DISCUSSION PAPER SERIES

DP16913

Optimal minimum wages

Gabriel Ahlfeldt, Duncan Roth and Tobias Seidel

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

LABOUR ECONOMICS

PUBLIC ECONOMICS

ISSN 0265-8003

Optimal minimum wages

Gabriel Ahlfeldt, Duncan Roth and Tobias Seidel

Discussion Paper DP16913
Published 17 January 2022
Submitted 12 January 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Optimal minimum wages

Abstract

We develop a quantitative spatial model with heterogeneous firms and a monopsonistic labour market to derive minimum wages that maximize employment or welfare. Quantifying the model for German micro regions, we find that the German minimum wage, set at 48% of the national mean wage, has increased aggregate worker welfare by about 2.1% at the cost or reducing employment by about 0.3%. The welfare-maximizing federal minimum wage, at 60% of the national mean wage, would increase aggregate worker welfare by 4%, but reduce employment by 5.6%. An employment-maximizing regional wage, set at 50\% of the regional mean wage, would achieve a similar aggregate welfare effect and increase employment by 1.1%.

JEL Classification: J31, J58, R12

Keywords: Applied general equilibrium model, minimum wage, employment, unemployment, minimum wage policy, Minimum Wages, Inequality, Germany, monopsony

Gabriel Ahlfeldt - g.ahlfeldt@lse.ac.uk
London School of Economics and CEPR

Duncan Roth - duncan.roth@iab.de
Institute for Employment Research

Tobias Seidel - tobias.seidel@uni-due.de
University of Duisburg-Essen

Acknowledgements
We are grateful to comments and suggestions we received at workshops and conferences in Amsterdam, Düsseldorf (UEA), Berlin, (DIW conference on the evaluation of minimum wages, minimum wage commission conference), Kassel (Regionaloekonomischer Ausschuss), Bonn (German Economists Abroad), London (LSE Labour workshop), Rome (Bank of Italy), the IAB (online) and in particular to Gaetano Basso, Francesca Carta, Emanuele Ciani, Wolfgang Dauth, Gilles Duranton, Maximilian von Ehrlich, Alexandra Fedorets, Bernd Fitzenberger, Steve Gibbons, Simon Jaeger, Philip Jung, Attila Lindner, Andreas Mense, Guy Michaels, Henry Overman, Michael Pflüger, Hans Koster, Joachim Möller, Gregor Singer, Michel Serafinelli, Jens Südekum, Erik Verhoef, Jos van Ommeren, Eliana Viviano, Jens Wrona, and Atsushi Yamagishi. The usual disclaimer applies.

Optimal minimum wages*

Gabriel M. Ahlfeldt   Duncan Roth   Tobias Seidel

This version: January 4, 2022
First version: December 17, 2021

Abstract

We develop a quantitative spatial model with heterogeneous firms and a monopsonistic labour market to derive minimum wages that maximize employment or welfare. Quantifying the model for German micro regions, we find that the German minimum wage, set at 48% of the national mean wage, has increased aggregate worker welfare by about 2.1% at the cost or reducing employment by about 0.3%. The *welfare-maximizing federal* minimum wage, at 60% of the national mean wage, would increase aggregate worker welfare by 4%, but reduce employment by 5.6%. An *employment-maximizing regional* wage, set at 50% of the regional mean wage, would achieve a similar aggregate welfare effect and increase employment by 1.1%.

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), Berlin, (DIW conference on the evaluation of minimum wages, minimum wage commission conference), Kassel (Regionaloekonomischer Ausschuss), Bonn (German Economists Abroad), London (LSE Labour workshop), Rome (Bank of Italy), the IAB (online) and in particular to Gaetano Basso, Francesca Carta, Emanuele Ciani, Wolfgang Dauth, Gilles Duranton, Maximilian von Ehrlich, Alexandra Fedorets, Bernd Fitzenberger, Steve Gibbons, Simon Jaeger, Philip Jung, Attila Lindner, Andreas Mense, Guy Michaels, Henry Overman, Michael Pflüger, Hans Koster, Joachim Möller, Gregor Singer, Michel Serafinelli, Jens Südekum, Erik Verhoef, Jos van Ommeren, Eliana Viviano, Jens Wrona, and Atsushi Yamagishi. The usual disclaimer applies.

1 Introduction

On few questions do economists disagree so passionately as on the desirability of minimum wages. The controversy is primarily an empirical one since there is arguably a theoretical consensus that a sufficiently high minimum wage will reduce employment. That a sufficiently low minimum wage may increase employment and welfare in a monopsonistic labour market also seems consensual. The open policy question is which minimum wage level maximizes employment or welfare. We develop a quantitative model that offers an answer.

