How do Establishments Choose their Location? Taxes, Monopsony, and Productivity

Catherine van der List *

University of British Columbia †

August 21, 2022

Updated frequently. Click here for the latest version

Abstract

To better understand the distribution of economic activity across space and the effects of place-based policies, I develop a partial-equilibrium model of the location choice of new establishments incorporating tax rates, monopsonistic labor markets, and spillovers. Results of the estimating equation with German establishment data indicate that establishments have a preference for lower taxes and worker outside options in their location choice, though more productive (as measured by a higher AKM fixed effect) and larger establishments are less likely to account for reference wage levels in their decision making process. Establishments dislike taxes, though the degree of distaste varies between economic sectors. The degree to which various types of spillovers, or the benefits firms receive from other firms in a location, are an important influence on the location decision varies greatly between sectors. I quantify the effects of a theoretical place-based policy in each commuting zone in Germany. I find that commuting zones display highly heterogeneous responses to the same policy, with some areas experiencing wage and economic activity growth and others declines. The distribution of predicted effects of the place-based policy empirical estimates of spillovers from place-based policies in previous research.

*I would like to thank my advisors Thomas Lemieux, Florian Hoffmann, and Raffaele Saggio for their guidance and feedback. I would also like to thank Victor Couture, Anand Chopra, David Green, Joshua Gottlieb, Torsten Jaggard, and Marit Rehavi, as well as seminar participants at CEA 2022 and the Federal Reserve Bank of San Francisco for their helpful comments, and Mirko Titze and Urban Janisch for their assistance with geocode harmonizations. Any errors are mine.

†Vancouver School of Economics, University of British Columbia, 997-1873 East Mall, Vancouver, BC V6T 1Z1, Canada.

1 Introduction

Economic activity is distributed unevenly across geographic space - with some areas enjoying low unemployment and high wages while others have high unemployment rates and low wages. The location decisions of firms reinforce this: a large number of establishments choose to go to high wage high productivity areas, while a much smaller number of firms choose underdeveloped areas. How this location decision is undertaken by establishments is largely unknown despite being of critical interest to governments, who develop place-based policies in an attempt to revitalize underdeveloped regions. Some types of establishments may prefer to leverage monopsony power to be a dominant employer in a small market, while other types of firms likely prefer being located near specialized labor or in hubs of technological innovation. Furthermore, these incentives likely compete with one another: highly productive locations with specialized labor and technological innovation are larger labor markets where firms will be less likely to be able to exert monopsony power.

In this paper, I present evidence on the extent to which monopsony power, spillovers, and taxes influence the location decision of establishments, and test the effects of policies (in the form of tax incentives) designed to attract establishments to underdeveloped regions. Starting from a model of imperfect competition in segmented labor markets, I show that the importance of monopsony, spillovers, and taxes in the location decisions of new firms of discrete types can be estimated using methods from the differentiated products models of the industrial organization literature (see Berry 1994).

I implement this method using establishment-level data from Germany, combined with data proxying three Marshallian agglomerating forces[1], data on corporate-tax rates, and estimates of the outside option of workers. My estimating equation relates the share-ratio of new establishments choosing a particular location to a base location to measures of the extent to which other establishments in the location use similar types of goods and labor

[1]Establishments may locate near one another to be closer to goods sellers and/or buyers, to pool labor, or to receive knowledge from other nearby establishments.

inputs, the degree of innovation in the area, the average sectoral minimum wage in the location, and tax rates.

My empirical findings point to two main conclusions. First, establishments have heterogeneous preferences in the location characteristics they consider when they choose where to locate. Some sectors demonstrate a preference for locations where they are close to goods suppliers or buyers, and others locations where average wages are low. Second, establishments of all sectors with the exception of personal services have a distaste for taxes with the elasticity of the share of firms choosing a particular commuting zone with respect to taxes ranging from 5.35 to 13.65. In extensions, I show that firms with below-median productivity (as measured by the two-way fixed effects design of Abowd, Kramarz, and Margolis 1999, hereafter AKM) or below-median size respond more strongly to average wages in the commuting zone in their decision making process.

Using the estimated coefficients of the estimating equation, I simulate the effects of a policy designed to attract establishments to locations by setting corporate tax rates in these locations for new establishments to 7%, the lowest legal rate. I find that in the counterfactual world, there is a great deal of heterogeneity in the overall response to taxes, with the shape and magnitudes of the simulated changes in the number of establishments and wages closely matching reduced-form estimates in Greenstone et al. (2010). A case study of the effect of tax policies in former East Germany demonstrates the effect differences in initial conditions can have on the overall impact of a place-based policy.

2 Contribution

The literature on spatial sorting of establishments (Behrens et al. 2014, Gaubert 2018, Fajgelbaum et al. 2019, Bilal 2022) has traditionally focused on establishments entering perfectly-competitive labor markets. However, economists know from the recent literature on monopsony power (Ashenfelter et al. 2010; Falch 2010, 2011; Schmieder 2013; Azar et al. 2018) that markets are not perfectly competitive in-practice. Two papers have examined the location choices of establishments in the context of imperfectly competitive labor markets (Manning 2010, Bamford 2021), but both of these papers assume homogeneous firms. Gaubert (2018) assumes perfectly competitive labor markets and heterogenous firms to examine the location choice of firms, but concentrates on whether more productive firms are better able to take advantage of a single generic type of spillover with no structure. In contrast, this paper concentrates on heterogenous valuation of different types of spillovers and how that contributes to differences in productivity of individual industrial sectors across geographic space.

A strand of the urban economics literature (Glaeser and Gottlieb 2009, Ellison and Glaeser 1997, Dumais et al. 2002, Desmet and Rossi-Hansberg 2014, Glaeser et al 2015, Beaudry et al. 2018) has examined the development of labor demand in particular locations, and the ways in which industrial concentration may be influenced by the presence of spillovers, existing firms, and entrepreneurship. These papers develop frameworks for the measurement of industry concentration, as well as establish the different ways existing firms versus entrepreneurs influence labor demand in a particular location. What this literature does not do is examine the location decision of establishments - the literature typically does not allow establishments to make a decision between a menu of locations, but rather a discrete entrance decision in a particular location or empirically decomposes the sources of the change in labor demand over time. This excludes a key aspect of the development of labor demand in different regions - the fact that entrepreneurs and establishments are choosing comparatively between different locations.

Existing literature examining spillovers directly has concentrated on empirical measurement of the magnitude of spillovers or the different types of spillovers which may exist theoretically and how to distinguish between the them empirically (Moretti 2004; Rosenthal and Strange 2004; Glaeser and Gottlieb 2009; Greenstone, Hornbeck, and Moretti 2010; Ellison, Glaeser, and Kerr 2010; Hanlon and Miscio 2017). This literature primarily focuses on how to effectively measure and distinguish between these different types of spillovers, rather than how such spillovers may affect firm location choice. These papers suggest that there are multiple types of spillovers, and that establishments in different industrial sectors may value these spillovers differently.

While some papers (Kline and Moretti 2013, 2014; Gaubert 2018; Austin et al. 2018; Slattery and Zidar 2019) have examined how the presence of spillovers and different structures of labor demand in different regions may influence optimal government policy and macroeconomic outcomes, none of these papers develop a specific model of establishment choice weighing the relative importance of factors such as monopsony power and spillovers to firm location choice. Since in practice establishments are making a choice between various locations weighing a variety of factors, the location choices of establishments demonstrate an important aspect of the economic development of a region.

Existing research has examined the role of labor supply and structural issues in regional wage inequality, but comparatively little work examines the role played by labor demand (Gaubert 2018 and Bilal 2022). Existing research (Moretti 2013, Beaudry et al. 2013, Diamond 2015, Dauth et al. 2018, Boeri et al. 2019, Heise and Porzio 2018) has examined the importance of supply-side and structural factors to the development of diverging regional wages, particularly how these factors affect the relative well-being of low-skill and high-skill workers. None of these papers have rigorously modeled the behavior of establishments and how demand-side factors may influence the available jobs in a region. Dustmann et al. (2022) does examine how changes in the minimum wage may affect the types of firms which exist in a particular location, and the channel through which workers are affected by such types of structural changes. The results in their paper supports the existence of monopsonistic labor markets at the local level, and examines how structural factors may affect labor demand, and find that changes in labor demand from a structural change in the labor market has important implications for workers.

The remainder of this paper is organized as follows. Section 3 outlines the model of establishment location choice, Section 4 describes the data, Section 5 discusses the results, Section 6 examines the effects of a theoretical place-based policy, and Section 7 concludes.

3 Model

I model firms' location choice is a partial-equilibrium monopsonistic model where establishments choose wages to maximize their profits subject to the workers labor supply equations. As in Card, Cardoso, Heining, and Kline (2018), I use a static industrial-organization style differentiated products framework to describe how heterogeneous workers value jobs at different firms. Within this framework, I model the location decision of a new firm choosing which labor market to enter using a differentiated-products framework, where "products" are locations with different characteristics. Firms first determine their optimal wage for each labor market, and then choose which location to enter based on which market offers the firm the highest level of profit.

3.1 Allocation of Workers to Establishments Within Labor Markets

In a particular labor market c and time t, each establishment indexed by j in industrial sector sec posts a wage offer w_{cjt} which is fully and costlessly observed by all workers living in that market. Establishments are willing to hire any worker who will accept the job at the posted wage.

Workers have heterogeneous preferences over establishments, the utility function of worker i at firm j is given by:

$$u_{icjt} = \mu_c ln(w_{cjt} - b_{ct}) + a_{sec} + v_{icjt} \tag{1}$$

where b_{ct} is the outside option of workers living in location c, a_{sec} are sector-specific amenities valued equally by all workers, and v_{icjt} is an idiosyncratic preference shock of workers for working at establishment j drawn independently from a type I extreme value distribution. Such preference shocks could be, for example, commuting costs. Workers supply inelastic labor hours normalized to one. By the standard arguments of the McFadden choice model (McFadden 1973) this leads to the logit choice equation of workers.

$$p_{icjt} = P(\underset{k\epsilon 1,\ldots,J}{argmax} = j) = \frac{exp(\mu_c ln(w_{cjt} - b_{ct}) + a_{sec})}{\sum_{k=1}^{J} exp(\mu_c ln(w_{ckt} - b_{ct}) + a_{sec})} \tag{2}$$

Assuming the number of establishments is sufficiently large in each location that firms are not strategically interacting in their wage setting, this logit-choice equation may be approximated by the exponential probability.

$$p_{icjt} \approx \lambda_{ct} exp(\mu_c ln(w_{cjt} - b_{ct}) + a_{sec}) \tag{3}$$

where λ_{ct} is constant for all establishments in market c. Since aa establishment's number of employees is the available pool of workers in the location times the probability a worker chooses the establishment, the labor supply equation of the establishment directly follows.

$$L_{cjt}(w_{cjt}) = \mathcal{L}_{ct} \lambda_{ct} exp(\mu_c ln(w_{cjt} - b_{ct}) + a_{sec}) \tag{4}$$

where $\mathcal{L}_{ct}$ is the size of the labor force in market c.

3.2 Optimal Establishment Behavior Within a Market

Individual establishments maximize their profits conditional on the market they operate within by posting an optimal wage subject to the labor supply behavior of workers outlined above. An individual establishment's profit equation for market c is given by:

$$Y_{jc} = (1 - \tau_{ct})(\beta_{ct,sec} L_{cjt}(w_{cjt}) - L_{cjt}(w_{cjt}) w_{cjt}) \tag{5}$$

Establishments have a marginal product of labor $\beta_{ct,sec}$ which differs by industrial sector and location. I remain agnostic about the exact form of this productivity, but it may be thought of as a function of agglomeration, spillovers, natural advantage, and worker characteristics available to the establishment in a given location. These benefits are allowed to differ by industrial sector since it is plausible that different types of establishments value different types of agglomeration and spillovers differently (i.e. tech companies likely value intellectual knowledge spillovers more highly than retail trade establishments). Corporate taxes for each market are denoted by τ_{ct}.

Establishments cannot observe workers' idiosyncratic preference shocks v_{icjt}, so they post a single optimal wage by maximizing their profit equation subject to the labor supply equation 4. Using the first order condition of the profit equation and the labor supply equation, the optimal wage posted by the establishment is:

$$w_{cjt} = \frac{\mu_c}{1+\mu_c}\beta_{ct,sec} + \frac{1}{1+\mu_c}b_{ct} \tag{6}$$

This wage equation takes the form of a weighted average of the marginal product of labor and the outside option available to workers in the establishments' location. The form of the equation demonstrates the manner in which monopsony power is exerted by firms, as $\mu_c \to \infty$ markets become perfectly competitive. Furthermore, as worker outside options increase, wage levels must also increase, with the relative importance of productivity and outside option in wage setting determined by the elasticity of labor supply to the establishment μ_c. Of note, since productivity $\beta_{ct,sec}$ varies at the sector and commuting-zone level and b_{ct} varies at the commuting-zone level, wages are sector and commuting-zone specific rather than establishment specific.

Substituting labor supply and wage equations and log-linearizing leads to the log-profit equation.

$$y_{ctj} = ln(1-\tau_{ct}) + ln(\mathcal{L}_{ct}\lambda_{ct}) + (1+\mu_c)ln[\frac{1}{1+\mu_c}(\beta_{ct,sec} - b_{ct})] + a_{sec} + \mu_c ln(\mu_c) + u_{ctj} \quad (7)$$

Where u_{ctj} is an idiosyncratic log-profit shock with a type I extreme value distribution.

This log-profit equation has several key components. The first is taxes $ln(1-\tau_{ct})$, the second is a market size term $ln(\mathcal{L}_{ct}\lambda_{ct})$, and the third is the relative productivity of workers compared to the outside option $(1+\mu_c)ln[\frac{1}{1+\mu_c}(\beta_{jct} - b_{ct})]$.

3.3 Modeling Location Choice

Within this framework, I model the location choice of new establishments entering the labor market[2]. Establishments solve their location choice problem by first solving for the optimal wage they would pay in each individual market, then choosing the location where log profit is highest. Since the idiosyncratic shocks to profit in equation 7 are drawn from a type I extreme value distribution, establishments have the standard logit probability of locating in location c. Since there is no establishment-level heterogeneity in the explanatory variables, this logit probability approximates the share of establishments of a particular industrial sector which locate in location c.

$$\begin{aligned} s_{ct,sec} &\approx p_{ct,sec} \\ &= \frac{exp[ln(1-\tau_{ct}) + ln(\mathcal{L}_{ct}\lambda_{ct}) + (1+\mu_c)ln[\frac{1}{1+\mu_c}(\beta_{ct,sec} - b_{ct})] + a_{sec} + \mu_c ln(\mu_c)]}{\sum_{k=1}^{C} exp[ln(1-\tau_{kt}) + ln(\mathcal{L}_{kt}\lambda_{kt}) + (1+\mu_k)ln[\frac{1}{1+\mu_k}(\beta_{kt,sec} - b_{kt})] + a_{sec} + \mu_k ln(\mu_k)]} \end{aligned} \quad (8)$$

Dividing this share equation by the share of establishments choosing a base location $s_{0t,sec}$ and taking logs leads to the structural share-ratio equation (Berry 1994).

