a weighted mean of the workers in this district, using the LIAB sample weights to make the distribution representative of the population in the district.

Figure 7: OOI by Training Occupation



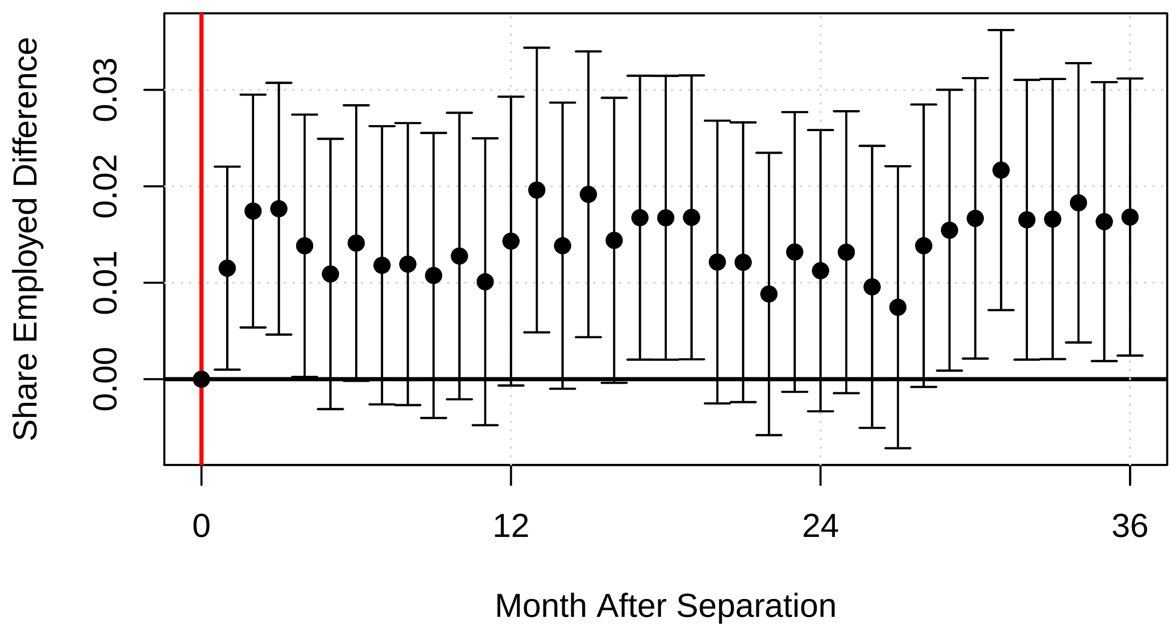
Note: This figure plots the mean residualized outside options index and log wages by training occupation as calculated for the population of German workers as of June 30th, 2014. The OOI was calculated using the procedure described in Section 4.2. Residuals for the OOI and log wages were taken from a regression on gender, a quadratic in age, education category, citizenship status and district of residence. Means are calculated using the LIAB sample weights to make the distribution representative of the population in the occupation. See Section 3.1 for exact definition of a training occupation.

Figure 8: Mass-Layoffs - Differences in Relative Income Between High/Low OOI Workers



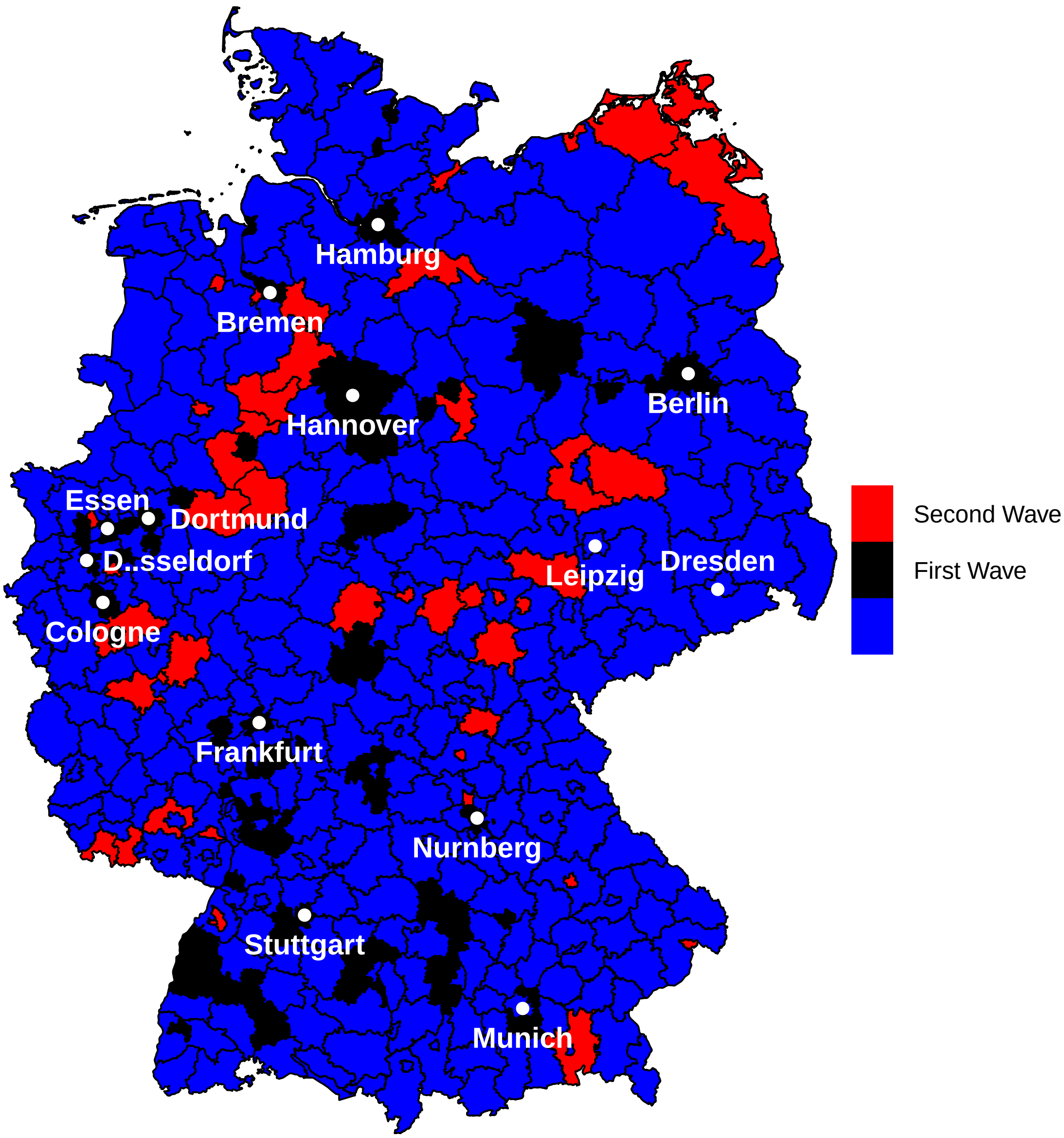
Note: This figure shows the difference in relative income for workers with OOI above and below the establishment OOI. Relative income is defined as the current daily income in that month divided by the last daily income before the layoff. Mass layoffs are defined as an establishment with at least 50 workers that reduced its workforce by at least 30% in a given year. The sample includes only workers who have worked for at least three years before the layoff and are below the age of 55. The median OOI is calculated based on the pool of laid-off workers in a given establishment and year. The coefficients are taken from a regression of relative income on an indicator for above median OOI, interacted with indicator for each month after separation (plotted), with fixed effects for establishment×month (Equation 14).

