Optimal minimum wages in spatial economies*

Gabriel M. Ahlfeldt Duncan Roth Tobias Seidel

This version: May 23, 2023
First version: December 17, 2021

Abstract

We develop a quantitative general-equilibrium framework for the normative evaluation of minimum wages in spatial economies with monopsonistic labour markets. We quantify the model for German micro-regions and successfully over-identify its predictions against the effects of the 2015 German minimum wage observed in data. Simulating the model, we find that at low levels, spatially blind national minimum wages can increase welfare and spatial equity simultaneously. At higher levels, however, welfare gains are traded against employment losses and spatial inequality. Because regional minimum wages are not spatially blind, they can increase employment and welfare in a spatially neutral manner.

Key words: General equilibrium, minimum wage, monopsony, employment, Germany, inequality
JEL: J31, J58, R12

*Ahlfeldt: London School of Economics and Political Sciences (LSE) and CEPR, CESifo, CEP; g.ahlfeldt@lse.ac.uk; Roth: Institute for employment research (IAB), Duncan.Roth@iab.de; Seidel: University of Duisburg-Essen, CESifo, CRED; tobias.seidel@uni-due.de. We are grateful to comments and suggestions we received at workshops and conferences in Amsterdam, Düsseldorf (UEA), Basel (VfS), Berlin, (DIW conference on the evaluation of minimum wages, minimum wage commission conference), Barcelona (Urban Economics Workshop), Bonn (German Economists Abroad & MacroFinance and Macro-History seminar), Boston (MIT), Cleveland (Fed, online), Cologne (Econtribute seminar), Helsinki (GSE), Kassel (Regionaloekonomischer Ausschuss), Konstanz (Uni), London (CEP, LSE), Madrid (Carlos III), Marburg (MACIE research seminar), Melbourne (WEAI, online), Milano (EEA), Minneapolis (SOLE), Montreal (UQAM), Munich (ifo), Nuremberg (FAU & IAB), Rome (Bank of Italy), Rotterdam (Erasmus University) Paris (OECD, PSE), Washington D.C. (UEA), and in particular to Gaetano Basso, Kristian Behrens, Mario Bossler, Francesca Carta, Emanuele Ciani, Wolfgang Dauth, Gilles Duranton, Maximilian von Ehrlich, Alexandra Fedorets, Sebastian Findeisen, Bernd Fitzenberger, Aart Gerritsen, Mathias Kredler, Baris Kaymak, Steve Gibbons, Simon Jaeger, Philip Jung, Pat Kline, Piotr Lewandowski, Attila Lindner, Julien Martin, Florian Mayneris, Andreas Mense, Guy Michaels, Enrico Moretti, Henry Overman, Roberto Pinheiro, Michael Pflüger, Frédéric Robert-Nicoud, Hans Koster, Joachim Möller, Michael Reich, Matti Sarvimäki, Gregor Singer, Michel Serafinelli, Sebastian Siegloch, Jens Südekum, Jan Stuhler, Erik Verhoef, Jos van Ommeren, Eliana Viviano, Felix Wellschmied, Matthias Wrede, Jens Wrona, and Atsushi Yamagishi. The usual disclaimer applies.

1 Introduction

Rising spatial inequalities have become a source of political polarization and a significant policy concern.[1] Yet, many policies remain spatially blind, although they are not necessarily spatially neutral. Federal minimum wages are a typical example. Whether they increase efficiency or reduce employment in a region of a country critically depends on their level relative to local productivity. In identifying their optimal minimum wage, policy-makers, therefore, face a host of questions that are inherently spatial: How does a federal minimum wage redistribute employment and welfare between regions of different productivity? What is the role of goods and factor mobility in shaping aggregate and distributive effects? Is there a trade-off between spatial equity, aggregate employment and welfare? And can regional minimum wages mitigate such trade-offs?

To answer these questions, we develop a framework for the quantitative evaluation of minimum wages in spatial economies with imperfectly competitive labour markets. Our quantitative spatial model can rationalize positive and negative employment effects at the aggregate and regional level.[2] It is the first to simultaneously account for aggregate and regional effects of a host of outcomes that have been documented in reduced-form, including effects on labour force participation (Lavecchia, 2020), tradable goods prices (Harasztosi and Lindner, 2019), housing rents (Yamagishi, 2021), or commuting costs (Pérez Pérez, 2020) and worker-firm matching (Dustmann et al., 2022). Our model, therefore, is uniquely equipped to study the effects of minimum wages on aggregate employment and welfare as well as on the spatial distribution of economic activity. A novel insight that emerges from our quantitative analysis is that a minimum wage can theoretically *increase* aggregate employment and welfare and *reduce* spatial inequality at the same time. In a competitive labour market, a binding minimum wage inevitably leads to a reduction in labour demand in the least productive regions, lowering aggregate welfare and amplifying spatial disparities. In our model, an appropriately set minimum wage can reduce monopsony power and increase labour demand in regions of below-average productivity. Via migration, commuting and labour market entry, labour supply adjusts across regions to accommodate the increase in demand, resulting in a less polarized geography of jobs. We find, however, that minimum wages that serve this purpose are generally lower than minimum wages that maximize aggregate welfare, implying a trade-off between aggregate and spatial distributional effects of national minimum wages. Ambitious minimum wages in the range of 60-70% of the national median wage—which are currently debated in the EU, the UK, and the US—may increase welfare at the cost of sizable job loss, which will be concentrated in the economically most vulnerable regions. Moderate regional minimum wages offer an attractive alternative that can achieve similar welfare gains as ambitious federal minimum wages, plus job creation throughout the country.

[1] See Moretti (2012) and Gaubert et al. (2021) for trends in spatial inequalities and Fetzer (2019) and Autor et al. (2020) for the effects of local economic conditions on political polarization.

[2] For surveys of the empirical literature, see Manning (2021); Neumark and Shirley (2022).

Our point of departure is a growing literature that has established that in a monopsonistic labour market,[3] a minimum wage can, in theory, raise wages without reducing employment (Stigler, 1946; Manning, 2003a).[4] Our *theoretical contribution* is to extend this line of research by developing the intuition for the spatial reallocation effects of minimum wages in a setting where productivity varies across firms and regions. In this setting, the minimum wage has no effect on *unconstrained* firms, which voluntarily pay wages above the minimum wage. All other firms can no longer lower the wage below the minimum wage, which implies that they lose some of their monopsony power. Among those, the more productive firms will respond by hiring all workers they can attract at the minimum wage—the new marginal cost of labour—which is why we refer to them as *supply-constrained*. Consequentially, they will increase employment. Less productive firms will hire until the marginal revenue product of labour (MRPL) falls below the minimum wage level, which is why we term them *demand-constrained*. Any demand-constrained firm that initially produces at a MRPL below the minimum wage will have to reduce employment to stay in the market once a minimum wage is introduced. The aggregation of the employment response across all firms within regions of different average productivity delivers the prediction that the regional employment effect of a federal minimum wage is a hump-shaped function of regional productivity. The employment response peaks in regions where the minimum-wage effect is driven by supply-constrained firms. Despite paying higher wages, these firms employ more workers, some of which will be attracted from more productive regions as workers save commuting or living costs. This is an important insight because it suggests that policy can reduce spatial inequalities in prices (wages) and quantities (employment), by choosing the national minimum wage so that the employment response is maximized for regions of relatively low productivity.

Our *empirical contribution* is to substantiate the central prediction of a hump-shaped relationship between the regional employment response to a national minimum wage and the regional productivity using a transparent reduced-form approach. The first-time introduction of a relatively high nationally uniform minimum wage (54% of the national median wage)[5] in Germany in 2015 represents an ideal case in point. The granularity of the linked-employer-employee data covering the universe of 30M workers from the Institute for Employment Research (IAB) allows us to leverage rich heterogeneity in regional productivity across 4,421 municipalities. We find that the regional employment response is flat in the regional wage level for high-productivity regions, where the 2014 mean hourly wage exceeds €18.6. Compared to this group, regions with a mean hourly wage of more than €13.1 tend to gain employment whereas those with a lower mean wage tend to lose. These estimates of a theory-consistent regional distribution of minimum wage-induced em-

[3] Manning (2020) offers a recent review of the literature.

[4] Similarly, search models do not restrict the sign of the employment effect of a minimum wage (Brown et al., 2014; Blömer et al., 2018; Vergara, 2022).

[5] This number relates to the hourly wages of full-time and part-time workers (see Section 2.2 for further information). Based on the wages of full-time workers, the Minimum Wage Commission reports that the Kaitz Index was 46% in Germany in the year 2018 (Mindestlohnkommission, 2018).

ployment effects add to a literature that has mostly focused on average treatment effects (e.g. Card and Krueger, 1994; Dustmann et al., 2022) or point estimates of the effect of the minimum wage bite (e.g. Machin et al., 2003; Ahlfeldt et al., 2018). Indirectly, they provide evidence supporting the monopsonistic labour market model that is still scarce (Neumark, 2018). Importantly, we bring to light a sizable negative employment effect in the least productive micro regions that has gone unnoticed in previous studies analyzing larger spatial units (Ahlfeldt et al., 2018; Caliendo et al., 2018; Dustmann et al., 2022).

Encouraged by the novel reduced-form support for the spatial reallocation effect (e.g. Dustmann et al., 2022; Engbom and Moser, 2022), we proceed to the quantitative evaluation within a general equilibrium framework. For one thing, we wish to quantitatively account for how the minimum wage reallocates workers across regions via migration and, in particular, commuting. Indeed, it is well-documented in the literature that the commuting openness is an important moderator of how local labour markets respond to local labour demand shocks (Monte et al., 2018). For another, we wish to account for the transmission of minimum wage effects across regions of different productivity via domestic trade and spatially varying effects on consumer prices. This is crucial for the correct measurement of minimum wage effects on real spatial inequalities (Moretti, 2013). For the same reason, we wish to capture how the minimum-wage-induced wage and employment effects affect regional housing prices. In developing a suitable model, we can draw from a quantitative spatial economics literature that offers a canonical framework to account for spatial linkages (Redding and Rossi-Hansberg, 2017; Monte et al., 2018). Workers choose where to live, where to work and how much to consume of a composite tradable good and housing, trading expected wages and amenities against commuting cost, goods prices and housing rents. Goods are produced in a monopolistically competitive market and traded at a cost. Housing is supplied inelastically, creating a congestion force that restores the spatial equilibrium.[6]

Our *methodological contribution* is to extend this framework in three important ways to make it amenable to the evaluation of minimum wage effects. First, we borrow from the trade literature and introduce a Pareto-shaped productivity distribution of firms (Melitz, 2003; Redding, 2011; Gaubert, 2018). This extension is critical to enabling the minimum wage to reallocate workers to more productive establishments within a region. It is also the first ingredient we require to generate a wage distribution within regions. Second, we follow Egger et al. (2022), who build on Card et al. (2018),[7] and generate an upward-sloping labour supply curve to the firm via Gumbel-distributed idiosyncratic preferences for employers, in addition to allowing for idiosyncrasy in preferences for residence and workplace locations (Ahlfeldt et al., 2015). This extension is critical to awarding employers monopsony power. It is also the second ingredient we require to generate a wage

[6]This canonical framework draws from Allen and Arkolakis (2014) and Ahlfeldt et al. (2015) who, in turn, build on Eaton and Kortum (2002).

[7]Haanwinckel (2020) and Dustmann et al. (2022) model non-pecuniary aspects of job choice in a similar way.

distribution within regions. Third, we generate imperfectly elastic aggregate labour supply via a Gumbel-distributed idiosyncratic utility from abstaining from the labour market. This extension is critical to capturing incentives minimum wages can create for workers to become active on the labour market and search for jobs (Mincer, 1976; Lavecchia, 2020).

Our *quantitative contribution* is to use the model to provide the first evaluation of the aggregate and distributional effects of the German minimum wage in a spatial general equilibrium. To this end, we leverage on the the German matched worker-establishment micro data to estimate the structural parameters that govern the wage distribution within regions. Our estimates of the labour supply elasticity to the firm are within the typical range found in the literature and provide direct evidence of monopsonistic labour markets in Germany (Sokolova and Sorensen, 2020; Yeh et al., 2022). Taking advantage of a recent micro-geographic house price index developed by Ahlfeldt et al. (2023), we then invert the model in 2014—the year before the introduction of the minimum wage—at the level of 4,421 municipalities. This high level of spatial disaggregation is critical to capturing spatial adjustments that operate via the commuting margin. Solving the model under the minimum wage of 48% of the national mean that we observe in our data delivers the comparative statics from which we infer the minimum wage effect. We find similar regional wage levels that characterize the hump-shape of the regional employment response as in the reduced-form analysis. The important advantage of the model-based general-equilibrium approach is that we do not have to assume any group of firms, workers, or regions to be unaffected by the minimum wage, which allows us to establish the aggregate employment effect. While the hump-shape in the model resembles our reduced-form estimates, we gain the additional insight that employment increases in regions of intermediate productivity at the expense of the least *and* most productive regions. Our model-based counterfactuals also allow us to uncover that labour supply adjustments via the commuting margin alone can rationalize the hump-shape; migration is not a necessary facilitator. In the national aggregate, full-time equivalent employment decreases by about 0.3% or 100K jobs. Ancillary empirical analyses suggest that this decrease is driven by a reduction in working hours rather than an increase in unemployment, which is consistent with extant reduced-form evidence (Bossler and Gerner, 2019; Dustmann et al., 2022). In any case, the employment effects are small compared to the predictions derived from competitive labour market models (Knabe et al., 2014).[8]

For our purposes, the ability of our model to speak to welfare effects is, at least, as important as establishing aggregate employment effects. We find that the German minimum wage has increased welfare, as measured by the expected utility of a representative worker, by 2.1%. This estimate of the minimum wage welfare effect is unprecedented in the literature in that it accounts for changes in nominal wages, employment probabilities, goods prices, housing rents, the quality of the worker-firm match, the reallocation

[8] Our model does not distinguish between employment effects at the extensive (unemployment) and intensive (working hours) margins. Comparisons to data suggest that the employment effect is driven by the intensive margin (see Section 4.4.2).

of workers across firms, commuting destinations, residences, and the growing number of workers who decide to be active on the labour market. In other words, the increase in real wages—adjusted for changes in tradable goods prices, housing rents, and commuting costs—dominates the reduction in the employment probability in terms of effects on the expected wage. As as a result, about 180K workers become active on the labour market and start searching for jobs. Again, there is significant spatial heterogeneity. The net-winners are low-productivity regions such as in the eastern states, resulting in long-run incentives for workers to relocate to regions that have experienced sustained population loss over the past decades.

In a demanding over-identification test, we show that the model's predictions for regional minimum wage effects in wages, workplace employment, full-time employment probability, labour force participation, commuting distance, average establishment productivity and establishment size are closely correlated with observed before-after changes in data. We also show that our model predicts changes in the Gini coefficients of wage inequality across all workers and employment distribution across regions that are in line with before-after changes observed in data. This suggests significant out-of-sample predictive power, which is reassuring with respect to our key *normative contribution*: The derivation of optimal minimum wages in spatial economies.

To this end, we compute aggregate full-time equivalent employment effects and welfare effects for a broad range of federal and regional minimum-wage schedules. We also provide two equity measures that summarize the distribution of wages across workers as well as economic activity across regions. Hence, we equip our readers with the key ingredients to compute their own optimal minimum wage. Under canonical welfare functions, the optimal *federal* minimum wage will not be lower than the employment-maximizing minimum wage, at 38% of the national mean wage. Up to 58%, the minimum wage can be justified on the grounds of welfare effects. Higher levels require equity (among those in employment) as an objective. Ambitious minimum wages need to be defended against negative employment effects that start building up rapidly beyond 50% of the national mean wage. Against this background, it is important to note that the employment-maximizing *regional* minimum wage, at 50% of the regional mean wage, would deliver positive welfare effects that are similar to the federal welfare-maximizing minimum wage (3.9%), plus an increase in employment by 1.1%. Hence, a relatively easy-to-implement regional minimum wage can go a long way in addressing the long lasting concern that federal minimum wages—because they disregard productivity differences across firms—only realize a fraction of the potential efficiency gains (Stigler, 1946).

With these results, we contribute to the identification of turning points where the costs of minimum wages start exceeding the benefits (Manning, 2021). In doing so, we complement a large literature using reduced-form approaches that suggest that minimum wages may (Meer and West, 2016; Clemens and Wither, 2019) or may not (Dube et al., 2010;

Cengiz et al., 2019) have negative employment effects.[9] This includes a growing literature evaluating the labour market effects of the German minimum wage, which we review in more detail in the Online Supplement (e.g. Ahlfeldt et al., 2018; Bossler and Gerner, 2019; Caliendo et al., 2018; Dustmann et al., 2022). Another contribution to this literature is to show that the reallocation of workers across establishments of different productivity documented by Dustmann et al. (2022) can work in either direction, depending on the regional productivity level.

We also contribute to a smaller normative literature on minimum wages that considers distributional effects of minimum wages (see, e.g. Chen and Teulings, 2021; Lee and Saez, 2021; Simon and Wilson, 2021).[10] Three current working papers study aggregate and distributional effects of minimum wages within non-spatial models. Drechsel-Grau (2021) studies the effects of the German minimum wage in a search-and-matching model with frictional unemployment. Berger et al. (2022) study the distributional effects of minimum wages in the US focusing on firm heterogeneity. Hurst et al. (2022) study the distributional effects focusing on worker heterogeneity. In contrast, we focus on spatial heterogeneity. Our contribution to this literature is to provide a quantitative framework that accounts for spatial margins of adjustment (commuting, migration, and trade) and is amenable to the normative evaluation of spatially varying minimum wages as well as the spatially heterogeneous effects of spatially invariant minimum wages.

Most closely, we connect to a nascent literature that studies the effects of minimum wages in spatial equilibrium. Our contribution complements Monras (2019), Pérez Pérez (2020) and Simon and Wilson (2021) who consider a competitive labour market. To our knowledge, the only other model that nests a monopsonistic labour market in a spatial general equilibrium is in the current working paper by Bamford (2021). At a higher level of spatial aggregation and abstracting from frictional trade and commuting, he also provides an evaluation of the German minimum wage, but his primary contribution is to show that lower monopsony power acts as an important concentration force in the spatial economy (see also Azar et al., 2019).[11] Our contribution to this literature is to provide a new quantitative framework that accounts for monopsonistic labour markets, heterogeneity in productivity across firms and regions, endogenous local and aggregate labour supply, and spatial linkages via migration and costly commuting and trade.

The remainder of the paper is structured as follows. Section 2 introduces the institutional context and our data, and presents stylized evidence that informs our modelling choices. Section 3 introduces a partial equilibrium version of our model and provides transparent reduced-form evidence that is consistent with stylized predictions. Section

[9] A new wave of empirical minimum wage research, based on difference-in-differences designs, started with the seminal paper by Card and Krueger (1994) whose findings, subsequently challenged by Neumark and Wascher (2000), cast doubt on the competitive labour market model which predicts that binding minimum wages necessarily lead to job loss.

[10] Minimum wages also interact with the optimal tax system (Allen, 1987; Guesnerie and Roberts, 1987; Gerritsen, 2022).

[11] We generate larger employment elasticities (Monte et al., 2018) and less monopsony power in thicker labour markets as workers can substitute across commuting destinations (Manning, 2003b; Datta, 2021).

4 develops the full quantitative model and takes the analysis to the general equilibrium. Section 5 concludes.

2 Empirical context

In this section, we introduce the German minimum wage policy, the various sources of data we rely on, and some stylized facts that inform our modelling choices (see Online Supplement for details).

2.1 The German minimum wage

The first uniformly binding federal minimum wage in Germany was introduced in 2015. Since then, German employers had to pay at least €8.50 per hour corresponding to 48% of the mean salary of full-time workers. Because no similar regulation preceded the statutory wage floor, it represented a potentially significant shock to regions in the left tail of the regional wage distribution. Subsequently, the minimum wage has been raised to €8.84 in 2017, €9.19 in 2019 and €9.35 in 2020. In relative terms, it has fluctuated within a close range of 47% to 49% of the national mean wage, suggesting that it is reasonable to treat the introduction of the minimum wage as a singular intervention in 2015.

2.2 Data

We compile a novel data set for German micro regions that is unique in terms of its national coverage of labour and housing market outcomes at the sub-city level.

Employment, establishments, productivity and wages. We use the Employment Histories (BeH) and the Integrated Employment Biographies (IEB) provided by the Institute for Employment Research (IAB) which contain individual-level panel data containing workplace, residence, establishment, wage, and characteristics such as age, gender, and skill on the universe of about 30M labour market participants in Germany. We derive a measure of establishment productivity from a standard decomposition of wages observed before the minimum wage was introduced (Abowd et al., 1999, henceforth AKM).

Hours worked. We follow Ahlfeldt et al. (2018) and impute average working hours separately for full-time and part-time workers from an auxiliary regression that accounts for the sector of employment, federal state of employment, and various socio-demographic attributes and using the 1% sample from the 2012 census. We find that full-time employees work approximately 40 hours per week while the number is lower for regularly employed (21 hours) and for marginally employed part-time workers (10 hours). Combining working hours with average daily earnings delivers hourly wages.

Real estate. We use a locally-weighted regression approach proposed by Ahlfeldt et al. (2023) to generate a municipality-year housing cost index. The raw data comes from Immoscout24, accessed via the FDZ-Ruhr (Boelmann and Schaffner, 2019). It covers nearly 20 million residential observations between 2007 and 2018.

Trade. Trade volumes are taken from the Forecast of Nationwide Transport Relations in Germany which are provided by the Clearing House of Transport Data at the Institute of Transport Research of the German Aerospace Center. The data set contains information about bilateral trade volumes between German counties in the year 2010 for different product groups. Following Henkel et al. (2021), we aggregate trade volumes across all modes of transport (road, rail and water). To convert volumes (measured in metric tonnes) into monetary quantities, we use information on national unit prices for the different product groups. Finally, we aggregate the value of trade flows across all product groups.

Spatial unit. The primary spatial unit of analysis are 4,421 municipal associations (*Verbandsgemeinden*) according to the delineation from 31 December 2018. Municipal associations are spatial aggregates of 11,089 municipalities (*Gemeinden*) that ensure a more even distribution of population and geographic size. Henceforth, we refer to municipal associations as municipalities for simplicity. On average, a municipality hosts 541 establishments employing 6,769 workers on less than 80 square kilometers, making it about a tenth of the size of an average county. For each pair of municipalities, we compute the Euclidean distance using the geographic centroids.

2.3 Stylized facts

The effects of the German minimum wage have been analyzed in a rapidly growing literature that we discuss in the Online Supplement. Here, we highlight the main insights that inform our modelling choices. The minimum wage had significant bite as it compressed the wage distribution, in particular in low-productivity regions where many workers earned less than the minimum wage in 2014. Contradicting ex-ante predictions, there is little evidence for a substantial reduction in aggregate employment. There is, however, substantial heterogeneity in regional employment growth, with regions with the highest *and* lowest bite underperforming relative to regions in the middle of the distribution. Commuting appears to play a significant role in reallocating labour supply to regions where employment has grown. Average establishment productivity has increased more where the minimum wage bit harder. Yet, a substantial share of the increase in labour cost has been passed on to consumers via higher prices of tradable and non-tradable goods. In the remainder of this paper, we develop a quantitative spatial general equilibrium model that can rationalize these stylized facts.

3 Partial equilibrium analysis

In this section, we develop a model of optimal behaviour of heterogeneous firms in a monopsonistic labour market with a minimum wage. We first use the model to develop the intuition for why the minimum wage reallocates workers to firms of intermediate productivity *within* a region. We then derive the novel prediction that in a comparison *between* regions, the aggregate employment response is a hump-shaped function of productivity.

This prediction is key to understanding how a minimum wage can result in a more balanced spatial distribution of employment. Therefore, we provide novel reduced-form estimates of the employment effect of the minimum wage that confirm this prediction before we turn to the quantitative general equilibrium analysis within the model.

3.1 Model I

For now, we take upward-sloping labour supply to the firm as well as downward-sloping product demand as exogenously given. We nest the firm problem introduced here into a quantitative spatial model in Section 4. The extended model will provide the micro-foundations for the labour supply and product demand functions and allow us to solve for the spatial general equilibrium of labour, goods, and housing markets.

3.1.1 Optimal firm behaviour

A firm in location $j \in J$ sells its product variety at monopolistically competitive goods markets across all locations $i \in J$. Because one firm produces only one variety, we use ω_j to denote both a firm and its variety. Given a productivity φ_j, firm ω_j hires $l_j(\omega_j)$ units of labour in a monopsonistically competitive labour market which it uses to produce output $y_j(\omega_j) = \varphi_j(\omega_j) l_j(\omega_j)$.

Labour supply. Firm ω_j faces an iso-elastic labour supply function

$$h_j(\omega_j) = S_j^h \left[\psi_j(\omega_j) w_j(\omega_j)\right]^{\varepsilon} \tag{1}$$

of the expected wage $\psi_j(\omega_j) w_j(\omega_j)$ that a worker earns in this firm, with $w_j(\omega_j) > 0$ being the firm's wage rate and $\psi_j(\omega_j) \in (0, 1]$ being the expected ratio of hours worked over full-time working hours. For convenience, we refer to this fraction as the *hiring probability.*[12] Unless otherwise indicated, we assume $\psi_j(\omega_j) = 1$ to ease notations. Notice that each worker can only be matched to one firm. We denote the firm's constant labour supply elasticity by $\varepsilon > 0$ and introduce $S_j^h > 0$ as an aggregate shift variable that summarizes all general equilibrium effects operating through location j's labour market (specified in more detail below and solved in general equilibrium in Section 4).

Goods demand. Similarly, there is iso-elastic demand for variety ω_j in location i

$$q_{ij}(\omega_j) = S_i^q p_{ij}(\omega_j)^{-\sigma}, \tag{2}$$

which depends inversely on the variety's consumer price $p_{ij}(\omega_j)$ with a constant price elasticity of demand $\sigma > 1$, and which is directly proportional to an aggregate shift variable $S_i^q > 0$ that summarizes all general equilibrium effects operating through location i's goods market (specified in more detail below and solved in general equilibrium in

[12] If a worker offering labour to a firm expects to remain unemployed and earn a zero wage during 40% of a year's working days, the hiring probability will be 60%.

Section 4). Under profit maximization and goods market clearing, we can express the revenue function as

$$r_j(\omega_j) = \sum_i p_{ij}(\omega_j) q_{ij}(\omega_j) = \left(S_j^r\right)^{\frac{1}{\sigma}} [y_j(\omega_j)]^\rho, \tag{3}$$

where $\rho = \frac{\sigma-1}{\sigma} \in (0,1)$. Intuitively, a greater market access $S_j^r \equiv \sum_i \tau_{ij}^{1-\sigma} S_i^q > 0$ implies that a smaller fraction of output melts away due to iceberg trade costs $\tau_{ij} \geq 1$, leading to relatively larger revenues (see Appendix A.1).

Minimum wage. In deriving the effects of a statutory minimum wage $\underline{w}$ on price, output, and labour input, it is instructive to distinguish between three firm-types: unconstrained firms (indexed by superscript u), for which the minimum wage $\underline{w}$ is non-binding; supply-constrained firms (indexed by superscript s), whose labour demand exceeds labour supply at the binding minimum wage $\underline{w}$; and demand-constrained firms (indexed by superscript d), that attract more workers than they require when the minimum wage $\underline{w}$ is binding. We present the key results for the three firm types below and refer to Appendix A.2 for further derivations. As each firm can be fully characterized by its productivity level and its firm-type, we drop the firm index ω_j in favour of a more parsimonious notation, combining the firm's productivity level φ_j with superscript $z \in \{u, s, d\}$.