Our approach differs from a vast literature using reduced-form methods to study employment effects of minimum wages summarized by Manning (2021) and Neumark and Shirley (2021). Instead, we develop a quantitative spatial model with heterogeneous firms and a monopsonistic labour market to study minimum wage effects in a spatial general equilibrium. Our model is uniquely equipped to derive optimal minimum wage schedules. For one thing, our model accounts for qualitatively and quantitatively heterogeneous employment responses in regions of distinct productivity (Christl et al., 2018). This allows us to predict both regionally differentiated and aggregate employment effects. For another, our model accounts for a broad range of minimum-wage effects that have recently been documented in the literature, including effects on labour force participation (Lavecchia, 2020), tradable goods prices (Harasztosi and Lindner, 2019), housing rents (Yamagishi, 2021), or commuting costs (Pérez Pérez, 2018) and worker-firm matching (Dustmann et al., 2021). This allows us to derive a worker welfare measure that incorporates all of those general equilibrium channels along with the effects on wages and employment probabilities.

We use our model to derive bounds for optimal minimum wages that are novel to the literature and of immediate policy interest. For Germany, we find an employment-maximizing *federal* minimum wage of 38% of the national mean wage, corresponding to 42% of the median wage. At less than 0.5%, however, the positive employment effect is small, and so is the impact on welfare. The welfare-maximizing federal minimum wage level is 60% of the national mean wage, corresponding to 66% of the median wage. While welfare increases by about 4%, there is a reduction in aggregate employment by 5.5%, driven by low-productivity regions. One important conclusion from our analysis is that within the bounds of the employment-maximizing and welfare-maximizing federal minimum wages, an increase in welfare can only be achieved at a cost of reducing employment. The implication is that ambitious federal 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. Moderate *regional* minimum wages offer an attractive alternative that can achieve similar welfare gains as ambitious federal minimum wages, plus significant job creation.

While our model is sufficiently tractable to be implemented in arbitrary empirical contexts that satisfy the data requirements, choosing Germany as our case in point comes with three advantages. First, the first-time introduction of a relatively high nationally uniform minimum wage (54% of the national median wage)[1] as of 2015 provides an opportunity to contrast theoretical predictions with evidence. Second, we are able to quantify the model at unprecedented spatial coverage and detail at the level of 4,421 micro regions (*Verbandsgemeinden*), owing to the availability of linked linked-employer-employee data covering the universe of 30M workers from the Institute for Employment Research (IAB) and a micro-geographic property price index recently developed by Ahlfeldt et al. (2021). Third, our ability to observe the spatial economy in the recent past before a minimum wage was introduced greatly simplifies the quantification since we can treat observed labour market outcomes as undistorted market outcomes.

This data set paves the way for our methodological contribution, which is to develop a quantitative spatial model with heterogeneous firms that possess monopsony power. We start from a canonical setup in the spirit of Redding and Rossi-Hansberg (2017). 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. We extend this canonical framework in three important respects. First, we borrow from the trade literature and introduce a Pareto-shaped productivity distribution of firms (Redding, 2011; Gaubert, 2018). This extension is critical to generating a wage distribution within regions and enabling the minimum wage to reallocate workers to more productive establishments. Second, we follow Egger et al. (2021), who build on Card et al. (2018), 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).[2] This extension is critical to awarding employers monopsony power. Third, we generate imperfectly elastic aggregate labour supply via a Gumbel-distributed idiosyncratic utility from non-employment. 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).

To develop the intuition for the regional employment response in our model, it is instructive to consider three firm types. The minimum wage has no effect on the most productive firms, which we term *unconstrained* because they voluntarily pay wages above the minimum wage. These firms still exercise their full monopsony power after the minimum wage is introduced. For all less productive firms, the minimum wage is binding. These firms can no longer lower the wage below the minimum wage, which implies that they lose some of their monopsony power. The more productive among the constrained firms will respond by hiring all workers they can attract at the minimum wage—the new

[1]This quantity is based on 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)

[2]Dustmann et al. (2021) model non-pecuniary aspects of job choice in a similar way.

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 falls below the minimum wage level, which is why we term them *demand-constrained*. While some demand-constrained firms will react to the introduction of the minimum wage by increasing employment, any demand-constrained firm that 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 a region delivers the prediction that the regional employment effect of a federal minimum wage is a hump-shaped function of regional productivity. We substantiate this predictions in two complementary approaches.