Figure 9: Mass-Layoffs - Differences in Search Time Between High/Low OOI Workers



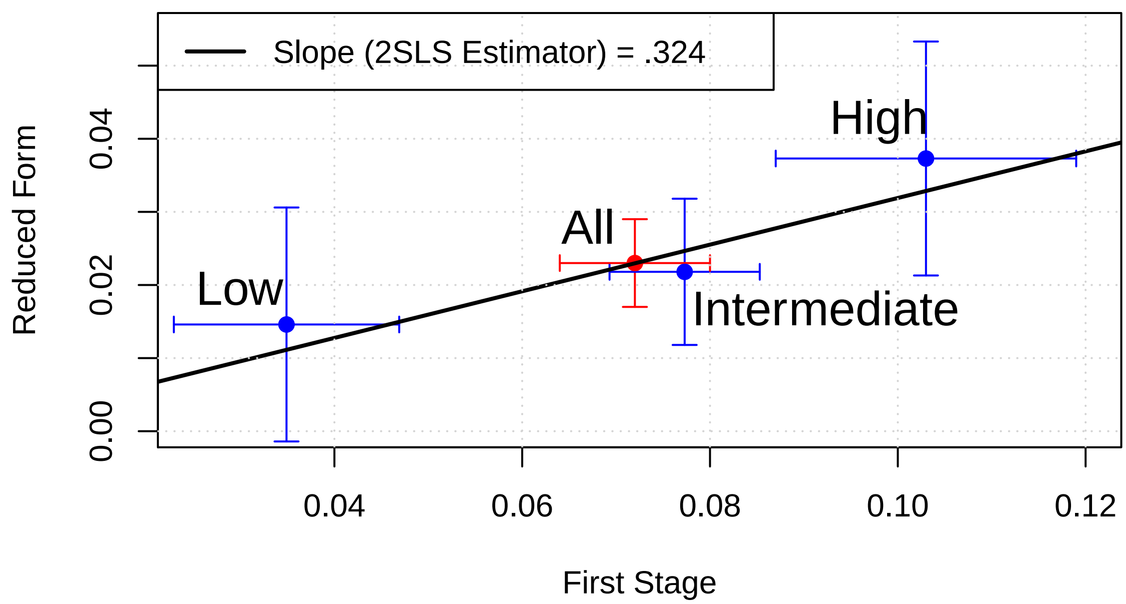
Note: This figure shows the difference in employment for workers with OOI above and below the establishment OOI. Employment is defined as any income greater than zero. Mass layoffs are defined as an establishment with at least 50 workers that reduced its workforce by at least 30% in a given year. The sample includes only workers who have worked for at least three years before the layoff and are below the age of 55. The median OOI is calculated based on the pool of laid-off workers in a given establishment and year. The difference is calculated using a regression of employment on an indicator for above median OOI, interacted with indicator for each month after separation, with fixed effects for establishment×month (Equation 15).

Figure 10: ICE Stations



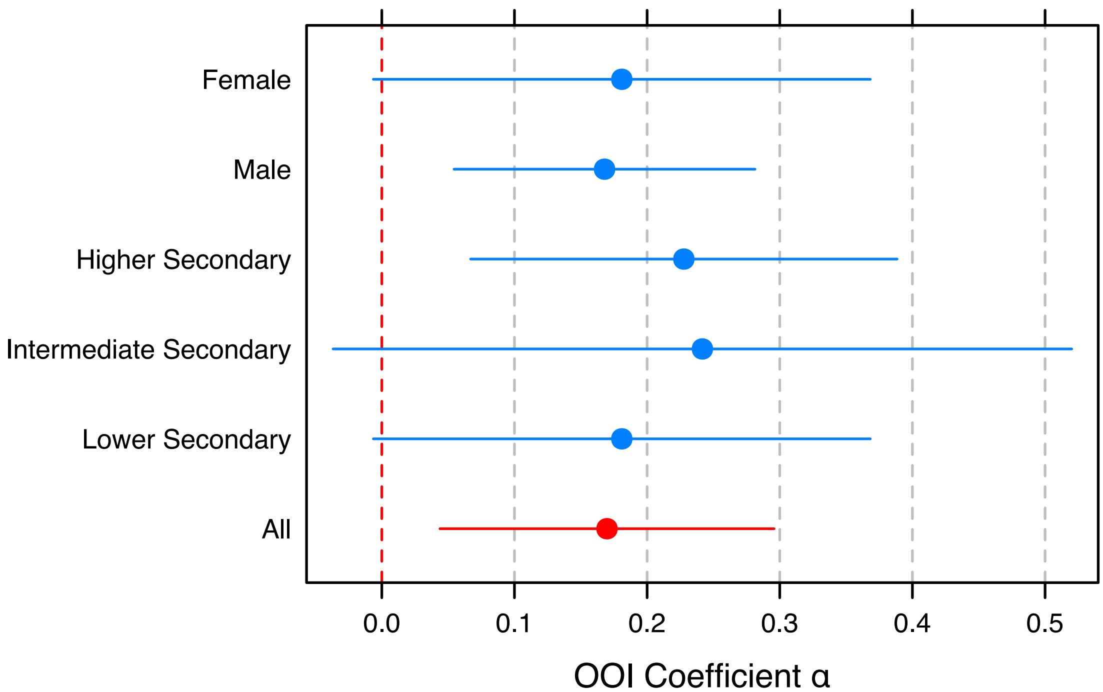
Note: This figure shows the locations of ICE train stations by districts. The first wave includes all stations that were opened pre-1999. The second wave includes all stations that were opened post-1999.

Figure 11: Impact of Express Trains by Schooling Level



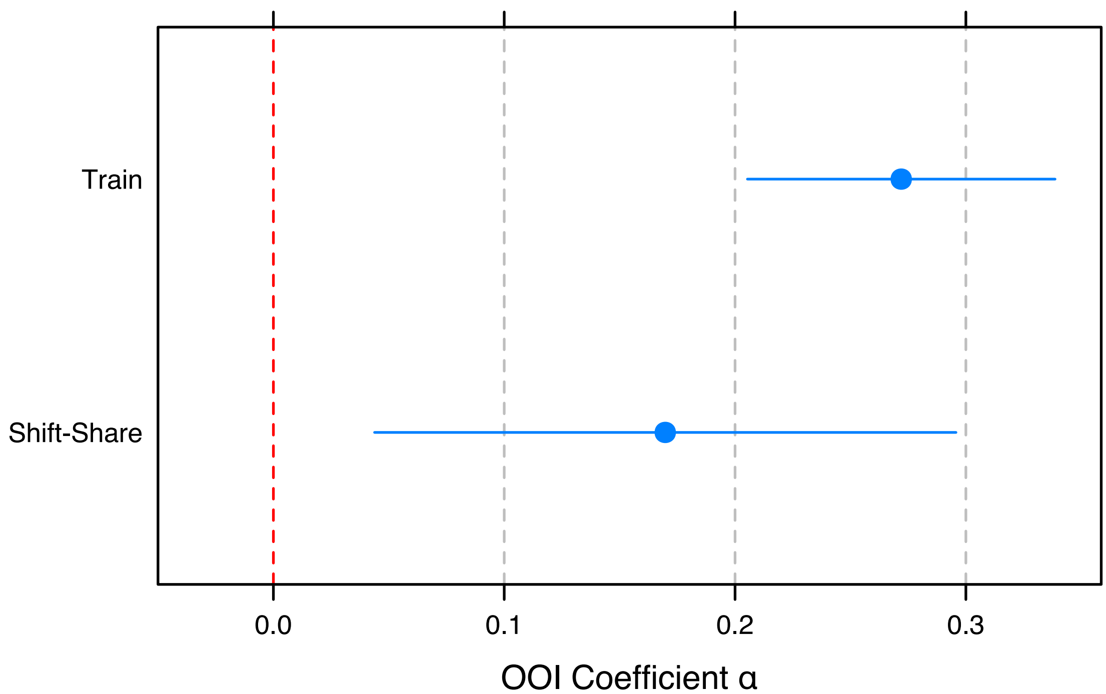
Note: This figure plots the first-stage and reduced-form results for three education groups, and their combination. First stage is the treatment effect on OOI. Reduced form is the treatment effect on log wages. Both were calculated using nearest-neighbor matching with replacement. Treatment is defined as workers that in 1999 lived in districts that got ICE stations post-1999. The control group includes workers that in 1999 lived in districts that never got ICE stations. Matching is done exactly on gender, education group, citizenship status, state and 2-digit training occupation and continuously on age, and PCA components for training occupation. Confidence intervals are at the 95% level, and are calculated based on standard errors derived from a method by Abadie and Imbens (2006). The black line represents the 2SLS point estimate for the entire sample.