[2] I concentrate on new establishments because it's extremely rare for establishments to change location (in my data, 1.60% of establishments do so in their lifetime), while new establishments are by definition choosing a location

$$
\begin{aligned}
ln(\frac{s_{ct,sec}}{s_{0t,sec}}) =& ln(1-\tau_{ct}) + ln(\mathcal{L}_{ct}\lambda_{ct}) + (1+\mu_c)ln[\frac{1}{1+\mu_c}(\beta_{ct,sec}-b_{ct})] + \mu_c ln(\mu_c) \\
& - ln(1-\tau_{0t}) - ln(\mathcal{L}_{0t}\lambda_{0t}) - (1+\mu_0)ln[\frac{1}{1+\mu_0}(\beta_{0t,sec}-b_{0t})] - \mu_0 ln(\mu_0)
\end{aligned} \tag{9}
$$

3.4 Estimating Equation

I run the following two-way-fixed-effects specification separately for each industrial sector:

$$
\begin{aligned}
ln(\frac{s_{ct,sec}}{s_{0t,sec}}) - \mu_c ln(\mu_c) =& \beta_0 + \beta_1 ln(1-\tau_{ct}) + \beta_2 ln(\mathcal{L}_{ct}\lambda_{ct}) + \beta_3 Spillovers_{ct,sec} + \beta_4 b_{ct} \\
& + \beta_5 X_{ct,sec} + \gamma_c + \zeta_t + u_{ct,sec}
\end{aligned} \tag{10}
$$

The latter four terms of equation 9, the base-location utilities, are cleanly captured by the time-fixed effect ζ_t. In order to control for the market-size term $ln(\mathcal{L}_{ct}\lambda_{ct})$ directly, I pre-estimate it in a first step, which I outline in Appendix B. I also estimate μ_c separately using the method of Bassier, Dube, and Naidu (2022), also described in Appendix B.

Given data on corporate tax rates, τ_{ct}, the only remaining term of the structural equation is $(1+\mu_c)ln[\frac{1}{1+\mu_c}(\beta_{ct,sec}-b_{ct})]$. This term incorporates two forces which vary across locations and time within sector: the productivity of an establishment in sector sec and the outside option available to workers. For the empirical specification, I construct an empirical proxy for outside option b_{ct} which will be explained in detail in section 4. I also construct empirical proxies for three types of agglomerating forces, which will also be discussed in detail in section 4.

I additionally control for commuting zone characteristics which may be correlated with productivity, such as share of highly educated workers in the commuting zone[3]. Natural advantage is captured by the location fixed effect γ_c. After estimation, I back out estimated

[3]Specifically, all empirical specifications include controls for share highly/medium qualified workers, share women, share full-time workers, share prime aged workers, and share German workers

values for the sector-location level productivity using the structural equation.

4 Data and Summary Statistics

In this section, I outline the various data sources I use to estimate my model. In this section, I first describe my main source of data on individual firms, then the empirical proxies I use for three types of spillovers, and finally additional necessary data. Once all of these sources of data are combined, I have a panel spanning the years 1999 to 2017.

4.1 Firm Data

The core source of data for this project is the Establishment History Panel (BHP) of the Institute for Employment Research (IAB) of the Federal Employment Agency (BA) of Germany. The data are a random 50% sample of all establishments[4] in West Germany from 1975 onward and East Germany from 1992 onward on an annual basis. Appendix A describes adjustments to the structural equations and bounding exercises to account for the fact that I cannot link establishments to parent firms. The data covers all establishments with at least one employee eligible for social security. The data consists of information about the establishments themselves such as industry as well as information about employee characteristics such as number of highly-qualified workers at the firm and median daily wages. Critically, the data also includes information on establishment location. My preferred location definition is the commuting zone, of which there are 141 in Germany.

Table 1 shows summary statistics concerning the size of establishments and industrial sector of establishments over the sample period. The majority of establishments have fewer than five employees over the entire period, though the average size is growing larger over time. The number of overall entrants has also declined over time, accounting for some of

[4]Establishments cannot be linked across locations to a parent firm, but all establishments of the same firm in one location are combined. For example, all of the McDonald's in Berlin are combined into one line in the data, but I cannot link McDonald's in Berlin and McDonald's in Munich as being associated with the same firm.

the change in composition. There are fewer small businesses being started, but more similar number of very large establishments. The largest industrial sectors are, broadly defined, the trade and service sectors (other comprises mostly personal services).

There are a number of other data sources merged to the BHP for analysis. An overview of the variables and their sources is shown in Table 2. I will discuss the construction of key variables in detail below.

4.2 Proxies for Agglomerating Forces

I incorporate empirical measures of the Marshallian agglomerative forces in my analysis. In constructing these empirical proxies, I follow Ellison, Glaeser, and Kerr (2010, hereafter EGK). EGK develop empirical proxies for the strength of these three forces between two industries. Establishments may locate close to one another to reduce the costs of sourcing inputs from suppliers or the costs of shipping outputs to customers. This efficient moving of goods is proxied by the strength of input-output relationships between a pair of industries. Establishments may also locate close to one another if they use similar types of labor to ensure that workers the workers they need to hire are located nearby. Efficient moving of labor is proxied by a vector correlation of occupational shares between two industries. Establishments may also locate near one another to share ideas and innovations. Efficient moving of ideas is proxied by the strength of patent citation relationships between two industries.

The key difference between my own measure and those in EGK is that their paper develops a set of pairwise-industry-level measures of agglomeration which is used to assess how important each factor is to the coagglomeration of various industries. However, I require a location-level measure of agglomerative forces as my analysis concerns locations rather than pairwise industries. To convert pairwise-sector-level measures of coagglomeration to an sector- and location-level measure I use a weighted average of the pairwise-sector measures weighted by industrial sector establishment shares in a particular location. As a simple example, imagine there are two sectors 1 and 2. Sector 1 has a pairwise-level agglomeration relationship of .75 with sector 1, and .25 with sector 2. Then, for a establishment in sector 1 in a location with an equal share of establishments from sector 1 and 2 their establishment-level agglomeration is .5*.75 + .5*.25 = .5. Below I describe the construction of these proxies formally.

4.3 Goods Agglomeration

To proxy efficient moving of goods, I first construct a measure of the strength of input-output linkages between sectors. $Input_{sec \leftarrow m}$ is the share of sector sec's inputs which come from sector m, and $Output_{sec \rightarrow m}$ is the share of sector sec's output which goes to sector m. For an establishment in sector sec, their input-output relationship with sector m is defined as the maximum of these two values. Then the weighted average is calculated as described above to obtain the input-output agglomeration benefits for locating in each possible location. For an establishment in sector sec in location c and time t their input-output agglomeration benefits are therefore:

$$\text{IO Agglom}_{ct,sec} = \sum_{i=1}^{I} \frac{N_{ctm}}{\mathcal{N}_{ct}} max(Input_{sec \leftarrow m,t}, Output_{sec \rightarrow m,t}) \tag{11}$$

Thus, the measure of input-output agglomeration benefits is measured for a particular sector, location and year cell. Figure 1 shows the strength of input-output linkages by industrial sector and commuting zone for the year 1999. The figure shows that the strength of input-output linkages varies across space, with different areas providing stronger potential input-output linkages for different sectors.

4.4 Labor Agglomeration

To proxy efficient pooling of labor, I first construct a measure of the similarity of labor used by a sector pair. For each sector, I construct a vector of the shares of industrial employment of each three-digit occupation. Then, for sector sec and m the vector correlation is the labor correlation of industry pair sec,m, $LC_{sec,m}$. Then the weighted average is calculated to obtain the labor correlation agglomeration benefits for locating in each possible location. For a firm in sector sec in location c and time t their labor correlation agglomeration benefits are therefore:

$$\text{LC Agglom}_{ct,sec} = \sum_{i=1}^{I} \frac{N_{mct}}{\mathcal{N}_{ct}} LC_{sec,mt} \tag{12}$$

As with input-output agglomeration above, this measure of agglomeration benefits varies by sector, location and year. Figure 2 shows the strength of labor linkages by industrial sector and commuting zone for the year 1999. The figure shows that the strength of labor linkages varies across space, with different areas providing stronger potential labor linkages for different sectors.

4.5 Knowledge Agglomeration

My measure of knowledge spillovers in a location comes from Jaffe et al. (1993). For each patent, I define a control patent as the patent with the closest publication date in the same 3-digit IPC patent class as the main patent. For each patent I also identify the commuting zone where the patent originates, as well as the region where each cited patent introduced by the applicant originates, both excluding and including self-citations. Then, for each location I define the knowledge agglomeration as the probability a cited patent comes from the same region (pr_{cite}) minus the probability that the control patent comes from the same region (pr_{cntrl}). This proxy measures knowledge spillovers since it measures the percentage of citations in a patent that are from the same location over-and-above the level that you would expect from the distribution of patents.

4.6 Construction of Additional Variables for Analysis

In addition to agglomeration, taxes also play a role in establishment location choice (Fajgelbaum et al. 2019). Corporate tax rates in Germany are set at a base level by the federal government, but individual municipalities are permitted to set their own corporate tax rates in the form of a multiplier on the federal rate of 3.5%. Changes in these rates are frequent, and are largely exogenous to local economic conditions (Fuest, Peichl, and Siegloch 2018). I have obtained data on these municipal tax rates and aggregated them to the average commuting zone level using a weighted-average with municipal population as the weight.

Figure 3 shows the average corporate tax rates for the years 2000, 2010, and 2017. Tax rates are generally increasing throughout the time period of my panel, and are highest in northwestern Germany, while the lowest in the south and parts of the east.

I also construct an empirical proxy for the outside option of workers within a commuting zone. I follow the approach of Card, Deviciente, and Maida (2013) and construct the average outside option as an employment-share weighted average of sectoral union minimum wage rates. I approximate the union minimum wage rate with the 20th percentile of the establishment-level distribution of low-qualification employees' mean wages.

Union minimum wages rates are not straightforward to obtain for Germany, so I have developed a data-based method to approximate the union minimum wage rate. There is no central repository of union contracts available for Germany, but I obtained the 2019 union contracts for the state of North-Rhine Westfalia. In the contracts, the minimum wage rate and the effective date is specified, typically as a monthly rate. There are different rates for different skill levels, I concentrate on the low-skill level for my analysis. Union contracts oftentimes do not map cleanly into a single industry code in the BHP data. I chose four two-digit industry codes (retail trade, wholesale trade, chemical industry, and transportation/logistics) which map into a single union contract, and have large enough employment in 2019 in North-Rhine Westfalia to analyze wage distributions. I have union contracts for 2019, but my main analysis data is only available through 2017. To approximate the union rates in 2017, I average the minimum wage growth rates for 2020 and 2021 to approximate the average wage increases year-over-year. On average, the minimum wage grows around 100 Euros per year, so I use this to back out an estimated 2017 minimum wage.

The BHP data has information about the wages of low-qualification workers (high school or less, no vocational qualification) at the establishment. For establishments with twenty or more employees, I plot the wage distribution of these low-qualification employees in Figure 4 along with the estimated minimum wage rates described above. As can be seen in Figure 4, with the exception of wholesale trade the union minimum wage rate falls at approximately the 20th percentile of the low-qualification wage distribution. Thus, I proxy the union minimum wage rate as the 20th percentile of the low-qualification wage distribution in a particular two-digit industry-state-year cell. These union minimum wage rates are aggregated to the location-level using establishment-sector-share weights as with the agglomerating forces described above. Figure 5 shows the geographic distribution of the outside option proxy aross space. There is a clear delineation between former East- and West-Germany, with former East Germany having persistently lower outside option compared to former West Germany.

Finally, I pre-estimate μ_c using the method of Bassier, Dube, and Naidu (2022) and $\mathcal{L}_{ct}\lambda_{ct}$ in a first stage labor-supply regression outlined in Appendix B.

5 Empirical Strategy and Results

5.1 Empirical Strategy

Revisiting the estimating equation:

$$\begin{aligned} ln(\frac{s_{ct,sec}}{s_{0t,sec}}) - \mu_c ln(\mu_c) =& \beta_0 + \beta_1 ln(1-\tau_{ct}) + \beta_2 ln(\mathcal{L}_{ct}\lambda_{ct}) + \beta_3 Spillovers_{ct,sec} + \beta_4 b_{ct} \\ & + \beta_5 X_{ct,sec} + \gamma_c + \zeta_t + u_{ct,sec} \end{aligned} \tag{10}$$

There may be bias in the estimates of β_3. Specifically, there may be unobserved demand or productivity shocks which impact the distribution of incumbent establishments in the commuting zones used as weights in the spillover proxies. These unobserved demand or productivity shocks may also make the location more attractive to new establishments, affecting the share-ratio.

In order to correct for this, I construct a shift-share instrument. Specifically, I instrument the weights in the spillover proxy weighted averages with:

$$\hat{N}_{mct} = N_{mc,1998} * growth_{mt,-c} \tag{13}$$

where $N_{mc,1998}$ is the number of incumbent establishments in sector m and commuting zone c in the pre-period 1998, and $growth_{m,-c}$ is the leave-out growth rate in sector m in similarly sized[5] commuting zones between 1998 and t.

5.2 Results

Figure 6 shows the results of the main specification. Each panel shows point estimates of a coefficient of interest as well as 95% confidence intervals. Appendix figure $D.1$ shows the corresponding OLS results. Table 3 shows the exact point estimates and F statistics for each specification.

Establishments in all economic sectors have a distaste for taxes, though the degree to which this distaste influences the location decision varies[6]. Agriculture, forestry, and fishing establishments display the highest aversion with a point estimate of the elasticity of 13.65, while the education and health sector, as well as the other (mostly services) sector do not demonstrate statistically significant responses to tax rates.

The results also indicate that establishments in all economic sectors either prefer lower outside-options or are indifferent. As discussed in Section 3, higher outside options force

[5] Specifically, I split the 141 commuting zones into quartiles (35 commuting zones each) and construct the growth rates as the leave-out growth rates within these quartiles.

[6] Estimated coefficients are positive because the independent variable is $ln(1 - \tau_c)$

establishments to pay higher wages. Thus, a negative coefficient on outside option may be thought of as demonstrating a preference on the part of their ability to markdown wages they would pay to their workforce. Sectors which display such a preference are the mining, utilities, and construction sector, the trade and transportation sector, and the education and health sector.

Utilization of spillovers differs substantially by economic sector, particularly the valuation by establishments of locating near establishments that either supply their inputs or buy their outputs. Four sectors (Agriculture, forestry, and fishing; mining utilities, and construction; professional services; and the other, mostly personal services, sector) prefer to be located nearly to establishments which supply their inputs or buy their output, with the other services sector displaying the strongest response to this spillover.

Preferences for locating close to firms which utilize similar types of labor displays less heterogeneity. The only sector which displays a preference for locating near to or away from establishments using similar types of labor is the other services sector. This sector prefers to locate away from establishments using similar types of labor, which suggests that this sector may find wages being big up in the presence of of competitors or experience congestion effects in relation to locating nearby one another.

Though preferences surrounding labor pooling are largely irrelevant in the main specifications, this may be because my definition of economic sector is overly broad to capture dynamics surrounding the use of specialized labor. For example, though manufacturing is a single sector the skills used in auto manufacturing may be very different from those used in pharmaceutical manufacturing. Table 4 shows the results of the main IV regression for a more disaggregated definition of industrial sector. These results are suggestive that this may be the case, with much more heterogeneity in the sign and size of the coefficient on the labor correlation measure in comparison to the main results in Table 3. However, the results for the finer industry disaggregation have smaller sample sizes and the instrument is weaker.

Furthermore, the underlying data is more sparse[7], so results should be taken as suggestive.

Knowledge spillovers are only statistically significant in the case of the manufacturing sector. It is possible that this is because the measure of knowledge is not relevant for non-manufacturing firms. I cannot match patent technology to a sector in the BHP data, so the measure if only at the commuting-zone level. The measure is also relatively sparse, with roughly 15% of patents being filed in Munich each year, and more than 40% of those patents being filed by Siemens, a manufacturing conglomerate. This is suggestive that the coefficient is significant only for the manufacturing sector because the measure is the most relevant for this sector.

5.3 Tradability

One might expect that there would be differences in how establishments producing tradable goods value a location's characteristics. Table 5 shows the share of establishments in the data which are classified as tradable using the definition of Dauth et al. (2017), based on the two-digit-industry-level import penetration and export opportunities using UN Comtrade data.