In the first step, we employ a reduced-form methodology that uses high-productivity regions in which firms are mostly unconstrained as a counterfactual in the spirit of a dynamic difference-in-difference model. This approach allows us to test a central prediction of the model without imposing the full structure of the quantitative model. Consistent with model predictions, we find that the 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 employment effects adds to a literature that has mostly focused on average effects for selected spatial units (e.g. Card and Krueger, 1994; Dustmann et al., 2021) 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., 2021).

In the second step, we quantify the full model to take the analysis into the general equilibrium. We exploit our matched worker-establishment micro data to estimate the structural parameters that govern the wage distribution within regions. We then invert the model in 2014—the year before the introduction of the minimum wage. 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 almost exactly the same 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. In the national aggregate, employment decreases by about 0.3% or 100K jobs, which is less than predicted by the competitive labour market model (Knabe et al., 2014).

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 by 2%. 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 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 wage—adjusted for changes in tradeble goods prices, housing rents, and commuting costs—dominates the reduction in the employment probability. 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.

Given the absence of a credible counterfactual, we cannot over-identify the aggregate effects of the German minimum wage our model delivers. We show, however, that the model's predictions for minimum wage effect in wages, employment, housing rents and commuting distances are closely correlated with observed before-after changes in the data at the regional level. We also show that our model predicts changes in the Gini coefficient of wage inequality across all workers in all 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 wage schedules.

To this end, we compute aggregate employment and welfare effects for a broad range of federal and regional minimum-wage schedules. We also provide an equity measure based on the Gini coefficient of wage inequality. 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 60%, 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 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 (4%), plus increase employment by 1.1%, suggesting that regional minimum wages are targeted policy instruments that warrant more attention.

With these results, we contribute to the identification of turning points where the costs of minimum wages start exceeding the benefits, a challenge that allegedly lies ahead of the field (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 have negative employment effects (Dube et al., 2010; Cengiz et al., 2019).[3] This includes a growing literature evaluating the labour market effects of the German minimum wage, which we review in more detail in Appendix A (e.g. Ahlfeldt et al., 2018; Bossler and Gerner, 2019; Caliendo et al., 2018; Dustmann et al., 2021). We also contribute to a smaller normative literature on minimum wages that considers distributional effects of minimum wages (Lee and Saez, 2021; Simon and Wilson, 2021).[4] For Germany, Drechsel-Grau (2021) studies the macroeconomic and distributional effects of minimum wages within a dynamic macroeconomic model in a current working paper.

Our theoretical contribution builds on a literature showing that, in a monopsonistic labour market,[5] a minimum wage can raise the wage without reducing employment (Stigler, 1946; Manning, 2003a).[6] We also draw from a literature on quantitative spatial models, which have recently emerged as general-purpose tools for policy evaluation that can account for mobility of residents across residence and workplace locations (Allen and Arkolakis, 2014; Ahlfeldt et al., 2015).[7] In particular, our model nests Monte et al. (2018) as a special case in which the dispersion of firm productivity approaches infinity, there is no idiosyncrasy in worker tastes for employers, and workers supply labour inelastically.

Most closely, we connect to a small literature that studies the effects of minimum wages in a spatial equilibrium. Our contribution complements Monras (2019) 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), who also provides an evaluation of the German minimum wage. Similar to us, he uses worker-firm-specific idiosyncratic utility to generate an upward-sloping labour-supply curve to the firm. By making the labour supply elasticity dependent on the density of nearby firms within relatively large regions that roughly correspond to local labour markets, he shows that lower monopsony power acts as an important concentration force in the spatial economy (see also Azar et al. (2019)). In contrast, our focus is the development of a comprehensive welfare measure for the normative evaluation of minimum wages. Hence, we develop our model at the micro-regional level to capture minimum wage effects on commuting costs explicitly.[8] We also account for how frictional trade shapes the spatial distribution of minimum wage effects and account for employment and welfare effects that arise when minimum wages incentivize workers to become active on the labour market, which is important to enable

[3] 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.

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

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

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

[7] Other recent models that quantitatively account for commuting include Tsivanidis (2019); Heblich et al. (2020); Almagro and Domínguez-Iino (2021).

[8] 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).

aggregate employment gains.

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 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.

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 euros 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. We provide a detailed discussion of the institutional context in Section 2.1.

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 sub-city level. We provide a brief summary of the various data sources here and refer to Appendix B.2 for details.

Employment, establishments 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.

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. (2021) to generate an area-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 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 (see Figure 1 for a map). 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.