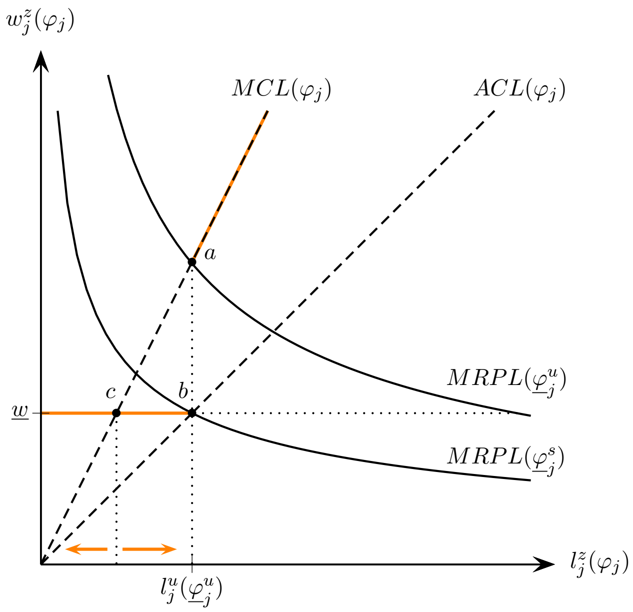
Unconstrained firms choose profit-maximizing wages that are larger or equal to the minimum wage level. Therefore, we can use the labour supply function to the firm in Eq. (1) to derive the relevant cost function

$$c_j^u(\varphi_j) = w_j^u(\varphi_j) l_j^u(\varphi_j) = \left(S_j^h\right)^{-\frac{1}{\varepsilon}} l_j^u(\varphi_j)^{\frac{\varepsilon+1}{\varepsilon}}. \tag{4}$$

Facing an upward-sloping labour supply function, firms can only increase their employment by offering higher wages. Hence, the average cost of labour $ACL(\varphi_j) = c_j^u(\varphi_j)/l_j^u(\varphi_j)$ is upward-sloping as illustrated in Figure 1. The marginal cost of labour $MCL(\varphi_j) = \partial c_j^u(\varphi_j)/\partial l_j^u(\varphi_j) = \frac{\varepsilon+1}{\varepsilon} ACL(\varphi_j)$ is also upward-sloping and strictly greater than $ACL(\varphi_j)$. Since demand for any variety is downward-sloping, an expansion of production and labour input is associated with a lower marginal revenue product of labour $MRPL(\varphi_j) = \partial r_j^u(\varphi_j)/\partial l_j^u(\varphi_j)$. Unconstrained firms find the profit-maximizing employment level by setting $MRPL(\varphi_j) = MCL(\varphi_j)$ which corresponds to point a in Figure 1. Since a higher productivity shifts the $MRPL(\varphi_j)$ function outwards, more productive firms hire more workers at higher wages (Oi and Idson, 1999). Unconstrained firms simultaneously act as monopolists in the goods market and monopsonists in the labour market, setting their prices as a constant mark-up $\sigma/(\sigma-1) > 1$ over marginal revenues and their wages as a constant mark-down $\varepsilon/(\varepsilon+1) < 1$ below marginal costs. The combined mark-up/mark-down factor is $1/\eta \equiv [\sigma/(\sigma-1)][(\varepsilon+1)/\varepsilon] > 1$.

We refer to $\underline{\varphi}_j^u$ as the least-productive unconstrained firm that is identified by setting

Figure 1: Optimal firm employment



$w_j(\underline{\varphi}_j^u) = \underline{w}$, so we obtain

$$\underline{\varphi}_j^u(\underline{w}) = \left(\frac{1}{\eta}\right)^{\frac{\sigma}{\sigma-1}} \left(\frac{S_j^h}{S_j^r}\right)^{\frac{1}{\sigma-1}} \underline{w}^{\frac{\sigma+\varepsilon}{\sigma-1}}. \tag{5}$$

All firms with $\varphi_j < \underline{\varphi}_j^u$ are constrained by the minimum wage. Any increase in the minimum wage level will lead to a firm with a greater productivity becoming the marginal unconstrained firm.

Supply-constrained firms face a binding minimum wage, resulting in $MRPL(\varphi_j) = \underline{w}$. At this wage, workers are willing to supply no more than $h_j^s(\varphi_j) = S_j^h \underline{w}^\varepsilon$ units of labour, which corresponds to $l_j^u(\underline{\varphi}_j^u)$ in Figure 1. Employment is constrained by labour supply because supply-constrained firms would be willing to hire more workers as the MRPL function intersects with $\underline{w}$ at an employment level greater than $l_j^u(\underline{\varphi}_j^u)$. In the absence of the minimum wage, supply-constrained firms would set a wage below $\underline{w}$ to equate MRPL and MCL. At this wage, workers would supply less than $l_j^u(\underline{\varphi}_j^u)$ units of labour. By removing the monopsony power, the mandatory wage floor raises employment for all firms with $\underline{\varphi}_j^s \leq \varphi_j < \underline{\varphi}_j^u$, where $\underline{\varphi}_j^s$ defines the least-productive supply-constrained firm given by

$$\underline{\varphi}_j^s(\underline{w}) = \left(\frac{\eta}{\rho}\right)^{\frac{\sigma}{\sigma-1}} \underline{\varphi}_j^u(\underline{w}) < \underline{\varphi}_j^u(\underline{w}) \quad \text{with} \quad \frac{\eta}{\rho} = \frac{\varepsilon}{\varepsilon+1} < 1. \tag{6}$$

Notice that all supply-constrained firms set the same wage (i.e. the minimum wage) and hire the same number of workers $l_j^s(\varphi_j) = h^s(\varphi_j) = \underline{w}^\varepsilon S_j^h = l_j^u(\underline{\varphi}_j^u)$, (determined by b in Figure 1).

Demand-constrained firms also face a binding minimum wage, resulting in $MRPL(\varphi_j) = \underline{w}$. For these firms with productivities $\varphi_j < \underline{\varphi}_j^s(\underline{w})$, however, employment is constrained by labour demand because at a wage of $\underline{w}$ firms demand less units of labour than workers are willing to supply. To see this, consider the MRPL curve for any firm with productivity $\varphi_j < \underline{\varphi}_j^s$ in Figure 1, which will be below $MRPL(\underline{\varphi}_j^s)$. Since $\underline{w}$ intersects with the MRPL before it intersects with ACL, there is job rationing with a hiring probability $\psi_j^d(\varphi_j) = l_j^d(\varphi_j)/h_j^d(\varphi_j) < 1$. Yet, demand-constrained firms do not necessarily reduce employment. As long as a demand-constrained firm is sufficiently productive for its MRPL curve to be above point c, the MRPL in the monopsony market equilibrium exceeds $\underline{w}$. Therefore, the intersection of MRPL and $\underline{w}$ is necessarily to the right of the intersection of MRPL and MCL, implying greater employment under the minimum wage. The opposite is true, however, for any firm whose productivity is sufficiently small for the MRPL curve to be below point c. Because the MRPL in the monopsony market equilibrium is smaller than $\underline{w}$, the firm has to reduce output and labour input to raise the MRPL to the minimum wage level.

3.1.2 Aggregate outcomes

Having characterized the optimal behaviour of the three firm types, we now explore how the introduction of a minimum wage affects aggregate outcomes at the regional level. To this end, we assume that firm productivity follows a Pareto distribution with shape parameter $k > 0$ and lower bound $\underline{\varphi}_j > 0$. For the following discussion, it is instructive to introduce the critical minimum wage levels $\underline{w}_j^z \;\forall\; z \in \{s, u\}$ as a function of $\underline{\varphi}_j$. They are implicitly defined through $\underline{\varphi}_j^z(\underline{w}_j^z) = \underline{\varphi}_j \;\forall\; z \in \{s, u\}$ and have the following interpretation: For a sufficiently small minimum wage, $\underline{w} < \underline{w}_j^u$, location j features only unconstrained firms. For higher minimum wages, $\underline{w} < \underline{w}_j^s$, location j also features supply-constrained, but no demand-constrained firms. Using Eq. (5), we obtain

$$\underline{w}_j^u = w_j^u(\underline{\varphi}_j) = \left(\eta^\sigma \underline{\varphi}_j^{\sigma-1} \frac{S_j^r}{S_j^h}\right)^{\frac{1}{\sigma+\varepsilon}} \tag{7}$$

as an implicit solution to $\underline{\varphi}_j^u(\underline{w}_j^u) = \underline{\varphi}_j$. Using Eq. (5) in Eq. (6) and solving $\underline{\varphi}_j^s(\underline{w}_j^s) = \underline{\varphi}_j$ for $\underline{w}_j^s$ results in

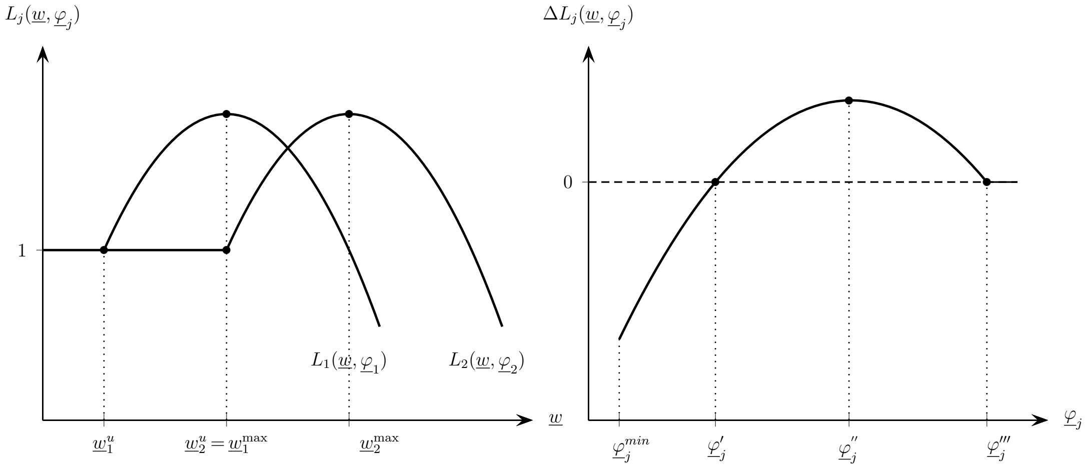
$$\underline{w}_j^s = \left(\rho^\sigma \underline{\varphi}_j^{\sigma-1} \frac{S_j^r}{S_j^h}\right)^{\frac{1}{\sigma+\varepsilon}}, \tag{8}$$

implying $\underline{w}_j^s/\underline{w}_j^u = (\rho/\eta)^{\frac{\sigma}{\sigma+\varepsilon}} > 1$. Using these critical minimum wages we derive the following proposition:

Proposition 1. *Aggregate employment L_j, aggregate labour supply H_j and aggregate revenues R_j are hump-shaped in the minimum wage level. Aggregate profits, Π_j, are declining in $\underline{w}$.*

Proof see Appendix A.4.

Figure 2: Regional employment, minimum wages and productivity



(a) Minimum wages (b) Regional productivity

Note: In this partial-equilibrium illustration, we assume constant general equilibrium terms $\{S_j^r, S_j^h\}$ that are invariant across regions and not affected by the minimum wage.

To develop the intuition, let's first consider the region indexed by $j = 1$ in panel a) of Figure 2. Any minimum wage $\underline{w}_1 \leq \underline{w}_1^u$ will have no effect because all firms in the region are unconstrained as they voluntarily set higher wages. A marginal increase in $\underline{w}_1$ turns some unconstrained firms into supply-constrained firms, whose response to the loss of monopsony power is to hire all workers who are willing to supply their labour at wage $\underline{w}_1^u$. Hence, regional employment increases. Once $\underline{w}_1 > \underline{w}_1^s$, some firms become demand-constrained. The marginal effect of an increase in $\underline{w}_1$ remains initially positive even beyond $\underline{w}_1^s$ because demand-constrained firms still increase the labour input as long as their MRPL exceeds $\underline{w}_1$ in the monopsony market equilibrium. At some point, however, $\underline{w}_1$ will exceed the MRPL of the least productive firms in the market equilibrium and these firms will respond by reducing output and labour input. The marginal effect of $\underline{w}_1$ declines and becomes zero at the employment-maximizing minimum wage $\underline{w}_1^{\max}$. Further increases have negative marginal effects and, eventually, the employment effect will turn negative. Henceforth, we refer to full-time equivalent employment simply as employment for convenience. The generalizable insight is that for given fundamentals $\{S_j^r, S_j^h\}$ and regional productivity summarized by $\underline{\varphi}_j$, aggregate (full-time equivalent) employment $L_j(\underline{w}_j, \underline{\varphi}_j)$ is hump-shaped in the minimum wage level $\underline{w}_j$.

Since the hiring probability for unconstrained and supply-constrained firms is $\psi_j^z = 1 \ \forall \ z \in \{s, u\}$, labour supply defined in Eq. (1) must increase at the extensive margin for low but binding minimum wage levels $\underline{w}_j^u \leq \underline{w}_j < \underline{w}_j^s$ to accommodate the increase in labour demand. In the spatial general equilibrium introduced in Section 4.1, this can happen via the margins of commuting, migration, and labour market entry. At a higher minimum wage level, labour supply negatively responds to the job rationing of demand-constrained firms since workers correctly anticipate the hiring rate ($\psi_j^d(\varphi_j) = l_j^d(\varphi_j)/h_j^d(\varphi_j)$). Thus, the hump-shaped pattern carries through to labour supply. Notice that at a high minimum wage, we can obtain an equilibrium with an aggregate regional hiring probability $\psi_j < 1$, which implies underemployment.[13] A further effect of a high minimum wage is that the reduction in employment of low-productivity demand-constrained firms results in an increase in average productivity, which is the reallocation effect emphasized by Dustmann et al. (2022). For sufficiently low minimum wages, however, the effect can be the opposite as the share of workers employed by higher-productivity demand-constrained firms and supply-constrained firms may increase at the expense of unconstrained firms.[14]

Let us now compare the effect of a uniform minimum wage in region $j = 1$ to a region $j = 2$ in which firms are generally more productive, for example, due to better infrastructure or institutions. To ease the comparison, we normalize initial employment to unity. A low minimum wage $\underline{w}_1^u < \underline{w} \leq \underline{w}_1^{\max}$ leads demand- and supply-constrained firms to hire more workers in region $j = 1$, whereas there is no employment effect in region $j = 2$ since all firms remain unconstrained. At a higher level $\underline{w}_1^{\max} < \underline{w} < \underline{w}_2^{\max}$, an increase in the minimum wage reduces employment in region $j = 1$ because the MRPL of the marginal firm falls below $\underline{w}$, whereas employment increases in region $i = 2$ owing to the loss of monopsony power of formerly unconstrained firms. Hence, the same increase in the minimum wage level can have qualitatively different employment effects in different regions because the employment-maximizing minimum wage depends on regional productivity. This is an important theoretical result that rationalizes why a large empirical literature has failed to reach consensus regarding the employment effects of minimum wage rises (Manning, 2021).

Of course, regional productivity not only affects the marginal effect of a minimum wage increase, but also the aggregate effect relative to the situation without a minimum wage. In panel (a) of Figure 2, the aggregate effect is given by $\Delta L_j(\underline{w}, \underline{\varphi}_j) = L_j(\underline{w}_j) - L_j(\underline{w}_j^u)$. In panel (b) of Figure 2, we plot $\Delta L_j(\underline{w}, \underline{\varphi}_j)$ against $\underline{\varphi_j}$, which directly maps into the average regional productivity given the Pareto-shaped firm productivity distribution. We consider a continuum of regions with heterogeneous productivity, but only one universal national minimum wage $\underline{w}$, which resembles the empirical setting in many countries. We can distinguish between three types of regions. The minimum wage has no effect in regions where even the least productive firm is unconstrained ($\underline{\varphi_j} \geq \varphi_j^{'''}$). In the least productive regions, there are negative aggregate employment effects driven by demand-constrained firms ($\underline{\varphi_j} < \underline{\varphi_j^{'}}$). In between, there are positive employment effects driven by supply-constrained firms (and some demand-constrained firms) that peak at the regional productivity level $\underline{\varphi_j^{''}}$. Hence, the regional employment effect of a national minimum wage

[13] Abstracting from unemployment benefits and defining the hiring probability as the probability with which a worker obtains a full-time job, there is an isomorphic interpretation as unemployment.

[14] See Online Supplement for details. For a discussion of the effect of the minimum wage on firm profits and revenues, we refer to Appendix A.4.

is hump-shaped in regional productivity. This is a novel theoretical prediction which we take to the data using a transparent reduced-form methodology before we return to the model to establish the spatial general equilibrium.

3.2 Reduced-form evidence

To empirically evaluate the central prediction that the regional employment effect of the German national minimum wage is hump-shaped in regional productivity, we require estimates of the minimum wage effect by spatial units that are sufficiently small to exhibit sizable variation in average productivity. The empirical challenge in establishing the regional minimum wage effect is that the counterfactual outcome in the absence of the minimum wage is unlikely to be independent of the regional productivity level $\underline{\varphi}_j$. Consider the following data generating process (DGP):

$$\ln L_{j,t} = \left[\overline{f} + f(\underline{\varphi}_j)\right] I(t \geq \mathcal{J}) + \boldsymbol{a_j} + t\boldsymbol{b_j} + \epsilon_{j,t}, \tag{9}$$

where $\mathcal{J} = 2015$ is the year of the minimum wage introduction, $L_{j,t}$ is employment in area j in year t, $\boldsymbol{a_j}$ is a $1 \times J$ vector of regional fixed effects and $\boldsymbol{b_j}$ is a vector of parameters that moderate regional-specific time trends of the same dimension. $\boldsymbol{a_j}$ is likely positively correlated with employment since more productive regions attract more workers. Conditional on $\boldsymbol{a_j}$, $\boldsymbol{b_j}$ can be positively or negatively correlated with employment depending on whether the economy experiences spatial convergence or divergence. $\epsilon_{j,t}$ is a random error term. Unless we hold $\boldsymbol{a_j}$ and $t\boldsymbol{b_j}$ constant, we will fail to recover the correct conditional expectation $\mathbb{E}[\ln L_{j,t}|\varphi_j], t \geq \mathcal{J}] - \mathbb{E}[\ln L_{j,t}|\varphi_j], t < \mathcal{J}]$. To address this concern, we take differences in Eq. (9) over n periods:

$$\ln L_{j,t} - \ln L_{j,t-n} = \Delta \ln L_j = \overline{f} + f(\underline{\varphi}_j) + (\Delta t)\boldsymbol{b_j} + \Delta\epsilon_{j,t}, \tag{10}$$

where $t \geq \mathcal{J}$ and $t - n < \mathcal{J}$ and Δ is the difference operator. This removes unobserved heterogeneity in levels ($\boldsymbol{a_j}$), but unobserved heterogeneity in trends ($\boldsymbol{b_j}$) remains a threat to identification. In programme evaluations of minimum wages, it is conventional to address this concern by controlling for trends observed before the minimum wage introduction under the assumption that these can be extrapolated to the post-policy period (Ahlfeldt et al., 2018; Monras, 2019; Dustmann et al., 2022). We follow this convention by assuming that the pre-policy trend can be described by $(\Delta t)\boldsymbol{b_j} \equiv \ln L_{j,t-n} - \ln L_{j,t-m} + \Delta\tilde{\epsilon}_j$, where $m < n$, $t - n \equiv n - m$, and $\Delta\tilde{\epsilon}_j$ is another white-noise error term. We can then subtract the pre-policy change in the outcome, $\ln L_{j,t-n} - \ln L_{j,t-m}$, from Eq. (10) to obtain

$$[\ln L_{j,t} - \ln L_{j,t-n}] - [\ln L_{j,t-n} - \ln L_{j,t-m}] = \Delta^2 \ln L_j = \overline{f} + f(\underline{\varphi}_j) + \Delta^2\epsilon_{i,t}, \tag{11}$$

where $\Delta^2\epsilon_{j,t} = \Delta\epsilon_{j,t} - \Delta\tilde{\epsilon}_{j,t}$. Guided by the theoretical predictions summarized in Figure 2, we define the relative (up to the constant $\overline{f}$) before-after minimum wage effect as a polynomial spline function

$$f(\underline{\varphi}_j) = \mathbb{E}\left[\Delta^2 \ln L_j | w_j^{\text{mean}}, (w^{\text{mean}_j} \leq \alpha_0)\right] - \mathbb{E}\left[\Delta^2 \ln L_j | w_j^{\text{mean}}, (w^{\text{mean}_j} > \alpha_0)\right]$$
$$= \mathbb{1}\left(w_j^{\text{mean}} \leq \alpha_0\right) \times \left[\sum_{g=1}^{2} \alpha_g \left(w_j^{\text{mean}} - \alpha_0\right)^g\right], \tag{12}$$

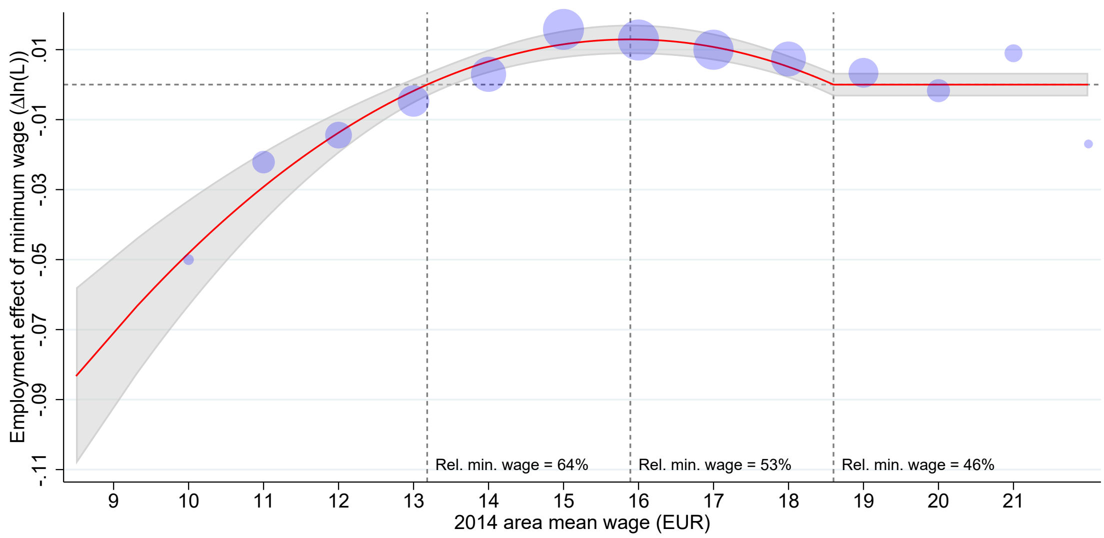
with the theory-consistent parameter restrictions $\{\alpha_0 > \frac{\alpha_1}{2\alpha_2}, \alpha_1 < 0, \alpha_2 < 0\}$. Since higher fundamental productivity maps to higher wages in our model, we use the 2014 mean wage w_j^{mean} as a proxy for regional productivity. Notice that the interpretation of $f(\underline{\varphi}_j)$ is akin to the treatment effect in an intensive-margin difference-in-difference setting in which regions populated solely by unconstrained firms form a control group to establish a counterfactual.

Substituting in Eq. (12), we are ready to estimate Eq. (11) for given years $\{t, t-n, t-m\}$. To obtain parameter estimates for $\{\alpha_0, \alpha_1, \alpha_2\}$, we nest an OLS estimation of $\{\alpha_1, \alpha_2\}$ in a grid search over a parameter space $\alpha_0 \in [\underline{\alpha}_o, \bar{\alpha}_o]$ and pick the parameter combination that minimizes the sum of squared residuals. From the identified parameters $\{\alpha_0, \alpha_1, \alpha_2\}$, there is a one-to-one mapping to regional mean wage levels that correspond to regional productivity levels $\{\underline{\varphi}_j^{'}, \underline{\varphi}_j^{''}, \underline{\varphi}_j^{'''}\}$ in Figure 2 (see Online Supplement for details). Note that consistent with the partial-equilibrium nature of the analysis, Eq. (12) lends a difference-in-difference interpretation to the predicted employment effect $\hat{f}(\underline{\varphi}_j)$ as regions dominated by unconstrained firms $(\underline{\varphi}_j \geq \underline{\varphi}_j^{'''}\})$ serve as the counterfactual.

The assumption that counterfactual area-specific trends extend from the $[t-m, t-n]$ to the $[t-n, t]$ period is more plausible over shorter study periods. Hence, we set $\{t = 2016, m = 4, n = 2\}$ in Figure 3, which restricts the comparison to two years before and after the minimum wage introduction. The results with one- or three-year windows are very similar. A dynamic difference-in-differences estimation using our estimate of $f(\underline{\varphi}_j)$ as a treatment measure further substantiates that we successfully addressed pre-trends. Conventional difference-in-difference estimates by groups of regions experiencing different minimum wage bites also substantiate the results discussed below (see Online Supplement).

Consistent with theory, we find an employment effect that is hump-shaped in the 2014 mean wage. Since the upward-sloping labour supply is key to generating the hump shape in our theory, Figure 3 provides indirect evidence of monopsonistic labour markets. The greatest positive employment effect is predicted for an area with a 2014 mean wage of about €16, which corresponds to $\underline{\varphi}_j^{''}$. The implication is that the regional employment effect is maximized for the area where the relative minimum wage amounts to €8.5/€16=53% of the mean wage. Municipalities with a lower mean wage, where the relative minimum wage is higher, have smaller predicted employment effects. At a relative minimum wage of 64% the predicted employment effect turns negative, a point that corresponds to $\underline{\varphi}_j^{'}$ in Figure 2. The empirical correspondent to productivity level $\underline{\varphi}_j^{'''}$—beyond which the minimum wage has no bite— is a regional mean wage of €18.6, which corresponds to a relative minimum wage of 46%.

Figure 3: Regional minimum wage effects: Reduced-form evidence



Note: Dependent variable is the second difference in log employment over the 2012-14 and 2014-16 periods. Markers give averages within one-euro bins, with the marker size representing the number of municipalities within a bin. The last bin (22) includes all municipalities with higher wages because observations are sparse. The red solid line illustrates the quadratic fit, weighted by bin size. Two outlier bins are excluded to improve readability, but they are included in the estimation of the quadratic fit. Confidence bands (gray-shaded area) are at the 95% level. The relative minimum wage is the ratio of the 2015 minimum wage level $\underline{w} = 8.50$ over the 2014 mean wage (when there was no minimum wage).

The 46-64% range for the relative minimum wage derived in this section represents a first point of reference for those wishing to ground the minimum-wage setting in transparent reduced-form evidence of employment effects. Yet, the reduced-form approach constrains us to identifying relative employment effects. By assumption, we do not capture any general equilibrium effects that affect the control group (unconstrained regions). Moreover, the reduced-form approach naturally does not allow us to derive the welfare effect, which not only depends on the effects on wages and employment probabilities, but also on changes in commuting costs, tradable goods and housing prices. We, therefore, take the analysis to the spatial general equilibrium in the next section, in which we will also evaluate a broader set of model predictions against observed changes in data.

4 General equilibrium analysis

We develop the model in Section 4.1 and discuss the quantification in Section 4.2 before laying out how to use the model for quantitative counterfactual analyses in Section 4.3. Then, we proceed to a three-step application. First, we use the model to quantitatively evaluate the general equilibrium effects of the German minimum wage introduced in 2015 in a counterfactual analysis in Section 4.4.1. Second, we treat the model's predictions of changes in endogenous outcomes as forecasts that we subject to over-identification tests by comparing them to observed before-after changes in the data in Section 4.4.2. Third, we find the optimal minimum wage in a series of counterfactuals in which we consider a range of national and regional minimum wages in Section 4.5.

4.1 Model II

Building on the partial equilibrium framework introduced in Section 3.1, we now expand the model to account for the interaction of goods and factor markets, free entry of firms and an endogenous choice of workers to enter the labour market. We refer to $\bar{N}$ as the working-age population and denote the labour force measured at the place of residence i by N_i and the labour force measured at the workplace j by H_j. L_j represents employment (at the workplace) and can generally be smaller than the labour force when minimum wages are binding. It is measured in full-time equivalent terms, so underemployment can arise at the intensive (working hours) or extensive (unemployment) margin.

4.1.1 Preferences and endowments

Workers are geographically mobile and have heterogeneous preferences to work for firms in different locations. Given the choices of other firms and workers, each worker maximizes utility by choosing a residence location i and a (potential) employer φ_j – thereby pinning down the (potential) workplace location j. The preferences of a worker ν who lives and consumes in location i and works at firm φ_j in location j are defined over final goods consumption $Q_{i\nu}$, housing $T_{i\nu}$, an idiosyncratic amenity shock $\exp[b_{ij\nu}(\varphi_j)]$, and commuting costs $\kappa_{ij} > 1$, according to the Cobb-Douglas form

$$U_{ij\nu}(\varphi_j) = \frac{\exp[b_{ij\nu}(\varphi_j)]}{\kappa_{ij}} \left(\frac{Q_{i\nu}}{\alpha}\right)^{\alpha} \left(\frac{T_{i\nu}}{1-\alpha}\right)^{1-\alpha}. \tag{13}$$

The household budget that can be spent on consumption goods and housing consists of expected income $\psi_j(\phi_j)w_j(\varphi_j)$. The amenity shock captures the idea that workers can have idiosyncratic reasons for living in different locations and working in different firms (Egger et al., 2022). We assume that $b_{ij\nu}(\varphi_j)$ is drawn from an independent Type I extreme value (Gumbel) distribution

$$F_{ij}(b) = \exp(-B_{ij}\exp\{-[\varepsilon b + \Gamma'(1)]\}), \quad \text{with} \quad B_{ij} > 0 \quad \text{and} \quad \varepsilon > 0, \tag{14}$$

in which B_{ij} is the scale parameter determining the average amenities from living in location i and working in location j, ε is the shape parameter controlling the dispersion of amenities, and $\Gamma'(1)$ is the Euler-Mascheroni constant (Jha and Rodriguez-Lopez, 2021).