Figure 12: Shift-Share Results by Gender and Education



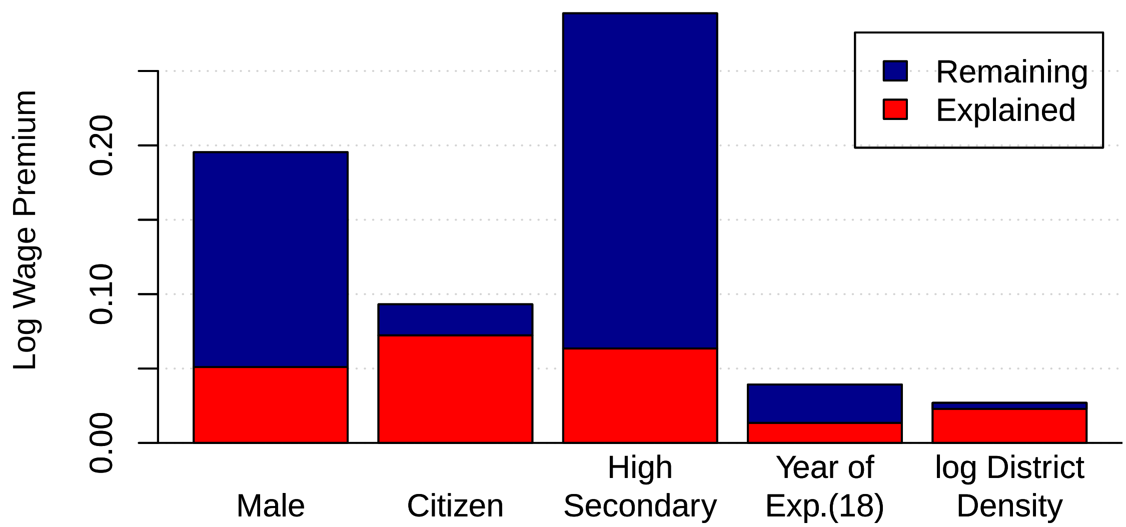
Note: Every category displays the estimate for coefficient $\widehat{\alpha}$ from Equation 18, ran separately for each education group or gender (blue), and for the entire population (red). This captures the effect of changes in OOI on changes in log wages between 2004-2014, when we instrument for the changes in OOI with the shift-share instrument. The instrument is constructed from an average of a 3-digit industry national employment growth weighted by the initial share of every industry in a region (see Section 6.2). Standard errors are clustered within the unit of treatment, which is regions. Confidence intervals are at the 95% level.

Figure 13: Estimates of Elasticity Between OOI and Wages



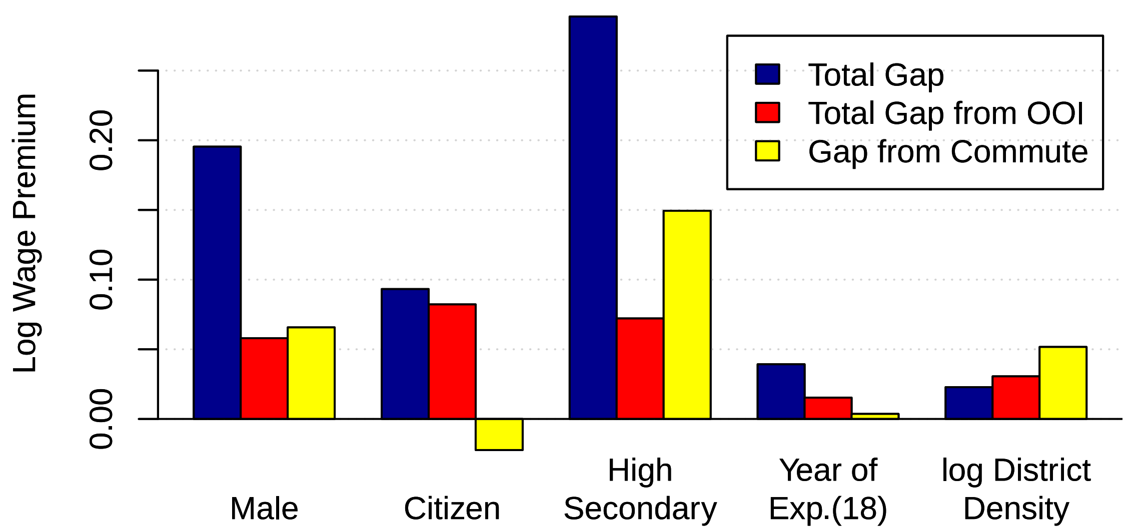
Note: This figure compares the elasticity between OOI and wages from the two different sources of quasi-random variations that we used. Train includes the results for parameter $\widehat{\alpha}$ estimated using the introduction of high-speed commuter rails (see Section 6.1 and notes for Figure 11 for more details) . Shift-Share uses an instrument based on national industry employment trends (see Section 6.2 and notes for Figure 12 for more details). Confidence intervals are at the 95% level. The difference between the point estimates is .156 (.081).

Figure 14: Overall Effect on Wage Inequality



Note: Every bar in this plot is the coefficient on the corresponding category in a regression of log wages on Male, Citizen, indicator for secondary-education category, a quadratic in age, district density and an indicator for part-time job. The blue portion of the bars (remaining gap) is the coefficient from the same regression, controlling for the OOI with a coefficient fixed to .26, which was estimated with the two quasi-random variations. The part in red (explained gap) is the difference between the two coefficients. The reference workers is a female, non-citizen, with intermediate secondary education and 18 years old.

Figure 15: Effect of Commuting/Moving Costs



Note: Blue bars (Total gap) are derived from the coefficient on the corresponding category in a regression of log wages on Male, Citizen, indicator for secondary-education category, a quadratic in age, district density and an indicator for part-time job. The red bars (Total gap from OOI) is the difference in coefficient between the same regression and one that control for the OOI with a coefficient fixed to .26, which was estimated with the two quasi-random variations. The yellow bars (Gap from Commute) is calculate from a similar regression, replacing the dependent variable with minus the gains from reducing commuting costs to their minimal level (see Section 7.2 for more details). The reference workers is a female, non-citizen, with intermediate secondary education and 18 years old.

Table 1: Descriptive Statistics

(a) Workers

	All		Females		Males	
	Mean	SD	Mean	SD	Mean	SD
Female	.46	(.50)	1	(.00)	0	(.00)
Age	45.05	(12.49)	45.53	(12.27)	44.66	(12.67)
German Citizen	.97	(.17)	.98	(.15)	.96	(.19)
Education: Higher Secondary	.29	(.45)	.30	(.46)	.28	(.45)
Education: Intermediate Secondary	.31	(.46)	.34	(.47)	.28	(.45)
Education: Lower Secondary	.22	(.41)	.16	(.37)	.26	(.44)
Education: Intermediate/Lower	.19	(.39)	.20	(.40)	.18	(.38)
N	450,917		162,780		288,137	

(b) Jobs

	All		Females		Males	
Name	Mean	SD	Mean	SD	Mean	SD
Part - Time	.31	(.46)	.53	(.50)	.13	(.34)
Fixed Contract	.11	(.31)	.11	(.32)	.10	(.30)
Temporary Agency	.02	(.12)	.01	(.08)	.02	(.15)
Establishment Size	1,552.8	(7,679)	827.2	(5,014)	2,166.2	(9,313)
Annual Sales per worker (Euro)	163,286	(185,651)	130,414	(163,955)	191,026	(197,953)
%Female in Management	.26	(.31)	.36	(.35)	.17	(.24)
N	450,917		162,780		288,137	
N Establishments	8,792					

Note: This table shows summary statistics of all workers and jobs in our sample on June 30th 2014. Sampling weights are used to make this a representative sample of the German population.

Table 2: OOI by Demographics

	Dep Var: Outside Option Index			
	(1)	(2)	(3)	(4)
Female	-.237***	-.231***	-.227***	-.167***
	(0.009)	(0.011)	(0.008)	(0.008)
School	-.660***	-.620***	-.596***	-.617***
Lower-Secondary	(0.013)	(0.013)	(0.009)	(0.010)
School	-.279***	-.277***	-.211***	-.216***
Intermediate	(0.011)	(0.011)	(0.007)	(0.008)
Non-Citizen	-.307***	-.295***	-.459***	-.501***
	(0.031)	(0.027)	(0.021)	(0.019)
Age	.099***	.107***	.107***	.099***
	(0.004)	(0.003)	(0.002)	(0.002)
Age^2	-.001***	-.001***	-.001***	-.001***
	(4e-05)	(4e-05)	(3e-05)	(2e-05)
District	.112***	.104***		
Density	(0.004)	(0.005)		
Training Occ FE		X	X	
District FE			X	
Establishment FE				X
R^2	0.13	0.29	0.66	0.56
N	380,109	380,109	380,109	380,109

Notes: This table shows the results of a regression of OOI on basic demographics (Equation 12). The sample includes all workers employed on June 30th 2014. Sampling weights are used to make this a representative sample of the German population. Training occupation fixed effects are at the 3-digit levels.