As the table shows, for the purposes of analysis, two sectors may be thought of as tradable (agriculture, forestry and fishing; manufacturing). Both of these sectors have all establishments in the underlying data in industries classified as medium or high tradability in the data. Three sectors may be though of as non- or low-tradable (mining, utilities, and construction; trade and transportation; and the education and health sector), with more than 90% of establishments producing non- or low-tradability goods. Two sectors (professional services and other services) are mostly non- or low-tradability with a sizable minority (roughly 20%) of establishments within-sector producing medium-tradability goods.

Comparing the relationship between the tradability index and estimated regression coefficients, there is no clear pattern with the exception that sectors that are non- or low-

[7] I.E. there is a larger proportion of commuting zones with only one establishment in the sector-year pair picking it, so results are less stable.

tradability seem to show more response to the outside option. Perhaps this is due to tradable sectors being able to move jobs to Eastern Europe or China in response to wage movements, making the local outside option less of a consideration.

5.4 Heterogeneity by AKM Firm-Effect and Establishment Size

I perform several extensions of my main results. The first is splitting the sample by above-median and below-median AKM effect. Although AKM effects are not a measure of firm productivity, the AKM firm-effect is correlated with worker value-added (Card, Cardoso, and Kline 2016). Previous research (Gaubert 2018) has demonstrated that more productive firms are able to better utilize the benefits of agglomeration economies. This leads to sorting of more productive firms to larger cities. In order to test whether firms with different levels of productivity consider different characteristics of locations when making their location decision, I repeat the specification in equation 10 separately for above- and below-median AKM effect establishments.

Figure 7 and Table 6 shows the results. There seem to be differences in how the outside option is valued between higher and lower productivity establishments, with below-median AKM establishments being more likely to place value on the outside option in their decision making process. This result makes sense, since less-productive establishments will, by definition, have a productivity closer to the outside option[8]. So, in order to make more profits lower outside options benefit them more than higher productivity establishments.

I also split the sample by establishment size to see if there are meaningful differences in behavior between larger and smaller employers, with results shown in Figure 8 and Table 7. The results show that smaller establishments may be more responsive to taxes in their location decision, though the results are noisy. As in the case of less productive establishments, smaller establishments also show more sensitivity to outside options in their location decisions.

[8]Recalling one of the terms of the log-profit equation, $(1+\mu_c)ln[\frac{1}{1+\mu_c}(\beta_{ct,sec} - b_{ct})]$, $\beta_{ct,sec}$ will be lower for the below-median AKM effect establishments

Taken together, these two results suggest that policies increasing wages in low-wage regions may make the location less attractive to smaller, less productive establishments. Analyzing net welfare effects of such a change is beyond the scope of this paper, but would likely depend on whether phasing out such establishments would lead to workers having jobs at higher wages, or if unemployment would increase.

5.5 Rental Prices

Rental prices are another important factor which may influence establishment location choice (Glaeser and Gottlieb 2008). In Appendix C I show changes which would be made to the model equations to account for rental prices, as well as the results of an empirical regression controlling for rental prices. Controlling for rents does not change the point estimates of the coefficients of interest, and only the mining, utilities, and construction sector demonstrates a statistically significant sensitivity to rental prices in their location choice.

5.6 Model Fit

I test how well the model fits the data by calculating the estimated commuting-zone-sector-year wage using Equation 6 and the backed out productivity estimates from the structural Equation 9, and regressing the actual commuting-zone-sector-year wage on the imputed wage. The results are shown in Table 8. The coefficient on imputed wage is 1.33, and the intercept is -3.52 euros/day.

6 Counterfactual Firm Distributions

In this section I simulate a place-based policy designed to attract establishments to a commuting zone. I simulate setting taxes to 7%, the lowest legally permitted rate. I first show that my model produces heterogeneous responses to an identical place-based policy in different locations within Germany. I then undertake a case study on the effects of a policy setting tax rates to 7% five commuting zones in former Eastern Germany to illuminate the importance of initial conditions and sectoral heterogeneity in explaining the heterogeneous effects.

6.1 Effects of Tax Policy - Theory

Intuitively, a change in tax policy, or any place-based policy more generally, will have both immediate and secondary effects. Immediately, the tax policy will attract new establishments, some of which would not have otherwise gone to the treated location. Subsequently, the composition of establishments in the area has changed and spillovers available to establishments by going to the area have also changed. Figure 9 shows a simplified version of this process.

The figure illustrates the effects of a tax change implemented in time t=0. Between time 0 and 1, the taxes have decreased, so the overall share ratio increases as shown in the leftmost graph. Now, in subsequent years the new entrants have moved the input-output spillovers, shown in the second graph. In this example, the input-output spillovers have decreased, and the dashed line shows an alternative possible path. As is clear in the figure, the overall movement in the share-ratio as a result of the tax policy is unclear over multiple years. It is possible for negative spillovers to be induced by the tax change and actually make the location overall less attractive to new establishments.

More formally, the movement in the share ratio due to the tax rate between t=0 and 1 is:

$$ln(\frac{s_{jct}}{s_{j0t}})_{CF} - ln(\frac{s_{jct}}{s_{j0t}})_{actual} = \beta_1[ln(1-\tau_{CF,t}) - ln(1-\tau_{ct})] \tag{14}$$

Where $\tau_{CF,t}$ is the counterfactual tax rate of 7%. Subsequently, the number of counterfactual entrants attracted by the tax policy may be calculated directly. Since the sum of shares for each location-year must sum to one in both the actual and counterfactual world, and rearranging implies that:

$$s_{j0t,CF} = \frac{1}{\sum_{ct} \frac{s_{1t,CF}}{s_{j0t,CF}} + ... + \frac{s_{Ct,CF}}{s_{j0t,CF}}} \tag{15}$$

Combining equations 14 and 15 leads directly to the expression for the counterfactual share of establishments choosing the treated location. I impose the additional assumption that the pool of establishments entering the entire German market each year is fixed in order to be able to calculate the counterfactual numbers of establishments going to each location.

I use this new distribution in firms to calculate the counterfactual spillovers establishments receive in the treated locations, which translates to the counterfactual productivity of firms using the structural equation.

$$(1+\mu_c)ln[\frac{\beta_{ct,sec,CF} - b_{ct,CF}}{\beta_{ct,sec} - b_{ct}}] = \beta_3(Spillovers_{ct,sec,CF} - Spillovers_{ct,sec}) + \beta_4(b_{ct,CF} - b_{ct}) \tag{16}$$

Using this counterfactual productivity and outside option I additionally calculate counterfactual wages using model Equation 6. I continue this process iteratively to examine the dynamic secondary effects induced by the tax policy change.

6.2 Counterfactual Results for All Commuting Zones

Figure 10 shows the estimated effects of tax policy if implemented in each individual commuting zone beginning in 1999, both after 9 years in 2008 and 18 years in 2017. Panel A shows the percent difference in establishments under the counterfactual tax policy compared to the actual tax policy, and panel B the percent difference in wages. The figure shows that the effects of a place-based policy are extremely heterogenous, with some locations experiencing large increases in the number of establishments and wages, while others actually experience declines in the number of establishments and wages. Furthermore, it seems counterproductive to keep the tax policy in place for a long period of time in the majority of cases, with the exception of a few winning commuting zones, outcomes are actually worse after 18 years of tax policy compared to 9 years.

The shape of the distribution and size of these effects is strikingly similar to the estimated TFP spillovers found by Greenstone et al. (2010, figure 2), who examined the spillover effects of million dollar plant openings in US counties empirically. This suggests that my model estimates can reproduce and explain heterogeneity in the effects of place-based policies between different locations, and provide insight to policymakers who wish to create effective place-based policies.

6.3 Case Study: Underdeveloped Areas in Former East Germany

In order to more deeply examine the sources of the heterogeneity shown in Figure 10, I will now examine the dynamic effects of a place-based policy designed to develop underdeveloped regions in former East Germany. I simultaneously set corporate tax rates to 7% the five commuting zones in Eastern Germany with the highest average unemployment rates over my sample period: Uckermark, Mecklenburgische Seenplatte, Nordvorpommen, Sudvorpommen, and Stendal. These commuting zones have average unemployment rates of 10-15% and are located in the northeast of the country near the Polish border.

Figure 11 shows the results of this counterfactual exercise. Panel A shows the results for all treated commuting zones, while Panel B shows only Uckermark and Stendal, which I will concentrate on for the remainder of this section. As is visible in Figure 11, Uckermark did not have a net change in establishment counts until roughly 2009, when the overall number of establishments began to decline. This strongly contrasts Stendal, where overall establishment counts were declining by 2002.

Figure 12 shows the changes in establishment counts by industrial sector, which sheds some light on the dynamics at play behind the between-commuting-zone heterogeneity. In Uckermark, the trade and transportation sectors and the professional services sectors were growing or remaining steady for the first half of the panel period, while in Stendal both of these sectors were in decline. In order to further understand these effects, recall that the trade and transportation sector had estimated $\beta_{tax} = 5.35$, $\beta_{outop} = -.0378$, while the professional services sector had estimated $\beta_{tax} = 8.07$, $\beta_{IO} = 4.05$.

Therefore, since the trade and transportation sector began to decline in Stendal almost immediately after the tax policy was implemented, I can conclude that the new establishments attracted to the commuting zone by the tax policy increased the outside option, which made the trade and transportation sector less productive in Stendal, and ultimately the sector began declining. In Uckermark, the new establishments either decreased the outside option, or the increases in the outside option were small enough that the direct effects of the tax changes dominated as illustrated in the solid line in Figure 9.

In the case of the professional services sector, a similar dynamic occurred with input-output spillovrs. In Stendal, new establishments attracted by the policy decreased the input-output spillovers available to establishments in the professional services sector, while in Uckermark they either increased the IO spillovers for professional service establishments, or decreases were dominated by the direct tax effects.

Table 9 shows the shares in each industrial sector in Uckermark and Stendal in 1999, as well as the outside option and tax rates for the same year. The table demonstrates how small differences in initial conditions can lead to very different effects on a policy overall. Stendal had a lower share of establishments in the professional services sector, and a higher outside option and tax rates. Though the commuting zones were not initially drastically different, they experienced very different trajectories as as a result of the same policy.

6.4 Distributional Impacts of Monopsony Power

In Table 10 I show the effects of reducing monopsony power and equalizing monopsony power across geographic space on the outcomes of these same five commuting zones. The results of this counterfactual suggest that if underdeveloped areas were more perfectly competitive, they would actually have less economic activity than in the presence of monopsony power.

This is consistent with the findings of Bamford (2021).

7 Conclusions

I develop a partial-equilibrium model of establishment-location choice incorporating corporate taxes, monopsonistic labor markets, and differential location productivity. I show that establishments in different economic sectors display different sensitivities to taxes, and differently value spillovers in their location decision. Most types of establishments prefer to pay lower wages as measured by the outside option, but some types of establishments are indifferent. All types of establishments have a distaste for taxes, though the strength of distaste varies between firm types. Different types of spillovers are also valued differently by different types of establishments.

The effects of place-based policies are highly heterogeneous across space due to the secondary impact of spillovers. The distribution of effects in the wage changes induced by a tax policy predicted by my model closely matches the empirical findings of Greenstone et al. (2010). As demonstrated by the counterfactuals, small differences in initial conditions can leads to very different effects of the same place-based policy in different locations, making effective policy challenging.

References

Abowd, J., Kramarz, F., and Margolis, D., 1999. "High Wage Workers and High Wage Firms," Econometrica 67, no. 2: 251-333.

Antoni, Manfred; Berge, Philipp vom; Graf, Tobias; Grießemer, Stephan; Kaimer, Steffen; Köhler, Markus; Lehnert, Claudia; Oertel, Martina; Schmucker, Alexandra; Seth, Stefan; Seysen, Christan (2019): "Weakly anonymous Version of the Sample of Integrated Labour Market Biographies (SIAB) – Version 7517 v1". Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB). DOI: 10.5164/IAB.SIAB7517.de.en.v1

Ashenfelter, O., Farber, H., and Ransom, M. 2010. "Labor Market Monopsony," Journal of Labor Economics 28, no. 2: 203-210.

Austin, B., Glaeser, E., and Summers, L., 2018. "Jobs for the Heartland: Place-Based Policies in 21st Century America," NBER Working Paper 24548.

Azar, J., Marinescu, I., Steinbaum, M., and Taska, B., 2018. "Concentration in US Labor Markets: Evidence From Online Vacancy Data," Working paper.

Bamford, I., 2021. "Monopsony Power, Spatial Equilibrium, and Minimum Wages," Working Paper.

Beaudry, P., Green, D., and Sand, B., 2013. "Spatial equilibrium with unemployment and wage bargaining: Theory and estimation," Journal of Urban Economics 79: 2-19.

Beaudry, P., Green, D., and Sand, B., 2018. "In Search of Labor Demand," American Economic Review 108, no. 9: 2714–2757.

Behrens, k., Duranton, G., and Robert-Nicoud, F., 2014. "Productive Cities: Sorting, Selection, and Agglomeration," Journal of Political Economy 122, no. 3: 507–553.

Berry, S., 1994. "Estimating Discrete-Choice Models of Product Differentiation," The RAND Journal of Economics 25, no. 2: 242–262.

Berry, S., Levinsohn, J., and Pakes, A., 1995. "Automobile Prices in Market Equilibrium," Econometrica 63, no. 4: 841–890.

Bilal, A., 2021. "The Geography of Unemployment," NBER Working Paper 29269.

Boeri, T., Ichino, A., Moretti, E., and Posch, J., 2019. "Wage Equalization and Regional Misallocation: Evidence from Italian and German Provinces," NBER Working Paper 25612.

Card, D., Cardoso, A.R., Heining, J., and Kline, P., 2018. "Firms and Labor Market Inequality: Evidence and Some Theory," Journal of Labor Economics 36, no. 1: S13-S70.

Card, D., Cardoso, A.R., and Kline, P., 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," The Quarterly Journal of Economics 131, no. 2: 633–686.

Card, D., Devicienti, F., and Maida, A., 2014. "Rent-sharing, Holdup, and Wages: Evidence from Matched Panel Data," Review of Economic Studies 81: 84–111.

Dauth, W., Findeisen, S., Moretti, E., and Suedekum, J., 2018. "Matching in Cities," NBER Working Paper 25227.

Dauth, W., Findeisen, S., and Suedekum, J., 2017. "Trade and Manufacturing Jobs in Germany." American Economic Review, 107, no. 5: 337-42.

Desmet, K. and Rossi-Hansberg, E., 2014. "Spatial Development," American Economic Review 104, no. 4: 1211–1243.

Diamond, R., 2016. "The Determinants and Welfare Implications of US Workers Diverging Location Choices by Skill: 1980-2000," American Economic Review 106, no. 3: 479-524.

Dumais, G., Ellison, G., and Glaeser, E., 2002. "Geographic Concentration as a Dynamic Process," The Review of Economics and Statistics 84, no. 2: 193-204.

Dustmann, C., Lindner, A., Schonberg, U., Umkehrer, M., and vom berge, P., 2022. "Reallocation Effects of the Minimum Wage" The Quarterly Journal of Economics 137, no. 2: 267–328.

Ellison, G. and Glaeser, E., 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," Journal of Political Economy 105, no. 5: 889-927.

Ellison, G., Glaeser, E. and Kerr, W., 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," American Economic Review 100: 1195–1213.

Fajgelbaum, O., Morales, E., Serrato, J.C., and Zidar, O., 2019. "State Taxes and Spatial Misallocation," Review of Economic Studies 86: 333-376

Falch, T., 2010. "The Elasticity of Labor Supply at the Establishment Level," Journal of Labor Economics 28, no. 2: 237-266.