The goods consumption index Q_i in location i is a constant elasticity of substitution (CES) function of a continuum of tradable varieties

$$Q_i = \left[\sum_j \int_{\varphi_j} q_{ij}(\varphi_j)^{\frac{\sigma-1}{\sigma}} d\varphi_j\right]^{\frac{\sigma}{\sigma-1}} \tag{15}$$

with $q_{ij}(\varphi_j) > 0$ denoting the quantity of variety φ_j sourced from location j and $\sigma > 1$ as the constant elasticity of substitution. Utility maximization yields $q_{ij}(\varphi_j) = S_i^q p_{ij}(\varphi_j)^{-\sigma}$ with $S_i^q \equiv E_i^Q \left(P_i^Q\right)^{\sigma-1}$ as defined in Eq. (2), in which E_i^Q is aggregate expenditure in location i for tradables, P_i^Q is the price index dual to Q_i in Eq. (15), and $p_{ij}(\varphi_j)$ is the consumer price of variety φ_j in location i.

The economy is further endowed with a fixed housing stock $\bar{T}_i$. Denoting by E_i^T total expenditure for housing in location i, we can equate supply with demand, $T_i^D = E_i^T/P_i^T$, to derive the market-clearing price for housing:

$$P_i^T = \frac{E_i^T}{\bar{T}_i}. \tag{16}$$

4.1.2 Free entry and goods trade

Firms learn their productivity φ_j only after paying market entry costs, $f_j^e P_j^T$, which consist of some start-up space f_j^e acquired at housing rent P_j^T. The investment is profitable whenever expected profits exceed these costs and we refer to this relation as the free-entry condition given by

$$\tilde{\pi}_j = \frac{\Pi_j}{M_j} = f_j^e P_j^T. \tag{17}$$

Using the facts that $\Pi_j = (1-\eta)[\Phi_j^\Pi(\underline{w})/\Phi_j^R(\underline{w})]R_j$ and that also the aggregate wage bill is proportional to revenues, $\tilde{w}_j L_j = [1-(1-\eta)\Phi_j^\Pi(\underline{w})/\Phi_j^R(\underline{w})]R_j$, we can reformulate Eq. (17) to get

$$M_j = \frac{\Phi_j^\Pi(\underline{w})(1-\eta)}{\Phi_j^R(\underline{w}) - \Phi_j^\Pi(\underline{w})(1-\eta)} \frac{\tilde{w}_j L_j}{P_j^T f_j^e}, \tag{18}$$

where

$$\tilde{w}_j = \frac{R_j - \Pi_j}{L_j} = \frac{1-(1-\eta)\Phi_j^\Pi(\underline{w})/\Phi_j^R(\underline{w})}{\eta} \frac{\chi_R \Phi_j^R(\underline{w})}{\chi_L \Phi_j^L(\underline{w})} w_j^u(\underline{\varphi}_j) \tag{19}$$

denotes the average wage rate in location j which is proportional to the cut-off wage $w_j^u(\underline{\varphi}_j)$ of an unconstrained firm with productivity $\underline{\varphi}_j$ given that $w_j^u(\varphi_j) l_j^u(\varphi_j)/\eta = r_j^u(\varphi_j) = \pi_j^u(\varphi_j)/(1-\eta)$.

With firm entry costs being paid in terms of housing and assuming that land owners spend their entire income on the tradable good, we can state that total housing expenditure in location i is given by $E_i^T = (1-\alpha)\tilde{v}_i N_i + \Pi_i$ and aggregate expenditure on tradable goods results as

$$E_i^Q = \alpha \tilde{v}_i N_i + E_i^T = \tilde{v}_i N_i + \Pi_i, \tag{20}$$

where $\tilde{v}_i$ is the average labour income of the residential labour force N_i across employment locations.

Building on optimal firm behaviour derived in Section 3.1, our model implies a gravity equation for bilateral trade between locations. Using the CES expenditure function and the measure of firms M_j, the share of location i's expenditure on goods produced in location

j is given by

$$\theta_{ij} = \frac{M_j \int_{\varphi_j} p_{ij}(\varphi_j)^{1-\sigma} dG(\varphi_j)}{\sum_{k\in J} M_k \int_{\varphi_k} p_{ik}(\varphi_k)^{1-\sigma} dG(\varphi_k)}, \\ = \frac{M_j \Phi_j^P(\underline{w}) \left(\left\{ \Phi_j^L(\underline{w}) / [\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^\Pi(\underline{w})] \right\} \tau_{ij} \tilde{w}_j / \underline{\varphi}_j \right)^{1-\sigma}}{\sum_{k\in J} M_k \Phi_k^P(\underline{w}) \left(\left\{ \Phi_k^L(\underline{w}) / [\Phi_k^R(\underline{w}) - (1-\eta)\Phi_k^\Pi(\underline{w})] \right\} \tau_{ik} \tilde{w}_k / \underline{\varphi}_k \right)^{1-\sigma}}. \tag{21}$$

To derive Eq. (21) we take advantage of the ideal price index $P_{ij} \equiv [\int_{\varphi_j} p_{ij}(\varphi_j)^{1-\sigma} d\varphi_j]^{1/(1-\sigma)}$ for the subset of commodities that are consumed in location i and produced in location j. As formally shown in Appendix A, it can be computed as

$$P_{ij} = \chi_P^{\frac{1}{1-\sigma}} \Phi_j^P(\underline{w})^{\frac{1}{1-\sigma}} M_j^{\frac{1}{1-\sigma}} p_{ij}^u(\underline{\varphi}_j), \tag{22}$$

with $\chi_P > 1$ as a constant and $\Phi_j^P(\underline{w}) > 0$ as a term that captures the aggregate effect of the minimum wage $\underline{w}$ on the price index P_{ij}. Notice that $\Phi_j^P(\underline{w}) = 1$ if the minimum wage $\underline{w}$ is not binding in location j. If the minimum wage is binding, $\Phi_j^p(\underline{w})$ can be larger or smaller than one, reflecting two opposing forces: Supply-constrained firms and highly productive demand-constrained firms lose their monopsony power and therefore set lower prices, which reduces the average price of firms from location j. At the same time, a binding minimum wage raises the costs – in particular for unproductive demand-constrained firms, which pass through this increase to their consumers in form of higher prices. The expenditure share θ_{ij} declines in bilateral trade costs τ_{ij} in the numerator ("bilateral resistance") relative to the trade costs to all possible sources of supply in the denominator ("multilateral resistance").

Using optimal prices together with Eqs. (19) and (22) to substitute for P_{ij}, $p_{ij}^u(\underline{\varphi}_j)$, and $w_j^u(\underline{\varphi}_j)$, into the price index $(P_i^Q)^{1-\sigma} \equiv \sum_j P_{ij}^{1-\sigma}$ dual to the consumption index in Eq. (15) we obtain

$$P_i^Q = \frac{\chi_L}{\chi_R} \chi_P^{\frac{1}{1-\sigma}} \left\{ \sum_j M_j \Phi_j^P(\underline{w}) \left[\frac{\Phi_j^L(\underline{w})}{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^\Pi(\underline{w})} \frac{\tau_{ij}\tilde{w}_j}{\underline{\varphi}_j} \right]^{1-\sigma} \right\}^{\frac{1}{1-\sigma}}, \\ = \frac{\chi_L}{\chi_R} \chi_P^{\frac{1}{1-\sigma}} \left[\frac{M_i \Phi_i^P(\underline{w})}{\theta_{ii}} \right]^{\frac{1}{1-\sigma}} \frac{\Phi_i^L(\underline{w})}{\Phi_i^R(\underline{w}) - (1-\eta)\Phi_i^\Pi(\underline{w})} \frac{\tau_{ii}\tilde{w}_i}{\underline{\varphi}_i}, \tag{23}$$

which we can rewrite in terms of location i's own expenditure share θ_{ii}.

Location j's aggregate labour income $\tilde{w}_j L_j$ is proportional to aggregate revenue R_j in location j, which equals total expenditure on goods produced in this location:

$$\tilde{w}_j L_j = \frac{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^\Pi(\underline{w})}{\Phi_j^R(\underline{w})} \sum_i \theta_{ij} \left(\tilde{v}_i N_i + \Pi_i \right). \tag{24}$$

4.1.3 Labour mobility, commuting, and labour supply

A worker's decision where to live, whether to enter the labour market and where to work depends on the indirect utility function $V_{ij\nu}(\varphi_j)$ dual to $U_{ij\nu}(\varphi_j)$ in Eq. (13) given by

$$V_{ij\nu}(\varphi_j) = \frac{\exp[b_{ij\nu}(\varphi_j)]}{\kappa_{ij}} \frac{\psi_j(\varphi_j)w_j(\varphi_j)}{\left(P_i^Q\right)^\alpha \left(P_i^T\right)^{1-\alpha}}, \tag{25}$$

in which the expected income of those seeking employment at firm φ_j in location j is the firm's wage rate $w_j(\varphi_j)$ evaluated at the hiring probability $\psi_j(\varphi_j)$. The probability that a worker chooses to live in location i and work in firm φ_j in location j then can be derived as

$$\lambda_{ij}(\varphi_j) = \frac{B_{ij}\left[\kappa_{ij}\left(P_i^Q\right)^\alpha \left(P_i^T\right)^{1-\alpha}\right]^{-\varepsilon}\left[\psi_j(\varphi_j)w_j(\varphi_j)\right]^\varepsilon}{\sum_r\sum_s B_{rs}\left[\kappa_{rs}\left(P_r^Q\right)^\alpha \left(P_r^T\right)^{1-\alpha}\right]^{-\varepsilon}\int_{\varphi_s}\left[\psi_s(\varphi_s)w_s(\varphi_s)\right]^\varepsilon d\varphi_s}. \tag{26}$$

The idiosyncratic shock to preferences $\exp[b_{ij\nu}(\varphi_j)]$ implies that individual workers choose different bilateral commutes and different employers when faced with the same prices and location characteristics. Other things equal, workers are more likely to live in location i and work for firm φ_j in location j, the lower the prices for consumption and housing P_i^Q and P_i^T in i; the higher the expected income $\psi_j(\varphi_j)w_j(\varphi_j)$ from working for firm φ_j in j; the more attractive average amenities B_{ij}; and the lower the commuting costs κ_{ij}. Summing across all residential locations i yields the probability that a worker is seeking employment at firm φ_j, $\lambda_j(\varphi_j) = \sum_i \lambda_{ij}(\varphi_j) = h_j(\varphi_j)/N$ with $N = \sum_i N_i$. The labour supply $h_j(\varphi_j)$ to firm φ_j therefore is given by Eq. (1) with

$$S_j^h \equiv \frac{\sum_i B_{ij}\left[\kappa_{ij}\left(P_i^Q\right)^\alpha \left(P_i^T\right)^{1-\alpha}\right]^{-\varepsilon}}{\sum_r\sum_s B_{rs}\left[\kappa_{rs}\left(P_r^Q\right)^\alpha \left(P_r^T\right)^{1-\alpha}\right]^{-\varepsilon} W_s^\varepsilon} N, \tag{27}$$

in which $W_j \equiv \{\int_{\varphi_j}\left[\psi_j(\varphi_j)w_j(\varphi_s)\right]^\varepsilon d\varphi_j\}^{\frac{1}{\varepsilon}}$ denotes an index of (expected) wages. In Appendix A.3, we demonstrate that W_j can be rewritten as a function of location j's cut-off wage $w_j^u(\underline{\varphi}_j)$, which according to Eq. (19) is proportional to the average wage $\tilde{w}_j$ in location j

$$\begin{aligned} W_j &= \chi_W^{\frac{1}{\varepsilon}} \Phi_j^W(\underline{w})^{\frac{1}{\varepsilon}} M_j^{\frac{1}{\varepsilon}} w_j^u(\underline{\varphi}_j) \\ &= \Omega_j(\underline{w})\tilde{w}_j \frac{\chi_L}{\chi_R}\chi_W^{\frac{1}{\varepsilon}} M_j^{\frac{1}{\varepsilon}}, \text{where} \\ \Omega_j(\underline{w}) &\equiv \frac{\eta\Phi_j^W(\underline{w})^{\frac{1}{\varepsilon}}\Phi_j^L(\underline{w})}{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^\Pi(\underline{w})} \end{aligned} \tag{28}$$

is a composite adjustment factor that captures various channels through which the minimum wage affects the wage index. Henceforth, we refer to $\Omega_j(\underline{w})\tilde{w}_j$ as *expected wage* for convenience. If the minimum wage $\underline{w}$ is not binding in location j, we have $\Phi_j^{X\in\{W,L,R,\Pi\}}(\underline{w}) = 1$ and, hence, $\Omega_j(\overline{w}) = 1$. If the minimum wage is binding, $\Omega_j(\underline{w})$ can be larger or smaller than one, reflecting two opposing forces: On the one hand, there is a direct effect captured by $\Phi_j^W(\underline{w})$. Because a binding minimum wage $\underline{w}$ exceeds the wages that supply- and demand-constrained firms would pay otherwise, the wage index increases. On the other hand, a binding minimum wage $\underline{w}$ causes demand-constrained firms to practice job rationing, such that the employment probability at these firms $\psi_j^d(\varphi_j)$ falls below one. If there are enough low-productivity demand-constrained firms, the employment response captured by $\Phi_j^L(\underline{w})$ will be negative (dominating the positive response by supply-constrained firms). It is possible that a lower hiring rate more than compensates for rising wages so that the minimum wage causes the expected wage index to fall.

Aggregating $\lambda_{ij}(\varphi_j)$ across all firms φ_j in workplace j, we obtain the overall probability that a worker living in i applies to a firm in j, to which we refer as unconditional commuting probability:

$$\lambda_{ij} = \int_{\varphi_j} \lambda_{ij}(\varphi_j) d\varphi_j = \frac{B_{ij} M_j \left[\frac{\Omega_j(\underline{w}) \tilde{w}_j}{\kappa_{ij} \left(P_i^Q\right)^{\alpha} \left(P_i^T\right)^{1-\alpha}} \right]^{\varepsilon}}{\sum_r \sum_s B_{rs} M_s \left[\frac{\Omega_s(\underline{w}) \tilde{w}_s}{\kappa_{rs} \left(P_r^Q\right)^{\alpha} \left(P_r^T\right)^{1-\alpha}} \right]^{\varepsilon}}. \tag{29}$$

From Eq. (29), we obtain the residential choice probability λ_i^N and the workplace choice probability λ_i^H as $\lambda_i^N = \frac{N_i}{N} = \sum_j \lambda_{ij}$ and $\lambda_j^H = \frac{H_j}{N} = \sum_i \lambda_{ij}$, with $\sum_i \lambda_i^N = \sum_j \lambda_j^H = 1$. In order to solve for location j's aggregate employment L_j, we have to account for the fact that not all workers H_j, who are willing to work in j, will necessarily find a job. This is a novel feature in the context of quantitative spatial models and results in a labour-market clearing condition that equates the full-time equivalent employment at j, L_j to the absolute number of workers *working or searching* in j, $\lambda_j^H N$, discounted by the employment probability Φ_j^L/Φ_j^H (which is equal to one in the absence of the minimum wage):

$$L_j = \frac{L_j}{H_j} \lambda_j^H N = \frac{\Phi_j^L(\underline{w})}{\Phi_j^H(\underline{w})} \lambda_j^H N, \tag{30}$$

with the second equality following from results derived in Appendix A.3 and $h_j^u(\underline{\varphi}_j) = l_j^u(\underline{\varphi}_j)$.

The average income of a worker living in location i depends on the expected wages in all employment locations. To construct this average income of residents, note first that the probability that a worker commutes to location j conditional on living in location i is given by:

$$\lambda_{ij|i}^N \equiv \frac{\int_{\varphi_j} \lambda_{ij}(\varphi_j) d\varphi_j}{\lambda_i^N} = \frac{B_{ij} M_j \left[\Omega_j(\underline{w}) \frac{\tilde{w}_j}{\kappa_{ij}} \right]^{\varepsilon}}{\sum_s B_{is} M_s \left[\Omega_s(\underline{w}) \frac{\tilde{w}_s}{\kappa_{is}} \right]^{\varepsilon}}, \tag{31}$$

in which ε can be interpreted as the elasticity of commuting flows with respect to commuting costs. Using these conditional commuting probabilities, we obtain the following condition that equates full-time equivalent employment in j, L_j,with the employment-probability-adjusted sum of all workers commuting from i to j, namely,

$$L_j = \frac{L_j}{H_j} \sum_i \lambda^N_{ij|i} N_i = \frac{\Phi^L_j(\underline{w})}{\Phi^H_j(\underline{w})} \sum_i \lambda^N_{ij|i} N_i. \tag{32}$$

Expected worker income conditional on living in location i is then equal to the expected income in all workplaces weighted by the employment probabilities in those locations conditional on living in i:

$$\tilde{v}_i = \sum_j \lambda^N_{ij|i} \frac{L_j}{H_j} \tilde{w}_j = \sum_j \lambda^N_{ij|i} \frac{\Phi^L_j(\underline{w})}{\Phi^H_j(\underline{w})} \tilde{w}_j. \tag{33}$$

The expected utility, conditional on being active on the labour market, is

$$\overline{V} = \left\{ \sum_i \sum_j B_{ij} M_j \left[\frac{\Omega_j(\underline{\omega}) \tilde{w}_j}{\kappa_{ij} \left(P^Q_i\right)^{\alpha} \left(P^T_i\right)^{1-\alpha}} \right]^{\varepsilon} \right\}^{\frac{1}{\varepsilon}}. \tag{34}$$

4.1.4 Labour market entry

Workers have the discrete choice between entering the labour market and abstaining. Since workers do not observe the idiosyncratic residence-workplace-employer shock $b_{ijv}(\varphi_i)$ when deciding on entering the labour market, they compare the correctly anticipated expected utility from working in Eq. (34) to the expected leisure utility. Following the conventions in the discrete choice literature (McFadden, 1974), we assume that individuals have Gumbel-distributed idiosyncratic preferences for the two alternatives. As we formally derive in Appendix B.2, we can express the labour force participation rate as

$$\mu = \frac{\overline{V}^{\zeta}}{\overline{V}^{\zeta} + A}, \tag{35}$$

where ζ is the Gumbel shape parameter that is a transformation of the Hicksian extensive-margin labour supply elasticity, and A is the shift parameter that captures the leisure amenity. Intuitively, workers are more likely to abstain from the labour market if there are greater leisure amenities and if the utility from entering the labour market is lower. Naturally, the labour force participation rate plays a key role in the aggregate labour market clearing condition

$$\sum_j H_j = \mu \overline{N}, \tag{36}$$

where the left-hand side represents the national labour force and $\overline{N}$ is the working-age population. Finally, the Gumbel distribution of idiosyncratic taste shocks implies that expected welfare across all workers (working, searching, and abstaining) takes the following form:

$$\overline{\mathcal{V}} = \left(A + \overline{V}^{\zeta}\right)^{\frac{1}{\zeta}} \tag{37}$$

4.1.5 General equilibrium

The general equilibrium of the model can be referenced by the following vector of seven variables $\{\tilde{w}_i, \tilde{v}_i, M_j, P_i^T, L_i, N_i, P_i^Q\}_{i=1}^J$ and the scalars $\{\mu, \overline{V}\}$. Given the equilibrium values of these variables and scalars, all other endogenous objects can be determined conditional on the model's primitives. This equilibrium vector solves the following seven sets of equations: income equals expenditure from Eq. (24); average residential income from Eq. (33); firm entry from Eq. (18); housing market clearing from Eq. (16); aggregate local employment from Eq. (32); $N_i = \lambda_i^N N$ based on Eq. (29) and the price index from Eq. (23). The conditions needed to determine the scalars $\{\mu, \overline{V}\}$ are labour force participation from Eq. (35) and the labour market clearing condition from Eq. (36). Notice that in the equilibrium without a binding minimum wage there is an endogenous number of workers $(1-\mu)\overline{N}$ who voluntarily abstain from the labour market. With a binding minimum wage, there is, in addition, also underemployment that is involuntary in the sense that workers would strictly prefer to work full-time. The full-time equivalent of this underemployment amounts to $\sum_j \frac{\Phi_j^L(\underline{w})}{\Phi_j^H(\underline{w})} H_j$.

4.2 Quantification

The primitives of the model consist of the structural parameters $\{\underline{w}, k, \alpha, \sigma, \epsilon, \zeta, \mu\}$ and the structural fundamentals $\{\tau_{ij}, \kappa_{ij}, B_{ij}, \underline{\varphi}_j, \bar{T}_i, f_j^e, A\}$. If these primitives are given alongside the endowment $\{\bar{N}\}$, we can solve for the variables $\{\tilde{w}_i, \tilde{v}_i, P_i^T, L_i, N_i, P_i^Q, M_i\}_{i=1}^J$ and the scalars $\{\mu, \overline{V}\}$ that reference the general equilibrium. We quantify the model using data from Germany in 2014, the year before the minimum wage introduction. Therefore, we can treat all firms as unconstrained and set $\underline{w} = 0$ in the quantification, which implies that $\Phi_j^{X \in \{L,H,R,P,W,\Pi\}} = 1$. We borrow $\{\alpha, \zeta\}$ from the literature and set σ such that all parameter restrictions of the model are satisfied. We infer all other primitives from the data using observed values of $\{P^T, \lambda_{ij}, N_i, M_j, w_j(\omega), \tilde{w}_j, (p_{ij}q_{ij}), \mu\}$. We provide a brief discussion below and refer to Appendix B.3 for details.

Expenditure share on housing $(1-\alpha)$. We set the housing expenditure share to $1-\alpha = 0.33$, which is in line with a literature summarized in Ahlfeldt and Pietrostefani (2019) and official data from Germany (Statistisches Bundesamt, 2020).

Labour force participation rate (μ). We use the 2014 employment rate of $\mu = 73.6\%$ reported by the German Federal Statistical Office.

Working-age population $(\bar{N})$. Based on the labour force participation rate and total employment in 2014, we get $\bar{N} = N/\mu$.

Leisure utility heterogeneity (ζ). As we show in Appendix B.2, we can express the heterogeneity of idiosyncratic shocks to the utility from non-employment ζ as a function of the Hicksian extensive-margin labour supply elasticity $\tilde{\zeta}$ and the labour force participation rate μ. Setting the former to the canonical value of $\tilde{\zeta} = 0.2$ in the literature (Chetty et al., 2011) and the latter to the value observed in German data, we obtain $\zeta = 0.8$.

Preference heterogeneity (ε). We exploit that the firm-level wage and firm size scale in firm productivity at elasticities that differ by multiplicative factor ε (see Table A1). This allows us to obtain a theory-consistent estimate of ε from an establishment-level regression of the log of wage against the log of employment, controlling for all supply shifters emphasized by the model via municipality-year fixed effects. To address establishment-level supply shocks that could confound our estimates, we restrict the identifying variation to within municipality-year demand shocks using a shift-share instrument that lets establishment employment grow at the national rate of the respective industry sector. Our estimate of $\varepsilon = 5.5$ is in between Monte et al. (2018), who use larger spatial units, and Ahlfeldt et al. (2015), who use smaller spatial units. It implies that workers earn $\varepsilon/(\varepsilon+1) = 85\%$ of their MRPL, which is in the middle of the range of extant estimates (Sokolova and Sorensen, 2020; Yeh et al., 2022). We refer to Appendix B.3.1 for details.

Productivity heterogeneity and elasticity of substitution (k, σ). Intuitively, we identify k by fitting a Pareto cumulative distribution function (CDF) of wages as is conventional in the trade literature (Arkolakis, 2010; Egger et al., 2013). We take a structural approach to the estimation of k because $\{k, \sigma, \epsilon\}$ jointly determine the dispersion of wages and the regional lower-bound wage, conditional on observed values of $\tilde{w}_j$. Taking our estimate of $\varepsilon = 5.5$ as given, we nest the estimation of k using a GMM estimator into a grid search over σ values. We choose $\sigma = 1.9$ as the value that is closest to the conventions in the literature and still satisfies all parameter restrictions of the model. Conditional on these values for $\{\epsilon, \sigma\}$, we obtain an estimate for k of 0.9. These values are smaller than the typical values found in the trade literature (Egger et al., 2013; Simonovska and Waugh, 2014), but they ensure that we obtain a decent fit of the wage distribution in the left tail. We refer to Appendix B.3.2 for details.

Minimum wage ($\underline{\mathrm{w}}$). Since we use the worker-weighted mean wage as the numeraire in our model, it is straightforward to define the minimum wage in relative terms as $\underline{w} = 0.48$, which is the share of the minimum wage at the national mean wage observed in the data (across full time and part-time workers). Notice that this share remains remarkably constant over time, suggesting that the adjustments to the absolute minimum wage level made in 2017, 2019, and 2020 aimed at keeping the relative level constant.

Trade cost (τ_{ij}). We estimate a gravity equation of bilateral trade volumes ($p_{lk}q_{lk}$) between county pairs lk within Germany allowing for a direction-specific inner-German border effect and origin-specific distance effects. Using the estimated reduced-form parameters and our set value of σ we predict τ_{ij} for pairs of municipalities in a theory-consistent way. We refer to Appendix B.3.3 for details. With this approach, we account for the legacy of German cold war history and the centrality bias in inter-city trade (Mori and Wrona, 2021).

Fundamental productivity ($\underline{\varphi_j}$). Given observed values of $\{L_j, N_i, \lambda^N_{ij|i}, \tilde{w}_j, M_j\}$, the set or estimated values of $\{\varepsilon, \sigma\}$, the predicted values of τ_{ij}, and exploiting that $\tilde{v}_j = \sum_j^J \lambda^N_{ij|i}\tilde{w}_j$, we can invert $\underline{\varphi}_j$ from Eq. (24) (substituting in Eq. (21)) using a conventional fixed-point solver. We refer to Appendix Section B.3.4 for details.

Ease of commuting ($B_{ij}\kappa_{ij}^{-\varepsilon}$). Following Monte et al. (2018), we refer to the composite term $B_{ij}\kappa_{ij}^{-\varepsilon}$ as ease of commuting since, conditional on a given residence i, it captures the attractiveness of commuting to a destination j holding the number of firms M_j and workplace wages $\tilde{w}_j$ constant. Given values of $\{\alpha, \varepsilon, \sigma, k, \tau_{ij}, \underline{\varphi}\}$ and observed values of $\{\lambda_{ij|i}^N, M_j, \tilde{w}_j, P_i^T\}$, we invert $B_{ij}\kappa^{-\epsilon}$ using the unconditional commuting probabilities λ_{ij} using Eq. (29) and a conventional fixed-point solver.

Start-up space (f_j^e). Given values of $\{\varepsilon, \sigma\}$ and observed values of $\{M_j, P_j^T, \tilde{w}_j, L_j\}$ it is straightforward to invert the start-up space firms need to acquire to enter the market, f_j^e, using the firm-entry condition in Eq. (17).

Housing supply ($\overline{T}_i$). For given values of $\{\lambda_{ij|i}^N, \tilde{w}_i, L_i\}$, we can exploit that Π_i scales at known parameters in $w_i L_i$ along with $\tilde{v}_j = \sum_j^J \lambda_{ij|i}^N \tilde{w}_j$ and $E_i^T = (1-\alpha)\tilde{v}_i + \Pi_i$ to infer housing supply $\overline{T}_i$ using the housing market clearing condition in Eq. (16).

Leisure amenity (A). Using observed values of $\{\mu_i, M_j, \tilde{w}_j, P_i^T\}$, inverted values of $\{\underline{\varphi}_j, \tau_{ij}, B_{ij}\kappa_{ij}^{-\varepsilon}\}$ and the estimated set of parameter values for $\{\alpha, \varepsilon, \sigma, \zeta, k\}$, we invert fundamental utility A using Eqs. (23), (34) and (35).

4.3 Quantitative analysis

Given the fully quantified model, the evaluation of the effects of an exogenous change in the minimum wage $\underline{w}$ on the vector of endogenous outcomes that references the general equilibrium $\mathbf{X} = \{\tilde{w}_i, \tilde{v}_i, P_i^T, L_i, N_i, P_i^Q, M_i, \mu, \overline{V}\}$ can be established by solving the model under different values of $\underline{w}$, holding all other primitives constant. We model the solution as a fixed point for which we solve using a conventional numerical procedure that we discuss in Appendix B.4.

We first solve the model for $\underline{w} = 0$ expressing all endogenous goods and factor prices in terms of the worker-weighted mean, which becomes the numeraire. This delivers equilibrium values of the vector of endogenous outcomes which we denote by $\mathbf{X^0}$. Up to a multiplicative constant, these are identical to the observed values in the data, $\mathbf{X^D}$.[15] We then set $\underline{w}$ to the desired value (in units of the numeraire) and solve the model for a vector of counterfactual outcomes, $\mathbf{X^C}$. With this approach, we acknowledge that policy makers set minimum wages that are routinely adjusted to maintain purchasing power. In line with conventional *exact hat algebra* notations (Dekle et al., 2007), we can express the relative change in endogenous outcomes as $\widehat{\mathbf{X}} = \frac{\mathbf{X^C}}{\mathbf{X^0}}$ and the absolute change as $\Delta\mathbf{X} = \widehat{\mathbf{X}} \cdot \mathbf{X^D}$. Unless otherwise indicated, we refer to welfare as the expected utility $\overline{\mathcal{V}}$ from Eq. (37), which takes into account workers inside (working and searching) and outside the labour market.