Table 3: Relative Income by OOI After Mass Layoff

OOI_i coefficient	(1)	(2)	(3)
3 months (λ_3)	.061**	.062**	.068**
	(.029)	(.029)	(.031)
6 months (λ_6)	.068**	.069**	.082**
	(.030)	(.030)	(.033)
12 months (λ_{12})	.061*	.064*	.079**
	(.034)	(.034)	(.038)
24 months (λ_{24})	.033	.039	.064
	(.042)	(.042)	(.048)
Mass-Layoff×Month FE	Y	Y	Y
Tenure		Y	Y
Age			Y
Education			Y
Gender			Y
No. of Observations	558,686	558,686	558,686
No. of Workers	13,707	13,707	13,707
No. of Mass-Layoffs×Month	26,561	26,561	26,561

Note: This table shows the results of regressing relative income on OOI for workers that lost their jobs in a mass-layoff, for different times after the separation. Relative income is defined as the current daily income in that month divided by the last daily income before the layoff. Mass layoffs are defined as an establishment with at least 50 workers that reduced its workforce by at least 30% in a given year. The sample includes only workers who have worked for at least three years before the layoff and are below the age of 55. We include monthly income for the 36 months following the separation. The regression is based on Equation 16. Tenure includes a quadratic polynomial for days at the previous establishment. Age includes a quadratic polynomial. Education is a categorical variable for the type of secondary education (see section 3.2 for details).

Table 4: Impact of Express Trains on Options and Wages

First-Stage 1999-2012 (1)	Reduced-Form 1999-2012 (2)	2SLS 1999-2012 (3)	Reduced-Form 1993-1999 (4)	OLS 1999-2012 (5)
.073***	.024***	.324***	.002	.004
(.003)	(.004)	(.048)	(.002)	(.007)

Number of observations:	143,313
Number of treated observations:	37,695

Notes: This table shows the results of the impact of express trains on outside options, and wages. Columns 1-4 use nearest-neighbor matching with replacement. Matching is done exactly on gender, education group, citizenship status, state and 2-digit training occupation and continuously on age, and PCA components for training occupation (the third digit). The outcome variables are change in OOI 1999-2012 (column 1), change in log income 1999-2012 (columns 2,3,5) and change in log income 1993-1999 (column 4). Standard errors in matching are calculated using Abadie and Imbens (2006). 2SLS estimator is the division of the estimates in column 1 and 2 (Equation 17). Standard errors in column 3 are calculated using a method building on Abadie and Imbens (2006) (see Appendix B.4 for details). OLS (column 5) estimates the regression of log wages on OOI with match fixed effects. Observations from the control group that appear in multiple matches also appear multiple times in the OLS. Standard errors are clustered for workers with the same variables we match on exactly to account for the replacement (see Appendix B.4 for details).

Table 5: Effect of OOI on Wages Using Shift-Share (Bartik) Instrument

	First-Stage (1)	Reduced-Form (2)	2SLS (3)
	.622*** (.241)	.106*** (.056)	.170*** (.064)
Industry FE	X	X	X
N	408,792	408,792	408,792
Number of Clusters	38	38	38

Notes: This table shows the results of the impact of a shift-share instrument (Bartik) on outside options, and wages (column 1 and 2). This captures the effect of changes in OOI on changes in log wages between 2004-2014, when we instrument for the changes in OOI with the shift-share instrument (column 3). The instrument is constructed from an average of a 3-digit industry national employment growth weighted by the initial share of every industry in a region (see Section 6.2). The outcome variables are the change in OOI 2004-2014 (column 1) and change in log daily wages (columns 2 and 3). All columns control for industry (in 2004) and age. Standard errors are clustered within the unit of treatment, which is regions.

Table 6: Shift-Share (Bartik) Results by Exporting Share of Sales

	Export>33% (1)	33%≥Export≥1% (2)	1%>Export (3)
OOI	.105**	.593**	.132
	(.052)	(.266)	(.141)
Industry FE	X	X	X
N	119,645	146,217	142,930

Notes: This table shows the results of the impact of OOI on wages, instrumented with a shift share instrument, calculated separately by share of export in the industry. Share of export is calculated for every 3-digit industry based on the establishment panel survey in 2014. The sample is split based on the worker industry in 2004. Outcome variable is change in log wages between 2004-2014. The dependent variable is change in OOI between 2004-2014. The instrument is constructed from an average of a 3-digit industry national employment growth weighted by the initial share of every industry in a region (see Section 6.2). All columns control for industry (in 2004), and age. Standard errors are clustered within the unit of treatment, which is regions.

Table 7: Shift-Share (Bartik) Results by Stayers and Movers

	(1)	(2)
OOI	.170***	.257***
	(.064)	(.092)
$OOI\times$ $Stay$		-.159***
		(.062)
Industry FE	X	X
N	408,792	408,792

Notes: This table shows the results of the impact of OOI on wages, instrumented with a shift share instrument, interacted with whether a worker stayed at the same establishment. Outcome variable is change in log wages between 2004-2014. The dependent variable is change in OOI between 2004-2014. The instrument is constructed from an average of a 3-digit industry national employment growth weighted by the initial share of every industry in a region (see Section 6.2). The indicator for stay is 1 if the worker works at the same establishment on June 30th of both 2004 and 2014. All columns control for industry (in 2004), and age. Standard errors are clustered within the unit of treatment, which is regions.