Falch, T., 2011. "Teacher Mobility Responses to Wage Changes: Evidence from a Quasi-Natural Experiment," AER: Papers & Proceedings 101, no. 3: 460–465.

Fuest, C., Peichl, A. and Siegloch, S., 2018. "Do Higher Corporate Taxes Reduce Wages? Micro Evidence from Germany," American Economic Review 108, no. 2: 393–418.

Ganzer, Andreas; Schmucker, Alexandra; Stegmaier, Jens; Wolter, Stefanie (2018): "Establishment History Panel (BHP) – Version 7517 v1". Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB). DOI: 10.5164/IAB.BHP7517.de.en.v1

Gaubert, C., 2018. "Firm Sorting and Agglomeration," American Economic Review 108, no. 11: 3117–3153.

Glaeser, E. and Gottlieb, J., 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," Journal of Economic Literature 47, no. 4: 983–1028

Glaeser, E., Kerr, S., and Kerr, W., 2015. "Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines," The Review of Economics and Statistics 97, no. 2: 498–520

Greenstone, M., Hornbeck, R., and Moretti, E., 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," Journal of Political Economy 118, no. 3: 536–598

Heise, S. and Porzio, T., 2018. "Why Do Spatial Wage Gaps Persist? Evidence from the Enduring Divide between East and West Germany," Working Paper.

Kline, P. and Moretti, E., 2013. "Place Based Policies with Unemployment," AER: Papers & Proceedings 103, no. 3: 238–243.

Kline, P. and Moretti, E., 2014. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," The Quarterly Journal of Economics 129, no. 1: 275–331.

Manning, A., 2010. "The plant size-place effect: agglomeration and monopsony in labour markets," Journal of Economic Geography 10, no. 1: 717–744.

McFadden, D., 1973. "Conditional Logit Analysis of Qualitative Choice" Be. In: Zarembka, P., Ed., Frontiers in Econometrics, Academic Press, New York, 105-142.

Moretti, E., 2004. "Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data," Journal of Econometrics 121: 175-212.

Moretti, E., 2013. "Real Wage Inequality," American Economic Journal: Applied Economics 5, no. 1: 65–103.

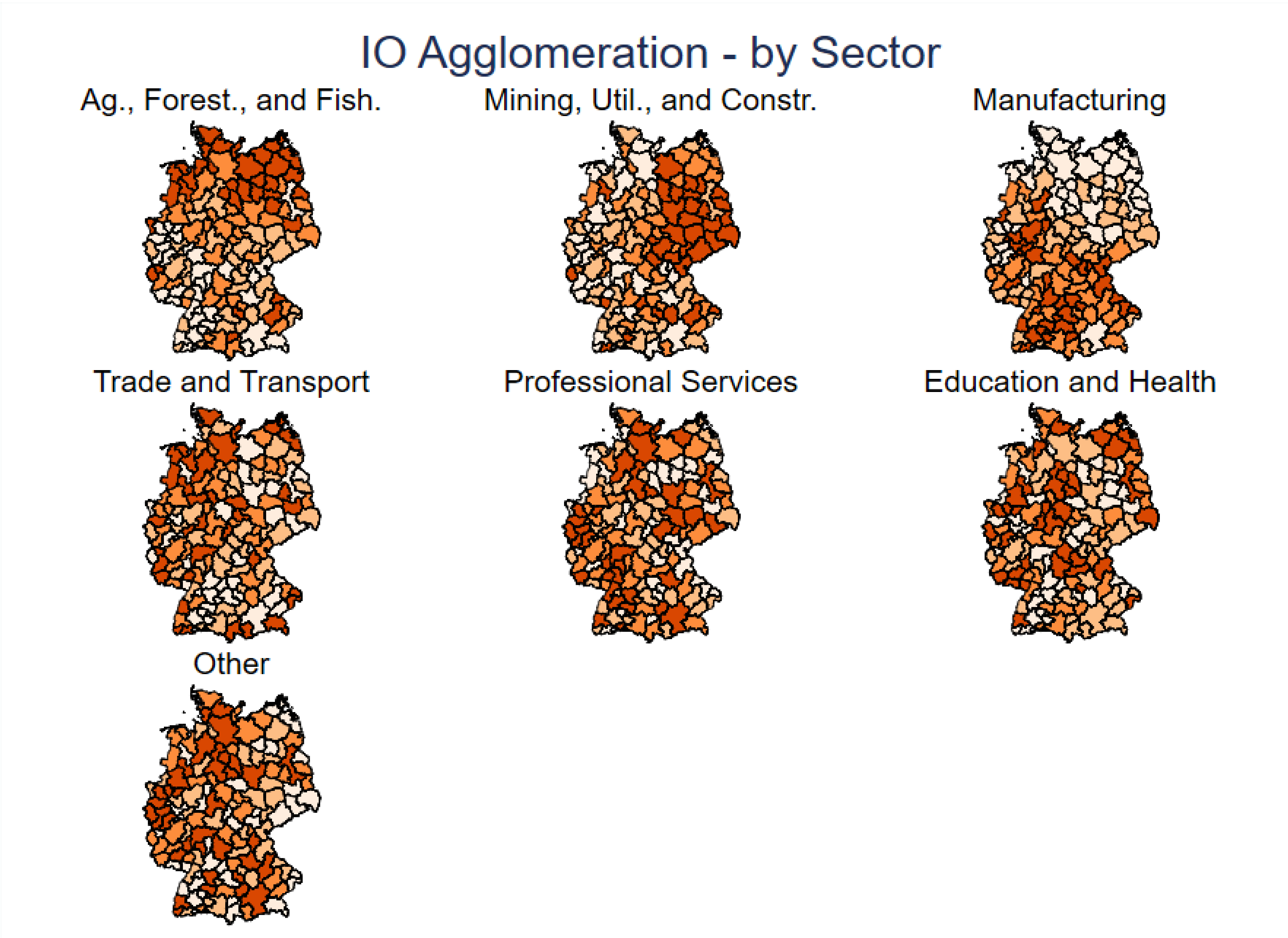
Rosenthal, S. and Strange, W. 2004. "Evidence on the Nature and Sources of Agglomeration Economies," Handbook of Regional and Urban Economics 4: 2119-2171.

Schmieder, J., 2013. "What Causes Wage Dispersion? Evidence from New Firms," Working paper.

Slattery, C. and Zidar, O., 2019. "Evaluating State and Local Business Tax Incentives". Journal of Economic Perspectives (in preparation) (Working Paper).

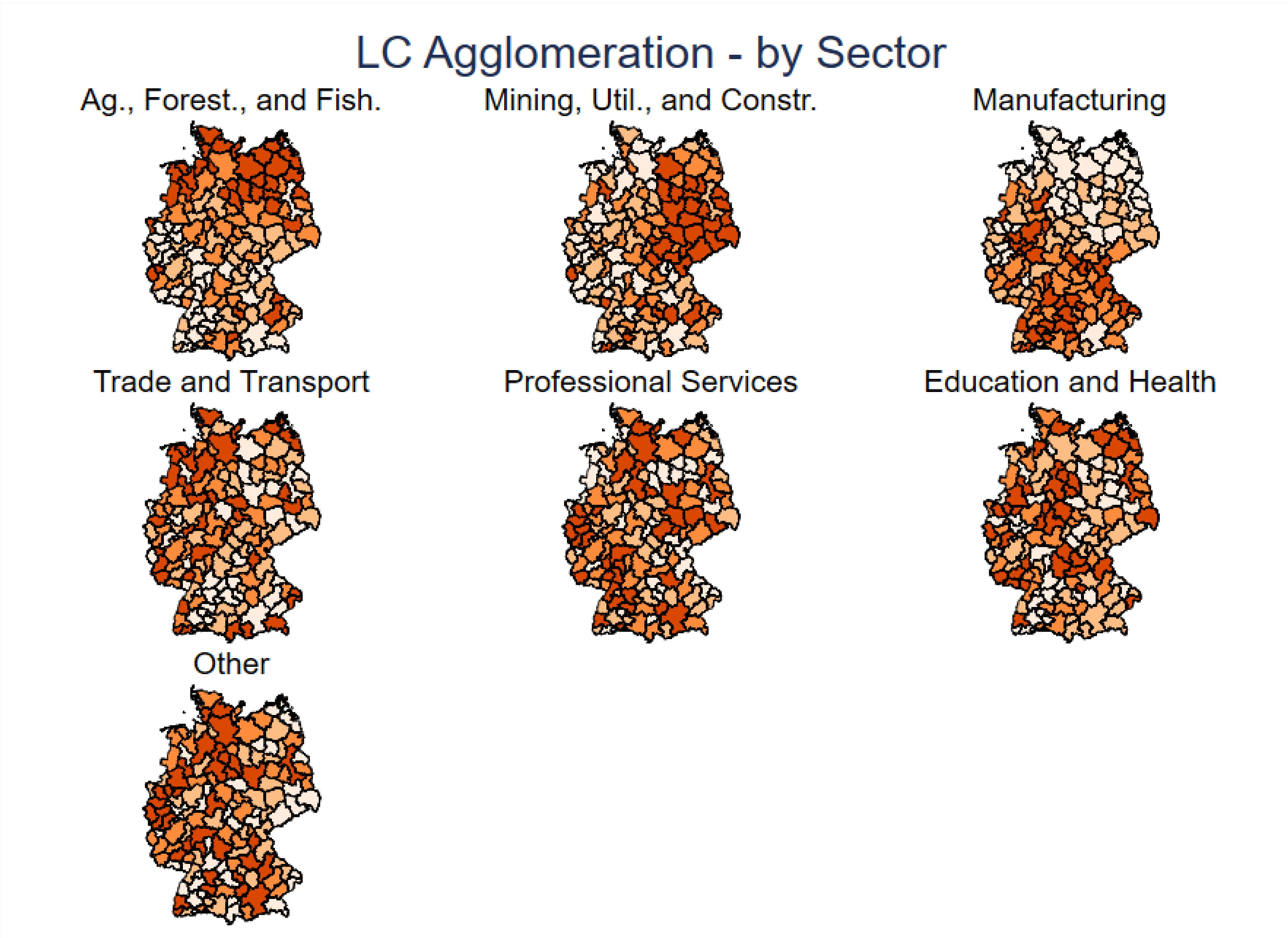
Figures

Figure 1: Input-Output Linkages Across Space, 1999



Notes: See Section 4 for data sources and variable definitions. Figure shows quantiles of spillovers within a commuting zone. Darker colors are higher spillovers.

Figure 2: Labor Linkages Across Space, 1999



Notes: See Section 4 for data sources and variable definitions. Figure shows quantiles of spillovers within a commuting zone. Darker colors are higher spillovers.

Figure 3: Average Corporate Tax Rates

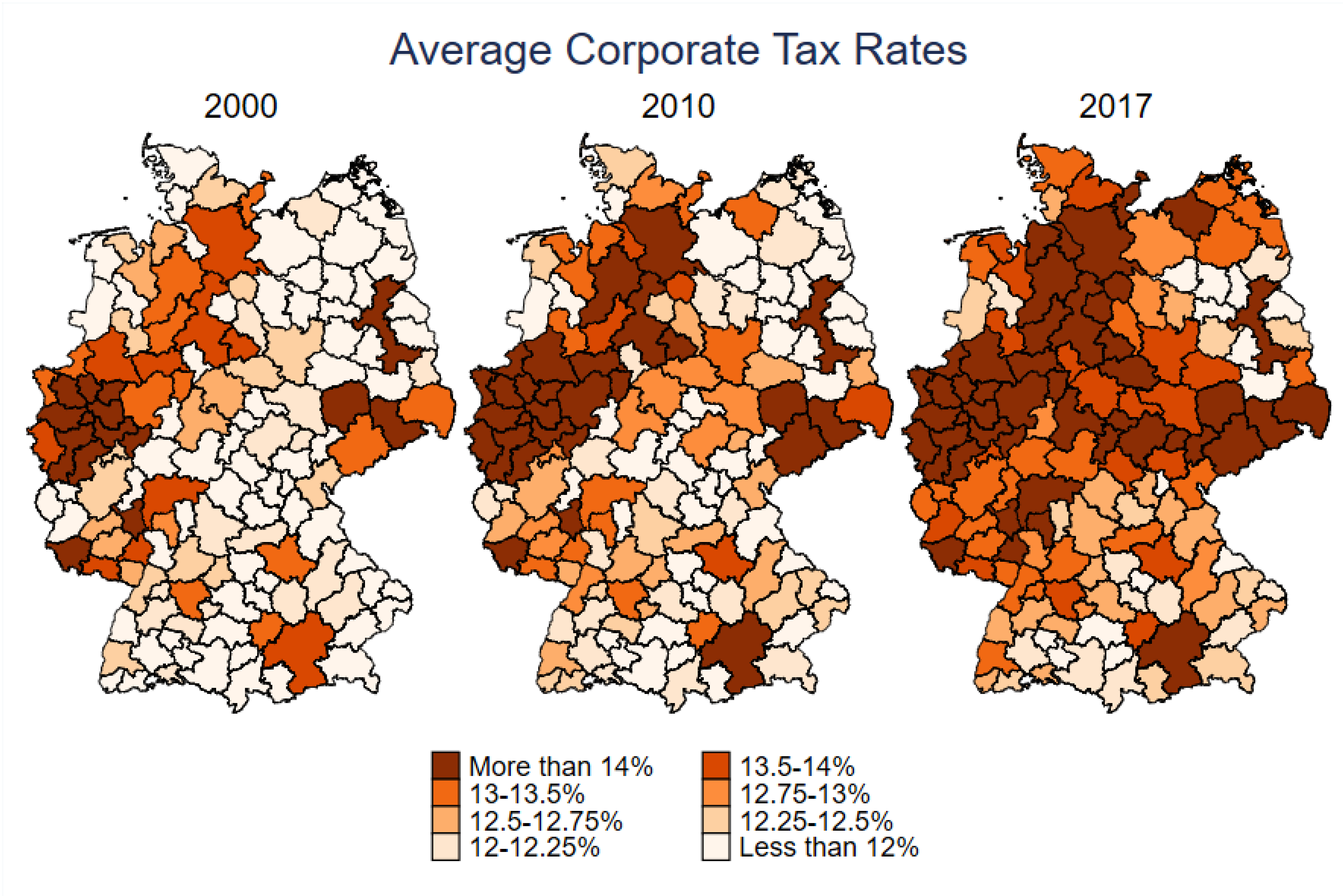
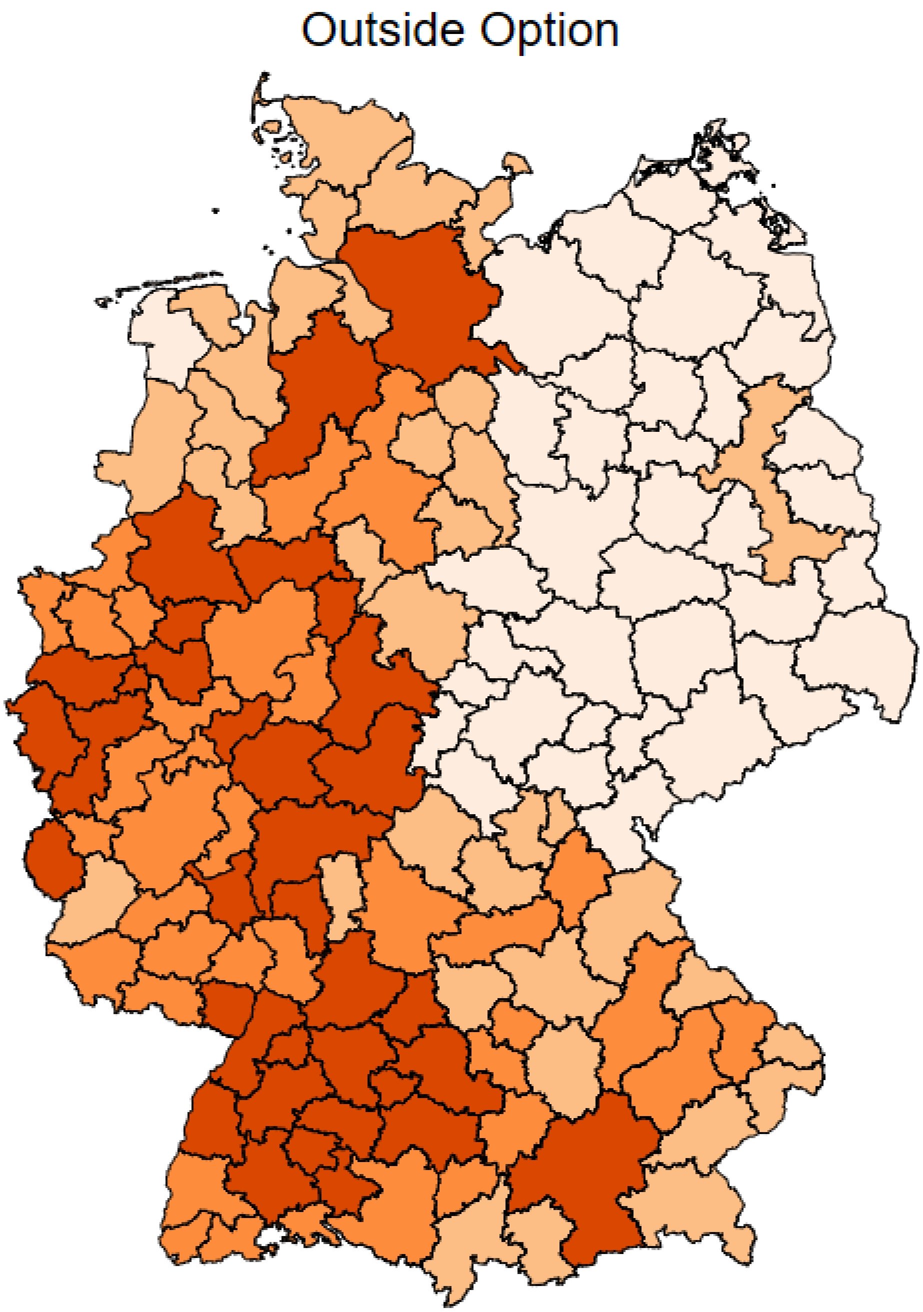