We follow the canonical approach in the spatial equilibrium literature and pin down

[15] The normalization is required because the lower-bound fundamental productivity $\underline{\phi}_j$ is identified up to a constant. The normalization of nominal wages does not affect the interpretation of real wages, which are relevant for welfare.

residential location choices by assuming perfect mobility, which results in a spatially invariant welfare $\overline{\mathcal{V}}$ (Roback, 1982). However, the assumption that residents are perfectly mobile across residential locations is obviously more plausible in the long run than in the short run. Therefore, we also evaluate a special case that approximates short-run spatial equilibrium adjustments. To this end, we make workers immobile across residences. This restriction is straightforward to implement in our counterfactual by solving the model conditional on holding $\{\bar{N}_i\}$ constant at the values observed in data. For further details on the short-run evaluation, we refer to Appendix Section B.4.2.

4.4 The German minimum wage

We now use the model to quantitatively evaluate the effects of the German minimum wage in the spatial general equilibrium. In Section 4.4.1, we use the procedure outlined in Section 4.3 to predict the effects a federal minimum wage of 48% of the national mean wage (the value we observe in data) has on endogenous model outcomes. Because migration costs are high (Koşar et al., 2021), relocations across local labour markets are rare events (Ahlfeldt et al., 2020). Since it is unlikely that workers have fully re-optimized their residential location choices within a few years, we provide a *short-run* evaluation in which residents are immobile across residential areas (but mobile across workplaces) and a *long-run* evaluation in which residents are fully mobile. In Section 4.4.2, we compare the predicted effects to observed before-after changes in our data. Note that our model-based counterfactuals deliver forecasts in the sense that they are based solely on data observed before the introduction of the minimum wage. Hence, the comparison of the model's predictions to observed changes in data represents an over-identification test that allows us to evaluate the out-of-sample predictive power of our model.

4.4.1 Model-based counterfactuals

In Table 1, we summarize the simulated short-run and long-run effects of the German minimum wage on various endogenous outcomes. We report the worker-weighted average across regions as well as the regional minimum and maximum values. Given a workforce of approx. 30M, the 0.3%-reduction in employment amounts to the full-time equivalent of about 100k jobs. Even if this reduction came 100% from the extensive margin (reduction of jobs), this would be much less than suggested by ex-ante predictions based on competitive labour market models (Knabe et al., 2014). Considering that at least some of the reduction in full-time equivalent employment originates from the intensive margin (reduction in hours worked, see Section 4.4.2 and Bossler and Gerner (2019)), it seems fair to conclude that the minimum wage did not cause much unemployment, a finding that is consistent with reduced-form evidence (Ahlfeldt et al., 2018; Dustmann et al., 2022).

Applying the relative welfare effect (for those working) of about 3% to the 2018 average annual wage of €34.4K to about 30M workers, we can monetize the aggregate welfare effect as equivalent to an increase in annual worker income of about €30BN. This increase

Table 1: Short-run and long-run effects of the German minimum wage

	Short run			Long run		
	Mean	Min	Max	Mean	Min	Max
Panel a: Employment						
Employment at workplace (L)	-0.250	-21.31	5.350	-0.350	-25.91	5.810
Labour supply at residence (N)	0.590	0.120	1.550	0.590	-6.540	14.23
Employment probability (L/H)	-0.820	-19.99	0	-0.880	-21.15	0
Employment-weighted productivity ($\tilde{\varphi}_j$)	-1.030	-3.910	36.83	-1.040	-3.910	31.74
Panel b: Wage and prices						
(Normalized) wage ($\tilde{w}$)	0.320	-1.360	25.72	0.390	-1.310	24.70
Real tradables price index (P^Q)	-3.040	-4.620	-2.200	-2.930	-5.630	-1.600
Real housing rent (P^T)	-1.040	-7.170	1.100	-1.070	-5.390	2.520
Panel c: Welfare components						
Exp. real wage $\tilde{v}\left[(P^Q)^\alpha(P^T)^{(1-\alpha)}\right]$	1.620	-0.260	5.510	1.630	0.370	4.350
# establishments (M)	-0.100	-7.290	0.920	-0.120	-16.43	2.770
Ease of commuting ($B\kappa^{-\epsilon}$)	1.160	-4.290	7.090	0.880	-14.04	8.440
Panel d: Welfare						
Worker welfare | working (V)	2.910	0.560	7.830	2.860	2.860	2.860
Worker welfare, all ($\mathcal{V}$)	2.150	0.410	5.770	2.110	2.110	2.110

Notes: All outcomes are given in terms of % changes. Mean is the mean outcome across municipalities, weighted by initial workplace or residence employment. Min and max are minimum and maximum values in the distribution across municipalities. Short run gives simulation results when workers are immobile across residences whereas long-run results allow workers to be fully mobile. Outcomes are normalized by the mean wage across all municipalities. Expected real wage effect captures the direct (positive) effect of the minimum wage on wages and the effect on the hiring probability that can be negative in municipalities with sufficient demand-constrained firms.

in welfare is driven by an increase in *expected* real wage, i.e. higher real wages more than compensate for lower employment probabilities. This is why the aggregate labour force increases by about 0.6%, corresponding to about 180k workers. Notice that the near-zero effect on the mean wage is an artifact of the choice of the numeraire in our model: the worker-weighted average of regional wages. Since, expressed in units of this numeraire, tradable goods prices and real housing rents decrease, real wages actually increase. Ease of commuting is another source of the positive welfare effect. Since the share of commuters increased where the minimum wage had greater bite (see Online Supplement for details), this effect is likely driven by the intensive margin, i.e. commuters find jobs in more convenient reach. In contrast, the reduction in the number of establishments has negative welfare effects because the chance of a good worker-firm match decreases. The mean welfare effects on all workers, at 2.1%, ($\mathcal{V}$) is smaller than the 2.9%-effect on those working (V) because about a quarter of the working-age population abstains from the labour market and, hence, experiences no welfare effect.

Table 1 also reveals that the national averages mask striking spatial heterogeneity. Some municipalities experience substantial reduction in employment, whereas employment increases in others. This mirrors a highly heterogeneous increase in real wages. Some (low-productivity) municipalities experience a significant increase in average productivity as low-productivity demand-constrained firms reduce their labour input, which is consistent with the reallocation effect documented in reduced-form by Dustmann et al. (2022). However, the general equilibrium effect is negative since in many of the more populated regions, demand-constrained and supply-constrained firms, as a consequence of lower monopsony power, expand employment at the expense of more productive unconstrained firms.[16] While the regional spread in short-run and long-run effects is mostly similar, there are two important exceptions. When we fix worker residences in the short run, we essentially switch off an important margin of the spatial arbitrage process. Within each area, the size of the labour force can only change due to workers entering or exiting the labour market. This rules out migration-induced adjustments in wages and rents that would equalize utility. As a result, there is significant spatial heterogeneity in the welfare incidence. When we allow for free residential choices in the long run, migration-induced spatial arbitrage equalizes the welfare incidence, but we observe much greater changes in the spatial distribution of the labour force.

We dig deeper into the spatial heterogeneity in minimum wage effects in Figure 4, where we correlate our simulated relative changes in selected outcomes to the 2014 mean wage observed in our data, which is a proxy for regional productivity.[17] As expected, the employment effect follows the hump-shaped pattern that we have derived theoretically and substantiated empirically in partial equilibrium in Section 3. It is straightforward to infer the critical points introduced in Figure 2. For regions where the 2014 mean hourly wage exceeds €19, the employment effect is flat in the initial regional wage ($\underline{\phi}_j^{'''}$). We find the most positive employment response for regions where the wage is about €16 ($\underline{\phi}_j^{''}$). For regions where the wage is below €13 ($\underline{\phi}_j^{'}$), the employment effects tend to be more negative than for the high-productivity regions. Reassuringly, these critical points derived from model-based counterfactuals are close to those in Figure 3, which are based on a reduced-form before-after comparison.

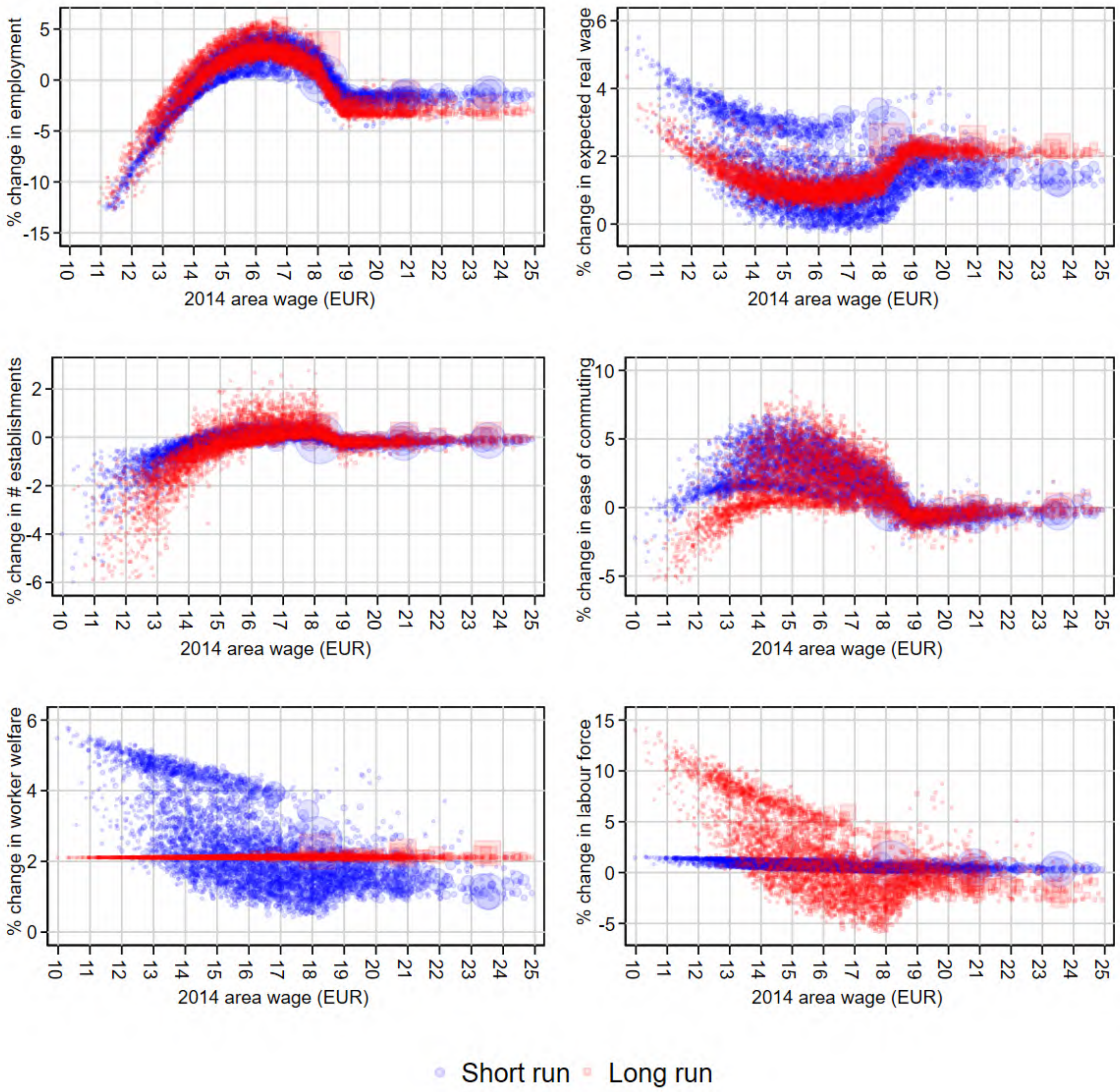
Our general equilibrium analysis adds the important insight that there is a negative level effect on unconstrained regions in the right tail of the regional productivity distribution, which is unidentifiable with the reduced-form approach in Figure 3. Intuitively, the expansion of employment and production in municipalities of intermediate productivity comes at the expense of the most productive municipalities as workers relocate to less productive regions to save commuting or living costs. Another important insight is that commuting is a sufficient margin of adjustment in labour supply to generate the hump shape; allowing, in addition, for migration only slightly reinforces the relocation of employment from high- to middle-wage regions. If, however, we rule out migration *and* commuting, the hump shape disappears and gives way to a concave relationship that is inconsistent with the data (see Online Supplement).

The effect on the number of establishments follows the employment effect qualitatively. The effect on the ease of commuting also has a hump shape. This is consistent with a

[16] For details on the reallocation effect, we relegate the interested reader to the Online Supplement.

[17] In the absence of the minimum wage when all firms are unconstrained, the mean wage $\tilde{w}$ maps to the lower-bound wage and productivity via Eqs. (19) and (7).

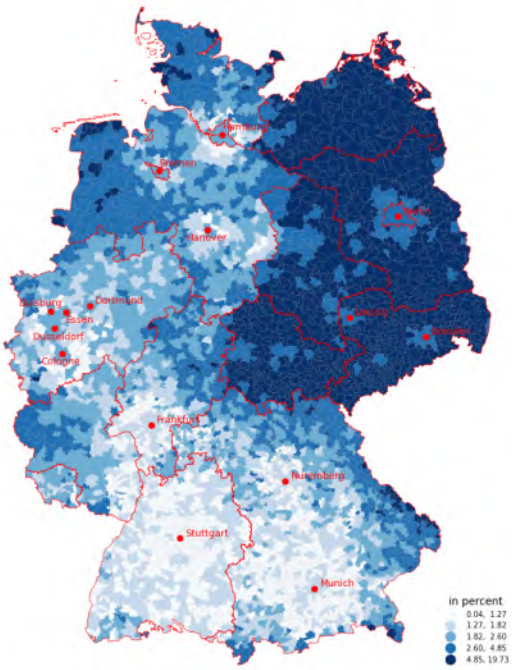
Figure 4: Short-run and long-run effects by regional productivity



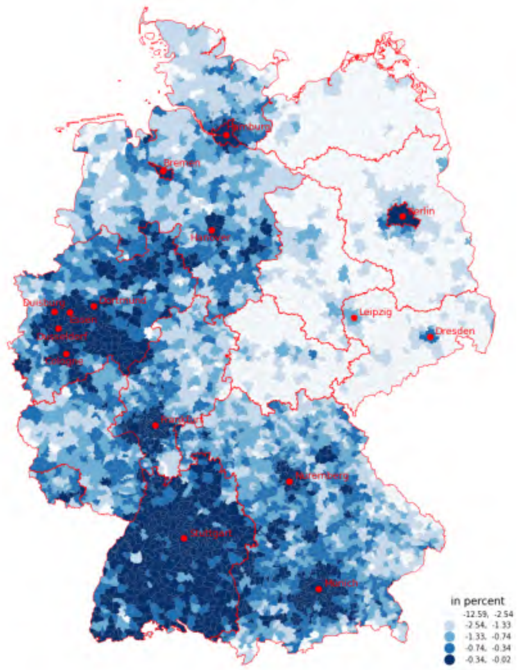
Note: Each icon represents one outcome for one municipality (*Verbandsgemeinde*). Results of model-based counterfactuals comparing the equilibrium under a federal minimum of 48% (the value observed in data) of national mean wage to the equilibrium with a zero minimum wage. *Blue circles* show outcomes when workers are immobile across residences (short run). *Red squares* show outcomes when workers are mobile across residence (long run). Expected real wage is measured at the residence and incorporates changes in (normalized) nominal wages at workplace, employment probabilities at workplace, bilateral commuting probabilities, housing rents at residence, and tradable goods prices at residence. For a more intuitive interpretation, we multiply the normalized regional mean wage on the x-axes by the 2014 national mean wage. To improve the presentation, we crop the right tail of the regional productivity distribution (about one percent).

reallocation effect of workers towards more productive establishments further away from their residences in low-productivity regions (Dustmann et al., 2022). In contrast, the ease of commuting increases in municipalities of intermediate productivity, revealing that workers find attractive employment opportunities that are more convenient to reach.

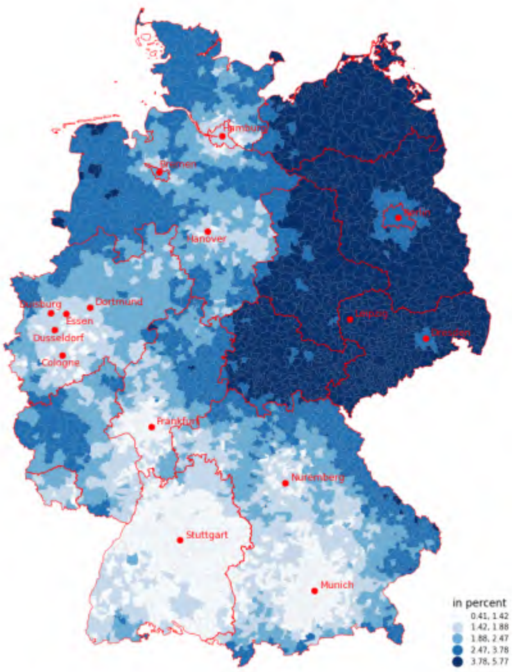
Figure 5: Regional effects of the German minimum wage



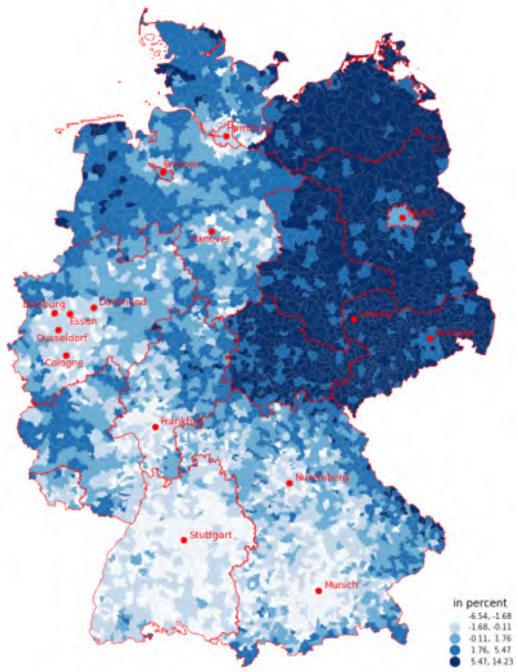
(a) Real wage, short run



(b) Employment probability, short run



(c) Welfare, short run



(d) Labour force, long run

Note: Unit of observation are 4,421 municipality groups. Results from model-based counterfactuals are expressed as percentage changes. All outcomes are measured at the place of residence. To generate the data displayed in panels a) and b), we break down residential income from Eq. (33) into two components. The first is the residential wage conditional on working $\sum_j \lambda_{ij|i}^N \tilde{w}_j$, which we normalize by the consumer price index (the weighted combination of goods prices and housing rent) to obtain the real wage. The second is the residential employment probability $\sum_j \lambda_{ij|i}^N L_j/H_j$, which captures the probability that a worker finds a job within the area-specific commuting zone.

While low-productivity regions are those that experience the largest decline in employment, they are also those where the minimum wage has had the greatest effect on wages. Figure 5 shows that municipalities experiencing real wage growth and a reduction in employment probability are over-represented in the east. Because the former dominates the latter, expected real wages increase (see top-middle panel in Figure 4). The bottom panels of Figures 4 and 5 illustrate how, as a result, welfare increases in the short run and the labour force increases in the long run. The important take-away for policy is that the German minimum wage has disproportionately improved welfare in economically weak municipalities, but the effect will become more uniform in the long run as workers re-optimize their location choices. That said, population growth in economically weaker regions may well represent a policy objective in its own right, especially in Germany where there has been substantive out-migration from former East Germany after the fall of the iron curtain.

4.4.2 Validation against data

By design, our model perfectly rationalizes the data in the initial equilibrium.[18] Therefore, we over-identify the model by testing its ability to forecast out-of-sample changes over time. To this end, we use the model-based spatial minimum wage effects discussed in Section 4.4.1 as treatment variables in a conventional difference-in-differences event study. Intuitively, this approach compares before-after changes in selected outcomes in the data to the corresponding before-after changes predicted by the model. We present the results in Figure 6. Before 2015 (the year of the minimum-wage introduction), we can interpret the estimated treatment effects as placebo effects which should be close to zero. From 2015 onward, a treatment effect of one would indicate that changes in the data scale proportionately in our model's prediction. In practice, it is unrealistic to expect a coefficient close to one since, unlike in the model, fundamentals in the real world change for reasons unrelated to the minimum wage, resulting in attenuation bias. Hence, positive coefficients are all the more reassuring of the model's ability to forecast minimum wage effects. For further details on the empirical specification and the interpretation of the estimated parameters, we refer to the Online Supplement.

The first insight from Figure 6 is that the before-after changes in regional mean hourly wage and employment observed in data converge towards the predictions of the model over time. One interpretation is that compliance has been imperfect, but increasing over time. Imperfect compliance with minimum wage laws is a well-known phenomenon (Ashenfelter and Smith, 1979) that can mitigate employment effects (Garnero and Lucifora, 2022). While Germany is no exception (Mindestlohnkommission, 2020), evidence from labour force surveys suggests that compliance has increased over time (Weinkopf, 2020). Figure 6 also reveals out-of-sample predictive power for underemployment. The model's forecasts

[18] Following the conventions in quantitative spatial economics, we invert the model's exogenous structural fundamentals to match observed values of endogenous variables, taking structural parameters as given (Redding and Rossi-Hansberg, 2017).

of the regional distribution of changes in employment probability are positively correlated with observed changes in the share of full-time workers. In contrast, we find no positive correlation with observed changes in the unemployment rate (not shown). This is consistent with a reduced-form literature that has found evidence for a negative effect of the German minimum wage on hours worked (Bossler and Gerner, 2019), but no evidence for an increase in unemployment (Dustmann et al., 2022).

Similarly, there is a trend in the labour force (measured at the residence) to converge to the short-run predictions of the model. Thus, workers become active on the labour market where the model predicts the utility from work to increase in the short run. In contrast, workers do not seem to have started to relocate to regions with positive short-run welfare gains within the first four years of the policy, which is consistent with high migration costs in Germany (Ahlfeldt et al., 2020). This may explain why the model's predictions of house price effects are only weakly correlated with observed changes in the data.[19]

Figure 6: Regional minimum wage effects: Model vs. data

Note: For each panel, we run a regression of the log of an outcome variable against the log of the relative change (the ratio of the model-predicted outcome with the minimum wage over the baseline) interacted with year dummies (omitting 2014 as the baseline), controlling for area and year-by-zone (former East and West Germany) dummies. Prior to this regression, we adjust all area-level time series for the pre-minimum wage time trend following Monras (2019). In the upper-right panel, the empirically observed variable is the share of full-time workers as we do not observe the exact number of hours worked. In all other panels we use the actual empirical counterparts to the variables in the model. Average establishment productivity is the worker-weighted average of time-invariant establishment fixed effects from an AKM wage decomposition. Icons denote point estimates. Error bars give 95% confidence bands. In the bottom-right panel, we pool over all outcomes, using area-by-outcome and year-by-zone-by-outcome fixed effects.

[19] Yamagishi (2021) shows that desirable minimum wages increase housing rents in Japan.

Given the emphasis of the model on the underlying mechanisms, it is reassuring that the model predicts the reallocation effect of the German minimum wage (Dustmann et al., 2022). To understand how well the model predicts the reallocation of workers to establishments in different commuting destinations, we compute the commuting distance at the residence in the model and the data. To understand how well the model predicts the reallocation of workers across establishments of different productivity, we compute the worker-weighted average of establishment productivity at the workplace in the model and the data. In both cases, the variation over time originates from time-varying worker weights exclusively while bilateral distance and establishment productivity are time-invariant. In both cases, the model's short-run and long-run predictions of changes are positively correlated with before-after changes observed in data. Since establishment size and productivity are positively correlated, it is no surprise that the model also predicts the regional minimum wage effect on average establishment size well. A pooled analysis of all outcomes confirms the impression that the before-after changes observed in data generally converge to the model's predictions over time.

We also find support for the model's prediction in another metric that is of first-order relevance in the context of minimum wage laws: The Gini coefficient of nominal wage inequality (across all workers in all regions).[20] Our model predicts a short-run reduction of the Gini coefficient of about two percentage points (from 32.7% to 30.7%). This is qualitatively and quantitatively in line with an empirically observed steady decline in the Gini coefficient from 30.7% to 29.1% during the first three years of the minimum wage. Similarly, the model correctly predicts that the minimum wage had virtually no effect on the Gini coefficient of the distribution of employment across regions.

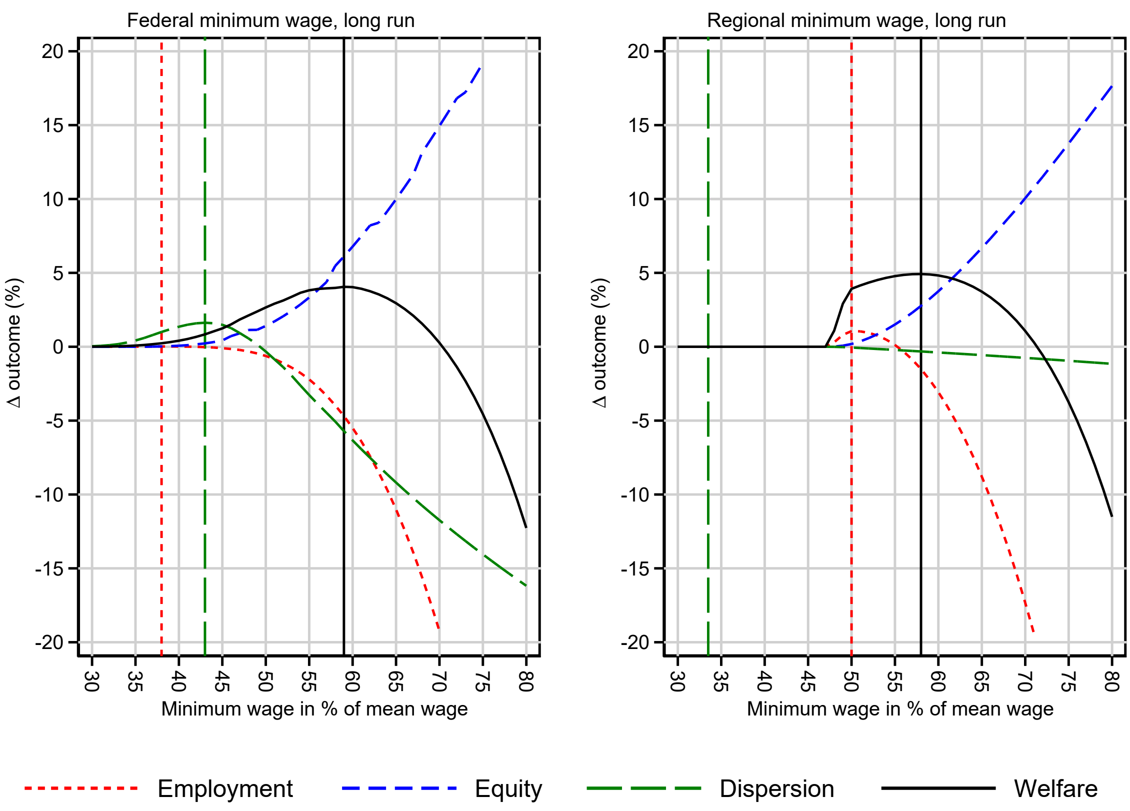
4.5 Optimal minimum wages

We now turn from the positive evaluation of the effects of the German minimum wage to a normative evaluation of optimal minimum wages. To this end, we conduct a series of counterfactual exercises using the procedure outlined in Section 4.3. We evaluate two alternative minimum wage schedules that are fairly straightforward to implement from a policy perspective. For one thing, we consider a $1 \times \mathcal{N}$ vector of uniform relative *national* minimum wages $\underline{\mathbf{w}}^{\mathbf{n}} \in (0.3, 0.31, ...0.8)$ that correspond to a fraction of the national mean wage, the numeraire in our model. For another, we consider a $J \times \mathcal{N}$ vector of *regional* minimum wages $\underline{\mathbf{w}}_{\mathbf{j}}^{\mathbf{r}} = \mathbf{w}_{\mathbf{j}}^{\mathbf{m}} \cdot \underline{\mathbf{w}}^{\mathbf{n}}$ that represents the regional minimum wage as a fraction of the $J \times 1$ vector of regional mean wages $\mathbf{w}_{\mathbf{j}}^{\mathbf{m}}$.

One obvious optimality criterion for a successful minimum wage policy in the context of our model is expected worker welfare as defined in Eq. (37). Since the literature on minimum wages is very much concerned with employment effects, we also evaluate the aggregate employment effect. In practice, one of the main policy objectives associated with minimum wages is a reduction in income inequality. Therefore, we also report an

[20] See Online Supplement for details.