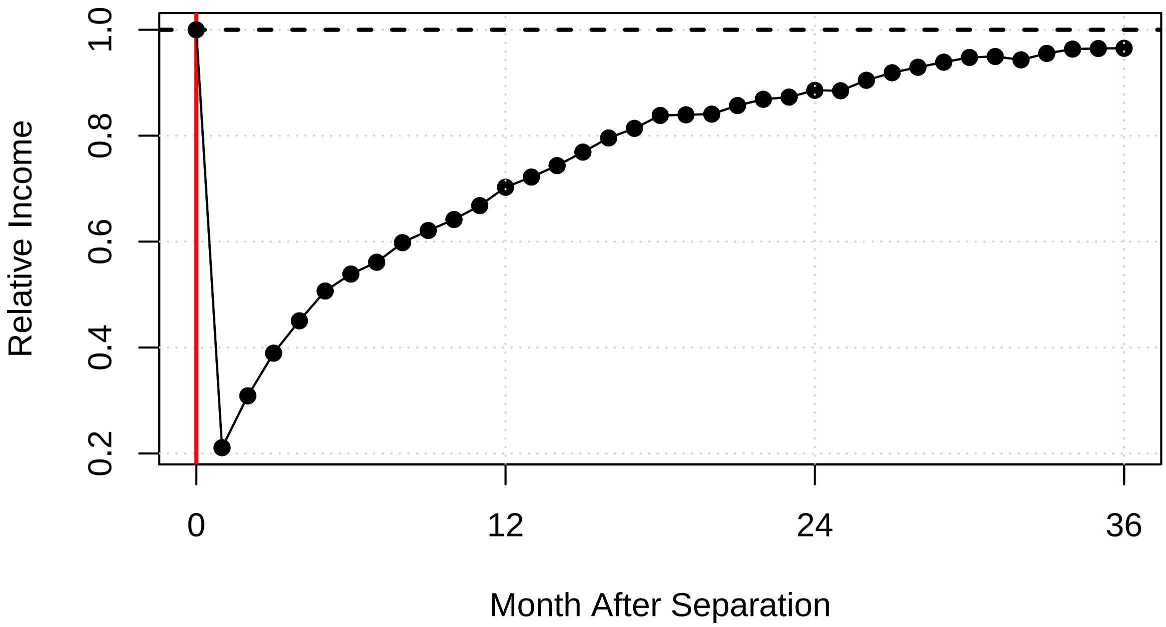
Table 8: Mincer Equation with OOI

	Dep Var: Log Daily Income			
	OLS	OLS	Tobit	Tobit
	(1)	(2)	(3)	(4)
OOI (fixed)		-.269		-.269
Female	-.171***	-.111***	-.195***	-.137***
	(0.011)	(0.012)	(0.011)	(0.012)
School Lower-Secondary	-.351***	-.174***	-.404***	-.230***
	(0.011)	(0.012)	(0.011)	(0.012)
School Intermediate	-.245***	-.170***	-.289***	-.217***
	(0.010)	(0.010)	(0.010)	(0.011)
German	.089**	.007	.093***	.011
	(0.031)	(0.034)	(0.032)	(0.034)
Age	.057***	.030***	.061***	.035***
	(0.003)	(0.004)	(0.003)	(0.004)
Age^2 (x 10^-3)	-.573***	-.261***	-.608***	-.297***
	(0.039)	(0.042)	(0.040)	(0.042)
District Density (log)	0.022***	-0.008**	0.023	-0.008**
	(0.003)	(0.003)	(0.003)	(0.003)
Part-Time	-.913***	-.905***	-.928***	-.921***
	(0.014)	(0.015)	(0.015)	(0.015)
N	378,776	378,776	378,776	378,776

Notes: This table shows the results from a regression of log wages on demographics and OOI. The coefficient of the OOI is fixed to be its point estimate from the 2SLS estimate based on the high-speed commuter rail introduction (Table 4). A Tobit model is used in Columns 3-4 to account for top coding of daily income at 195 Euros per day. OLS results use winsorized log income. Sampling weights are used to make this a representative sample of the German population

A Appendix

Figure A1: Relative Income Following Mass-Layoff



Note: This figure shows the relative income for workers who lost their jobs in mass-layoffs, for each month in the three years after the layoff. Relative income is defined as the current daily income in that month divided by the last daily income before the layoff. Mass layoffs are defined as an establishment with at least 50 workers that reduced its workforce by at least 30% in a given year. The sample includes only workers who have worked for at least three years before the layoff and are below the age of 55. The values are calculated using a regression of relative income on months after separation, with a fixed effect for every mass-layoff (Equation 13).

Table A1: Most Weighted Question in PCA - Establishment 2014 Survey

Name	N	Comp 1	Comp 2
Business Performance	8,792	Member of chamber of industry	Profit category
Investment & Innovation	8,792	IT investment	Total investment
Hours	8,792	Long leaves policy	Flextime
In-Company Training	8,792	Internal courses	Share workers in training
Vocational Training	8,792	Offer apprenticeship	Ability to fill
General	8,792	Family managed	Staff representation

This table shows the survey question that received the most weight in this principal component. We take the first two principal component from each survey category.

Table A2: Most Weighted Question in PCA - BIBB

Name	N	Comp 1	Comp 2
Hours	11,021	Sundays and public holidays	hours per week like to work
Type of Task	15,035	responsibility for other people	Cleaning, waste, recycling
Requirements	10,904	Acute pressure & deadlines	Highly specific Regulations
Physical	20,036	Oil, dirt, grease, grime	pathogens, bacteria
Mental	17,790	Support from colleagues	Often missing information

This table shows the survey question that received the most weight in this principal component. We take the first two principal component from each survey category.

Table A3: Commuting Distance by Gender and Education

	Distance from Job (Miles)	
	Mean	SD
All	15.5	41.9
Female	12.1	37.1
Male	17.4	44.3
Lower-Secondary	9.4	27.9
Intermediate-Secondary	11.4	34.4
Higher-Secondary	26.2	56.1

Values are mean distance in miles between workers previous place of residence and their job.

Table A4: Summary Stats for Mass-Layoff Workers by Treatment Status

	Above Median OOI		Below Median OOI	
	Mean	SD	Mean	SD
Female	.36	.48	.44	.50
Age	40.0	9.7	37.2	11.3
Higher-Secondary Education	.21	.41	.14	.35
Tenure in Establishment (days)	2316.3	1272.3	2167.4	1197.8
Daily Income	63.8	43.1	57.5	42.2
N	6.839		6,887	

Note: This table shows the summary stats for workers that lost their jobs in a mass-layoff above and below the establishment median OOI. Mass layoffs are defined as an establishment with at least 50 workers that reduced its workforce by at least 30% in a given year. The sample includes only workers who have worked for at least three years before the layoff and are below the age of 55. We include monthly income for the 36 months following the separation.

Table A5: Correlation Between OOI and $\log wage$ - OLS Results

	Dep.Var: $\log \text{wage}_i$		
OOI	.107***	.027***	-.010*
	(.005)	(.005)	(.005)
Demographics		X	X
District FE			X
N	378,776		

Note: Demographics include gender, education group, a quadratic in age and citizenship status. The sample includes all workers employed on June 30th 2014. Sampling weights are used to make this a representative sample of the German population.

Table A6: Balance Table - High Speed Train

	Treatment		Control	
	Mean	SD	Mean	Sd
$\log wage$ (1993)	3.26	2.49	3.27	2.49
$\log wage$ (1999)	4.25	.62	4.27	.58
Female	.356	.479	.356	.479
Age	36.4	6.8	36.4	6.7
Citizen	.995	.073	.995	.073
Low-Secondary	.257	.437	.257	.437
Intermediate-Secondary	.508	.500	.508	.500
High-Secondary	.235	.424	.235	.424
N	37,695		26,963	

Note: This table shows the summary stats for workers used to estimate the impact of high-speed trains on OOI and log wages. Treated group includes workers who lived in districts in which a new station was introduced between 1999-2012. Control group was chosen from a pool of workers living in districts that never got a station. Control workers were chosen through nearest-neighbor matching with replacement on gender, age, citizenship, education level, training occupation, state (Bundesländer) and lagged income. We require the match to be exact on gender, education, state and 2-digit occupation.

Table A7: Top Standardized Values of A

Variable (X)	Variable (Z)	A_{xz}
Distance		-4.15
Train Occ - Physical Cond. 1	Occ - Physical Cond. 1	1.477
Train Occ - Task Type 2	Occ - Task Type 2	1.077
Train Occ - Task Type 2	Occ - Physical Cond. 1	-.93
Train Occ - Physical Cond. 1	Occ - Task Type 2	-.82
Lower Secondary Education	Distance	-.74
Intermediate Education	Distance	-.61
Train Occ - Contract 2	Occ - Contract 2	.56
Train Occ - Task Type 1	Occ - Task Type 1	.55
Lower/Intermediate Education	Distance	-.54

Results from logistic regression for dummy variable on real vs. simulated match, on interaction of worker and job characteristics (Equation 10). Results are standardized, such that each variable has standard deviation of 1.

Table A8: Distance Coefficient by Demographics

Baseline	-.141
Female	-.024
Non-Citizen	-.019
Lower-Secondary	-.037
Intermediate-Secondary	-.012
Age	.002
Age^2 ($\times 10^{-3}$)	-.026

Results from logistic regression for dummy variable on real vs. simulated match, on interaction of worker and job characteristics (Equation 10). Baseline category is a forty years old high-secondary male.

B Theoretical Appendix

B.1 Continuous Logit Distribution

We follow Dagsvik (1994) in defining the continuous logit that produces ε_{i,z_j} and ε_{j,z_i}. In this section we define the distribution of ε_{i,z_j} and the distribution of ε_{j,x_i} is defined similarly.

Every worker $i \in \mathcal{I}$ draws ε_{i,z_j} shocks from a Poisson process on $\mathcal{Z} \times \mathbb{R}$ with intensity

$$f(z)\, dz \times e^{-\epsilon} d\epsilon$$

This is different from the Poisson process used in Dupuy and Galichon (2014) as the density $f(z)$ also affects the intensity, which allows this distribution to be properly defined over a larger class of functions for $\tau(x,z)$, including a constant, or simple polynomials. Denoting by P_i the infinite but countable points chosen in the process, every worker has a set

$$\left\{\varepsilon_{iz_j} = \alpha\epsilon \middle| (z,\epsilon) \in P_i\right\}$$

This process yields a distribution of ε_{i,z_j} that has several similarities to finite extremum value type-1 distribution. These similarities are all derived from one basic property of this point process.