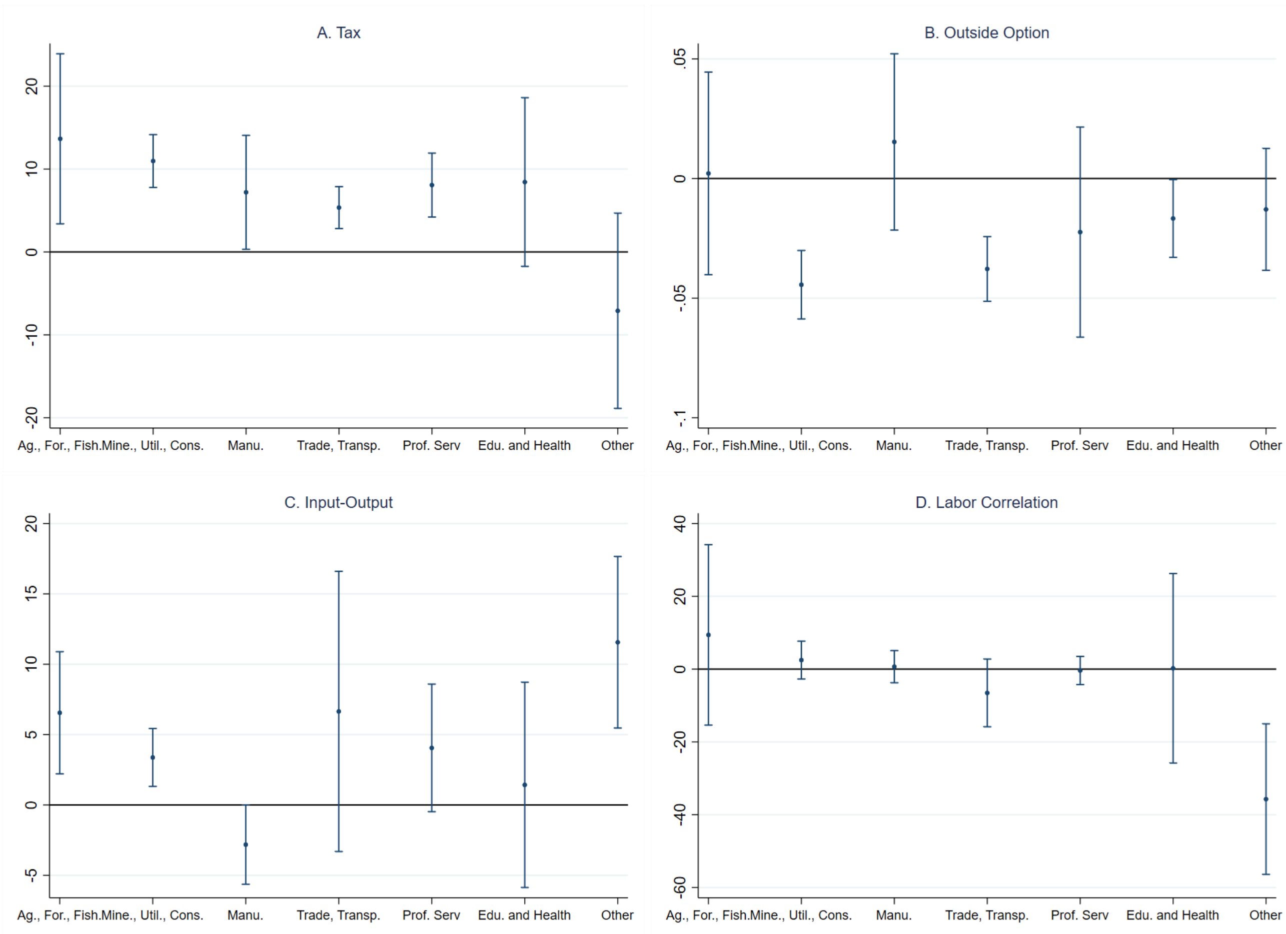
Figure 4: Union Minimum Wage Rates

Figure 5: Geographic Distribution of Outside Option



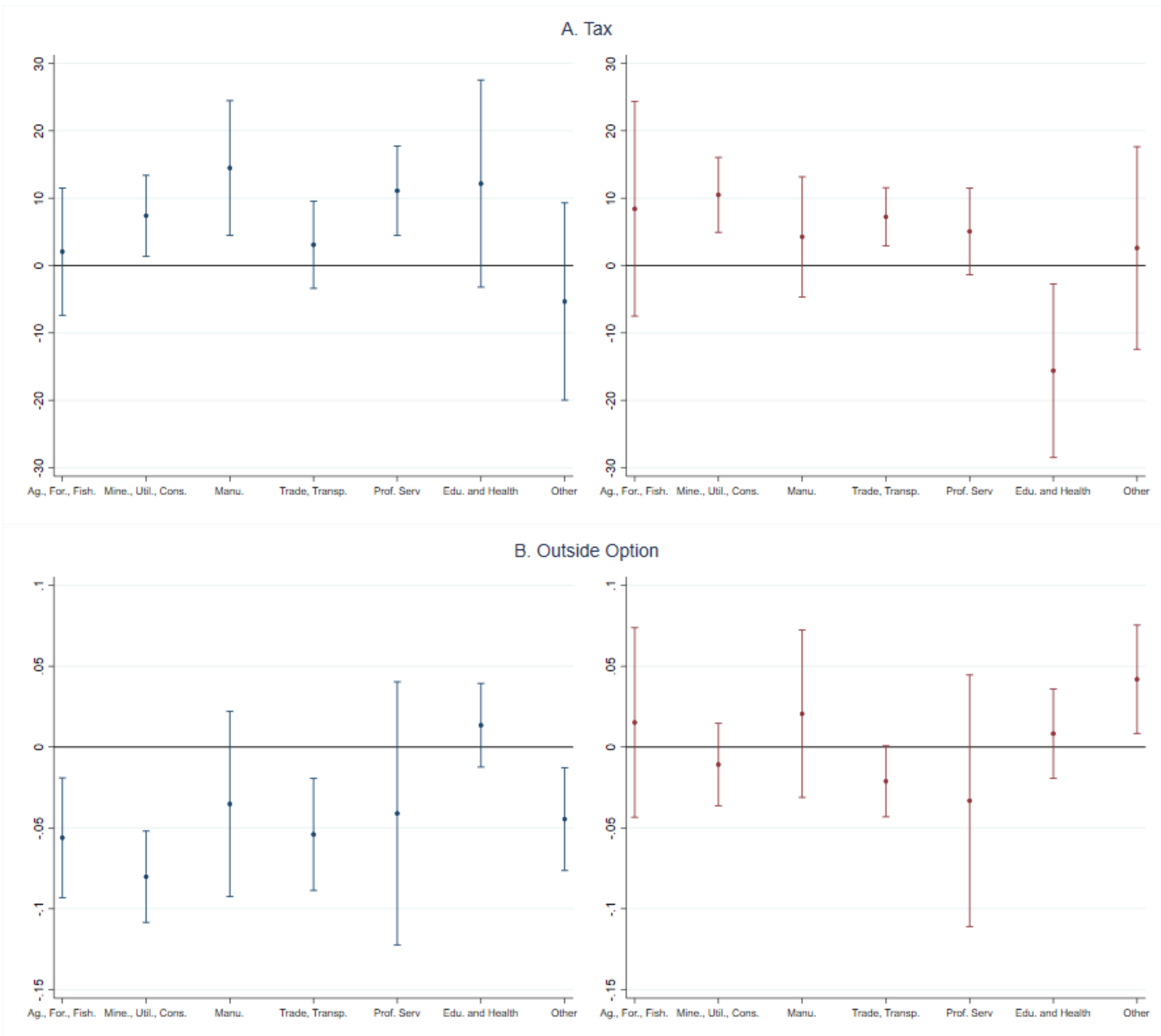
Notes: Figure shows quantiles of outside option within a commuting zone. Darker colors are higher outside options.

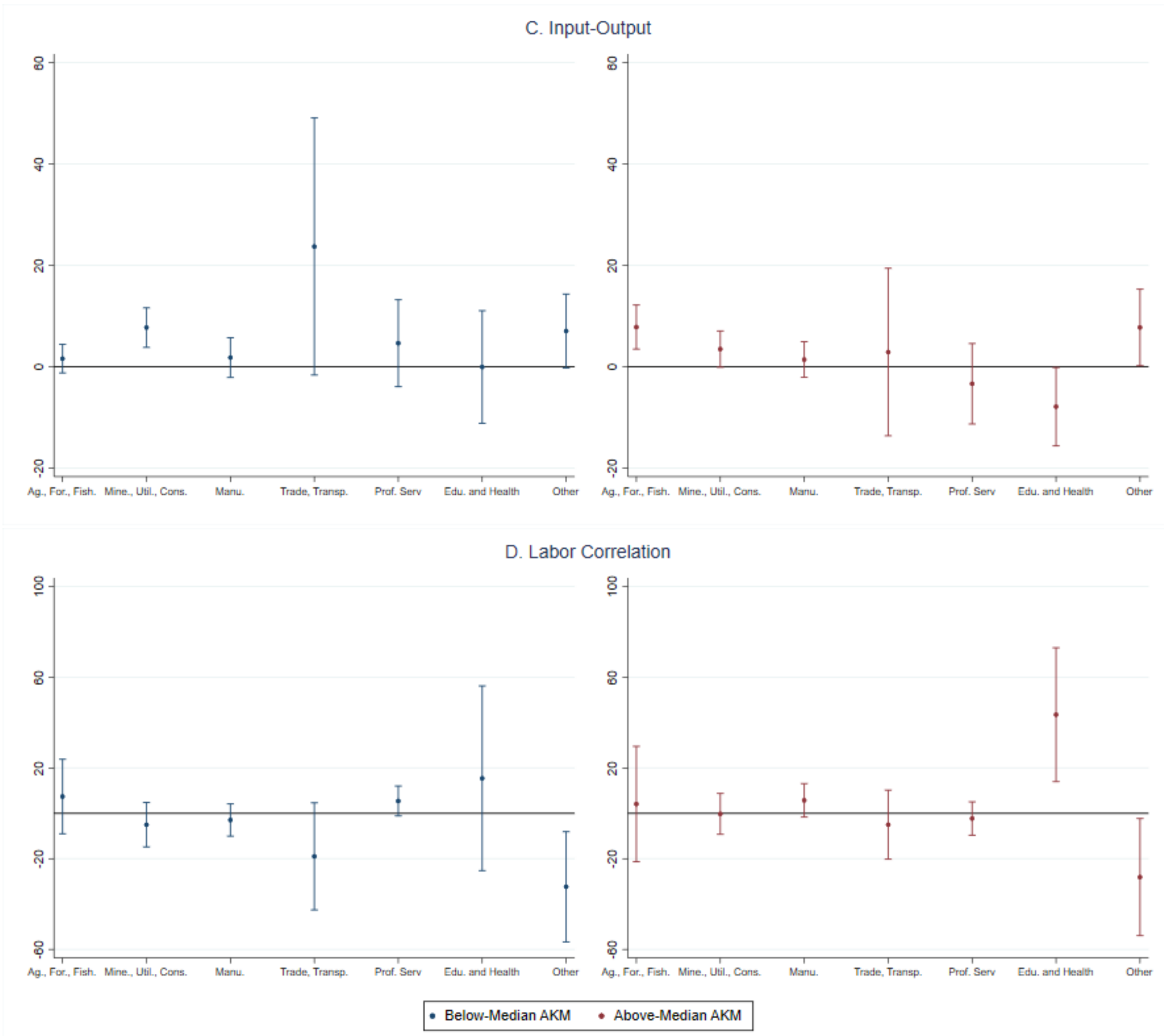
Figure 6: Coefficients by Sector



Notes: These figures show the point estimates of coefficients of the variables of interest for each industrial sector in the instrumental variable regression 10, as well as the 95% confidence intervals. See Section 4 for data sources and variable definitions.