Figure 7: Minimum wage effects on employment, equity, dispersion, and welfare



Note: Results of model-based counterfactuals. Employment is the total number of workers in employment. Equity is measured as 1-$\mathcal{G}$, where $\mathcal{G}$ is the Gini coefficient of real wage inequality across all workers in employment. Dispersion is measured as 1-$\mathcal{S}$, where $\mathcal{S}$ is the Gini coefficient of the distribution of employment across regions. Welfare is the expected utility as defined by Eq. (37). It captures individuals who are active on and absent from the labour market and accounts for minimum wage effects on employment probabilities, wages, tradable goods prices, housing rents, commuting costs, and worker-firm matching qualities. In the short run, workers are immobile across residence locations whereas workers re-optimize their residential location choice in the long run.

equity measure $1 - \mathcal{G}$, where $\mathcal{G}$ is the Gini coefficient of nominal wage inequality across all workers in all regions. To capture the effect on the spatial distribution of economic activity, we compute a spatial dispersion measure $1 - \mathcal{S}$, where $\mathcal{S}$ is the Gini coefficient of the distribution of employment across regions (see Online Supplement for derivations).

Figure 7 summarizes how employment, equity, dispersion and welfare effects vary in the level of a federal or regional minimum wage. We focus on the long-run scenario in which workers are fully mobile to ease the presentation. We document in the Online Supplement that the short-run results are very similar, confirming that commuting is a sufficient spatial margin of adjustment for the reallocation of labour supply across regions if there is great heterogeneity in regional productivity. We compare the effects of regional and federal minimum wages that maximize employment, dispersion or welfare in Table 2. With these ingredients at hand, the interested reader will be able to infer a social welfare effect according to their preferred social welfare function.

The first insight is that the welfare effect is hump-shaped in the minimum wage level, whether the minimum wage is nationally uniform or regionally differentiated. The intuition is that up to the welfare-maximizing minimum wage, the positive effect on real wages dominates the negative effect on employment probabilities due to efficiency gains, such that expected wages and welfare increase. With a federal minimum wage, this point is reached at a level of 58% of the national mean wage. Beyond this point, the negative effect on employment probabilities dominates at the margin. At 70%, the absolute welfare effect turns negative.

Since minimum wages mechanically compress the nominal wage distribution, it is no surprise that our measure of equity increases monotonically in the level of the minimum wage. Under conventional social welfare functions that discount aggregate welfare by the Gini coefficient of income inequality (Newbery, 1970), an increase beyond the welfare-maximizing minimum wage can be justified. Yet, policy makers may wish to take into account that beyond a minimum wage of 50% of the national mean wage, negative employment effects start building up as more and more firms must reduce their labour input in order to raise their MRPL to the minimum wage level. Since the reduction in employment is concentrated in low-productivity regions, it is no surprise that the spatial distribution of jobs becomes less equitable at high federal minimum wages. At moderate levels, however, a federal minimum wage can result in a more dispersed spatial distribution because supply-constrained firms in lower-productivity regions expand employment. Indeed, the minimum wage that results in the most even geography of jobs, at 43%, is significantly lower than the welfare-maximizing minimum wage.

Intuitively, the employment-maximizing minimum wage must also be lower than the welfare-maximizing minimum wage since, unlike the latter, the former does not take into account positive welfare effects from higher wages earned by those who remain in (full-time) employment. Indeed, the long-run employment-maximizing federal minimum wage is as low as 38%. While this moderate minimum wage does increase employment, the effect is very small, and so are the effects on equity and welfare. The employment-maximizing minimum wage is almost identical to the *efficiency maximizing minimum wage* of 37% of the mean wage simulated by Berger et al. (2022) for the US within a dynamic macroeconomic model. Like our employment-maximizing minimum wage, their *efficiency maximizing minimum wage* is only concerned with correcting for the inefficiencies that originate from the employer monopsony. Similar to ours, their model predicts higher optimal minimum wages once worker welfare effects are taken into account.[21] The important takeaway is that, in setting federal minimum wages, policy makers trade positive aggregate welfare effects and progressive between-worker distributional effects (within employed workers) against negative employment and regressive spatial distributional effects.

This trade-off can be mitigated by setting minimum wages at the regional level.

[21] Their optimal minimum wage of $15/h under utilitarian welfare weights would correspond to 77% of the national mean wage, resulting in an employment loss of about 3.5%. We thank the authors for converting the absolute relative minimum wages reported in Berger et al. (2022) into the relative minimum wages reported here.

Table 2: Minimum wage schedules

Objective	Scheme	Level rel. to Mean	p50	Empl. SR	LR	Equity SR	LR	Dispersion SR	LR	Welfare SR	LR
Actual	Federal	48.0	52.8	-0.3	-0.3	1.2	1.1	0.4	0.6	2.1	2.1
Employment	Federal	38.0	41.8	0.0	0.0	0.1	0.0	0.9	1.0	0.2	0.2
Dispersion	Federal	43.0	47.3	0.0	0.0	0.3	0.2	1.4	1.6	0.8	0.8
Welfare	Federal	58.0	63.8	-3.9	-4.0	5.5	5.5	-4.9	-5.1	4.0	4.0
Employment	Regional	50.0	55.0	1.1	1.1	0.2	0.2	-0.1	-0.1	3.9	3.9
Dispersion	Regional	33.0	36.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Welfare	Regional	58.0	63.8	-1.5	-1.5	2.8	2.8	-0.3	-0.3	4.9	4.9

Notes: All values are given in %. *Objective* describes if the minimum wage is employment-maximizing or welfare-maximizing. To compute the minimum wage relative to the median, we multiply the minimum wage relative to the mean by the inverse of the ratio of the median wage over the mean wage. In Germany, this ratio was 0.908 in 2015, with remarkably little variation over time. *Federal* indicates a uniform minimum wage, where the minimum wage *level* is given as a percentage of the national mean wage. *Regional* indicates a minimum wage that is set the respective *level* of the municipality mean. Results are from model-based counterfactuals. Employment is the total number of workers in employment. Equity is measured as 1-$\mathcal{G}$, where $\mathcal{G}$ is the Gini coefficient of real wage inequality across all workers in employment. Dispersion is measured as 1-$\mathcal{S}$, where $\mathcal{S}$ is the Gini coefficient of employment across regions. Welfare is the expected utility of as defined by Eq. (37). It captures individual who are active on and absent from the labour market and accounts for minimum wage effects on employment probabilities, wages, tradable goods prices, housing rents, commuting costs, and worker-firm matching qualities. In the short run, workers are immobile across residence locations whereas workers re-optimize their residential location choice in the long run. We strictly select the long-run maximizing minimum wages.

The employment-maximizing regional minimum wage—at 50% of the municipality mean wage—delivers a similar welfare gain as the welfare-maximizing federal minimum wage, plus a positive employment effect of 1.1%. Intuitively, the regional minimum wage is a more targeted policy instrument that avoids the main problem of the federal minimum wage: Reducing the monopsony power of supply-constrained firms in high productivity municipalities comes at the cost of increasing the wage beyond the MRPL of low-productivity firms in low-productivity regions. Instead, the regional minimum wage, by accounting for regional productivity heterogeneity, affects mostly supply-constrained firms in all regions. Since the regional minimum wage affects regions similarly, there are generally small effects on the spatial distribution of jobs. Notice that we find similar effects if we set the minimum wage at the county level (*Kreise*) whereas a state (*Bundesland*) minimum wage has effects that resemble the federal minimum wage (see Online Supplement). This confirms the intuition that a minimum wage needs to be sufficiently localized to account for productivity differentials *between* commuting zones but not necessarily *within* commuting zones. The reason is that workers can relatively easily re-optimize to heterogeneous labour demand shocks via the commuting margin (Monte et al., 2018). In this context, it is worth highlighting that the employment effects we simulate for the municipality and county regional minimum wages are closer to Drechsel-Grau (2021) than our simulations for federal minimum wages. This is intuitive since, in relative terms, the regional minimum wage is uniform within our spatial economy, similar to the federal minimum wage in Drechsel-Grau's macroeconomic model with only one region.

Another insight from Figure 7 and Table 2 is that long-run and short-run welfare effects are generally similar in the national aggregate. It is important, however, to recall that there is substantial regional heterogeneity in the welfare effect of federal minimum wages in the short run, which is equalized through migration in the long run (see Figures 4 and 5). How this regional heterogeneity plays out very much depends on the level of the minimum wage and the regional productivity distribution. While the actual German minimum wage benefits many low-productivity municipalities in the eastern states in terms of short-run welfare and long-run migration (see Figure 5), the regional fortunes reverse under a 25% higher welfare-maximizing federal minimum wage.[22] Short-run welfare increases more in the more productive west, resulting in a long-run increase in labour force at the expense of the east. In contrast, because the regional minimum wage "bites" similarly in all regions, there is little spatial heterogeneity in the short-run effects on welfare and the long-run effects on the labour force.

5 Conclusion

Minimum wage policies have been popular policy tools to reduce wage inequality. In light of the success of the monopsony model and a growing body of reduced-form evidence, they have also become more popular among economists as the fear of catastrophic employment effects is fading. As a result, more ambitious minimum wages are now being debated in many countries. The European Commission advocates an *adequate minimum wage* of 60% of the median wage. A recent report published by *HM Treasury* recommends a similar level. The German government has recently implemented a minimum wage of €12 that was close to 70% of the median wage in October 2022. The *Raise the Wage Act* would increase the U.S. federal minimum wage to $15 per hour by 2025, putting it in a similar ballpark, in relative terms.[23] We inform this debate in a concrete, yet nuanced fashion.

Our simulations within a quantitative model calibrated to German micro-regional data suggest that the optimal minimum wage depends on the policy objective. The welfare-maximizing national minimum wage is as high as 64% of the median wage. Aversion to between-worker wage inequality can motivate even higher minimum wages. The minimum wage that leads to the most even spatial distribution of jobs, at 47%, is much lower. The employment-maximizing minimum wage, which marks the lower-bound for optimal national minimum wages is as low as 42%. Thus, policy makers trade positive aggregate welfare effects and progressive between-worker distributional effects (among employed workers) against negative employment and regressive between-regional distributional effects. While these trade-offs may appear frustrating from a policy perspective, our analysis also reveals some more encouraging news. Instead of going down the route of ever higher federal minimum wages, policy makers have the alternative of implementing regional minimum wages. We find that regional minimum wages—if set for spatial units no larger than counties—are targeted policy instruments that mitigate the trade-off of

[22] We relegate details to the Online Supplement.

[23] For background on these initiatives, see European Commission (2020); Dube (2019); Deutscher Bundestag (2020a,b); H. R. 603 (2021).

negative employment effects and positive welfare effects. To illustrate the potential, the employment-maximizing minimum wage, at 50% of the municipality mean wage, could increase welfare by 3.9%—as much as the welfare-maximizing federal minimum wage—*and* generate a sizable positive employment effect of 1.1%, in all regions.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

Ahlfeldt, Gabriel M. and Elisabetta Pietrostefani, "The economic effects of density: A synthesis," *Journal of Urban Economics*, 2019, *111* (February), 93–107.

_, Duncan Roth, and Tobias Seidel, "The regional effects of Germany's national minimum wage," *Economics Letters*, 2018, *172*, 127–130.

_, Fabian Bald, Duncan Roth, and Tobias Seidel, "Quality of life in a dynamic spatial model," *CEPR Discussion Paper*, 2020, *15594.*

Ahlfeldt, Gabriel M, Stephan Heblich, and Tobias Seidel, "Micro-geographic property price and rent indices," *Regional Science and Urban Economics*, 2023, *98*, 103836.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf, "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (4), 2127–2189.

Allen, Stephen P., "Taxes, redistribution, and the minimum wage: A theoretical analysis," *The Quarterly Journal of Economics*, 1987, *102* (3), 477–490.

Allen, Treb and Costas Arkolakis, "Trade and the topography of the spatial economy," *The Quarterly Journal of Economics*, 2014, *129* (3), 1085–1140.

Arkolakis, Costas, "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 2010, *118* (6), 1151–1199.

Ashenfelter, Orley and Robert S Smith, "Compliance with the Minimum Wage Law," *Journal of Political Economy*, 1979, *87* (2), 333–350.

Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi, "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure," *American Economic Review*, 2020, *110* (10), 3139–3183.

Azar, José, Emiliano Huet-Vaughn, Ioana Marinescu, Bledi Taska, and Till von Wachter, "Minimum Wage Employment Effects and Labor Market Concentration," *National Bureau of Economic Research Working Paper Series*, 2019, *No. 26101.*

Bamford, Iain, "Monopsony Power, Spatial Equilibrium, and Minimum Wages," *Working paper*, 2021.

Berger, David W., Kyle F. Herkenhoff, and Simon Mongey, "Minimum Wages, Efficiency and Welfare," *National Bureau of Economic Research Working Paper Series*, 2022, *No. 29662.*

Blömer, Maximilian J, Nicole Guertzgen, Laura Pohlan, Holger Stichnoth, and Gerard J van den Berg, "Unemployment Effects of the German Minimum Wage in an Equilibrium Job Search Model," *Center for European Economic Research Discussion Paper*, 2018, *18-032.*

Boelmann, Barbara and Sandra Schaffner, "Real-Estate Data for Germany (RWI-GEO-RED v1) - Advertisements on the Internet Platform ImmobilienScout24 2007-03/2019," Technical Report, RWI Leibniz-Institut für Wirtschaftsforschung 2019.

Bossler, Mario and Hans-Dieter Gerner, "Employment Effects of the New German Minimum Wage: Evidence from Establishment-Level Microdata," *ILR Review*, 2019, *73* (5), 1070–1094.

Brown, Alessio J G, Christian Merkl, and Dennis J Snower, "The minimum wage from a two-sided perspective," *Economics Letters*, 2014, *124* (3), 389–391.

Caliendo, Marco, Alexandra Fedorets, Malte Preuss, Carsten Schröder, and Linda Wittbrodt, "The short-run employment effects of the German minimum wage reform," *Labour Economics*, 2018, *53* (August), 46–62.

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

_ **and Alan B Krueger**, "Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania," *American Economic Review*, 1994, *84* (4), 772–793.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Chen, Yujiang and Coen Teulings, "What is the Optimal Minimum Wage?," *Working paper*, 2021.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber, "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 2011, *101* (3), 471–475.

Clemens, Jeffrey and Michael Wither, "The minimum wage and the Great Recession: Evidence of effects on the employment and income trajectories of low-skilled workers," *Journal of Public Economics*, 2019, *170*, 53–67.

Datta, Nikhil, "Local Monopsony Power," *Working paper*, 2021.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Unbalanced Trade," *American Economic Review*, 2007, *97* (2), 351–355.

Deutscher Bundestag, "Gesetzlichen Mindestlohn in einmaligem Schritt auf 12 Euro erhöhen," *Drucksache*, 2020, *19/20030.*

_ , "Mindestlohn erhöhen, durchsetzen und die Mindestlohnkommission reformieren," *Drucksache*, 2020, *19/22554.*

Drechsel-Grau, Moritz, "Macroconomic and distributional effects of higher minimum wages," *Working paper*, 2021.

Dube, Arindrajit, *Impacts of minimum wages wages: Review of the international evidence*, London, UK: HM Treasury, 2019.

__ , **T William Lester, and Michael Reich**, "Minimum wage effects across state borders: Estimates using contiguous counties," *Review of Economics and Statistics*, 2010, *92* (4), 945–964.

Dustmann, Christian, Attila Lindner, Uta Schönberg, Matthias Umkehrer, and Philipp vom Berge, "Reallocation Effects of the Minimum Wage," *The Quarterly Journal of Economics*, 2022, *137* (1), 267–328.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Egger, Hartmut, Peter Egger, and Udo Kreickemeier, "Trade, wages, and profits," *European Economic Review*, 2013, *64*, 332–350.

__ , **Udo Kreickemeier, Christoph Moser, and Jens Wrona**, "Exporting and offshoring with monopsonistic competition," *Economic Journal*, 2022, *132* (644), 1449–1488.

Engbom, Niklas and Christian Moser, "Earnings Inequality and the Minimum Wage: Evidence from Brazil," *American Economic Review*, 2022, *112* (12), 3803–3847.

European Commission, "Proposal for a directive of the european parliament and of the council on adequate minimum wages in the european union. COM(2020) 682 final," 2020.

Fetzer, Thiemo, "Did Austerity Cause Brexit?," *American Economic Review*, 2019, *109* (11), 3849–3886.

Garnero, Andrea and Claudio Lucifora, "Turning a 'Blind Eye'? Compliance with Minimum Wage Standards and Employment," *Economica*, 3 2022, *n/a* (n/a).

Gaubert, Cecile, "Firm Sorting and Agglomeration," *American Economic Review*, 2018, *108* (11), 3117–3153.

__ , **Patrick Kline, Damián Vergara, and Danny Yagan**, "Trends in US Spatial Inequality: Concentrating Affluence and a Democratization of Poverty," *AEA Papers and Proceedings*, 2021, *111*, 520–525.

Gerritsen, Aart, "To redistribute or to predistribute?," *Working paper*, 2022.

Guesnerie, Roger and Kevin Roberts, "Minimum wage legislation as a second best policy," *European Economic Review*, 1987, *31* (1-2), 490–498.

H. R. 603, "To provide for increases in the Federal minimum wage, and for other purposes," 2021.

Haanwinckel, Daniel, "Supply, Demand, Institutions, and Firms: A Theory of Labor Market Sorting and the Wage Distribution," *Working paper*, 2020.

Harasztosi, Peter and Attila Lindner, "Who Pays for the Minimum Wage?," *American Economic Review*, 2019, *109* (8), 2693–2727.

Henkel, Marcel, Tobias Seidel, and Jens Suedekum, "Fiscal Transfers in the Spatial Economy," *American Economic Journal: Economic Policy*, 2021, *13* (4), 433–468.

Hurst, Erik, Patrick J Kehoe, Elena Pastorino, and Thomas Winberry, "The Distributional Impact of the Minimum Wage in the Short and Long Run," *National Bureau of Economic Research Working Paper Series*, 2022, *No. 30294.*

Jha, Priyaranjan and Antonio Rodriguez-Lopez, "Monopsonistic labor markets and international trade," *European Economic Review*, 2021, *140*, 103939.

Knabe, Andreas, Ronnie Schöb, and Marcel Thum, "Der flächendeckende Mindestlohn," *Perspektiven der Wirtschaftspolitik*, 2014, *15* (2), 133–157.

Koşar, Gizem, Tyler Ransom, and Wilbert van der Klaauw, "Understanding Migration Aversion Using Elicited Counterfactual Choice Probabilities," *Journal of Econometrics*, 2021, p. forthcoming.

Lavecchia, Adam M, "Minimum wage policy with optimal taxes and unemployment," *Journal of Public Economics*, 2020, *190*, 104228.

Lee, David and Emmanuel Saez, "Optimal minimum wage policy in competitive labor markets," *Journal of Public Economics*, 2021, *96* (9-10), 739–749.

Machin, Stephen, Alan Manning, and Lupin Rahman, "Where the minimum wage bites hard: Introduction of minimum wages to a low wage sector," *Journal of the European Economic Association*, 2003, *1* (1), 154–180.

Manning, Alan, *Monopsony in Motion: Imperfect Competition in Labor Markets*, Princeton, New Jersey: Princeton University Press, 2003.

_ , "The real thin theory: monopsony in modern labour markets," *Labour Economics*, 4 2003, *10* (2), 105–131.

_ , "Monopsony in Labor Markets: A Review," *ILR Review*, 2020, *74* (1), 3–26.

_ , "The Elusive Employment Effect of the Minimum Wage," *Journal of Economic Perspectives*, 2021, *35* (1), 3–26.

McFadden, Daniel, "The measurement of urban travel demand," *Journal of Public Economics*, 1974, *3* (4), 303–328.

Meer, Jonathan and Jeremy West, "Effects of the Minimum Wage on Employment Dynamics," *Journal of Human Resources*, 2016, *51* (2), 500–522.

Melitz, Marc J, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 7 2003, *71* (6), 1695–1725.

Mincer, Jacob, "Unemployment Effects of Minimum Wages," *Journal of Political Economy*, 1976, *84* (4, Part 2), S87–S104.

Mindestlohnkommission, "Zweiter Bericht zu den Auswirkungen des gesetzlichen Mindestlohns," Technical Report, Berlin 2018.

_ , "Dritter Bericht zu den Auswirkungen des gesetzlichen Mindestlohns. Bericht der Mindestlohnkommission an die Bundesregierung nach §9 Abs. 4 Mindestlohngesetz," Technical Report, Berlin 2020.

Monras, Joan, "Minimum wages and spatial equilibrium: theory and evidence," *Journal of Labor Economics*, 2019, *37* (3), 853–904.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg, "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 2018, *108* (12), 3855–3890.

Moretti, Enrico, *The New Geography of Jobs*, Houghton Miin Harcourt, 2012.

_ , "Real Wage Inequality," *American Economic Journal: Applied Economics*, 2013, *5* (1), 65–103.

Mori, Tomoya and Jens Wrona, "Centrality Bias in Inter-city Trade," *RIETI Discussion Paper*, 2021, *E* (35).

Neumark, David, "The Employment Effects of Minimum Wages: Some Questions We Need to Answer," *Oxford Research Encyclopedia of Economics and Finance*, 2018.

_ **and Peter Shirley**, "Myth or measurement: What does the new minimum wage research say about minimum wages and job loss in the United States?," *Industrial Relations: A Journal of Economy and Society*, 4 2022, *n/a* (n/a).

_ **and William Wascher**, "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Comment," *American Economic Review*, 2000, *90* (5), 1362–1396.

Newbery, David, "A theorem on the measurement of inequality," *Journal of Economic Theory*, 1970, *2* (3), 264–266.

Oi, Walter Y. and Todd L. Idson, "Firm size and wages," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3, Elsevier, 1999, pp. 2165–2214.

Pérez, Jorge Pérez, "City Minimum Wages and Spatial Equilibrium Effects," *Working paper*, 2020, pp. 1–75.

Redding, Stephen J, "Theories of Heterogeneous Firms and Trade," *Annual Review of Economics*, 2011, *3* (1), 77–105.

Redding, Stephen J. and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9* (1), 21–58.

Roback, Jennifer, "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 1982, *90*, 1257–1278.

Simon, Andrew and Matthew Wilson, "Optimal minimum wage setting in a federal system," *Journal of Urban Economics*, 2021, *123*, 103336.

Simonovska, Ina and Michael E Waugh, "The elasticity of trade: Estimates and evidence," *Journal of International Economics*, 2014, *92* (1), 34–50.

Sokolova, Anna and Todd Sorensen, "Monopsony in Labor Markets: A Meta-Analysis," *ILR Review*, 10 2020, *74* (1), 27–55.

Statistisches Bundesamt, "Einkommens- und Verbrauchsstichprobe Konsumausgaben privater Haushalte," *Fachserie*, 2020, *15* (5).

Stigler, George J, "The Economics of Minimum Wage Legislation," *The American Economic Review*, 1946, *36* (3), 358–365.

Vergara, Damian, "Minimum Wages and Optimal Redistribution in the Presence of Taxes and Transfers," *Working paper*, 2022.

Weinkopf, Claudia, "Zur Durchsetzung des gesetzlichen Mindestlohns in Deutschland," *Aus Politik und Zeitgeschichte*, 2020, *39* (40).

Yamagishi, Atsushi, "Minimum wages and housing rents: Theory and evidence," *Regional Science and Urban Economics*, 2021, *87*, 103649.

Yeh, Chen, Claudia Macaluso, and Brad Hershbein, "Monopsony in the US Labor Market," *American Economic Review*, 2022, *112* (7), 2099–2138.

APPENDIX

This appendix provides complementary material on the theory and on the quantification of the model. It does not replace the reading of the main paper. Additional information on literature, data, the German minimum wage as well as on various empirical applications can be found in the Online Supplement available on the authors' webpages.

A Partial equilibrium

This section complements Section 3 in the main paper.

A.1 Derivation of Eq. (3)

Firm ω_j maximizes its profits

$$\max_{y_j(\omega_j),\, q_{ij}(\omega_j)} \sum_i (S_i^q)^{\frac{1}{\sigma}} q_{ij}(\omega_j)^{\frac{\sigma-1}{\sigma}} - w_j(\omega_j)\frac{y_j(\omega_j)}{\varphi_j(\omega_j)} \quad \text{s.t.} \quad y_j(\omega_j) = \sum_i \tau_{ij} q_{ij}(\omega_j), \tag{A.1}$$

with $\tau_{ij} \geq 1$ as the iceberg-type trade costs of serving location i from location j. Because firm ω_j's profit maximization problem is recursive, we can in a first step solve for the optimal allocation of sales quantities $q_{ij}(\omega_j)$ $\forall\ i \in J$ for a notionally fixed output level $\bar{y}_j(\omega_j)$, before determining in a second step the optimal level of production $y_j(\omega_j) \geq 0$. Using the corresponding first-order condition

$$\frac{q_{ij}(\omega_j)}{q_{\ell j}(\omega_j)} = \frac{S_i^q}{S_\ell^q}\left(\frac{\tau_{ij}}{\tau_{\ell j}}\right)^{-\sigma} \quad \forall\ \ell \in J, \tag{A.2}$$

to replace $q_{ij}(\omega_j)$ in the goods market clearing condition $y_j(\omega_j) = \sum_i \tau_{ij} q_{ij}(\omega_j)$ allows us to solve for $q_{ij} = (S_i^q/S_j^r)\tau_{ij}^{-\sigma} y_j(\omega_j)$, which we can substitute into the revenue equation $\sum_i p_{ij}(\omega_j) q_{ij}(\omega_j) = \sum_i (S_i^q)^{\frac{1}{\sigma}} q_{ij}(\omega_j)^{\frac{\sigma-1}{\sigma}}$ in order to obtain $r_j(\omega_j)$ in Eq. (3). ■

A.2 Firm-level outcomes

In this section, we derive the solutions for firm-level wages $w_j^z(\varphi_j)$, employment $l_j^z(\varphi_j)$, costs $c_j^z(\varphi_j)$, prices $p_{ij}^z(\varphi_j)$, quantities $q_{ij}^z(\varphi_j)$, and revenues $r_j^z(\varphi_j)$ for all firm types $z \in \{u, s, d\}$. While Table A1 collects the results, we provide derivation details for each firm type below.