Proposition 1. *Let $g : \mathcal{Z} \to \mathbb{R}$ be a function that satisfies*

$$\int_{\mathcal{Z}} e^{g(z)} f(z)\, dz < \infty$$

and let $S \subseteq \mathcal{Z}$ be some Borel measurable subset. Define

$$\psi_S^g = \max_{z \in S \cap P_Z} \left\{g(z) + \varepsilon_{i,z_j}\right\}$$

Then

$$\psi_S^g \sim EV_1\left(\alpha \log \int_S \exp \frac{g(z)}{\alpha} f(z)\, dz, \alpha\right)$$

and

$$S_1 \cap S_2 = \phi \iff \psi_{S_1}^{g_1} \perp \psi_{S_2}^{g_2}$$

Proof. This proposition stems from the fact that in a Poisson process, the amount of points chosen in two disjoint Borel measurable sets B_1, B_2 has an independent distribution $N(B_i) \sim$

$Poisson\left(\Lambda\left(B_I\right)\right)$ with

$$\Lambda\left(B_i\right)=\int_{B_i}\lambda\left(x\right)dx$$

Therefore, in our context the cumulative distribution function of ψ_S^g is

$$P\left(\psi_S^g\leq x\right)=P\left(N\left(\left(S\times\mathbb{R}\right)\cap\left\{g\left(z\right)+\alpha\epsilon>x\right\}\right)=0\right)$$

From the Poisson distribution this is

$$\begin{aligned}\log P\left(\psi_S^g<x\right) &= -\Lambda\left(S\times\left\{\epsilon>\tfrac{x-g(z)}{\alpha}\right\}\right)\\ &= -\int_S\int_{\frac{x-g(z)}{\alpha}}f\left(z\right)e^{-\epsilon}dzd\epsilon\\ &= -\int_S e^{-\frac{x-g(z)}{\alpha}}f\left(z\right)dz\\ &= -\exp\left[-\frac{x-\alpha\log\int_S\exp\frac{g(z)}{\alpha}f(z)dz}{\alpha}\right]\end{aligned}$$

which is exactly a cumulative distribution function of $EV_1\left(\alpha\log\int_S\exp\frac{g(z)}{\alpha}f\left(z\right)dz,\alpha\right)$.

Since every draw of points in a Poisson process is independent, $S_1\cap S_2\iff\psi_{S_1}^{g_1}\perp\psi_{S_2}^{g_2}$. $\square$

This Proposition has several important implications for our context. It implies that even though ε_{i,z_j} is not defined for every $z\in\mathcal{Z}$, it is defined infinitely often for every Borel measurable subset that includes z, and the maximum for that set ψ_S^1 has an extreme-value type-1 distribution.

Since workers in equilibrium are getting a sum of a continuous function (which we mark by $\omega\left(x,z\right)$) and ε_{i,z_j} (Lemma 2), then we get that the maximum value they receive also has an EV_1 distribution, for every Borel measurable set of jobs. Moreover, the probability density to choose a particular observables z_j is similar to the finite case and its exact value is

$$f\left(z_j|i\right)=f\left(z_j=\arg\max_{z\in S\cap P_Z}\left\{g\left(z\right)+\varepsilon_{i,z_j}\right\}\right)=\frac{\exp\left[\frac{1}{\alpha}\omega\left(x,z_j\right)\right]f\left(z_j\right)}{\int_{\mathcal{Z}}\exp\left[\frac{1}{\alpha}\omega\left(x,z\right)\right]f\left(z\right)dz}$$

Another link to the finite multinomial logit can be drawn if we divide $\mathcal{Z}$ into a finite number of disjoint sets $\mathcal{Z}=\bigcup_{i=1}^n S_i, S_i\cap S_j=\phi$. Then the value of the best job for worker i in each subset ($\psi_{S_i}^{\omega}$) is EV_1 distributed. The choice of the best job characteristics $z_{m(i)}$ would be made with a finite multinomial logit, over these n options. When we increase n, the sets become smaller, and the choice becomes closer to an infinite options choice.

Note that in a standard multinomial logit, increasing the number of options to infinity will yield an infinite compensation, but this is not the case here. This is because when the number of options n grow, the mean measure of S_i decreases in a rate of $\frac{1}{n}$. Therefore the location parameter of each one of the choices, decreases in a rate of $\frac{1}{n}$ as well from the proposition.

B.2 Proofs

Proof of Lemma 1

Part 1: We will start by formally defining the densities we are using. We will use $\mathcal{I}^{\phi}, \mathcal{J}^{\phi}$ to mark the set of unmatched workers and jobs.

Definition 1. Let $f(i,j) : \mathcal{I} \times \mathcal{J} \to \mathbb{R}_{\geq 0}$ be the density that satisfies for every Borel measurable subset of potential matches $B \subseteq \mathcal{I} \times \mathcal{J}$

$$\int_B f(i,j)\, didj = \frac{\mu(B \cap M)}{\mu(M) + \mu(\mathcal{I}^{\phi}) + \mu(\mathcal{J}^{\phi})}$$

where μ is the measure function.

Intuitively, this is the joint density of observing i and j matched in equilibrium. Similarly we define a density over the probability of observing worker and job with specific characteristics matched in equilibrium.

Definition 2. Let $f(x,z) : \mathcal{X} \times \mathcal{Z} \to \mathbb{R}_{\geq 0}$ be the density that satisfies

$$f(x,z) = \int_{X_i = x} \int_{Z_j = z} f(i,j)\, didj$$

From these definitions we can derive the conditional distribution of a match for a given worker.

Definition 3. Let f_j^i be

$$f_j^i = f(j|i) = \frac{f(i,j)}{f(i)} = \frac{f(i,j)}{I^{-1}}$$

Part 2: Let $i, i' \in \mathcal{I}$, $j, j' \in \mathcal{J}$ with $X_i = X_{i'}$ and $Z_j = Z_{j'}$ From Assumption 1 τ_{ij} has the same distribution as $\tau_{i'j'}$, and therefore $f(i,j) = f(i', j')$.

Hence, from Definition 2,

$$f(X_i, Z_j) = I f(X_i)\, J f(Z_j)\, f(i,j)$$

and from Definition 3

$$f_j^i = \frac{f(X_i, Z_j)}{f(X_i)\, f(Z_j)} J^{-1}$$

and we normalized $J = 1$.

Proof of Lemma 2

Let $i, i' \in \mathcal{I}$ with $X_i = X_{i'} = x_0$ and $j, j' \in \mathcal{J}$ with $Z_j = Z_{j'} = z_0$, where $m(i) = j$ and $m(i') = j'$. The sum of compensation equals the total surplus, hence

$$\omega_{ij} + \pi_{ij} = \tau(x_0, z_0) + \varepsilon_{i,z_0} + \varepsilon_{j,x_0}$$

$$\omega_{i'j'} + \pi_{i'j'} = \tau(x_0, z_0) + \varepsilon_{i',z_0} + \varepsilon_{j',x_0}$$

For stability, it must be that

$$\omega_{ij} + \pi_{i'j'} \geq \tau(x_0, z_0) + \varepsilon_{i,z_0} + \varepsilon_{j',x_0}$$

$$\omega_{i'j'} + \pi_{ij} \geq \tau(x_0, z_0) + \varepsilon_{i',z_0} + \varepsilon_{j',x_0}$$

Note that the sum of the two weak-inequalities is equal to the sum of the two equalities, therefore they must hold with equality (otherwise, the sum should hold both as an equality and strong inequality). Hence, we can rewrite

$$\omega_{ij} - \omega_{i'j'} = \varepsilon_{i,z_0} - \varepsilon_{i',z_0}$$

$$\pi_{ij} - \pi_{i'j'} = \varepsilon_{j,x_0} - \varepsilon_{j',x_0}$$

In other words, compensation for workers and employers in matches with the same characteristics is constant up to their value of ε, so we can write

$$\omega_{ij} = \omega(x_0, z_0) + \varepsilon_{i,z_0}$$

$$\pi_{ij} = \pi(x_0, z_0) + \varepsilon_{j,x_0}$$

$$\omega(x_0, z_0) + \pi(x_0, z_0) = \tau(x_0, z_0)$$

We can also pin down the alternative offers

$$\omega_{ij'} = \tau_{ij'} - \pi_{j'} = \tau(x_0, z_0) + \varepsilon_{i,z_0} + \varepsilon_{j',x_0} - \pi(x_0, z_0) - \varepsilon_{j',x_0} = \omega(x_0, z_0) + \varepsilon_{i,z_0} = \omega_{ij}$$