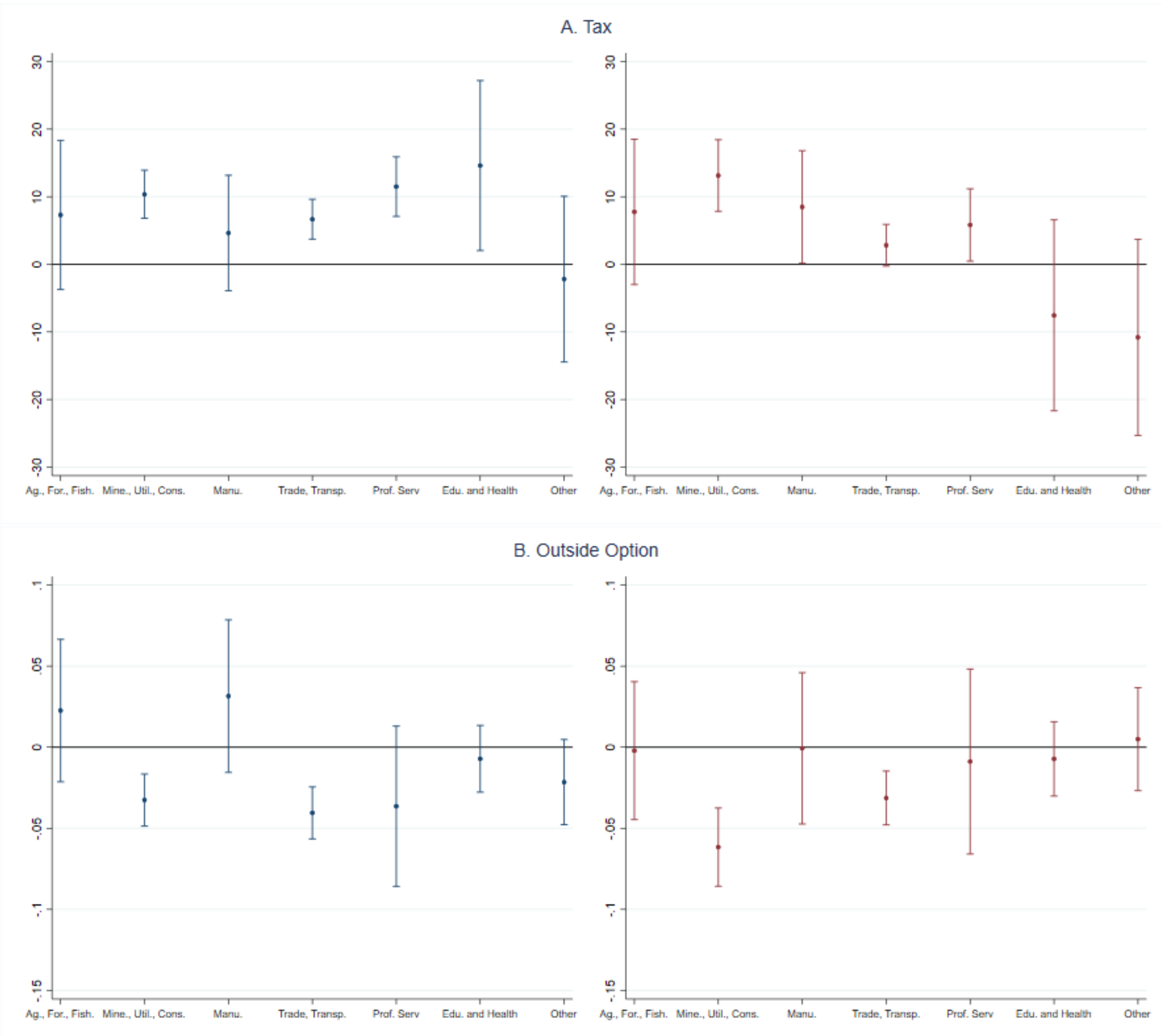
Figure 7: Coefficients: Splitting Sample by AKM Firm-Effect

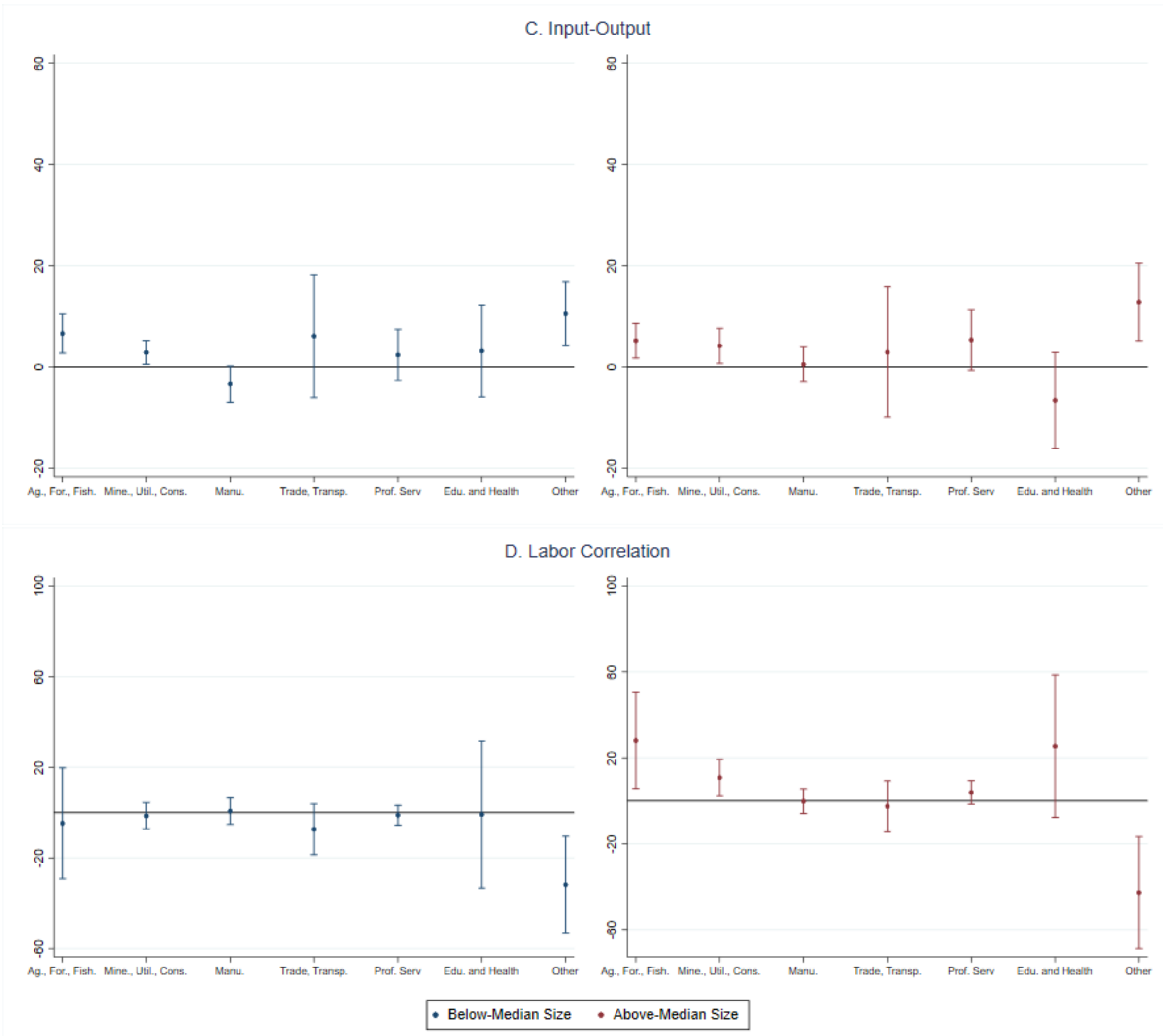




Notes: These figures show the point estimates of coefficients of the variables of interest for each industrial sector in the instrumental variable regression 10, as well as the 95% confidence intervals for establishments with below- and above-median AKM firm effects. See Section 4 for data sources and variable definitions.

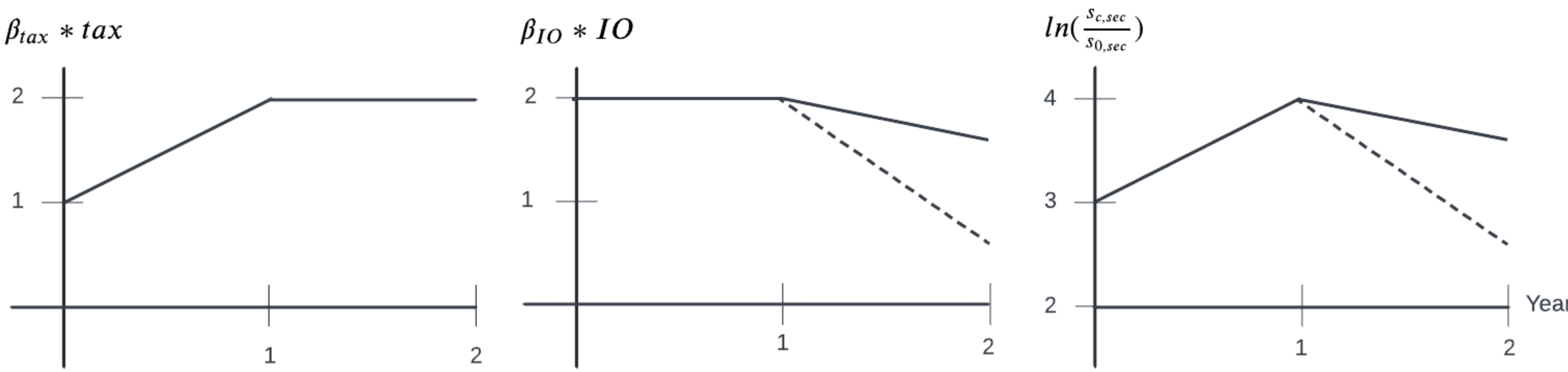
Figure 8: Coefficients: Splitting Sample by AKM Firm-Effect





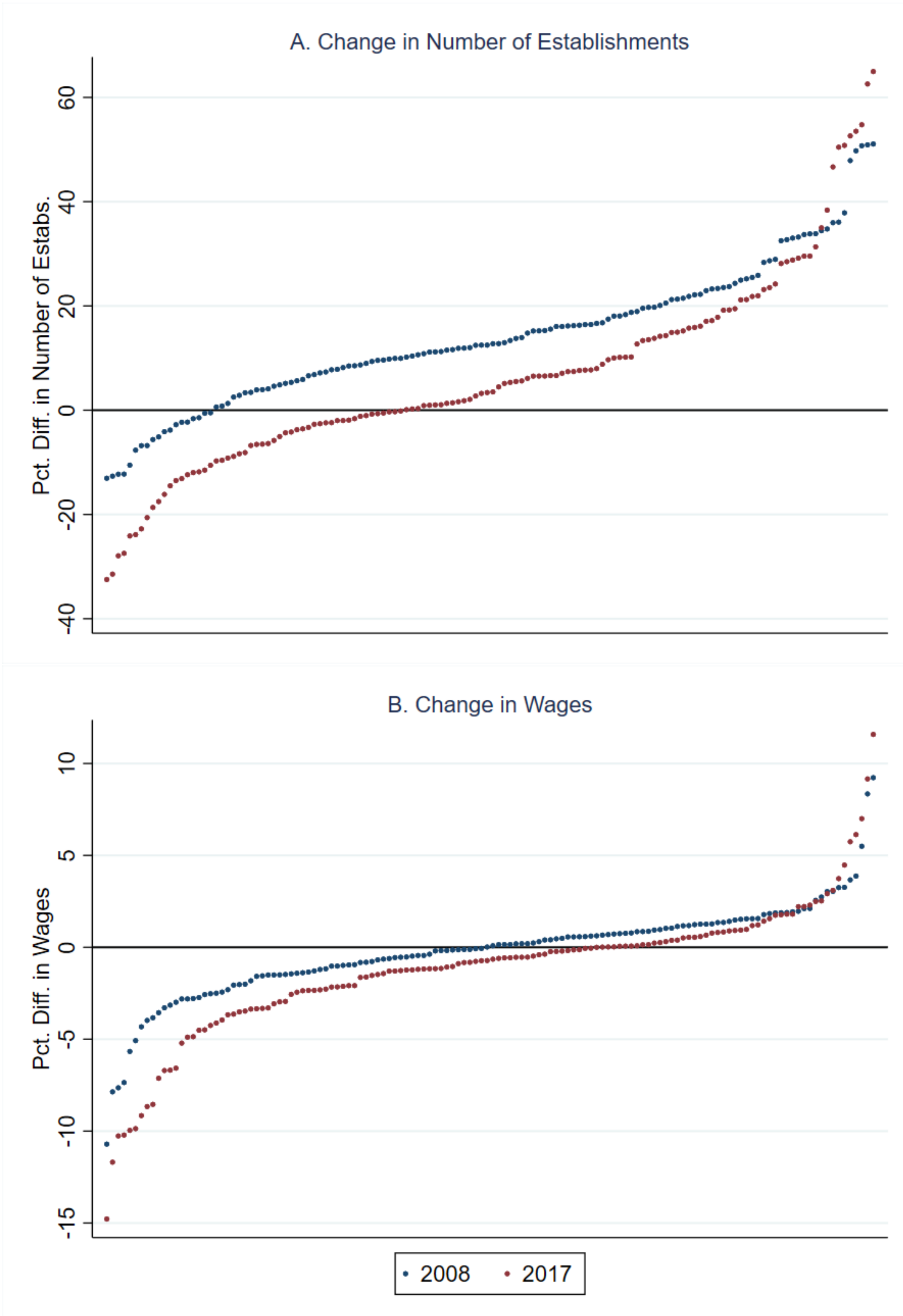
Notes: These figures show the point estimates of coefficients of the variables of interest for each industrial sector in the instrumental variable regression 10, as well as the 95% confidence intervals for establishments with below- and above-median sizes. Size is number of employees. See Section 4 for data sources and variable definitions.

Figure 9: Illustration of Primary and Secondary Tax Policy Effects



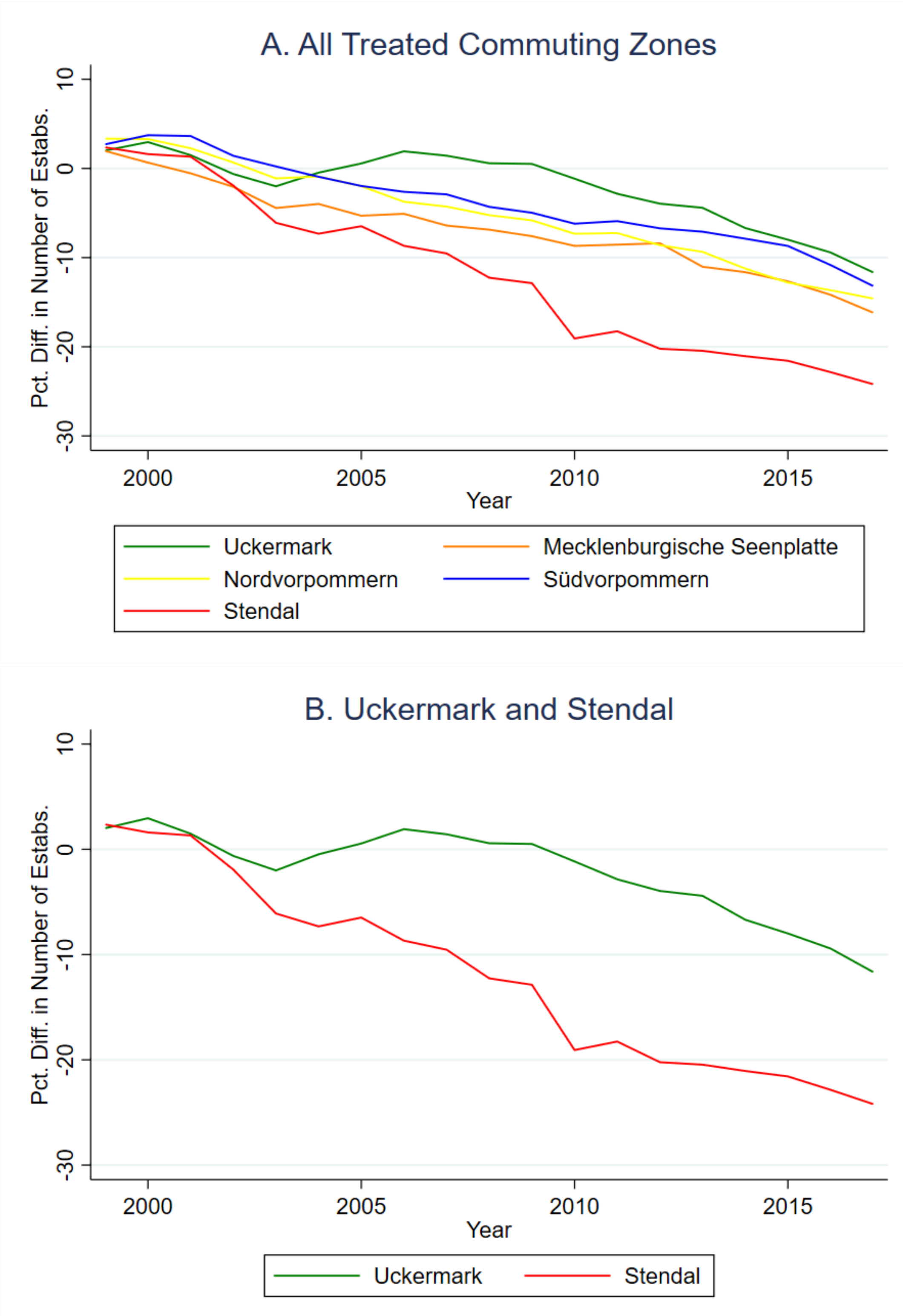
Notes: Time since a tax policy was enacted in t=0. When the tax rate decreases, the share ratio increases directly between time 0 and 1. Between time 1 and 2 the subsequent composition of establishments has changed, which leads to changes in the input-output spillovers. The solid and dashed lines show two possible paths, one of which leads to an overall decline in the share ratio despite the tax policy remaining in place.