Unconstrained firms. According to Eqs. (3) and (4) marginal revenues and marginal costs are proportional to average revenues $r_j(\omega_j)/l_j(\omega_j)$ and average costs $c_j(\omega_j)/l_j(\omega_j)$, where we have used $y_j(\omega_j) = \varphi_j(\omega_j) l_j(\omega_j)$ to express revenues as a function of the firm's total employment. We define the combined mark-up/mark-down factor by $1/\eta > 1$ and note that $\eta \equiv [(\sigma-1)/\sigma][\varepsilon/(\varepsilon+1)] \in (0, 1]$ is the share of revenues $r_j^u(\varphi_j)$ that corresponds to the firm's costs $c_j^u(\varphi_j)$, whereas $1-\eta$ is the share of revenues $r_j^u(\varphi_j)$ that corresponds to the firm's profits $\pi_j^u(\varphi_j)$. Evaluating $c_j^u(\varphi_j) = \eta r_j^u(\varphi_j)$ at $r_j^u(\varphi_j)$ from Eq. (3) and

Table A1: Firm-level outcomes

Unconstrained firms ($z = u$)
$w_j^u = \eta^{\frac{\sigma}{\sigma+\varepsilon}} \left(S_j^r\right)^{\frac{1}{\sigma+\varepsilon}} \left(S_j^h\right)^{-\frac{1}{\sigma+\varepsilon}} \varphi_j^{\frac{\sigma-1}{\sigma+\varepsilon}}$
$l_j^u = \eta^{\frac{\sigma\varepsilon}{\sigma+\varepsilon}} \left(S_j^r\right)^{\frac{\varepsilon}{\sigma+\varepsilon}} \left(S_j^h\right)^{\frac{\sigma}{\sigma+\varepsilon}} \varphi_j^{\frac{(\sigma-1)\varepsilon}{\sigma+\varepsilon}}$
$c_j^u = \eta^{\frac{\sigma(\varepsilon+1)}{\sigma+\varepsilon}} \left(S_j^r\right)^{\frac{\varepsilon+1}{\sigma+\varepsilon}} \left(S_j^h\right)^{\frac{\sigma-1}{\sigma+\varepsilon}} \varphi_j^{\frac{(\sigma-1)(\varepsilon+1)}{\sigma+\varepsilon}}$
$p_{ij}^u = \eta^{-\frac{\varepsilon}{\sigma+\varepsilon}} \tau_{ij} \left(S_j^r\right)^{\frac{1}{\sigma+\varepsilon}} \left(S_j^h\right)^{-\frac{1}{\sigma+\varepsilon}} \varphi_j^{-\frac{\varepsilon+1}{\sigma+\varepsilon}}$
$q_{ij}^u = \eta^{\frac{\sigma\varepsilon}{\sigma+\varepsilon}} \tau_{ij}^{-\sigma} \left(S_j^r\right)^{-\frac{\sigma}{\sigma+\varepsilon}} \left(S_j^h\right)^{\frac{\sigma}{\sigma+\varepsilon}} S_i^q \varphi_j^{\frac{\sigma(\varepsilon+1)}{\sigma+\varepsilon}}$
$r_j^u = \eta^{\frac{(\sigma-1)\varepsilon}{\sigma+\varepsilon}} \left(S_j^r\right)^{\frac{\varepsilon+1}{\sigma+\varepsilon}} \left(S_j^h\right)^{\frac{\sigma-1}{\sigma+\varepsilon}} \varphi_j^{\frac{(\sigma-1)(\varepsilon+1)}{\sigma+\varepsilon}}$
Supply-constrained firms ($z = s$)
$w_j^s = \underline{w}$
$l_j^s = S_j^h \underline{w}^{\varepsilon}$
$c_j^s = S_j^h \underline{w}^{\varepsilon+1}$
$p_{ij}^s = \tau_{ij} \left(S_j^r\right)^{\frac{1}{\sigma}} \left(S_j^h\right)^{-\frac{1}{\sigma}} \varphi_j^{-\frac{1}{\sigma}} \underline{w}^{-\frac{\varepsilon}{\sigma}}$
$q_{ij}^s = \tau_{ij}^{-\sigma} \left(S_j^r / S_j^h\right) S_i^q \varphi_j \underline{w}^{\varepsilon}$
$r_j^s = \left(S_j^r\right)^{\frac{1}{\sigma}} \left(S_j^h\right)^{\frac{\sigma-1}{\sigma}} \varphi_j^{\frac{\sigma-1}{\sigma}} \underline{w}^{\frac{(\sigma-1)\varepsilon}{\sigma}}$
Demand-constrained firms ($z = d$)
$w_j^d = \underline{w}$
$l_j^d = \rho^{\sigma} \varphi^{\sigma-1} S_j^r \underline{w}^{-\sigma}$
$c_j^d = \rho^{\sigma} S_j^r \varphi^{\sigma-1} \underline{w}^{1-\sigma}$
$p_{ij}^d = \frac{\tau_{ij}\underline{w}}{\rho\varphi_j}$
$q_{ij}^d = \left(\frac{\tau_{ij}\underline{w}}{\rho\varphi_j}\right)^{-\sigma} S_i^q$
$r_j^d = \left(\frac{\underline{w}}{\rho\varphi_j}\right)^{1-\sigma} S_j^r$

$c_j^u(\varphi_j)$ from Eq. (4) allows us to solve for the optimal employment level $l_j^u(\varphi_j)$, with the corresponding wage rate $w_j^u(\varphi_j)$ following from substitution into the (inverse) labor supply function. Revenues $r_j^u(\varphi_j)$, costs $c_j^u(\varphi_j)$, and profits $\pi_j^u(\varphi_j)$ then can be solved accordingly from Eq. (3) in combination with $c_j^u(\varphi_j)/\eta = r_j^u(\varphi_j) = \pi_j^u(\varphi_j)/(1-\eta)$. Further, defining $\gamma \equiv (\sigma-1)(\varepsilon+1)/(\sigma+\varepsilon) \in [\sigma-1, (\sigma-1)/\sigma]$ as the elasticity of revenues with respect to the firm-level productivity, we find that γ is smaller than its counterpart $\sigma-1$ in a perfectly competitive labor market (for $\varepsilon \to \infty$), because diseconomies of scale due to an upward-sloping labor supply function dampen the revenue-increasing effect associated with a higher productivity level φ_j.

The elasticities of employment $l_j^u(\varphi_j)$ and wages $w_j^u(\varphi_j)$ with respect to the productivity level are given by $[\varepsilon/(\varepsilon+1)]\gamma$ and $[1/(\varepsilon+1)]\gamma$, respectively, which highlights that the labor supply elasticity ε governs to what extent a rising productivity translates into wage and employment increases. For a perfectly elastic labour supply (i.e. $\varepsilon \to \infty$) the labour market converges to its competitive limit, in which all firms pay the same wage. If the supply of labour to the firm is perfectly inelastic (i.e. $\varepsilon = 0$), all firms in location j share the same employment level.

Supply-constrained firms. Firm-level outcomes can be obtained straightforwardly from the equations in Section 3.1. Notice that fixed labour supply at a given minimum wage fixes total firm output which will in turn be distributed across markets according to the splitting rule discussed in the context of Eq. (3). This delivers bilateral prices and quantities from which we obtain revenues and profits. Notice that the hiring probability ψ_j is equal to unity for this firm type.

Demand-constrained firms. According to Eq. (3) marginal revenues are by factor $\rho = (\sigma-1)/\sigma \in (0,1)$ lower than average revenues, which is why prices are set as constant mark-ups $1/\rho > 1$ over marginal costs $p_{ij}^d(\varphi_j) = (1/\rho)\tau_{ij}\underline{w}/\varphi_j$, implying that costs $c_j^d(\varphi_j)$ and profits $\pi_j^d(\varphi_j)$ are constant shares ρ and $1-\rho$ of the firm's revenues $r_j^d(\varphi_j)$. Having solved the optimal employment level $l_j^d(\varphi_j) = y_j^d(\varphi_j)/\varphi_j = \sum_i \tau_{ij} q_{ij}^d(\varphi_j)/\varphi_j = \rho^\sigma S_j^r \varphi_j^{\sigma-1} \underline{w}^{-\sigma}$ through substitution of the optimal price $p_{ij}^d(\varphi_j)$ into the demand function from Eq. (2), the firm-level revenues $r_j^d(\varphi)$ can be determined by evaluating Eq. (3) at $y_j^d(\varphi_j) = \varphi_j l_j^d(\varphi_j)$. Firm-level employment $l_j^d(\varphi_j)$ thereby is pinned down by the demand side of the labor market, which falls short of the labor supply $h_j^d(\varphi_j) = S_j^h \left[\psi_j(\varphi_j)\underline{w}\right]^\varepsilon$.

The hiring rate for demand-constrained firms is defined as $\psi_j^d(\varphi_j) = l_j^d(\varphi_j)/h_j^d(\varphi_j)$. Substituting employment $l_j^d(\varphi_j)$ and labour supply $h_j^d(\varphi_j)$, evaluated at the minimum wage $\underline{w}$, allows us to solve for

$$\psi_j^d(\varphi_j) = \rho^{\frac{\sigma}{\varepsilon+1}} (S_j^r)^{\frac{1}{\varepsilon+1}} (S_j^h)^{-\frac{1}{\varepsilon+1}} \varphi_j^{\frac{\sigma-1}{\varepsilon+1}} \underline{w}^{-\frac{\sigma+\varepsilon}{\varepsilon+1}}, \tag{A.3}$$

$$h_j^d(\varphi_j) = \rho^{\frac{\sigma\varepsilon}{\varepsilon+1}} (S_j^r)^{\frac{\varepsilon}{\varepsilon+1}} (S_j^h)^{\frac{1}{\varepsilon+1}} \varphi_j^{\frac{(\sigma-1)\varepsilon}{\varepsilon+1}} \underline{w}^{-\frac{(\sigma-1)\varepsilon}{\varepsilon+1}}. \tag{A.4}$$

A.3 Aggregation

In this appendix section, we derive aggregate employment L_j, aggregate labor supply H_j, aggregate revenues R_j, and aggregate profits Π_j as well as the price index P_j and the wage index W_j. To this end, we claim that firm-level productivity φ_j follows a Pareto distribution with shape parameter $k > 0$ and lower bound $\underline{\varphi}_j > 0$. The results of the aggregation process thereby can be summarized as

$$X_j = \chi_X \Phi_j^X(\underline{w}) M_j x_j^u(\underline{\varphi}_j), \tag{A.5}$$

in which $X_j \in \{L_j, H_j, R_j, \Pi_j\}$ serves as a placeholder for the respective aggregate outcomes, whereas $x_j^u(\underline{\varphi}_j) \in \{l_j^u(\underline{\varphi}_j), h_j^u(\underline{\varphi}_j), r_j^u(\underline{\varphi}_j), \pi_j^u(\underline{\varphi}_j)\}$ is a substitute for the respective firm-level variable of an unconstrained firm evaluated at the lower-bound productivity $\underline{\varphi}_j$. Aggregate outcomes X_j are proportional to the respective firm-level variables $x_j^u(\underline{\varphi}_j)$ with the factor of proportionality depending on the number of firms $M_j > 0$, a constant $\chi_X \geq 1$, that converges to $\chi_X = 1$ in a scenario with homogeneous firms (i.e. for $k \to \infty$), and a multiplier $\Phi_j^X(\underline{w}) > 0$, that captures the effect of a binding minimum wage $\underline{w}$ on location j's aggregate outcomes and which takes a value of $\Phi_j^X(\underline{w}) = 1$ if the minimum wage $\underline{w}$ is non-binding.

To compute the aggregate outcomes of our model for each location j as a function of the minimum wage $\underline{w}$ we can use the fact that

$$\frac{\underline{\varphi}_j^z(\underline{w})}{\underline{\varphi}_j} = \left(\frac{\underline{w}}{\underline{w}_j^z}\right)^{\frac{\sigma+\varepsilon}{\sigma-1}} \quad \forall\ z \in \{s, u\}. \tag{A.6}$$

Eq. (A.6) relates the critical productivity levels $\underline{\varphi}_i^z(\underline{w})\ \forall\ z \in \{s, u\}$ from Eqs. (5) and (6) (normalized by the lower bound of the productivity distribution $\underline{\varphi}_j$) to the minimum wage $\underline{w}$ (normalized by the critical minimum wage level $\underline{w}_j^z\ \forall\ z \in \{s, u\}$).

Aggregate employment in location j is defined as

$$\begin{aligned} L_j = M_j \Bigg\{ & l_j^d(\underline{\varphi}_j) \int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}} \frac{l_j^d(\varphi_j)}{l_j^d(\underline{\varphi}_j)} \frac{dG(\varphi_j)}{1 - G(\underline{\varphi}_j)} \\ & + l_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}) \frac{1 - G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})}{1 - G(\underline{\varphi}_j)} \\ & \times \int_{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}} \frac{l_j^s(\varphi_j)}{l_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1 - G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \\ & + l_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}) \frac{1 - G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})}{1 - G(\underline{\varphi}_j)} \\ & \times \int_{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}}^{\infty} \frac{l_j^u(\varphi_j)}{l_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1 - G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \Bigg\}. \end{aligned} \tag{A.7}$$

Using firm-level outcomes, Eq. (A.6), and Eq. (7) we can solve for L_j as defined by Eq. (A.5), with $\chi_L \equiv k/\{k-[\varepsilon/(\varepsilon+1)]\gamma\}$ and

$$\begin{aligned}
\Phi_j^L(\underline{w}) \equiv & \frac{l_j^d(\underline{\varphi}_j)}{l_j^u(\underline{\varphi}_j)} \frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k-(\sigma-1)} \left\{ 1 - \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}} \right\} \\
& + \frac{l_j^s(\underline{\varphi}_j)}{l_j^u(\underline{\varphi}_j)} \frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k} \left[\left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right. \\
& \left. - \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right] + \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)]\gamma\}(\sigma+\varepsilon)}{\sigma-1}}, \\
= & \left(\frac{\rho}{\eta} \frac{\underline{w}_j^u}{\underline{w}} \right)^{\sigma} \frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k-(\sigma-1)} \left\{ 1 - \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}} \right\} \\
& + \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{-\varepsilon} \frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k} \left[\left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right. \\
& \left. - \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right] + \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)]\gamma\}(\sigma+\varepsilon)}{\sigma-1}}.
\end{aligned}$$

Aggregate labour supply to location j is defined as

$$\begin{aligned}
H_j = M_j \Bigg\{ & h_j^d(\underline{\varphi}_j) \int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}} \frac{h_j^d(\varphi_j)}{h_j^d(\underline{\varphi}_j)} \frac{dG(\varphi_j)}{1-G(\underline{\varphi}_j)} \\
& + h_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}} \frac{h_j^s(\varphi_j)}{h_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \\
& + h_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}}^{\infty} \frac{h_j^u(\varphi_j)}{h_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \Bigg\}.
\end{aligned} \tag{A.8}$$

Using $h_j^u(\varphi_j) = h_j^s(\varphi_j) = 1$ and $h_j^d(\varphi_j)$ from Eq. (A.4) in combination with the Eqs. (A.6) and Eq. (7) allows us to solve for H_j as defined by Eq. (A.5), with $\chi_H \equiv k/\{k-[\varepsilon/(\varepsilon+1)]\gamma\}$

and

$$\Phi_j^H(\underline{w}) \equiv \frac{h_j^d(\underline{\varphi}_j)}{h_j^u(\underline{\varphi}_j)} \frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k-[\varepsilon/(\varepsilon+1)](\sigma-1)} \left\{1-\left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s,\underline{w}\}}\right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)](\sigma-1)\}(\sigma+\varepsilon)}{\sigma-1}}\right\}$$
$$+\frac{h_j^s(\underline{\varphi}_j)}{h_j^u(\underline{\varphi}_j)} \frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k} \left[\left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s,\underline{w}\}}\right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}}\right.$$
$$\left.-\left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u,\underline{w}\}}\right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}}\right]+\left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u,\underline{w}\}}\right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)]\gamma\}(\sigma+\varepsilon)}{\sigma-1}},$$
$$=\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{\varepsilon}{\varepsilon+1}(\sigma-1)}\left(\frac{\rho}{\eta}\right)^{\frac{\varepsilon}{\varepsilon+1}\sigma} \frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k-[\varepsilon/(\varepsilon+1)](\sigma-1)} \left\{1-\left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s,\underline{w}\}}\right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)](\sigma-1)\}(\sigma+\varepsilon)}{\sigma-1}}\right\}$$
$$+\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{-\varepsilon} \frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k} \left[\left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s,\underline{w}\}}\right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}}\right.$$
$$\left.-\left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u,\underline{w}\}}\right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}}\right]+\left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u,\underline{w}\}}\right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)]\gamma\}(\sigma+\varepsilon)}{\sigma-1}}.$$

Aggregate revenues in location j are defined as

$$R_j = M_j\left\{r_j^d(\underline{\varphi}_j)\int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\}} \frac{r_j^d(\varphi_j)}{r_j^d(\underline{\varphi}_j)} \frac{dG(\varphi_j)}{1-G(\underline{\varphi}_j)}\right.$$
$$+r_j^s(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})\frac{1-G(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)}$$
$$\times\int_{\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\}} \frac{r_j^s(\varphi_j)}{r_j^s(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})} \tag{A.9}$$
$$+r_j^u(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})\frac{1-G(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)}$$
$$\left.\times\int_{\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\}}^{\infty} \frac{r_j^u(\varphi_j)}{r_j^u(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})}\right\}.$$

Using firm-level outcomes, Eq. (A.6), and Eq. (7) we can solve for aggregate revenues R_j

as defined by Eq. (A.5), with $\chi_R \equiv k/(k-\gamma)$ and

$$
\begin{aligned}
\Phi_j^R(\underline{w}) \equiv{}& \frac{r_j^d(\underline{\varphi}_j)}{r_j^u(\underline{\varphi}_j)} \frac{k-\gamma}{k-(\sigma-1)} \left\{ 1 - \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}} \right\} \\
&+ \frac{r_j^s(\underline{\varphi}_j)}{r_j^u(\underline{\varphi}_j)} \frac{k-\gamma}{k-(\sigma-1)/\sigma} \left\{ \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)/\sigma](\sigma+\varepsilon)}{\sigma-1}} \right. \\
&\left. - \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)/\sigma](\sigma+\varepsilon)}{\sigma-1}} \right\} + \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}}, \\
={}& \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\sigma-1} \left(\frac{\rho}{\eta} \right)^{\sigma-1} \frac{k-\gamma}{k-(\sigma-1)} \left\{ 1 - \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}} \right\} \\
&+ \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{-\frac{\sigma-1}{\sigma}\varepsilon} \frac{k-\gamma}{k-(\sigma-1)/\sigma} \left\{ \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)/\sigma](\sigma+\varepsilon)}{\sigma-1}} \right. \\
&\left. - \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)/\sigma](\sigma+\varepsilon)}{\sigma-1}} \right\} + \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}}.
\end{aligned}
\tag{A.10}
$$

Aggregate profits can be computed as the difference between aggregate revenues and aggregate costs. For location j, the latter is defined as

$$
\begin{aligned}
C_j = M_j \Bigg\{ & c_j^d(\underline{\varphi}_j) \int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}} \frac{c_j^d(\varphi_j)}{c_j^d(\underline{\varphi}_j)} \frac{dG(\varphi_j)}{1-G(\underline{\varphi}_j)} \\
&+ c_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
&\times \int_{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}} \frac{c_j^s(\varphi_j)}{c_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \\
&+ c_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
&\times \int_{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}}^{\infty} \frac{c_j^u(\varphi_j)}{c_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \Bigg\}.
\end{aligned}
\tag{A.11}
$$

Using firm-level outcomes, Eq. (A.6), and Eq. (7) we can solve for aggregate costs

$C_j = \chi_C \Phi_j^C(\underline{w}) M_j c_j^u(\underline{\varphi}_j)$ with $\chi_C \equiv k/(k-\gamma)$ and

$$
\begin{aligned}
\Phi_j^C(\underline{w}) \equiv & \frac{c_j^d(\underline{\varphi}_j)}{c_j^u(\underline{\varphi}_j)} \frac{k-\gamma}{k-(\sigma-1)} \left\{ 1 - \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}} \right\} \\
& + \frac{c_j^s(\underline{\varphi}_j)}{c_j^u(\underline{\varphi}_j)} \frac{k-\gamma}{k} \left\{ \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} - \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right\} \\
& + \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}}, \\
= & \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\sigma-1} \left(\frac{\rho}{\eta} \right)^{\sigma} \frac{k-\gamma}{k-(\sigma-1)} \left\{ 1 - \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}} \right\} \\
& + \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{-(\varepsilon+1)} \frac{k-\gamma}{k} \left\{ \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right. \\
& \left. - \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right\} + \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}}.
\end{aligned}
\tag{A.12}
$$

Defining $\chi_\Pi \equiv k/(k-\gamma)$ and $\Phi_j^\Pi(\underline{w}) \equiv [\Phi_j^R(\underline{w}) - \eta \Phi_j^C(\underline{w})]/(1-\eta)$ we solve for the aggregate profits $\Pi_j = R_j - C_j$ as defined by Eq. (A.5).

In order to derive the **price index** in Eq. (22) we start out from the definition

$$
\begin{aligned}
P_{ij}^{1-\sigma} = M_j \Bigg\{ & [p_{ij}^d(\underline{\varphi}_j)]^{1-\sigma} \int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}} \left[\frac{p_{ij}^d(\varphi_j)}{p_{ij}^d(\underline{\varphi}_j)} \right]^{1-\sigma} \frac{dG(\varphi_j)}{1-G(\underline{\varphi}_j)} \\
& + [p_{ij}^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})]^{1-\sigma} \frac{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}} \left[\frac{p_{ij}^s(\varphi_j)}{p_{ij}^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \right]^{1-\sigma} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \\
& + [p_{ij}^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})]^{1-\sigma} \frac{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}}^{\infty} \left[\frac{p_{ij}^u(\varphi_j)}{p_{ij}^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \right]^{1-\sigma} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \Bigg\}.
\end{aligned}
\tag{A.13}
$$

Using firm-level outcomes, Eq. (A.6), and Eq. (7) we can solve for P_{ij} from Eq. (22), in which $\chi_P = \chi_R = k/(k-\gamma)$ and $\Phi_j^P(\underline{w}) = \Phi_j^R(\underline{w})$ with $\Phi_j^R(\underline{w})$ from Eq. (A.10). As a consequence, it follows that we have $\Phi_j^P(\underline{w})|_{\underline{w}<\underline{w}_j^u} = 1$ and $d\Phi_j^P(\underline{w})|_{\underline{w}<\underline{w}_j^u}/d\underline{w} = 0$ for $\underline{w} < \underline{w}_j^u$, $d\Phi_j^P(\underline{w})|_{\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s}/d\underline{w} > 0$ for $\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s$, and $d\Phi_j^P(\underline{w})|_{\underline{w}_j^s \leq \underline{w}}/d\underline{w} < 0$ for $\underline{w}_j^s \leq \underline{w}$. Finally, it is easily verified that $\lim_{\underline{w}\to\infty} \Phi_j^P(\underline{w})|_{\underline{w}_j^s \leq \underline{w}} = 0$.

In order to derive the (expected) **wage index** in Eq. (28) we start out from the definition

$$
\begin{aligned}
W_j^{\varepsilon} = M_j \Bigg\{ & [\psi_j^d(\underline{\varphi}_j) w_j^d(\underline{\varphi}_j)]^{\varepsilon} \int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}} \left[\frac{\psi_j^d(\varphi_j)}{\psi_j^d(\underline{\varphi}_j)} \frac{w_j^d(\varphi_j)}{w_j^d(\underline{\varphi}_j)} \right]^{\varepsilon} \frac{dG(\varphi_j)}{1 - G(\underline{\varphi}_j)} \\
& + \underline{w}^{\varepsilon} \frac{1 - G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})}{1 - G(\underline{\varphi}_j)} \int_{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}} \frac{dG(\varphi_j)}{1 - G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \\
& + [w_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})]^{\varepsilon} \frac{1 - G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})}{1 - G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}}^{\infty} \left[\frac{w_j^u(\varphi_j)}{w_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \right]^{\varepsilon} \frac{dG(\varphi_j)}{1 - G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \Bigg\}.
\end{aligned}
$$

Using firm-level outcomes, Eq. (A.6), and Eq. (7) we can solve for W_j from Eq. (22), in which $\chi_W = \chi_H = k/(k-\gamma)$ and $\Phi_j^W(\underline{w}) = \Phi_j^H(\underline{w})$ with $\Phi_j^H(\underline{w})$ from Eq. (A.9) As a consequence, it follows that we have $\Phi_j^W(\underline{w})|_{\underline{w}<\underline{w}_j^u} = 1$ and $d\Phi_j^W(\underline{w})|_{\underline{w}<\underline{w}_j^u}/d\underline{w} = 0$ for $\underline{w} < \underline{w}_j^u$ as well as $d\Phi_j^W(\underline{w})|_{\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s}/d\underline{w} > 0$ for $\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s$. For $\underline{w}_j^s \leq \underline{w}$ we have $d\Phi_j^W(\underline{w})|_{\underline{w}_j^s \leq \underline{w}}/d\underline{w} < 0$ for sufficiently large values of $\underline{w}$. If $\underline{w} > \underline{w}_j^s$ is small, $d\Phi_j^W(\underline{w})|_{\underline{w}_j^s \leq \underline{w}}/d\underline{w}$ can be positive or negative. Finally, it is easily verified that $\lim_{\underline{w}\to\infty} \Phi_j^W(\underline{w})|_{\underline{w}_j^s \leq \underline{w}} = 0$.

A.4 Proof of Proposition 1

In this appendix, we proof the results summarized in Proposition 1, holding the number of firms M_j fixed in partial equilibrium.

Aggregate employment is hump-shaped in $\underline{w}$. For $\underline{w} < \underline{w}_j^u$ we have $\Phi_j^L(\underline{w})|_{\underline{w}<\underline{w}_j^u} = 1$ and $d\Phi_j^L(\underline{w})|_{\underline{w}<\underline{w}_j^u}/d\underline{w} = 0$. For $\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s$ we have

$$
\Phi_j^L(\underline{w})|_{\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s} = \frac{1}{k} \left[\left(k - \frac{\varepsilon}{\varepsilon+1}\gamma \right) \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{-\varepsilon} + \frac{\varepsilon}{\varepsilon+1}\gamma \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)]\gamma\}(\sigma+\varepsilon)}{\sigma-1}} \right],
$$

and it is easily verified that

$$
\frac{d\Phi_j^L(\underline{w})|_{\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s}}{d\underline{w}} = \frac{\varepsilon}{k} \frac{1}{\underline{w}} \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{-\varepsilon} \left(k - \frac{\varepsilon}{\varepsilon+1}\gamma \right) \left[1 - \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right] > 0.
$$

For $\underline{w}_j^s \leq \underline{w}$ we have

$$
\begin{aligned}
\Phi_j^L(\underline{w})|_{\underline{w}_j^s \leq \underline{w}} = & \frac{k - [\varepsilon/(\varepsilon+1)]\gamma}{k - (\sigma-1)} \left(\frac{\rho}{\eta} \right)^{\sigma} \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\sigma} \\
& + \frac{\varepsilon(\sigma-1)}{k(\sigma+\varepsilon)} \left[1 - \frac{k - (\sigma-1) + \frac{\sigma k}{\varepsilon}}{k - (\sigma-1)} \left(\frac{\rho}{\eta} \right)^{\frac{\sigma k}{\sigma-1}} \right] \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)]\gamma\}(\sigma+\varepsilon)}{\sigma-1}}
\end{aligned}
$$

which is increasing in $\underline{w}$ for small values of the minimum wage and decreasing for higher values. Finally, it is easily verified that $\lim_{\underline{w}\to\infty} \Phi_j^L(\underline{w})|_{\underline{w}_j^s \le \underline{w}} = 0$. This completes the proof. ■

2. Aggregate labor supply is hump-shaped in $\underline{w}$. For $\underline{w} < \underline{w}_j^u$ we have $\Phi_j^H(\underline{w})|_{\underline{w}<\underline{w}_j^u} = 1$ and $d\Phi_j^H(\underline{w})|_{\underline{w}<\underline{w}_j^u}/d\underline{w} = 0$. For $\underline{w}_j^u \le \underline{w} < \underline{w}_j^s$ we have

$$\Phi_j^H(\underline{w})|_{\underline{w}_j^u \le \underline{w} < \underline{w}_j^s} = \frac{1}{k}\left[\left(k - \frac{\varepsilon}{\varepsilon+1}\gamma\right)\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{-\varepsilon} + \frac{\varepsilon}{\varepsilon+1}\gamma\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)]\gamma\}(\sigma+\varepsilon)}{\sigma-1}}\right],$$

and it is easily verified that

$$\frac{d\Phi_j^H(\underline{w})|_{\underline{w}_j^u \le \underline{w} < \underline{w}_j^s}}{d\underline{w}} = \frac{\varepsilon}{k}\frac{1}{\underline{w}}\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{-\varepsilon}\left(k - \frac{\varepsilon}{\varepsilon+1}\gamma\right)\left[1 - \left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}}\right] > 0.$$

For $\underline{w}_j^s \le \underline{w}$ we have

$$\begin{aligned}\Phi_j^H(\underline{w})|_{\underline{w}_j^s \le \underline{w}} = {} & \left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{(\sigma-1)\varepsilon}{\varepsilon+1}}\Bigg\{\left(\frac{\rho}{\eta}\right)^{\frac{\sigma\varepsilon}{\varepsilon+1}}\frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k-[\varepsilon/(\varepsilon+1)](\sigma-1)} \\ & - \frac{[\varepsilon/(\varepsilon+1)]\gamma}{k}\left\{\left[1+\frac{k[(\sigma-1)/(\varepsilon+1)]}{k-[\varepsilon/(\varepsilon+1)]\gamma}\right]\left(\frac{\rho}{\eta}\right)^{\frac{k\sigma}{\sigma-1}} - 1\right\}\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)](\sigma-1)\}(\sigma+\varepsilon)}{\sigma-1}}\Bigg\},\end{aligned}$$

and it is straightforward to show that

$$\begin{aligned}\frac{d\Phi_j^H(\underline{w})|_{\underline{w}_j^s \le \underline{w}}}{d\underline{w}} = {} & -\frac{\varepsilon}{\underline{w}}\frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k}\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{(\sigma-1)\varepsilon}{\varepsilon+1}}\Bigg[\frac{k}{k-[\varepsilon/(\varepsilon+1)](\sigma-1)}\frac{\sigma-1}{\varepsilon+1}\left(\frac{\rho}{\eta}\right)^{\frac{\sigma\varepsilon}{\varepsilon+1}} \\ & - \left\{\left[1+\frac{k[(\sigma-1)/(\varepsilon+1)]}{k-[\varepsilon/(\varepsilon+1)]\gamma}\right]\left(\frac{\rho}{\eta}\right)^{\frac{k\sigma}{\sigma-1}} - 1\right\}\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{\{k-[\varepsilon/(\varepsilon+1)](\sigma-1)\}(\sigma+\varepsilon)}{\sigma-1}}\Bigg].\end{aligned}$$