This implies that all employers with $Z_j = z_0$ who are matched with $X_{m^{-1}(j)} = x_0$ are willing to make the same offer. Therefore, both workers and employers are facing a continuous logit choice. Hence, we can link the values of $\omega(x_0, z_0)$ and $\pi(x_0, z_0)$ to their choice probabilities (see Appendix B.1):

$$f(x_0|z_0) = \frac{\exp\left[\frac{1}{\alpha}\pi(x_0, z_0)\right] f(x_0)}{\int_{\mathcal{X}} \exp\left[\frac{1}{\alpha}\pi(x, z_0)\right] f(x)\, dx}$$

The denominator is the expected value π_j, which is a function of $Z_j = z_0$ so we can rewrite it as $\pi(z_0)$. Taking logs we get

$$\alpha \log f(x_0|z_0) = \pi(x_0, z_0) + \log f(x_0) - \pi(z_0)$$

and with Lemma 1

$$\pi(x_0, z_0) = \alpha \log f_j^i + \alpha \log J + \pi(z_0)$$

therefore

$$\omega_{ij'} = \tau(x_0, z_0) - \pi(z_0) + \alpha \log f_j^i + \varepsilon_{i,z_0}$$

where J was pinned to 1.

Proof of Lemma 3

First equality is by definition, and because the first best and second best options are equivalent. We showed that $\omega_{ij} = \omega(x_i, z_j) + \varepsilon_{i,z_j}$. The expected compensation of worker i is

$$\omega(x_i) = E\left[\omega^*(x_i, z_j)\right] + E\left[\varepsilon^*_{i,z_j}\right]$$

From the continuous logit structure we know that (similar to the previous proof)

$$\omega(x_i, z_j) = \alpha \log f_j^i + \omega(x_i)$$

hence

$$\omega(x_i) = E\left[\alpha \log f_j^i + \omega(x_i)\right] + E\left[\varepsilon^*_{i,z_j}\right]$$

Therefore

$$E\left[\varepsilon^*_{i,z_j}\right] = -\alpha \int f_j^i \log f_j^i dj$$

Similarly for ε_{j,x_0} and combinedly:

$$E\left[\varepsilon^*_{i,z_j} + \varepsilon^*_{j,x_0}\right] = -\alpha \int f_j^i \log f_j^i dj$$

Proof of Theorem 2

Following the notations from the previous proofs. The ω_{ij} offer can be written as

$$\omega_{ij} = \omega(x_i, z_j) + \varepsilon_{i,z_j}$$

and ε_{i,z_j} is unaffected by λ hence

$$\frac{d\omega_{i,j}}{d\lambda_i} = \frac{d\omega(x_i, z_j)}{d\lambda_i}$$

In the previous proofs we showed that

$$\omega(x_i, z_j) = \alpha \log f_j^i + \omega(x_i)$$

$$\pi(x_i, z_j) = \alpha \log f_j^i + \pi(z_j)$$

hence

$$\omega(x_i, z_j) - \pi(x_i, z_j) = \omega(x_i) - \pi(z_j)$$

Adding $\tau(x_i, z_j)$ and dividing by 2:

$$\omega(x_i, z_j) = \frac{1}{2}\left(\tau(x_i, z_j) - \pi(z_j) + \omega(x_i)\right)$$

$\tau(x_i, z_j), \pi(z_j)$ don't change by the definition of λ. Hence the only effect is on $\omega(x_i)$.

$$\frac{d\omega_{i,j}}{d\lambda_i} = \frac{1}{2}\frac{d\omega(x_i)}{d\lambda_i}$$

We get the value for $\omega(x_i)$ from the decomposition in Equation 6. Since $\tau(x_i, z_{j'}), \pi(z_{j'})$ remain constant the remaining effect is on the OOI.

$$\frac{d\omega_{i,j}}{d\lambda_i} = \alpha\frac{dOOI_i}{d\lambda_i}$$

Proof of Theorem 2

This is similar to before, only that ε_{i,z_j} is allowed to change as well. Since $E\left[\varepsilon_{i,z_j}\right] = \alpha OOI$ we get the effect from the previous lemma, in addition to the effect on the OOI.

$$\frac{d\omega_{i,j}}{d\lambda_i} = 2\alpha\frac{dOOI_i}{d\lambda_i}$$

Alternative Definitions for λ

Assume workers and equally distributed across the real line (as in Section 2.5). Each worker is a 3-dimensional tuple (l_i, y_i, c_i) and τ_{ij} is defined as

$$\tau_{ij} = y_i - c_i \left|l_i - l_j\right| + \varepsilon_{ij}$$

Now workers log density is a triangular function, with its peak at l_i (Laplace). Hence,

$$f_j^i = \frac{c_i}{2} \exp -c_i \left|l_i - l_j\right|$$

The OOI is (shifting l_i to 0)

$$\int_0^\infty c \exp^{-cl} \left(\log \frac{c}{2} - cl\right) dl = \log \frac{c}{2} - 1$$

The mean value $E\left[\tau\left(x, z\right)\right]$ is

$$\int_0^\infty c \exp^{-cl} \left(y_i - cl\right) = y_i - 1$$

Hence, setting the commuting cost c only affects the OOI but not net productivity and can be served as λ. This will also work for more general settings, as long as worker and job locations are not correlated with locations.

Another example is to define λ as the intensity of the Poisson process for the continuous logit process. Higher λ will mean more options on average in every subset of jobs.

B.3 $f(x, z)$ Estimation

To estimate a logistic regression following Equation 10, we maximize the following likelihood

$$\max_\theta \sum_k \log P\left(y_k | x_k, z_k; \theta\right)$$

where θ are the parameters defined in this equation, including matrix A. We rewrite Equation 10 in a more general form. Note $p_k\left(\theta\right) = P\left(Y_k = 1 | X = x_k, Z = z_k\right)$:

$$\log \frac{p_k\left(\theta\right)}{1 - p_k\left(\theta\right)} = \sum_{j=1}^K \beta_j h_j\left(x_k, z_k\right)$$

where K is the number of moments h_j we control for in this regression.

Then the K FOC of this maximization converge asymptotically to

$$E\left[p_k\left(\theta\right)h_j\left(x_k,z_k\right)\right]=E\left[h\left(x_k,z_k\right)|y_k=1\right]s$$

where $s=P\left(Y=1\right)$ is the share of real data (in our case $\frac{1}{2}$). Using $\frac{p_k(\theta)}{1-p_k(\theta)}=\frac{f(x,z)}{f(x)f(z)}\frac{s}{1-s}$ we can write

$$E\left[\frac{f\left(x,z\right)}{sf\left(x,z\right)+\left(1-s\right)f\left(x\right)f\left(z\right)}h_j\left(x,z\right)\right]=E\left[h_j\left(x,z\right)|real\right]$$

The RHS is simply the moment of $h_j\left(x,z\right)$ in the real data. The LHS is the moment of $h_j\left(x,z\right)$ in the full data (real and simulated), weighted by the probability it is real.