Figure 10: Effects of Counterfactual Tax Policy on Commuting Zones



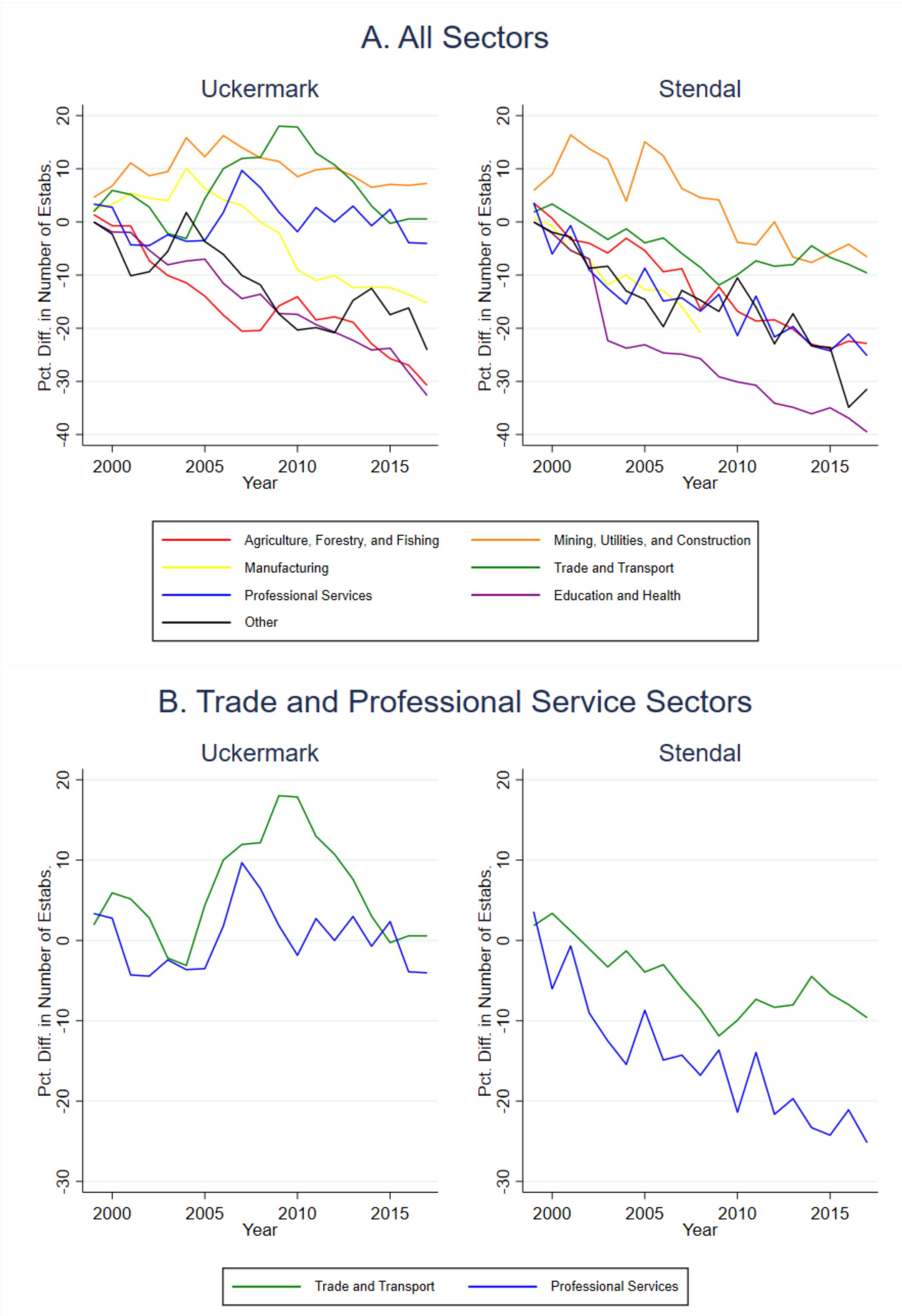
Notes: Each dot is a commuting zone. Panel A shows the percent difference in establishments under the counterfactual tax policy compared to the actual tax policy, and panel B the percent difference in wages. See Section 6 for details on calculations. Bottom and top 5% of commuting zones were trimmed for readability.

Figure 11: Case Study: Change in Establishment Counts



Notes: Panel A shows the change in establishment counts in all treated commuting zones, and panel B the change in Uckermark and Stendal. See Section 6 for details on calculations.

Figure 12: Case Study: Change in Establishment Counts by Industrial Sector



Notes: Panel A shows the change in establishment counts in all treated commuting zones by industrial sector, and panel B the change in Uckermark and Stendal. See Section 6 for details on calculations.

Table 1: BHP Summary Statistics

Summary Statistics: New Establishments			
Sector	1999	2008	2017
Agriculture, Forestry, and Fishing	1.83	1.85	1.8
Mining, Utilities, and Construction	14.44	11.25	14.22
Manufacturing	5.64	5.13	4.15
Trade and Transport	38.03	36.26	34.44
Professional Services	19.46	19.91	19.23
Education and Health	6.56	7.68	8.17
Other	14.05	17.92	17.98
Number of Employees			
1-4	82.15	76.86	69.96
5-9	10.72	13.73	16.84
10-19	4.44	5.75	7.98
20-49	1.96	2.63	3.94
50-99	0.51	0.72	0.94
100-199	0.16	0.22	0.23
200+	0.06	0.09	0.11
Total Entrants	94,537	61,888	37,972

Tables

Table 2: Overview of data

Additional variables merged to the establishment data

Target Variable	Proxied Using	Source
Spillovers	Follows Ellison, Glaeser, and Kerr (2010)	
Goods	Input-Output Tables	German statistical library
Labor	Vector correlation of occupational shares between two industries	Linked employer-employee data (SIAB)
Knowledge	Measure from Jaffe et al. (1993)	OECD Patent Database
Taxes	Weighted average of municipal corporate tax rates within the commuting zone. A multiplier of the federal tax rate.	German statistical library
Outside Option - Estimated Average Union Contract Minimum	20th percentile of establishment-median low-skill workers wages in an industry-state cell in the BHP	Derived using the BHP, union contracts obtained from Tarifregister Nordrhein-Westfalen

Table 3: Response of the Share Ratio to Taxes, Spillovers, and Outside Option

	Ag. For. Fish.	Mine., Util., Constr.	Manu.	Trade, Transport	Prof. Serv.	Edu., Health.	Other
Tax	13.65***	10.97***	7.197**	5.353***	8.069***	8.434	-7.097
Outside Option	0.00214	-0.0444***	0.0153	-0.0378***	-0.0224	-0.0167**	-0.0129
Input-Output	6.550***	3.375***	-2.819**	6.646	4.052*	1.427	11.56***
Labor Correlation	9.400	2.479	0.673	-6.542	-0.390	0.223	-35.72***
Knowledge	-0.00552	0.0316	0.0747**	0.00504	0.00867	0.00941	0.0379
N	2561	2660	2637	2660	2657	2654	2660
F	14.45	212.1	116.8	412.7	242.8	150.4	107.4

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results of the reduced-form specification in equation 10

Table 4: Response of the Share Ratio to Taxes, Spillovers, and Outside Option: Alternative Sector Definition

	Ag. For. Fish.	Mine., Util., Constr.	Goods Manu.	Chem., Pharm. Manu.	Metal. Manu.	Elec. Manu.	Mach. Manu.	Other Manu.
Tax	5.245	10.73***	11.34**	10.63	6.418*	-6.543	10.65**	-14.62
Outside Option	-0.0737	-0.0483***	-0.0724**	0.0126	-0.0516**	0.0673**	0.0137	0.149
Input-Output	-14.55*	4.866***	4.475**	-3.977	1.093	3.698***	1.108	-17.92
Labor Correlation	-22.89	0.0672	-24.28	46.79	51.30***	-258.7***	-6.115	187.4
Knowledge	-0.0227	0.0319	0.0297	0.0201	0.0682	0.0895	0.0184	0.141
N	2561	2660	2381	1978	2390	1624	2001	2004
F	4.326	187.5	30.96	20.50	28.74	13.26	24.89	4.654

	Trade	Trans., Logis.	Arts, Rec.	Media, Comm.	Tech. Serv.	Bus. Serv.	Edu., Health	Other
Tax	5.938**	6.085*	29.73**	1.134	5.480	-6.569	25.95*	-74.77
Outside Option	-0.0236	-0.0445***	-0.194**	0.0353*	-0.171*	0.0812	-0.0188	0.0485
Input-Output	22.57**	-15.42*	-3.838*	-1.057	-6.102	11.21	8.209	51.29
Labor Correlation	-25.05**	199.1**	365.1**	23.71	-142.1	13.86	-68.32	-88.74
Knowledge	-0.0281	-0.147**	0.0887	-0.0109	0.0565	0.0167	0.0371	0.162
N	2660	2615	2423	2098	2636	2645	2654	2659
F	97.80	50.77	27.99	44.99	35.90	43.09	70.58	7.154

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results of the reduced-form specification in equation 10

Table 5: Share of Establishments Within Sector by Tradability Group

	Non Tradable	Low Tradability	Medium Tradability	High Tradability
Ag. For. Fish.	0.00	0.00	10.55	89.45
Mine., Util., Constr.	15.06	77.86	7.03	0.05
Manufacturing	0.00	0.00	55.93	44.07
Trade, Transport	86.55	13.45	0.00	0.00
Prof. Serv.	69.07	12.53	18.40	0.00
Edu., Health.	100.00	0.00	0.00	0.00
Other	70.18	7.42	22.39	0.00

Notes: Tradability definitions from Dauth et al. (2017). The tradability index is level of import penetration and export opportunities at the two-digit-industry level. Non-tradable industries are those below the 10th percentile, low tradable those between the 10th and 40th percentile, medium tradable those between the 40th and 70th, and highly tradable those above the 70th percentile.

Table 6: Heterogeneity of Results by AKM Firm Effect

Panel A: Above-Median AKM Effect

	Ag. For. Fish.	Mine., Util., Constr.	Manu.	Trade, Transport	Prof. Serv.	Edu., Health.	Other
Tax	8.414	10.49***	4.248	7.234***	5.067	-15.61**	2.603
Outside Option	0.0152	-0.0108	0.0206	-0.0212*	-0.0332	0.00831	0.0419**
Input-Output	7.815***	3.472*	1.401	2.895	-3.370	-7.895**	7.739**
Labor Correlation	4.077	-0.248	5.722	-4.987	-2.282	43.48***	-28.10**
Knowledge	0.0748	0.0148	-0.0516	-0.0470	0.0386	-0.0493	0.0276
N	1426	2590	2179	2641	2456	2302	2551
F	5.037	66.76	40.92	164.0	92.55	50.56	72.63

Panel B: Below-Median AKM Effect

	Ag. For. Fish.	Mine., Util., Constr.	Manu.	Trade, Transport	Prof. Serv.	Edu., Health.	Other
Tax	2.062	7.383**	14.48***	3.072	11.10***	12.16	-5.339
Outside Option	-0.0561***	-0.0802***	-0.0352	-0.0540***	-0.0410	0.0135	-0.0446***
Input-Output	1.594	7.708***	1.794	23.71*	4.628	-0.0615	7.044*
Labor Correlation	7.416	-5.016	-2.951	-18.97	5.443	15.45	-32.38***
Knowledge	-0.0462	0.0526	0.0840	0.0312	-0.0234	0.0838*	0.0127
N	1479	2547	2332	2656	2569	2424	2550
F	14.61	62.28	36.04	64.51	74.32	46.03	67.80

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results of the reduced-form specification in equation 10, splitting the sample by AKM firm-effect.

Table 7: Heterogeneity of Results by Establishment Size

Panel A: Above-Median Size

	Ag. For. Fish.	Mine., Util., Constr.	Manu.	Trade, Transport	Prof. Serv.	Edu., Health.	Other
Tax	7.759	13.14***	8.486**	2.822*	5.826**	-7.549	-10.80
Outside Option	-0.00213	-0.0616***	-0.000687	-0.0313***	-0.00883	-0.00721	0.00495
Input-Output	5.147***	4.137**	0.500	2.912	5.309*	-6.631	12.80***
Labor Correlation	28.04**	10.77**	-0.223	-2.581	3.871	25.44	-42.73***
Knowledge	-0.0560	0.0101	0.0665	0.00958	0.00949	-0.0313	0.0410
N	2143	2634	2535	2659	2619	2586	2634
F	9.700	85.13	64.28	292.3	126.4	78.84	76.99

Panel B: Below-Median Size

	Ag. For. Fish.	Mine., Util., Constr.	Manu.	Trade, Transport	Prof. Serv.	Edu., Health.	Other
Tax	7.308	10.37***	4.629	6.679***	11.50***	14.62**	-2.196
Outside Option	0.0226	-0.0326***	0.0315	-0.0405***	-0.0364	-0.00715	-0.0215
Input-Output	6.575***	2.845**	-3.427*	6.068	2.342	3.128	10.49***
Labor Correlation	-4.739	-1.418	0.615	-7.371	-1.174	-0.873	-31.88***
Knowledge	0.0560	0.0422	0.0358	0.0163	0.0167	0.00632	0.0487
N	2385	2659	2575	2658	2653	2631	2653
F	11.23	164.1	65.26	303.9	179.9	91.80	96.79

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results of the reduced-form specification in equation 10, splitting the sample by establishment size, measured by the number of employees at the establishment.