By inspection of $d\Phi_j^H(\underline{w})|_{\underline{w}_j^s \le \underline{w}}/d\underline{w}$ it is easily verified that $d\Phi_j^H(\underline{w})|_{\underline{w}_j^s \le \underline{w}}/d\underline{w} < 0$ for large values of $\underline{w} > \underline{w}_j^s$. To show that $d\Phi_j^H(\underline{w})|_{\underline{w}_j^s \le \underline{w}}/d\underline{w} > 0$ is a possible outcome for small values of $\underline{w} > \underline{w}_j^s$ we evaluate $d\Phi_j^H(\underline{w})|_{\underline{w}_j^s \le \underline{w}}/d\underline{w}$ at $\underline{w}_j^s$

$$\begin{aligned}\left.\frac{d\Phi_j^H(\underline{w})|_{\underline{w}_j^s \le \underline{w}}}{d\underline{w}}\right|_{\underline{w}=\underline{w}_j^s} = {} & -\frac{\varepsilon}{\underline{w}_j^s}\frac{k-[\varepsilon/(\varepsilon+1)]\gamma}{k}\left(\frac{\rho}{\eta}\right)^{\frac{\sigma\varepsilon}{\sigma+\varepsilon}\left(1-\frac{\sigma-1}{\varepsilon+1}\right)} \\ & \times\left[\frac{\sigma-1}{\varepsilon+1}\left\{\frac{k}{k-[\varepsilon/(\varepsilon+1)](\sigma-1)} - \frac{k}{k-[\varepsilon/(\varepsilon+1)]\gamma}\right\} + \left(\frac{\rho}{\eta}\right)^{-\frac{k\sigma}{\sigma-1}} - 1\right],\end{aligned}$$

and note that

$$\lim_{k\to\infty}\left.\frac{d\Phi_j^H(\underline{w})|_{\underline{w}_j^s \le \underline{w}}}{d\underline{w}}\right|_{\underline{w}=\underline{w}_j^s} = \frac{\varepsilon}{\underline{w}_j^s}\left(\frac{\rho}{\eta}\right)^{\frac{\sigma\varepsilon}{\sigma+\varepsilon}\left(1-\frac{\sigma-1}{\varepsilon+1}\right)} > 0.$$

Finally, it is easily verified that $\lim_{\underline{w}\to\infty} \Phi_j^H(\underline{w})|_{\underline{w}_j^s \leq \underline{w}} = 0$. This completes the proof. ■

3. Aggregate revenues are hump-shaped in $\underline{w}$. For $\underline{w} < \underline{w}_j^u$ we have $\Phi_j^R(\underline{w})|_{\underline{w}<\underline{w}_j^u} = 1$ and $d\Phi_j^R(\underline{w})|_{\underline{w}<\underline{w}_j^u}/d\underline{w} = 0$. For $\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s$ we have

$$\Phi_j^R(\underline{w})|_{\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s} = \frac{1}{k-(\sigma-1)/\sigma} \left[(k-\gamma) \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{-\frac{\sigma-1}{\sigma}\varepsilon} + \frac{\sigma-1}{\sigma} \frac{\varepsilon}{\varepsilon+1} \gamma \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}} \right],$$

and it is easily verified that

$$\frac{d\Phi_j^R(\underline{w})|_{\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s}}{d\underline{w}} = \frac{\varepsilon(\sigma-1)/\sigma}{k-(\sigma-1)\sigma} \frac{1}{\underline{w}} \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{-\frac{\sigma-1}{\sigma}\varepsilon} (k-\gamma) \left[1 - \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\frac{\{k-[(\sigma-1)/\sigma]\}(\sigma+\varepsilon)}{\sigma-1}} \right] > 0.$$

For $\underline{w}_j^s \leq \underline{w}$ we have

$$\begin{aligned}\Phi_j^R(\underline{w})|_{\underline{w}_j^s \leq \underline{w}} = \frac{k-\gamma}{k-(\sigma-1)} \frac{\eta}{\rho} \Bigg[& \left(\frac{\rho}{\eta} \right)^{\sigma} \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\sigma-1} \\ & + \left\{ \frac{k-(\sigma-1)}{k-(\sigma-1)/\sigma} \left[\left(\frac{\rho}{\eta} \right)^{\frac{k\sigma}{\sigma-1}} + \frac{\sigma-1}{\sigma} \frac{\gamma}{k-\gamma} \right] - \left(\frac{\rho}{\eta} \right)^{\frac{k}{\sigma-1}} \right\} \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}} \Bigg]\end{aligned}$$

and it is straightforward to show that

$$\begin{aligned}\frac{d\Phi_j^R(\underline{w})|_{\underline{w}_j^s \leq \underline{w}}}{d\underline{w}} = -\frac{(k-\gamma)(\sigma-1)}{k-(\sigma-1)} \frac{\eta}{\rho} \frac{1}{\underline{w}} \Bigg[& \left(\frac{\rho}{\eta} \right)^{\sigma} \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\sigma-1} + \frac{k-\gamma}{\gamma} \frac{\varepsilon+1}{\sigma-1} \\ & \times \left\{ \frac{k-(\sigma-1)}{k-(\sigma-1)/\sigma} \left[\left(\frac{\rho}{\eta} \right)^{\frac{k\sigma}{\sigma-1}} + \frac{\sigma-1}{\sigma} \frac{\gamma}{k-\gamma} \right] - \left(\frac{\rho}{\eta} \right)^{\frac{k}{\sigma-1}} \right\} \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}} \Bigg] < 0.\end{aligned}$$

Finally, it is easily verified that $\lim_{\underline{w}\to\infty} \Phi_j^R(\underline{w})|_{\underline{w}_j^s \leq \underline{w}} = 0$. This completes the proof. ■

4. Aggregate profits are declining in $\underline{w}$. For $\underline{w} < \underline{w}_j^u$ we have $\Phi_j^\Pi(\underline{w})|_{\underline{w}<\underline{w}_j^u} = 1$ and $d\Phi_j^\Pi(\underline{w})|_{\underline{w}<\underline{w}_j^u}/d\underline{w} = 0$. For $\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s$ we have

$$\begin{aligned}\Phi_j^\Pi(\underline{w})|_{\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s} = \frac{1}{1-\eta} \Bigg\{ & \frac{k-\gamma}{k-(\sigma-1)/\sigma} \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{-\frac{(\sigma-1)\varepsilon}{\sigma}} + \eta \frac{\gamma}{k} \frac{(\sigma-1)/\sigma}{k-(\sigma-1)\sigma} \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}} \\ & - \eta \frac{k-\gamma}{k} \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{-(\varepsilon+1)} \Bigg\}\end{aligned}$$

and it is easily verified that

$$\frac{d\Phi_j^{\Pi}(\underline{w})|_{\underline{w}_j^u \le \underline{w} < \underline{w}_j^s}}{d\underline{w}} = \frac{\eta}{1-\eta}\frac{k-\gamma}{k}(\varepsilon+1)\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{-(\varepsilon+1)}\frac{1}{\underline{w}}\Bigg\{\frac{k}{k-(\sigma-1)/\sigma}\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{\sigma+\varepsilon}{\sigma}}$$
$$-\left[1+\frac{(\sigma-1)/\sigma}{k-(\sigma-1)\sigma}\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}}\right]\Bigg\}.$$

Note that $d\Phi_j^{\Pi}(\underline{w})|_{\underline{w}_j^u \le \underline{w} < \underline{w}_j^s} d\underline{w}|_{\underline{w}=\underline{w}_j^u} = 0$ and that

$$\frac{d^2\Phi_j^{\Pi}(\underline{w})|_{\underline{w}_j^u \le \underline{w} < \underline{w}_j^s}}{d\underline{w}^2} = \frac{d\Phi_j^{\Pi}(\underline{w})|_{\underline{w}_j^u \le \underline{w} < \underline{w}_j^s}}{d\underline{w}}\frac{\varepsilon}{\underline{w}} + \frac{\eta}{1-\eta}\frac{k-\gamma}{k}(\varepsilon+1)\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{-(\varepsilon+1)}\frac{1}{\underline{w}}$$
$$\times\frac{k}{k-(\sigma-1)/\sigma}\frac{\sigma+\varepsilon}{\sigma}\frac{1}{\underline{w}}\left[1-\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{-\frac{[k\sigma-(\sigma-1)](\sigma+\varepsilon)}{\sigma(\sigma-1)}}\right]\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{-\frac{k(\sigma+\varepsilon)}{\sigma-1}} < 0,$$

with $d\Phi_j^{\Pi}(\underline{w})|_{\underline{w}_j^u \le \underline{w} < \underline{w}_j^s} d\underline{w}|_{\underline{w}=\underline{w}_j^u} < 0$ following from the second line of the above equation for $\underline{w} > \underline{w}_j^u$. For $\underline{w}_j^s \le \underline{w}$ we have

$$\Phi_j^{\Pi}(\underline{w})|_{\underline{w}_j^s \le \underline{w}} = \left[\frac{k-\gamma}{k-(\sigma-1)}\frac{1-\rho}{1-\eta}\left(\frac{\rho}{\eta}\right)^{\sigma-1} - \Bigg\{\frac{k-\gamma}{k-(\sigma-1)}\frac{1-\rho}{1-\eta}\left(\frac{\rho}{\eta}\right)^{\sigma-1}\right.$$
$$-\frac{k-\gamma}{k-(\sigma-1)/\sigma}\frac{1}{1-\eta}\left[\left(\frac{\rho}{\eta}\right)^{\frac{[k-(\sigma-1)/\sigma]\sigma}{\sigma-1}}-1\right]$$
$$\left.+\frac{k-\gamma}{k}\frac{\eta}{1-\eta}\left[\left(\frac{\rho}{\eta}\right)^{\frac{k\sigma}{\sigma-1}}-1\right]\Bigg\}\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}}\right]\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\sigma-1},$$

and it is straightforward to show that

$$\frac{d\Phi_j^{\Pi}(\underline{w})|_{\underline{w}_j^s \le \underline{w}}}{d\underline{w}} = \left[\frac{(k-\gamma)(\sigma-1)}{k-(\sigma-1)}\frac{1-\rho}{1-\eta}\left(\frac{\rho}{\eta}\right)^{\sigma-1} - \frac{k-(\sigma-1)}{\sigma-1}(\sigma+\varepsilon)\right.$$
$$\times\Bigg\{\frac{k-\gamma}{k-(\sigma-1)}\frac{1-\rho}{1-\eta}\left(\frac{\rho}{\eta}\right)^{\sigma-1} - \frac{k-\gamma}{k-(\sigma-1)/\sigma}\frac{1}{1-\eta}\left[\left(\frac{\rho}{\eta}\right)^{\frac{[k-(\sigma-1)/\sigma]\sigma}{\sigma-1}}-1\right]$$
$$\left.+\frac{k-\gamma}{k}\frac{\eta}{1-\eta}\left[\left(\frac{\rho}{\eta}\right)^{\frac{k\sigma}{\sigma-1}}-1\right]\Bigg\}\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}}\right]\frac{1}{\underline{w}}\left(\frac{\underline{w}_j^u}{\underline{w}}\right)^{\sigma-1}.$$

It is worth noting that $\Phi_j^{\Pi}(\underline{w})|_{\underline{w}_j^s \leq \underline{w}}$ has at most one maximum in $\underline{w} \in (\underline{w}_j^u, \infty)$ at

$$\frac{\underline{w}_j^u}{\underline{w}_{\max}^{\Pi}} = \left[\frac{(k-\gamma)(\sigma-1)}{k-(\sigma-1)} \frac{1-\rho}{1-\eta} \left(\frac{\rho}{\eta}\right)^{\sigma-1} \Bigg/ \frac{k-(\sigma-1)}{\sigma-1}(\sigma+\varepsilon) \right.$$
$$\times \left\{ \frac{k-\gamma}{k-(\sigma-1)} \frac{1-\rho}{1-\eta} \left(\frac{\rho}{\eta}\right)^{\sigma-1} - \frac{k-\gamma}{k-(\sigma-1)/\sigma} \frac{1}{1-\eta} \left[\left(\frac{\rho}{\eta}\right)^{\frac{[k-(\sigma-1)/\sigma]\sigma}{\sigma-1}} - 1 \right] \right.$$
$$\left.\left. + \frac{k-\gamma}{k} \frac{\eta}{1-\eta} \left[\left(\frac{\rho}{\eta}\right)^{\frac{k\sigma}{\sigma-1}} - 1 \right] \right\} \right]^{\frac{\sigma-1}{[k-(\sigma-1)](\sigma+\varepsilon)}}.$$

For $\underline{w}_j^u/\underline{w}_{\max}^{\Pi} > \underline{w}_j^u/\underline{w}_j^s = (\eta/\rho)^{\frac{\sigma}{\sigma+\varepsilon}}$ the maximum is located to the right of the critical value $\underline{w}_j^s$, and we can conclude that $\Phi_j^{\Pi}(\underline{w})|_{\underline{w}_j^s \leq \underline{w}}$ is downward sloping in $\underline{w} \in [\underline{w}_j^s, \infty)$. This completes the proof. ■

Intuition. We have discussed the intuition for the employment effect being hump-shaped in the minimum wage level in Section 3.1.2. As firm-level revenues are an increasing function of the firm's employment level (see Eq. (3)), aggregate revenues also inherit their hump-shaped pattern for all $\underline{w} \geq \underline{w}_j^u$. At a low, but binding minimum wage $\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s$, supply-constrained firms increase labour input, which results in greater output at lower prices. Because prices decrease in quantity at an elasticity $-1/\sigma \geq -1$ (since $\sigma > 1$), the quantity effect dominates the price effect and revenues increase. For the same reason, a reduction in output to raise prices and increase the MRPL to $\underline{w}_j > \underline{w}_j^{\max}$ results in falling revenues for low-productivity demand-constrained firms. Given that firms are profit-maximizing, a binding minimum wage mechanically reduces firm profits. Intuitively, the profit margin $\pi_j^z = \frac{r_j^z - c_j^z}{r_j^z}$ declines from $\pi_j^z = (1-\eta)$ for unconstrained firms via $(1-\rho) < \pi_j^s < (1-\eta)$ for supply-constrained firms to $\pi_j^d = (1-\rho)$ for demand-constrained firms (where $\eta = \rho\frac{\varepsilon}{\varepsilon+1} < \rho$ given that $\varepsilon > 0$). Since a higher $\underline{w}_j$ turns some unconstrained into supply-constrained and supply-constrained into demand-constrained firms, the marginal effect on profits is strictly negative.

B General equilibrium

This section complements Section 4 in the main paper.

B.1 Location choice probabilities

Aggregating $\lambda_{ij}(\varphi_j)$ across all firms φ_j in all workplaces j for a given residence i, we obtain the overall probability λ_i^N that a worker resides in location i.

$$\lambda_i^N = \frac{N_i}{L} = \sum_j \int_{\varphi_j} \lambda_{ij}(\varphi_j) d\varphi_j,$$
$$= \frac{\sum_j B_{ij} \left[\kappa_{ij} \left(P_i^Q\right)^{\alpha} \left(P_i^T\right)^{1-\alpha}\right]^{-\varepsilon} \left[\frac{\eta \Phi_j^W(\underline{w})^{\frac{1}{\varepsilon}} \Phi_j^L(\underline{w})}{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^{\Pi}(\underline{w})}\right]^{\varepsilon} M_j \tilde{w}_j^{\varepsilon}}{\sum_r \sum_s B_{rs} \left[\kappa_{rs} \left(P_r^Q\right)^{\alpha} \left(P_r^T\right)^{1-\alpha}\right]^{-\varepsilon} \left[\frac{\eta \Phi_s^W(\underline{w})^{\frac{1}{\varepsilon}} \Phi_s^L(\underline{w})}{\Phi_s^R(\underline{w}) - (1-\eta)\Phi_s^{\Pi}(\underline{w})}\right]^{\varepsilon} M_s \tilde{w}_s^{\varepsilon}}. \tag{A.14}$$

Aggregating $\lambda_{ij}(\varphi_j)$ over all firms in workplace j and across all residences i, we obtain the overall probability λ_j^H that a worker applies to a firm in location j

$$\lambda_j^H = \frac{H_j}{L} = \sum_i \int_{\varphi_j} \lambda_{ij}(\varphi_j) d\varphi_j,$$
$$= \frac{\sum_i B_{ij} \left[\kappa_{ij} \left(P_i^Q\right)^{\alpha} \left(P_i^T\right)^{1-\alpha}\right]^{-\varepsilon} \left[\frac{\eta \Phi_j^W(\underline{w})^{\frac{1}{\varepsilon}} \Phi_j^L(\underline{w})}{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^{\Pi}(\underline{w})}\right]^{\varepsilon} M_j \tilde{w}_j^{\varepsilon}}{\sum_r \sum_s B_{rs} \left[\kappa_{rs} \left(P_r^Q\right)^{\alpha} \left(P_r^T\right)^{1-\alpha}\right]^{-\varepsilon} \left[\frac{\eta \Phi_s^W(\underline{w})^{\frac{1}{\varepsilon}} \Phi_s^L(\underline{w})}{\Phi_s^R(\underline{w}) - (1-\eta)\Phi_s^{\Pi}(\underline{w})}\right]^{\varepsilon} M_s \tilde{w}_s^{\varepsilon}}. \tag{A.15}$$

B.2 Labour market entry

This section complements Section 4.1.4 in the main paper.

B.2.1 Labor market entry rate, μ

Households decide between entering the labor market (*emp*) and not working (*non*) based on respective (expected) utility levels. We introduce shocks $\exp(a_{i\nu}^o)$ that affect worker utility according to

$$V_{i\nu}^o = V_i^o \exp(a_{i\nu}^o) \tag{A.16}$$

for all options $o \in \{emp, non\}$. The shocks are drawn from a Gumbel distribution with the cdf given by

$$G_i^o(a) = \exp(-A_i^o \exp[-\zeta a - \Gamma]), \tag{A.17}$$

where A_i^o is a region-option-specific average (location parameter), ζ governs the dispersion of shocks and Γ is the Euler-Mascheroni constant.

We refer to μ as the share of the labor force that decides to enter the labor market and search for jobs. It is given by

$$\begin{aligned}\mu &= Pr\left[ln(V_i^{emp}) + a_{i\nu}^{emp} \geq ln(V_i^{non}) + a_{i\nu}^{non}\right] \\ &= Pr\left[ln\left(\frac{V_i^{emp}}{V_i^{non}}\right) + a_{i\nu}^{emp} \geq a_{i\nu}^{non}\right].\end{aligned}$$

Using the probability density function

$$g_i^o = \zeta A_i^o \exp(-\zeta a - A_i^o \exp[-\zeta a])$$

we get

$$\begin{aligned}\mu &= \int_{-\infty}^{\infty} g_i^{emp}(a_{i\nu}) G_i^{non}(a_{i\nu}) da_{i\nu}^{emp} \\ &= \int_{-\infty}^{\infty} g_i^{emp}(a_{i\nu}) G_i^{non}\left(\ln\left(\frac{V_i^{emp}}{V_i^{non}}\right) + a_{i\nu}^{emp}\right) da_{i\nu}^{emp} \\ &= \int_{-\infty}^{\infty} \zeta A_i^{emp} \exp(-\zeta a_{i\nu}^{emp} - A_i^{emp} \exp\{-\zeta a_{i\nu}^{emp}\}) \\ &\quad \times \exp\left(-A_i^{non} \exp\left(-\zeta \ln\left(\frac{V_i^{emp}}{V_i^{non}}\right) - \zeta a_{i\nu}^{emp}\right)\right) da_{i\nu}^{emp} \\ &= \int_{-\infty}^{\infty} \zeta A_i^{emp} \exp(-\zeta a_{i\nu}^{emp}) \\ &\quad \times \exp\left(-\sum_o A_i^o \exp\left(-\zeta ln\left(\frac{V_i^{emp}}{V_i^o}\right) - \zeta a_{i\nu}^{emp}\right)\right) da_{i\nu}^{emp}\end{aligned}$$

We now define:

$$\begin{aligned}x_1 \equiv& \zeta a_{i\nu}^{emp} \\ x_2 \equiv& \ln\left(\sum_o A_i^o \exp\left(-\zeta \ln\left(\frac{V_i^{emp}}{V_i^o}\right)\right)\right) \\ y \equiv& x_1 - x_2\end{aligned}$$

Substituting these expressions, we obtain

$$
\begin{aligned}
\mu = & \int_{-\infty}^{\infty} \zeta A_i^{emp} \exp(-x_1) \exp(-\exp(x_2)\exp(-x_1)) \frac{1}{\zeta} dx_1 \\
= & \int_{-\infty}^{\infty} A_i^{emp} \exp(-y - x_2) \exp(-\exp(x_2)\exp(-y - x_2)) dy \\
= & A_i^{emp} \exp(-x_2) \int_{-\infty}^{\infty} \exp(-y - \exp(-y)) dy
\end{aligned}
$$

Using the fact that the derivative of $\exp(-\exp(-y))$ is $\exp(-y - \exp(-y))$ we can reformulate the above expression to

$$
\begin{aligned}
\mu = & A_i^{emp} \exp(-x_2) \left[\exp(-\exp(-y))\right]_{-\infty}^{\infty} \\
= & \frac{A_i^{emp} \left(V_i^{emp}\right)^{\zeta}}{\sum_o A_i^o (V_i^o)^{\zeta}}
\end{aligned}
$$

As A_i^o is only identified up to scale, we set $A_i^{emp} \equiv 1$. Further, we normalize the outside utility to $V^{non} = V_i^{non} \equiv 1$ from above to get

$$
\mu = \frac{(V_i^{emp})^{\zeta}}{(V_i^{emp})^{\zeta} + A_i^{non}} \tag{A.18}
$$

The labour supply elasticity can be computed as

$$
\begin{aligned}
\frac{d\mu}{dV}\frac{V}{\mu} = & \frac{\zeta V^{\zeta-1}\left(V^{\zeta} + A^{non}\right) - \zeta V^{\zeta-1} V^{\zeta}}{\left(V^{\zeta} + A^{non}\right)^2} \frac{V}{\mu} \\
= & \zeta \frac{V^{\zeta}(1-\mu)}{V^{\zeta} + A^{non}} \frac{V^{\zeta} + A^{non}}{V^{\zeta}} \\
= & \zeta(1-\mu)
\end{aligned}
$$

B.2.2 Expected utility

Apart from their optimal consumption choices, households decide (i) whether to enter the labor market, (ii) where to live and (iii) where to work. Using equalized utility $\bar{V}$ based on Eq. (34), we now compute the expected utility across entering the labor market and leisure.

Referring to average utility for each option as V^o, we assume that households receive shocks $\exp(a^o)$ that affect their utility as follows:

$$
V^o = \bar{V}^o \exp(a^o), \tag{A.19}
$$

where V^o represents the average utility from entering the labor market or leisure. Assuming shocks to follow an extreme value type-I distribution (Gumbel), we can use the fact that the distribution of V is given by

$$\begin{aligned} G(V) =& \Pi_o \exp\{-A^o \exp(-\zeta \ln(V/\bar{V}^o) - \Gamma)\} \\ =& \Pi_o \exp\{-A^o \exp(\zeta \ln(\bar{V}^o)) \exp(-\Gamma) V^{-\zeta}\} \\ =& \exp\left\{-\sum_o A^o (V^o)^\zeta \exp(-\Gamma) V^{-\zeta}\right\}. \end{aligned}$$

Based on the probability density function, the expected utility results as

$$\begin{aligned} E(V) =& \int_0^\infty V dG(V) \\ =& \int_0^\infty -\zeta \sum_o A^o (V^o)^\zeta \exp(-\Gamma) V^{-\zeta} \\ & \times \exp\left\{-\sum_o A^o (V^o)^\zeta \exp(-\Gamma) V^{-\zeta}\right\} dV \end{aligned} \tag{A.20}$$

Defining the following expressions:

$$\begin{aligned} \Psi =& \sum_o A^o (V^o)^\zeta \\ z =& \Psi V^{-\zeta} \\ dz =& -\zeta V^{-(\zeta+1)} \Psi dV, \end{aligned}$$

we obtain

$$
\begin{aligned}
E(V) = & \int_0^\infty -\zeta z \exp(-z) \exp(-\Gamma) \exp(\exp(-\Gamma)) dV \\
= & \int_0^\infty \frac{-\zeta z \exp(-z) \exp(-\Gamma) \exp(\exp(-\Gamma))}{-\zeta V^{-(\zeta+1)} \Psi} dz \\
= & \int_0^\infty \frac{-\zeta z \exp(-z) \exp(-\Gamma) \exp(\exp(-\Gamma))}{-\zeta V^{-1} z} dz \\
= & \int_0^\infty z^{-\frac{1}{\zeta}} \exp(-z) \Psi^{\frac{1}{\zeta}} \exp(-\Gamma) \exp(\exp(-\Gamma)) dz \\
= & \Psi^{\frac{1}{\zeta}} \int_0^\infty z^{-\frac{1}{z}} \exp(-z) \exp(-\Gamma) \exp(\exp(-\Gamma)) dz \\
= & \Psi^{\frac{1}{\zeta}} \int_0^\infty z^{-\frac{1}{z}} \exp(-z) \exp(-\Gamma) \exp(\exp(-\Gamma)) dz \\
= & \Psi^{\frac{1}{\zeta}} \Gamma \exp(-\Gamma) \exp(\exp(-\Gamma)) \\
= & \Psi^{\frac{1}{\zeta}} \left[\sum_o A^o (V^o)^\zeta \right]^{\frac{1}{\zeta}}
\end{aligned}
$$

B.3 Quantification

This section complements Section 4.2 in the main paper.

B.3.1 Preference heterogeneity (ε)

As discussed in detail in Section A.2, ε governs how, at the firm level, greater productivity φ translates into higher wages $w(\varphi)$ and larger employment $l(\varphi)$. As summarized in Table A1,

$$ w_j^u(\omega) = \eta^{\frac{\sigma}{\sigma+\varepsilon}} \left(S_j^r\right)^{\frac{1}{\sigma+\varepsilon}} \left(S_j^h\right)^{-\frac{1}{\sigma+\varepsilon}} \varphi(\omega)_j^{\frac{\sigma-1}{\sigma+\varepsilon}} \tag{A.21} $$

and

$$ l_j^u(\omega) = \eta^{\frac{\sigma\varepsilon}{\sigma+\varepsilon}} \left(S_j^r\right)^{\frac{\varepsilon}{\sigma+\varepsilon}} \left(S_j^h\right)^{\frac{\sigma}{\sigma+\varepsilon}} \varphi(\omega)_j^{\frac{(\sigma-1)\varepsilon}{\sigma+\varepsilon}} \tag{A.22} $$

must hold for unconstrained firms (all firms prior to the minimum-wage introduction) in equilibrium. Solving Eq. (A.22) for productivity φ_j and using it in Eq. (A.21) delivers an equilibrium relationship between $w(\omega)$ and $l(\omega)$ which is governed by ε:

$$ w_j(\omega) = l_j(\omega)^{\frac{1}{\varepsilon}} \mu_j^e, $$

where μ_j^e absorbs the effects of the region-specific demand and supply shifters $\{S_j^r, S_j^h\}$ as well as the general equilibrium constant. Taking logs, we obtain the empirical reduced-form relationship

$$ \ln w_{\omega,j,t} = \tilde{\varepsilon} \ln l_{\omega,j,t} + \tilde{\mu}_{j,t}^e + \epsilon_{\omega,j,t}^e, \tag{A.23} $$

where $\tilde{\varepsilon} \equiv 1/\varepsilon$ and $\tilde{\mu}^e_j \equiv \ln \mu^e_j$. We use subscript ω to index variation across establishments and add a subscript t since establishments are observed at different points in time. $\epsilon^e_{\omega,j,t}$ is an error term whose nature critically determines the estimation strategy. For one thing, we expect $\epsilon^e_{\omega,j,t}$ to capture measurement error in hourly wages since we observe wages, but impute hours worked (see B.2.1). Since this measurement error is plausibly uncorrelated with establishment employment, we obtain a theory-consistent estimate of $\tilde{\varepsilon}$ using OLS after controlling for municipality-time effects that capture all labour demand and supply shocks emphasized by the model. For another, $\epsilon^e_{\omega,j}$ may capture establishment-time-specific shocks to labour demand and supply from which we abstract in the model (all shocks are region-time specific). Such shocks impose a threat to identification since we wish to estimate the (inverse) labour supply elasticity solely from variation in labour demand. To address potentially correlated establishment-level supply shocks, we require an instrumental variable for labour demand. We use a shift-share approach that has a long tradition in the literature (Bartik, 1991; Severen, 2021).

$$\hat{l}_{\omega,s,t} = \frac{l_{\omega,s,t=\bar{t}}}{\sum_{\omega} l_{\omega,s,t=\bar{t}}} \times \sum_{\omega} l_{\omega,s,t}, \tag{A.24}$$

where $l_{\omega,s,t=\bar{t}}$ is the employment of an establishment ω in sector s in an initial year $t = \bar{t}$. We use the 88 2-digit sectors as defined by the *Klassifikation der Wirtschaftszweige, Ausgabe 2008*. Intuitively, we use Eq. (A.24) to predict an establishment-time employment measure $\hat{l}_{\omega,s,t}$ combining the share of an establishment at national employment in a given sector (the share component) with the national employment trend in this sector (the shift component). We then use $\ln \hat{l}_{\omega,s,t}$ as an instrument for $\ln l_{\omega,s,t}$ in an expanded version of Eq. (A.23) that adds establishment fixed effects $\tilde{\eta_\omega}$.

$$\ln w_{\omega,j,t} = \tilde{\varepsilon} \ln l_{\omega,j,t} + \tilde{\mu}^e_{j,t} + \tilde{\eta_\omega} + \epsilon^e_{\omega,j,t} \tag{A.25}$$

With this approach, we estimate $\tilde{\varepsilon}$ from within-municipality variation over time that is generated from national sector employment trends. The conventional identifying assumption is that these reflect changes in aggregate labour demand. Notice that unlike in the IV specification in Eq. (A.25), we can use cross-sectional and temporal variation under the assumptions made in the OLS-specification in Eq. (A.23), which is why the latter excludes establishment fixed effects.