If the model is correctly specified and the functional form assumption on $\frac{f(x,z)}{f(x)f(z)}$ is true, θ will be estimated consistently. This is because

$$E\left[\frac{f\left(x,z\right)}{sf\left(x,z\right)+\left(1-s\right)f\left(x\right)f\left(z\right)}h_j\left(x,z\right)\right]=$$

$$\int\frac{f\left(x,z\right)}{sf\left(x,z\right)+\left(1-s\right)f\left(x\right)f\left(z\right)}h_j\left(x,z\right)\left(sf\left(x,z\right)+\left(1-s\right)f\left(x\right)f\left(z\right)\right)dxdz=$$

$$=\int h\left(x,z\right)f\left(x,z\right)dxdz=E\left[h\left(x,z\right)|real\right]$$

If the model is misspecified, our estimate of $\frac{f(x,z)}{sf(x,z)+(1-s)f(x)f(z)}$ will not be converging to the real density rations. Instead we will equalize moments of some other weighted average of h_j

$$E\left[w\left(x,z,\theta\right)h_j\left(x,z\right)\right]=E\left[h_j\left(x,z\right)|real\right]$$

where

$$w\left(x_k,z_k,\theta\right)=s^{-1}\frac{\exp\sum_{j=1}^{K}\beta_jh_j\left(x_k,z_k\right)}{1+\exp\sum_{j=1}^{K}\beta_jh_j\left(x_k,z_k\right)}$$

We next analyze these weights as $s\to0$. We will mark $h_1\left(x,z\right)=1$, the offset of the regression. When $s\to0$, $\frac{p(\theta)}{1-p(\theta)}\to0$ as well, therefore $\exp\sum_{j=1}^{K}\beta_jh_j\left(x_k,z_k\right)\to0$. With some abuse of notation, we will redefine β_1 as $\beta_1-\log s$. Therefore

$$\lim_{s\to0}w\left(x,z,\theta\right)=\exp\sum_{j=1}^{K}\beta_jh_j\left(x_k,z_k\right)=\frac{\widehat{f\left(x,z\right)}}{f\left(x\right)f\left(z\right)}$$

The density of the full data approaches the density of the simulated data. Hence overall, we get

$$E\left[w\left(x,z,\theta\right)h_j\left(x,z\right)|sim\right]=E\left[h_j\left(x,z\right)|real\right]$$

In order to calculate the OOI, we simulate values from $f(x) f(z)$, and reweight them based on $\widehat{\frac{f(x,z)}{f(x)f(z)}}$. This is because we hold workers fixed, and simulate z values from $f(z)$. As $s \to 0$ we use weights that converge to $w(x, z, \theta)$. The above equation guarantees that we sample from a distribution with same moment value for every $h_j(x, z)$, even if the model is misspecified.

Dupuy and Galichon (2014) produce a distribution with the same second moments as the data, and same marginal distributions. Therefore, when $s \to 0$, and h include all X, Z interactions, and an indicator for every x_k, and every z_k value (that is, $h(x, z) = 1_{x=x_k}$ or $h(x, z) = 1_{z=z_k}$ for every k), we get the same distribution.

B.4 Standard Errors for a Wald Estimator with Matching

We want to estimate the standard errors of $\widehat{\alpha}$, defined in Equation 17. Both the nominator (reduced form), and the denominator (first stage) are standard matching estimators for average treatment effect on treated (ATET). Abadie and Imbens (2006) show how to estimate standard errors for ATET. But to estimate correctly the standard error for the Wald estimator, we also need to estimate the covariance of the first stage and reduced form. So we extend their approach for this case.

Mark the ATET on log wages (reduced form), and OOI (first stage) as:

$$\begin{aligned} \rho &= E\left[\log w(1) - \log w(0) \,|T = 1\right] \\ \gamma &= E\left[OOI(1) - OOI(0) \,|T = 1\right] \end{aligned}$$

where (1) means value when treated and (0) when not treated. T is treatment status (so this is the mean effect for treated).

The Wald estimator is then

$$\alpha = \frac{\rho}{\gamma}$$

For each match m (treated unit and one or more control unit), define

$$X_m = \begin{pmatrix} X_{1m} \\ X_{2m} \end{pmatrix} = \begin{pmatrix} \log w(1) - \overline{\log w(0)} \\ OOI(1) - \overline{OOI(0)} \end{pmatrix}$$

Our estimators are then simply

$$\begin{pmatrix} \widehat{\rho} \\ \widehat{\gamma} \end{pmatrix} = \overline{X_m}$$

Asymptotically

$$\overline{X_m} \sim N\left(\left(\begin{array}{c}\rho\\ \gamma\end{array}\right),\left(\begin{array}{cc}\sigma_{11} & \sigma_{12}\\ \sigma_{12} & \sigma_{22}\end{array}\right)\right)$$

With the Delta method

$$V\left(\widehat{\alpha}\right)=\frac{1}{\gamma^2}\left(\sigma_{11}-2\frac{\rho}{\gamma}\sigma_{12}+\frac{\rho^2}{\gamma^2}\sigma_{22}\right)$$

Abadie and Imbens (2006) tells us how to find σ_{11}, σ_{22} which are $V\left(\widehat{\rho}\right), V\left(\widehat{\gamma}\right)$. We want to extend their approach to σ_{12}.

The challenge in getting the variance correctly for matching with replacement, is that the matches are not independent. Some observations from the control pool appear in more than one match. Following Abadie Imbens we write

$$V\left(\overline{X_m}\right)=\frac{1}{N_1}\sum_m\left(X_m-\overline{X_m}\right)^T\left(X_m-\overline{X_m}\right)+\frac{1}{N_1}\sum_{T=0}\left(K_i\left(K_i-1\right)\right)\widehat{V_i}$$

with

$$\widehat{V_i}=\widehat{V}\left(\begin{array}{c}\log w_i\\ OOI_i\end{array}\right)$$

where N_1 is number of treated units,K_i is the number of times observation i from the control pool was used. $\widehat{V_i}$ is a 2x2 matrix of the variance for that particular observation. The first part is a standard variance calculation. The second part corrects for the covariance between the matches.

If an observation i is used $K_i > 1$ times, then there are $K_i\left(K_i-1\right) > 0$ pairs of matches that both use it, and so their covariance is not 0, but includes $\widehat{V_i}$.

To estimate $\widehat{V_i}$ we follow Abadie and Imbens (2006) and use nearest neighbor from the control group. So for every control observation we find a match from the control group as well and write

$$\widehat{V_i}=\frac{1}{2}\left(\begin{array}{c}\log w_i-\log w_{m(i)}\\ OOI_i-OOI_{m(i)}\end{array}\right)\left(\begin{array}{c}\log w_i-\log w_{m(i)}\\ OOI_i-OOI_{m(i)}\end{array}\right)^T$$

This is asymptotically unbiased.

In practice, the only difference from Abadie and Imbens (2006) is that we also have a covariance component.

$$COV\left(\widehat{\rho},\widehat{\gamma}\right)$$

Which we estimate with

$$\frac{1}{N_1}\sum\left(\log w_m(1)-\overline{\log w_m(0)}\right)\left(OOI_m(1)-\overline{OOI_m(0)}\right)-\widehat{\rho}\widehat{\gamma}+$$

$$2*\frac{1}{2}*\sum_{T=0}K_i(K_i-1)\left(\log w_i-\log w_{m(i)}\right)\left(OOI_i-OOI_{m(i)}\right)$$

If our two variables were the same ($\log w = OOI$) then this would be the standard Abadie and Imbens (2006) formula for variance, as expected.

C Data Appendix

C.1 LIAB

In this section we clarify the coding of some of the variables we use. Our panel data, allows us to observe some variables several times in the data, and correct for coding errors. In particular, we set German citizenship to one, if this worker was ever reported as a German citizen by her employer.

We also take the highest level of education we observed until every year. All upper secondary school certificates are coded as upper-secondary. In some years intermediate and lower secondary education are coded with the same value. In these cases, if we observe the worker in other years and can infer their schooling level we use that. Otherwise, we code these workers in a separate category for either lower or intermedia secondary education.

For training occupation, we use the occupation in which workers spent the longest time in training. The LIAB data specify whether a worker is in vocational training and their occupation. For the large majority of workers, there is only one occupation in which they perform their vocational training. In rare cases where workers have conducted training in more than one occupation, we use the occupation in which the training was longer. If the we never observe the worker during vocational training, we take the occupation in which they conducted an internship. If this is unobserved as well, we use the first occupation they were observed in, as long as at least ten years have passed since we first observed them.

We calculate distance at the district level. For each district, we calculate the district center, by taking the weighted average of the latitude and longitude coordination of each city in this district. We then calculate the distance between the districts, taking into account the concavity of the earth.

C.2 BIBB Survey

In this section we describe in more detail the BIBB survey and PCA analysis.

We use data from the 2011-2012 wave of the German Qualification and Career Survey conducted by the Federal Institute of Vocational Training (BIBB) and the Institute for Labor Market Research (IAB). The data cover 20,000 employed individuals between the ages of 16 and 65. We run PCA on this survey by questions category and aggregate the results by 2-digit industry and 3-digit occupations. We link the results to our main data. The top question in each category are shown in Table A2.