Table 8: Model Fit - Comparing Imputed and Actual Wages

	Average Wage
Imputed Wage	1.33***
constant	-3.52
N	18,622
R2	.383

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Outcome variable is the average daily wage in the sector-commuting zone-year cell

Table 9: Initial Conditions

	Uckermark	Stendal
Agriculture, Forestry, and Fishing	8.71	8.57
Mining, Utilities, and Construction	15.88	17.91
Manufacturing	7.17	8.83
Trade and Transport	32.76	32.69
Professional Services	10.01	8.32
Education and Health	12.67	11.45
Other	12.79	12.22
Outside Option	37.29	39.69
Tax	10.52	11.51

Notes: Author's calculations using the BHP for the year 1999

Table 10: Effects of Reducing Monopsony Power on Underdeveloped Regions

	Uckermark	Mecklenburgische Seenplatte	Nordvorpommern	Südvorpommern	Stendal
Number (Actual Elasticity)	1093	1049	2713	2518	2368
Number (Elasticity = Hamburg)	756	591	1656	1386	1445

A Adjustment for Establishment Observation

As discussed in Section 4, my unit of observation is an establishment rather than a firm. More precisely, I observe all of the branch offices in a particular municipality as a single line of data (hereafter "establishment"), and I cannot link establishments across municipalities. A simple application of Bayes rule shows how this could potentially bias my empirical result. In my data, a unit of observation is a new establishment that is observed in my dataset, where observed means that a new establishment that is not located in a municipality where the expanding firm is already operating. Therefore, my expression for the share of establishments in sector sec picking a particular location C may be more precisely written as:

$$\text{share}_{sec}(\text{pick C}|\text{observed in data}) \approx \text{pr}_{sec}(\text{pick C}|\text{observed in data}) \tag{1}$$

Bayes' rule implies

$$\text{pr}_{sec}(\text{pick C}|\text{observed in data}) = \frac{\text{pr}_{sec}(\text{pick C})\text{pr}_{sec}(\text{observed in data}|\text{pick C})}{\text{pr}_{sec}(\text{observed in data})} \tag{2}$$

Combining the two expressions and taking the ratio of shares for a base location 0 as in the main analysis

$$\frac{\text{share}_{sec}(\text{pick C}|\text{observed in data})}{\text{share}_{sec}(\text{pick 0}|\text{observed in data})} \approx \frac{\text{pr}_{sec}(\text{pick C})}{\text{pr}_{sec}(\text{pick 0})}\frac{\text{pr}_{sec}(\text{observed in data}|\text{pick C})}{\text{pr}_{sec}(\text{observed in data}|\text{pick 0})} \tag{3}$$

The first term $\frac{\text{pr}_{sec}(\text{pick C})}{\text{pr}_{sec}(\text{pick 0})}$ leads to the same unconditional logit share ratio equation as in the main specification. The second term is what could potentially bias my results. After substituting my functional forms of the logit into the share ratio equation and taking logs, I obtain the following structural equation:

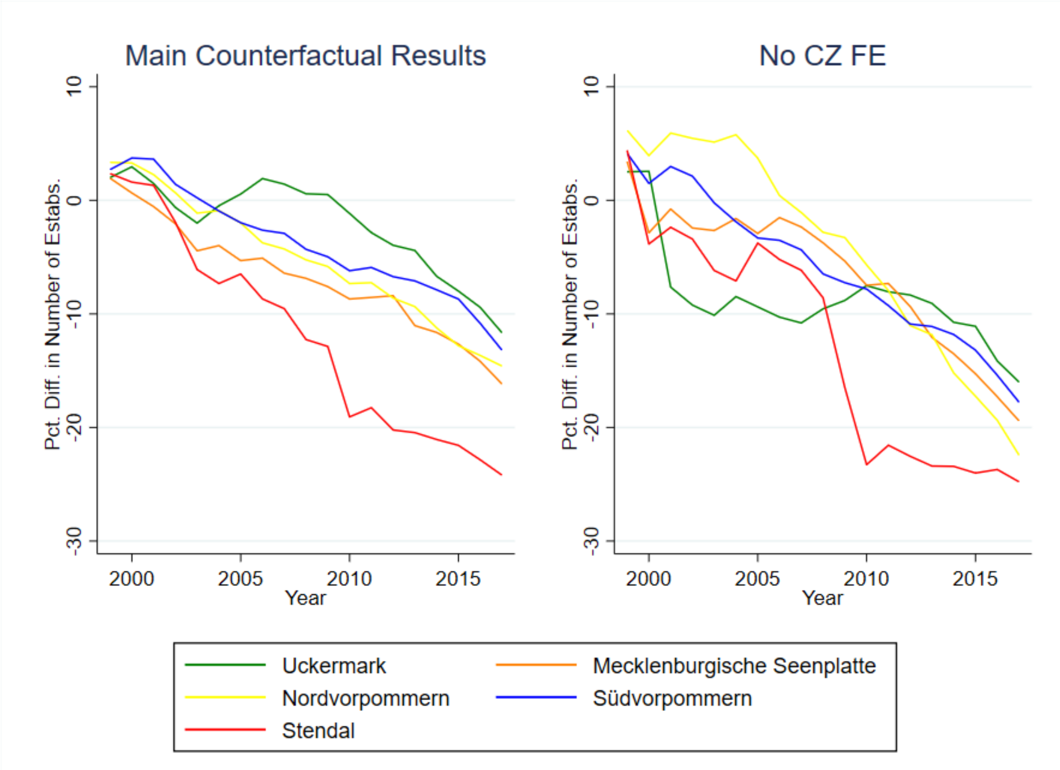
$$ln(\frac{s_{ct,sec}}{s_{0,sec}}) = y_{ct,sec} - y_{0t,sec} + ln(\text{pr}_{sec,t}(\text{observed in data}|\text{pick C})) - ln(\text{pr}_{sec,t}(\text{observed in data}|\text{pick 0})) \tag{4}$$

In my reduced-form analysis, the term $ln(\text{pr}_{sec}(\text{observed in data}|\text{pick 0}))$ is cleanly captured by the time fixed effect, and the term $ln(\text{pr}_{sec}(\text{observed in data}|\text{pick C}))$ is at least partially absorbed by the location fixed effect. To bias the coefficients of interest spillovers or taxes would need to be correlated with $ln(\text{pr}_{sec}(\text{observed in data}|\text{pick C}))$, but not in a way correlated within commuting zone or year.

Of greater concern is the fact that I am using the location fixed effect in order to back out the sector-location productivity for my counterfactual exercises, the inability to observe the location choices of the universe of new establishments could affect these estimates. I cannot directly control for this since, as discussed, there is no way to calculate the probability of observing a new establishment in the data conditionally or unconditionally. In order to test whether this is a problem in practice, I repeat my main counterfactual exercises without including the location fixed effect when I back out my measure of productivity.

The results of this bounding exercise are shown in Figure A.1. As is clear from the figure, the main results are similar whether or not I include the location fixed effect in

Figure A.1: Bounding Exercise: Adjustment for Establishment Observation



my measure of productivity, though the implications of the tax policy for some individual commuting zones may be sightly different in any particular time period. In the main results, I choose to include the fixed effect because it includes factors other than the adjustment for establishment observation, such as natural advantage.

B Pre-estimation of μ_c and $\mathcal{L}_{ct}\lambda_{ct}$

B.1 Estimation of μ_c

I estimate the elasticity of labor supply to the firm using the method of Bassier, Dube, and Naidu (2022). The estimating equation is a regression of an indicator for separating from an establishment:

$$s_{ijt} = \sum_j -(\frac{1}{\mu_c})\phi_j f_{jt}^i + X_{it} + v_{ijt}$$

Where s_{ijt} is an indicator for separation of individual i from establishment j at time t, ϕ_j is the AKM fixed effect of the establishment, and f_{jt}^i is an indicator variable for individual i working at establishment j in time t. Put simply. the coefficient of interest is on the AKM establishment effect.

B.2 Estimation of $\mathcal{L}_{ct}\lambda_{ct}$

With the estimate of μ_c in hand, I turn to estimation of the market size. Recall the labor supply equation of an establishment:

$$L_{ctj}(w_{ctj}) = \mathcal{L}_{ct}\lambda_{ct}exp(\mu_c ln(w_{ctj} - b_{ct}) + a_{sec}) \tag{4}$$

In a log regression, with the pre-estimate of μ_c, the market size may be estimated using a simple fixed-effects regression:

$$ln(L_{ct,sec}(w_{ct,sec})) - \hat{\mu}_c ln(w_{ctj} - b_{ct}) = ln(\mathcal{L}_{ct}\lambda_{ct}) + a_{sec} + \epsilon_{cjt} \tag{1}$$

C Including Rental Prices

Rental prices are a key component of classic spatial equilibrium models. Since I make the assumption that workers are immobile, the inclusion of rental prices does not change the workers' labor supply decision since rental prices they pay do not differ no matter which firm they choose to work at in their commuting zone. However, rental prices will enter the establishment's profit equation. Assume that establishments pay a fixed price r_c per square meter of space they rent. Each worker requires a fixed amount of space k that does not differ between locations. This leads to the profit equation:

$$Y_{jc} = (1 - \tau_{ct})[\beta_{ct,sec}L_{ct,sec}(w_{cjt}) - L_{cjt}(w_{cjt})w_{cjt} - kr_c L_{ct,sec}(w_{cjt})] \tag{1}$$

Taking first order conditions leads to the wage equation:

$$w_{ct,sec} = \frac{\mu_c}{1 + \mu_c}(\beta_{ct,sec} - r_c k) + \frac{1}{1 + \mu_c}b_{ct} \tag{2}$$

This wage equation is very similar to the wage in the main specification, but the productivity portion of the wage is marked down by the price that the establishment needs to pay in rental prices. When functional forms are substituted back into the wage equation and I log-linearize, the log-profits are:

$$y_{cjt} = \mu_c ln(\mu_c) + ln(\mathcal{L}_{ct}\lambda_{ct}) + (1+\mu_c)ln[\frac{1}{1+\mu_c}(\beta_{ct,sec} - r_c k - b_{ct})] + ln(1-\tau_{ct}) + u_{ctj} \quad (3)$$

The difference between this specification is that now rents appear in the productivity term of the equation. I can control for this directly in my reduced form with data on rental prices. I obtained data on rental prices for Germany from the RWI-GEO-REDX dataset maintained by RWI-Essen. Unfortunately, this data is only for residential housing prices rather than commercial real estate prices, but data on commercial prices is not available for Germany.

The dataset provides information on relative housing prices within each district (Klick and Schaffner 2020). I combine the reported fixed effects from the first cross-sectional regression 2 and the housing price growth rates from regression 3 of their paper to create a panel dataset of relative housing prices over time which is merged with my main dataset. I report the results of the main regression specification controlling for rents in Appendix Table 3.3. As this data is only available from 2008 forward, so including it as a control necessitates cutting my panel in half. Therefore, I additionally report the results of the main specification without controlling for rents for the same set of years 2008 to 2017.

The coefficient estimates are not significantly different when controlling for rental prices compared to not, and the only sector with a significant coefficient on rental prices is the Mining, Utlities, and Construction sector. This coefficient is actually positive, likely reflecting the construction industry being able to sell new buildings for more in areas where real estate is expensive.

Table 3.3: Affect of Rental Prices

Panel A: Rental Price Controls

	Ag. For. Fish.	Mine., Util., Constr.	Manu.	Trade, Transport	Prof. Serv.	Edu., Health.	Other
Tax	-0.875	6.747**	7.715	2.691	5.505*	-3.551	-3.321
Outside Option	0.0330	-0.0472***	0.00198	-0.0171	0.0305	-0.0121	-0.0268**
Input-Output	5.038**	5.543*	-1.468	11.03	6.626	-7.609	8.027*
Labor Correlation	8.216	7.240*	-3.682	-3.972	-0.918	4.876	-1.939
Knowledge	0.0131	-0.0192	0.0858	-0.00812	-0.0263	-0.00226	0.0502
Rental Prices	0.00688	0.00718**	0.00180	0.00276	-0.00107	-0.00825	-0.00316
N	1345	1400	1380	1400	1397	1396	1400
F	13.77	90.16	57.37	163.0	108.4	66.57	87.64

Panel B: No Rental Price Controls

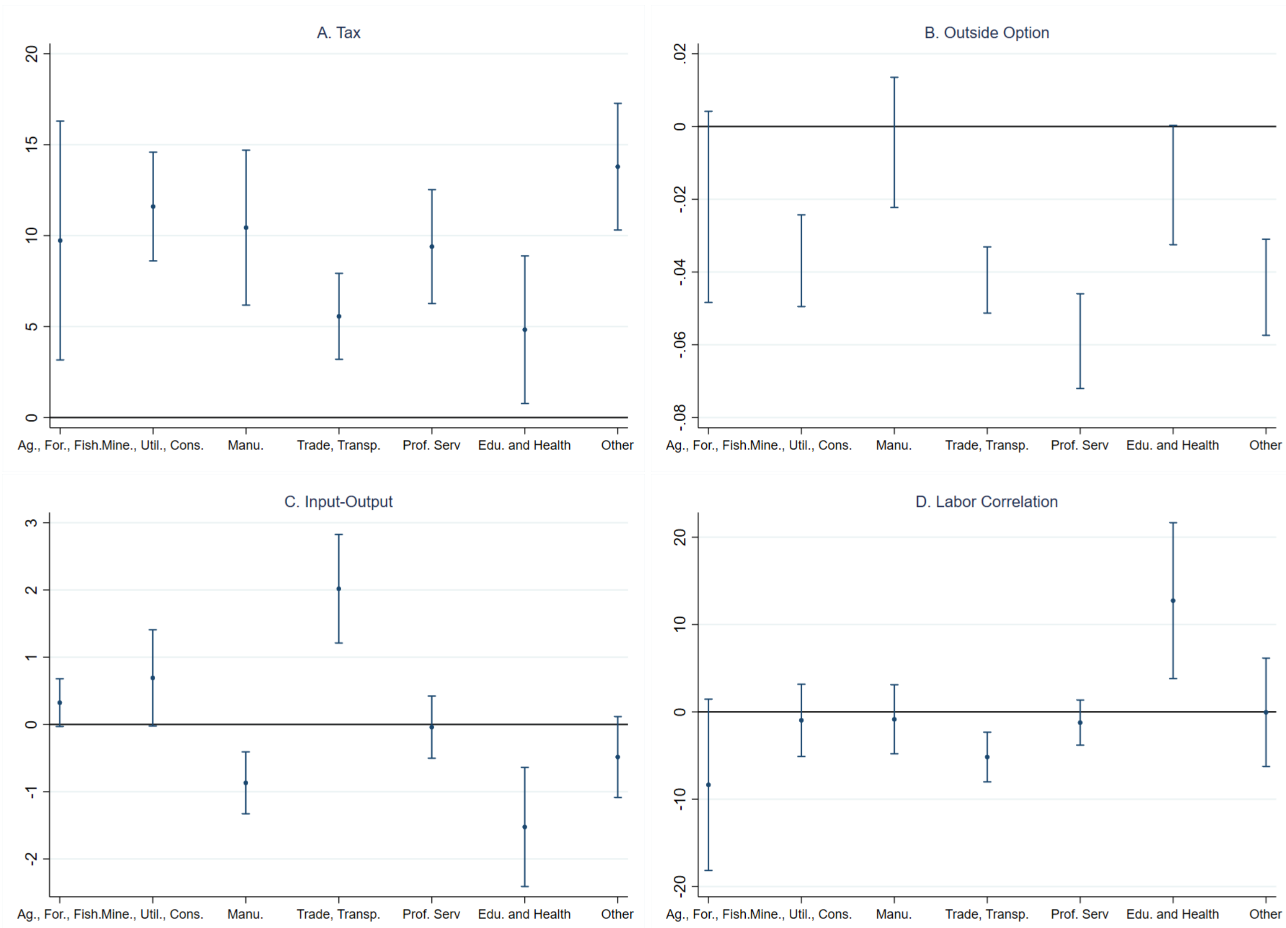
	Ag. For. Fish.	Mine., Util., Constr.	Manu.	Trade, Transport	Prof. Serv.	Edu., Health.	Other
Tax	2.863	8.789***	8.919	3.144	5.129	-7.867	-4.016
Outside Option	0.0224	-0.0453***	-0.00255	-0.0209	0.0302	-0.00797	-0.0256**
Input-Output	4.396**	4.760*	-1.030	8.830	6.512	-8.699	7.698*
Labor Correlation	5.632	9.062**	-3.215	-3.272	-0.777	8.829	-0.609
Knowledge	0.0107	-0.0177	0.0854	-0.00817	-0.0251	-0.00468	0.0489
N	1345	1400	1380	1400	1397	1396	1400
F	15.09	94.48	57.47	195.3	110.3	63.99	90.37

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results of the reduced-form specification in equation 10, also controlling for rental prices.

D Additional Results

Figure D.1: Coefficients by Sector - OLS



Notes: These figures show the point estimates of coefficients of the variables of interest for each industrial sector in the OLS regression 10, as well as the 95% confidence intervals. See Section 4 for data sources and variable definitions.