We present our estimates in Table A2. The OLS and IV approaches deliver estimates that are within close range. This is, perhaps, not surprising given that many labour supply shocks are already absorbed by municipality-year effects. We find that wages scale in firm-level employment at an elasticity of slightly below 0.2. Our preferred IV estimate of $\varepsilon = 5.5$ (IV) lies between the value of 3.3 estimated by Monte et al. (2018) and the value of 6.7 estimated by Ahlfeldt et al. (2015). It is worth noting that the size of the spatial units we use lies between those in Monte et al. (2018) (counties) and Ahlfeldt et al. (2015) (housing blocks). It is intuitive, that the dispersion of tastes for places increases in the size of the

Table A2: Preference heterogeneity

	(1)	(2)
	OLS	2SLS
Log employment	0.1471***	0.1811***
	(0.0001)	(0.0194)
Establishment effects	No	Yes
Region-Year effects	Yes	Yes
Observations	9,981,996	9,460,017
R^2	0.166	-
Preference heterogeneity (ϵ)	6.8	5.52
	(0.0063)	(0.59)

Notes: Dpendent variable is log wage. Unit of observation is establishment-year. The estimate of ε is defined as $1/\hat{\hat{\varepsilon}}$. Robust standard errors in parentheses. Instrument for log employment in (2) is log employment predicted by a shift-share approach where the shift component is the share of an establishment's employment at national employment in a given sector and the shit component is the national employment trend in that sector. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

considered spatial units. Our estimate implies that workers earn $\varepsilon/(\varepsilon+1) = 85\%$ of their MRPL, which is in the middle of the range of extant estimates (Sokolova and Sorensen, 2020; Yeh et al., 2022).

B.3.2 Productivity heterogeneity (k)

To estimate k_j, which monitors the within-regional distribution of firm productivity, we exploit that we can observe the distribution of worker wages in our micro data. While we provide a novel micro-economic foundation for our estimation approach in the context of our model, the empirical approach is related to a literature that has fitted Pareto distributions of firm productivities (Arkolakis, 2010; Egger et al., 2013).

To derive the estimation equation, we compute the share of employment, $\mathcal{S}_j^b$, with wages lower than a particular threshold, w_j^b. This is helpful because firm-level employment is a function of firm productivity. Using firm-level employment of unconstrained firms from

Table A1, $l_j^u(\varphi_j)$, delivers:

$$\mathcal{S}_j^b = 1 - \frac{\int_{\varphi_j^b}^{\infty} l_j^u(\varphi_j) dG(\varphi_j)}{\int_{\underline{\varphi}_j}^{\infty} l_j^u(\varphi_j) dG(\varphi_j)} = 1 - \frac{\int_{\varphi_j^b}^{\infty} \varphi_j^{\frac{(\sigma-1)\varepsilon}{\sigma+\varepsilon}} k_j \varphi_j^{-k_j-1} \underline{\varphi}_j^{k_j} d\varphi_j}{\int_{\underline{\varphi}_j}^{\infty} \varphi_j^{\frac{(\sigma-1)\varepsilon}{\sigma+\varepsilon}} k_j \varphi_j^{-k_j-1} \underline{\varphi}_j^{k_j} d\varphi_j} \tag{A.26}$$

$$= 1 - \frac{\left[-\frac{\sigma+\varepsilon}{k_j(\sigma+\varepsilon)-(\sigma-1)\varepsilon} \varphi_j^{-\frac{k_j(\sigma+\varepsilon)-(\sigma-1)\varepsilon}{\sigma+\varepsilon}} \right]_{\varphi_j^b}^{\infty}}{\left[-\frac{\sigma+\varepsilon}{k_j(\sigma+\varepsilon)-(\sigma-1)\varepsilon} \varphi_j^{-\frac{k_j(\sigma+\varepsilon)-(\sigma-1)\varepsilon}{\sigma+\varepsilon}} \right]_{\underline{\varphi}_j}^{\infty}}$$

$$= 1 - \left(\frac{\underline{\varphi}_j}{\varphi_j^b} \right)^{\frac{k_j(\sigma+\varepsilon)-(\sigma-1)\varepsilon}{\sigma+\varepsilon}}$$

Substituting $\underline{\varphi}_j$ and φ_j^b using Eq. (5) and the formular for average wages, Eq. (19), we obtain:

$$\mathcal{S}_j^b = 1 - \left(\frac{w_j(\underline{\varphi}_j)}{w_j^b(\varphi_j^b)} \right)^{\frac{k_j(\sigma+\varepsilon)-(\sigma-1)\varepsilon}{\sigma-1}}$$

$$= 1 - \left(\frac{k_j(\sigma+\varepsilon) - (\varepsilon+1)(\sigma-1)}{k_j(\sigma+\varepsilon) - \varepsilon(\sigma-1)} \frac{\tilde{w}_j}{w_j^b(\varphi_j^b)} \right)^{\frac{k_j(\sigma+\varepsilon)-(\sigma-1)\varepsilon}{\sigma-1}} \tag{A.27}$$

Our data allows us to observe the share of workers earning less than w^b in area j, $\tilde{S}_j^b$. We assume that our empirically observed $\tilde{S}_j^b$ is a good proxy for S_j^b, subject to a zero-mean random shock, e_j^b, that originates from forces outside our model.

$$\tilde{S}_j^b = S_j^b - e_j^b \tag{A.28}$$

Making the identifying assumption that these shocks are uncorrelated with the wage level,

$$\mathbb{E}\left(w^b e_j^b\right) = 0, \tag{A.29}$$

we can derive J moment conditions (for each area):

$$\mathbb{E}\left(w^b \left[1 - \tilde{S}_j^b - \left(\frac{k_j(\sigma+\varepsilon) - (\varepsilon+1)(\sigma-1)}{k_j(\sigma+\varepsilon) - \varepsilon(\sigma-1)} \frac{\tilde{w}_j}{w_j^b(\varphi_j^b)} \right)^{\frac{k_j(\sigma+\varepsilon)-(\sigma-1)\varepsilon}{\sigma-1}} \right] \right) = 0 \tag{A.30}$$

Note that our choice of k_j determines the dispersion of wages—via the exponent—as well as the lower-bound wage within an area j since $w_j(\underline{\varphi}_j) = \frac{k_j(\sigma+\varepsilon)-(\varepsilon+1)(\sigma-1)}{k_j(\sigma+\varepsilon)-\varepsilon(\sigma-1)} \tilde{w}_j$. Therefore, it is important to impose the full parametric structure when identifying k_j. Intuitively, a larger value of k_j, conditional on given values of $\{\varepsilon, \sigma\}$ and an observed average wage $\tilde{w}_j$, implies that the lower-bound wage $w_j(\underline{\varphi}_j)$ is higher and the distribution across workers is more dispersed (there is more inequality).

Note further that the choice of parameter values for $\{k, \varepsilon, \sigma\}$ is subject to the following constraints that follow from the aggregation of firm-level outcomes described in Section A.3:

$$\begin{aligned} k &> \sigma - 1 \\ k &> \frac{(-1)\varepsilon}{\sigma + \varepsilon} \\ k &> \frac{(\sigma - 1)\varepsilon}{\varepsilon + 1} \\ k &> \frac{\sigma - 1}{\sigma} \\ k &> \frac{(\sigma - 1)(\varepsilon + 1)}{\varepsilon + \sigma} \end{aligned} \tag{A.31}$$

Therefore, we set $\varepsilon = 5.5$ to the value estimated in Section B.3.1 and nest a GMM estimation of k using the moment condition in Eq. (A.30) into a grid search for a theory-consistent parameter value for σ. In particular, we start from a canonical parameter value $\sigma = 4$ and gradually reduce σ until we obtain an estimate of k that satisfies all parameter constraints. Since the left tail of the distribution is particularly relevant to us, we weigh observations in Eq. (A.30) using the binary weights returned by the indicator function $\mathbb{1}[\underline{b} \leq w_j^b \leq \bar{b}]$. We choose $\underline{b} = 7$ and $\bar{b} = 14$ as these appear like generous bounds of minimum wages to be considered by policy.

This procedure identifies a Pareto firm productivity shape parameter of $k = 0.90$ and an elasticity of substitution of $\sigma = 1.9$. Simonovska and Waugh (2014) report a typical range for σ from 2.79 to 4.46. However, our σ captures the elasticity of domestic trade rather than international trade. We present our GMM estimates of k for varying σ values in Table A3. These values are smaller than than those typically found in the trade literature. Egger et al. (2013) report a range from 4 to 6 for 4. However, unlike them, we focus on the left tail of the distribution.

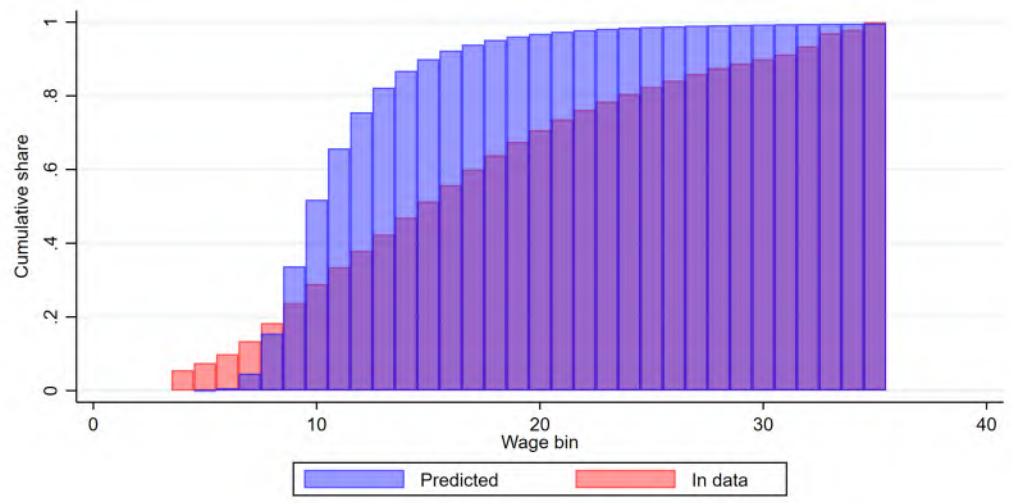
Table A3: Estimation of firm productivity distribution parameter k

	$\sigma = 1.9$	$\sigma = 3$	$\sigma = 4.5$	$\sigma = 6$
k	0.901***	1.7366***	1.7837***	3.1785***
	(0.0002)	(0.0004)	(0.0006)	(0.0008)
Observations	39,789	39,789	39,789	39,789

Notes: Unit of observation are the municipality group-specific shares of workers whose hourly wages are below specified thresholds given by 1-Euro bins in the range of 7 and 14 Euro per hour. Estimation by GMM. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In Figure A1, we compare the cumulative distribution the model generates at the national level to the distribution in the data. For our purposes, the important feature is that the model matches the minimum wage bite (i.e. the share of workers earning less than the minimum wage of €8.50) fairly well.

Figure A1: Cumulative wage distribution in model and data



Note: Cumulative wage distributions at the national level. Model-based distribution generated by employment-weighted aggregation of area distributions defined by Eq. (A.27).

B.3.3 Trade cost (τ_{ij})

We parameterize trade costs as a negative exponential function of the bilateral straight-line distance $DIST$ and an inner-German border effect:

$$\tau_{ij} = \exp\left(b_i^T DIST_{ij} + d^{T,EW} D_{ij}^{T,EW} + d^{T,WE} D_{ij}^{T,WE}\right), \tag{A.32}$$

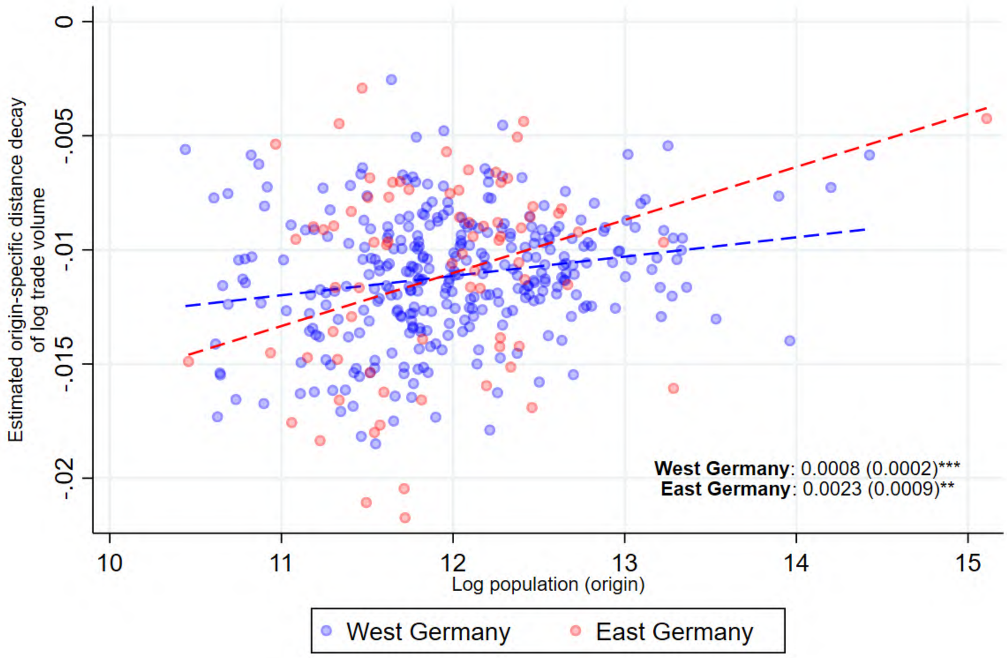
where $D^{T,EW}$ takes the value of one if i refers to a region in East Germany and j to a region in West Germany, while $D^{T,WE}$ takes the value of one for routes starting in West and ending in East Germany. Note that the distance effect on trade cost b_i^T is origin-specific. This allows some regions to export more locally than others, for example because they specialize on perishable products, and accounts for the centrality bias in inter-city trade (Mori and Wrona, 2021). Following conventions in the trade literature, we set the internal distance to $DIST_{ij=i} = \frac{1}{6}\sqrt{A_i/\pi}$, where A_i is the geographic area of i (Combes et al., 2005).

Using Eq. (A.32) in Eq. (21), we can derive a gravity equation of trade:

$$\ln(F_{lk}) = c^T + O_l^T + D_k^T + \tilde{b}_l^T DIST_{lk} + \tilde{d}^{T,EW} D_{lk}^{T,EW} + \tilde{d}^{T,WE} D_{lk}^{T,WE} + e_{lk}^T, \tag{A.33}$$

where we have described the trade share $\theta_{lk} = F_{lk} \exp^{(c^T + e_{lk}^T)}$ as a function of empirically observed trade flows F_{lk} between counties l and k, a stochastic zero-mean error term e_{lk}^T capturing measurement error, and a scaling constant c^T. $\{O_l^T, D_k^T\}$ capture all origin and destination effects. Moreover, we account for the possibility that for historical reasons the size of trade volumes may also still depend on whether routes cross the former inner-German border and in which direction they do so. For this reason we include the indicator variables $D^{T,EW}$ and $D^{T,WE}$. These variables capture the average difference in the size of trade volumes for routes that cross the former inner-German border (from East to West and from West to East, respectively) relative to routes between counties that are both in West or in East Germany. Estimation of Eq. (A.33) yields a reduced-form estimate of the average distance elasticity of $\frac{1}{L}\sum_l \tilde{b}_l^T = \frac{1}{L}\sum_l b_l^T(1-\sigma) = -0.01$. Compared to routes within East or West Germany, trade volumes are on average 54% (= (exp(−0.7872) − 1) ∗ 100%) smaller on routes that start in East and end in West Germany, while there is no statistically significant difference for routes running from West to East Germany (see column (3) in Table A4). Figure A2 illustrates the variation in the estimated origin-specific distance elasticities. On average, trade volumes are predicted to fall more slowly over distance for larger than for smaller origin counties, which is consistent with the centrality bias in inter-city trade (Mori and Wrona, 2021).

Figure A2: Estimated distance elasticity of trade volumes



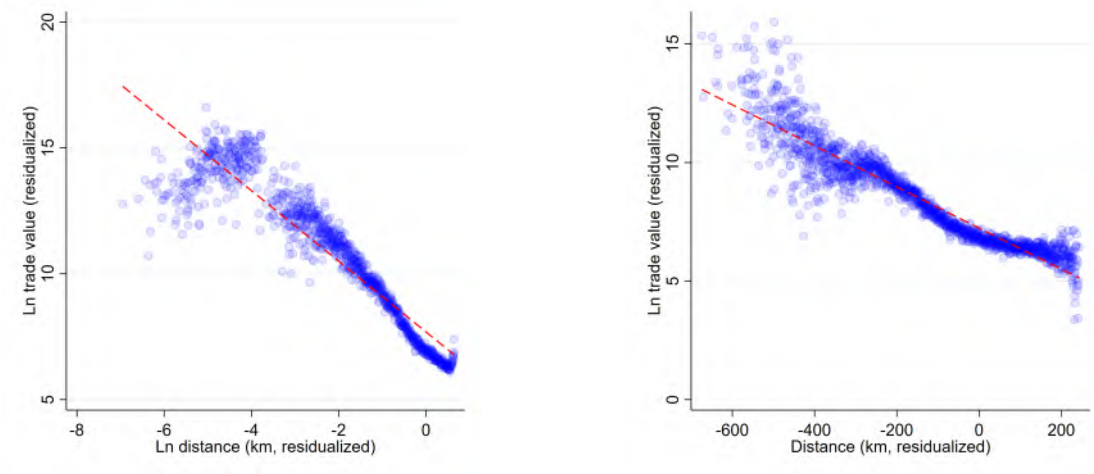
Note: Unit of observation is the county level. The figure plots the estimated origin-specific distance elasticities of trade volumes (given by $\tilde{b_l}^T$ in Eq. (A.33)) against log origin population size at the county level separately for East and West Germany. The dashed lines show the linear fit between the two variables. * p < 0.1, ** p < 0.05, *** p < 0.01

From these reduced-form estimates, we recover our measure of bilateral trade cost as

$$\tau_{ij} = \exp\left(\frac{\tilde{b}^T_{i(l)}}{1-\sigma}DIST_{ij} + \frac{\tilde{d}^{T,EW}}{(1-\sigma)}D^{T,EW}_{ij} + \frac{\tilde{d}^{T,WE}}{(1-\sigma)}D^{T,WE}_{ij}\right).$$

Figure A3 substantiates the choice of the negative exponential function as a reasonable approximation for the true functional relationship in our empirical setting. Moreover, a convenient property of the negative exponential form is that τ_{ij} takes a unit value by default at a zero distance, allowing for a straightforward interpretation as iceberg trade costs. At the mean bilateral distance, our implied estimates of the distance elasticity are -3.38 (trade volumes) and 6.22 (trade costs).

Figure A3: Trade Gravity



(a) Log-log

(b) Semi-log

Note: Units of observation are county-county pairs. All variables are residualised in regressions against origin fixed effects, destination fixed effects, an indicator for whether counties are in different states as well as an indicator for whether one county is in East Germany and the other in West Germany. Log trade residuals are averaged within bins: 0.005 log point bins in the left panel and 0.5km bins in the right panel. Averages are computed using the origin population of the county-county pair as a weight. The size of the markers reflect the population size of the origin population.

B.3.4 Fundamental productivity ($\underline{\varphi}$)

The minimum wage was introduced in 2015 in Germany. For $t < 2015$, we have $\underline{w} = 0$ and $\Phi_j^{X \in \{L,H,R,P,W,\Pi\}} = 1$. In this special case, we can use $\tilde{v}_i = \sum_j^J \lambda_{ij|i}^N \tilde{w}_j$ and Eqs. (21) in (24) to obtain:

$$\tilde{w}_j L_j = \sum_i \left[N_i \frac{M_j (\tau_{ij} \tilde{w}_j / \underline{\varphi}_j)^{1-\sigma}}{\sum_{k \in J} M_k (\tau_{ik} \tilde{w}_k / \underline{\varphi}_k)^{1-\sigma}} \sum_j^J \lambda_{ij|i}^N \tilde{w}_j \right] \quad \text{(A.34)}$$

Since we observe $\{\tilde{w}_j, L_j, M_j, \lambda_{ij|i}^N\}$ and have parameterized τ_{ij} in Section B.3.3, Eq. (A.34) provides a system of J equations that we can solve for a unique vector of J productivities $\underline{\varphi}_j$ using a fixed-point approach following Monte et al. (2018).

Table A4: Distance elasticity of trade volumes

	(1)	(2)	(3)
Distance (in km)	-0.0106***	-0.0116***	-0.0112***
	(0.0001)	(0.0001)	(0.0001)
West-to-East			-0.1058
			(0.2790)
East-to-West			-0.7872***
			(0.2853)
Origin FE	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes
Origin-specific distance elasticity	No	Yes	Yes
Observations	114,951	114,951	114,951
R^2	.391	.403	.405

Notes: Unit of observation are bilateral county-county trade values. Columns (2) and (3) show the estimated mean of the origin-specific distance elasticities and its standard error. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B.4 Quantitative analysis

This section complements Section 4.3 in the main paper by providing further details on the numerical procedure to solve the model.

B.4.1 Long run

Given the model's parameters $\{\underline{w}, k, \alpha, \sigma, \epsilon, \zeta, \mu\}$ and structural fundamentals $\{\tau_{ij}, \kappa_{ij}, B_{ij}, \underline{\varphi}_j, \bar{T}_i, f_j^e, A\}$, we describe in this section how we solve for the endogenous variables $\{\tilde{w}_i, \tilde{v}_i, P_i^T, L_i, N_i, P_i^Q, M_i, \mu, \overline{V}\}$. In the long run, we fix the nation-wide population $\bar{N}$ and determine one labor force participation rate. By solving for the (unconditional) probability of living in i and working in j, we determine labor supply and the employment for each location. In the sequel, we describe the procedure to solve the model:

1. Guess λ_{ij}; $\tilde{w}_j$; $N = \sum_i N_i = \sum_j H_j$; Φ_j^R; Φ_j^Π; Φ_j^H; Φ_j^L; Φ_j^P; Φ_j^W
   (a) Compute $\tilde{v}_i$ based on Eq. (33).
   (b) Compute L_j based on Eq. (32).
   (c) Compute residents according to $N_i = \lambda_i^N L$.
   (d) Compute house price index P_i^T according to Eq. (16).
   (e) Using the free-entry condition Eq. (18), compute expenditure shares θ_{ij} according to Eq. (21).
   (f) Compute the goods price index P_i^Q according to Eq. (23).
2. Derive new values of initially guessed variables:
   (a) Compute new value of $\tilde{w}_j$ according to Eq. (24) and normalize values with employment-weighted average wage.
   (b) Compute new value of λ_{ij} according to Eq. (29).

(c) Use the value of the minimum wage $\underline{w}$ ($\underline{w}_j$ for regional minimum wages) which is defined relative to the numeraire (employment-weighted average wage) together with $\underline{w}_j^u$ from Eq. (7) and $\underline{w}_j^s$ from Eq. (8) to compute new values of all Φ_j^X according to Appendix A.3.

(d) Compute new value of labor force participation rate μ according to Eq. (35) to get a new value for aggregate labor supply (measured at residence) N

3. Determine new initial guesses by a computing convex combinations of values from previous iteration with updated values.

4. Iterate until convergence.

B.4.2 Short run

Consistent with the perfect-mobility assumption the expected utility is equalized across origin-destination commuting pairs in our model as per Eq. (34). However, the expected utility conditional on being settled in a specific residence, V_i is not equalized. The intuition is that expected utility does not incorporate idiosyncratic Gumbel-distributed taste shocks, whereas the equilibrium allocation of workers across residences is the result of the realization of these taste shocks. The expected utility conditional on being in i is irrelevant in the long-run equilibrium since workers are perfectly mobile and re-optimize location choices such that they locate in places that suit a given realization of the shock. Our definition of the short run, however, is that workers are immobile. Therefore, the expected utility—-before drawing a taste shock—conditional on being in a residence i becomes the relevant benchmark for the a welfare evaluation.

Conditional expected utility. To derive the conditional expected utility V_i, we use Eq. (34) and exploit that the Gumbel distribution of residence-workplace-employer tastes shocks implies that we can rewrite unconditional expected utility as

$$\overline{V} = \left(\sum_i V_i^{\varepsilon} \right)^{\frac{1}{\varepsilon}}$$

from which it follows that

$$\tilde{V}_i = \left\{ \sum_j B_{ij} \left[\kappa_{ij} \left(P_i^Q\right)^{\alpha} \left(P_i^T\right)^{1-\alpha} \right]^{-\varepsilon} \left[\frac{\eta \Phi_j^W(\underline{w})^{\frac{1}{\varepsilon}} \Phi_j^L(\underline{w})}{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^{\Pi}(\underline{w})} \right]^{\varepsilon} M_j \tilde{w}_j^{\varepsilon} \right\}^{\frac{1}{\varepsilon}}.$$

Quantification. The immobility in the short run also implies that workers make their decision as to enter the labour market knowing the location in which they will enjoy the leisure amenity. Therefore, we obtain a variant of the Eq. (35), which determines the labour force participation rate, in which expected utility $\tilde{V}_i$ and leisure amenity $\tilde{A}_i$ are location specific. To rationalize the same uniform labor force participation rate as in the long-run equilibrium, we invert $\tilde{A}_i$ from

$$\mu = \frac{\tilde{V}_i^{\zeta}}{\tilde{V}_i^{\zeta} + \tilde{A}_i} \tag{A.35}$$

Consequentially, conditional welfare becomes location-specific:

$$\tilde{\mathcal{V}}_i = \left(\tilde{A}_i + \tilde{V}_i^{\zeta}\right)^{\frac{1}{\zeta}} \tag{A.36}$$

Quantitative analysis. In perfect analogy to the long-run evaluation, we solve for the unconstrained endogenous variables of the model in the absence and the presence of the minimum wage to establish the causal effect. The main difference is that we now have an exogenous endowment with working-age population at the local level $\overline{N}_i$. Since we obtain spatially varying changes in the conditional expected utility from work $\hat{\tilde{V}}_i$, we obtain spatially varying changes in labour force participation rates as per Eq. (A.35) and spatially varying changes in conditional welfare as Eq. (A.36). While the working-age population is a fixed endowment in the short run, the labour force remains an endogenous variable as per

$$N_i = \mu_i \overline{N}_i. \tag{A.37}$$

Against this background, we adjust the numerical procedure for the long run as follows. First, we guess $\lambda_{ij|i}$ instead of λ_{ij} and N_i instead of N. Second, under 2.(b), we compute new values for $\lambda_{ij|i}$ according to Eq. (31). Third, under 2.(d), we compute new values for local labor force participation rates based on Eq. (A.37) and thus new N_i. All other steps remain the same.

References

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf, "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (4), 2127–2189.

Arkolakis, Costas, "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 2010, *118* (6), 1151–1199.

Bartik, Timothy J., *Who Benefits from State and Local Economic Development Policies?*, Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 1991.

Combes, Pierre Philippe, Miren Lafourcade, and Thierry Mayer, "The trade-creating effects of business and social networks: Evidence from France," *Journal of International Economics*, 5 2005, *66* (1), 1–29.

Egger, Hartmut, Peter Egger, and Udo Kreickemeier, "Trade, wages, and profits," *European Economic Review*, 2013, *64*, 332–350.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg, "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 2018, *108* (12), 3855–3890.

Mori, Tomoya and Jens Wrona, "Centrality Bias in Inter-city Trade," *RIETI Discussion Paper*, 2021, *E* (35).

Severen, Christopher, "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification," *The Review of Economics and Statistics*, 9 2021, pp. 1–99.

Simonovska, Ina and Michael E Waugh, "The elasticity of trade: Estimates and evidence," *Journal of International Economics*, 2014, *92* (1), 34–50.

Sokolova, Anna and Todd Sorensen, "Monopsony in Labor Markets: A Meta-Analysis," *ILR Review*, 10 2020, *74* (1), 27–55.

Yeh, Chen, Claudia Macaluso, and Brad Hershbein, "Monopsony in the US Labor Market," *American Economic Review*, 2022, *112* (7), 2099–2138.