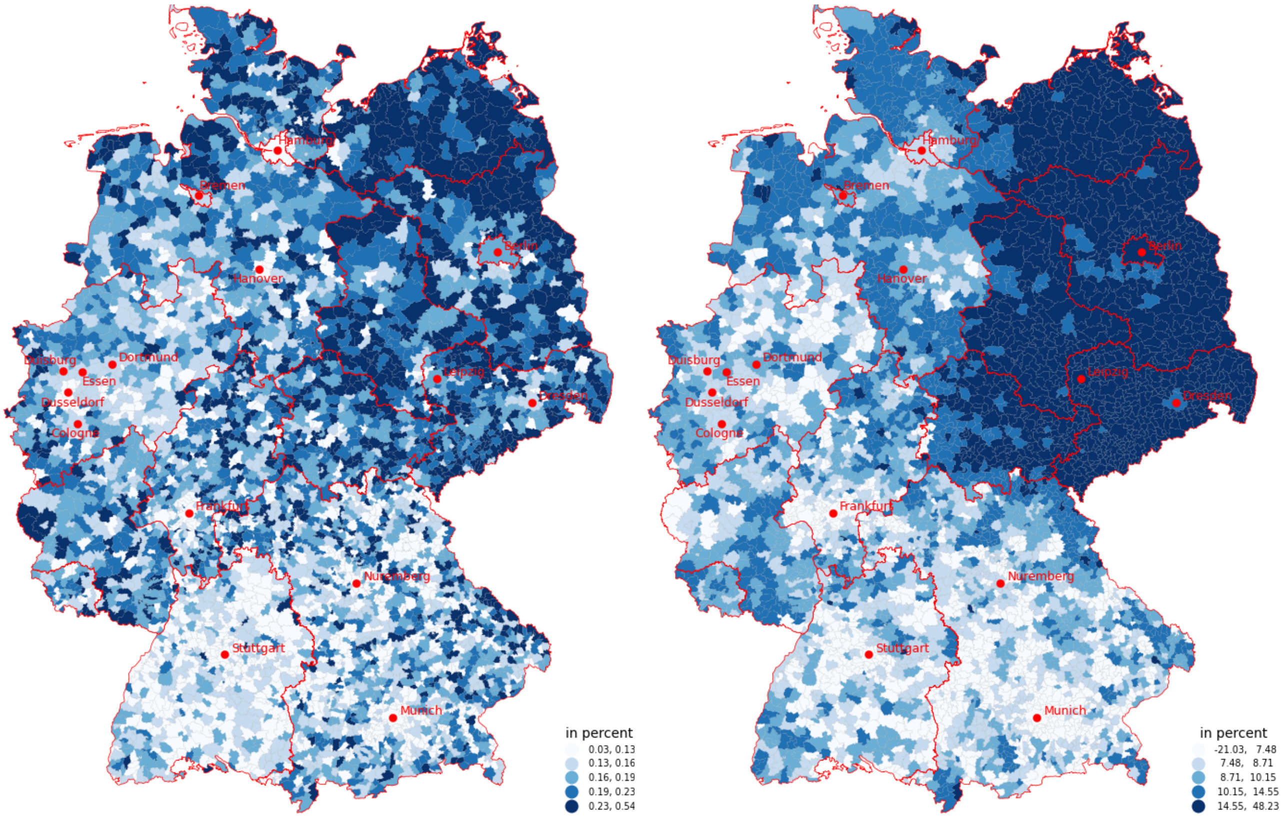
2.3 Stylized facts

Figure 1, illustrates a measure of the regionally differentiated "bite" of the national minimum wage, very much in the tradition of Machin et al. (2003). Concretely, we compute a bite exposure measure at the residence by taking the average over the shares of below-minimum-wage workers at the workplace across nearby municipalities, weighted by the bilateral commuting flows in 2014.[9] This way, we capture the bite within the actual commuting zone of a municipality. Evidently, the minimum wage had a greater bite in the east, in line with the generally lower productivity. Changes in low wages, defined as the 10^{th} percentile in the within-area wage distribution, from 2014 to 2016 closely follow the distribution of the bite, suggesting a significant degree of compliance. Together, the two maps suggest that the minimum wage contributed to the reduction of spatial wage disparities in Germany, an impression that we substantiate with further evidence in Appendix B.3.

The striking heterogeneity in the policy-induced wage increase between the eastern and the western states makes it instructive to compare how employment and other outcomes evolved in the respective parts of the country over time. We offer this purely descriptive

[9] Formally, we define the bite as $\mathcal{B}_i = \sum_j \frac{L_{i,j}}{\sum_j L_{i,j}} S_j^{MW}$, where $L_{i,j}$ is the number of employees who live in municipality i and commute into municipality j for work and S_j^{MW} is the share of workers compensated below the minimum wage in j.

Figure 1: Minimum wage bite and change in 10th pct. regional wages



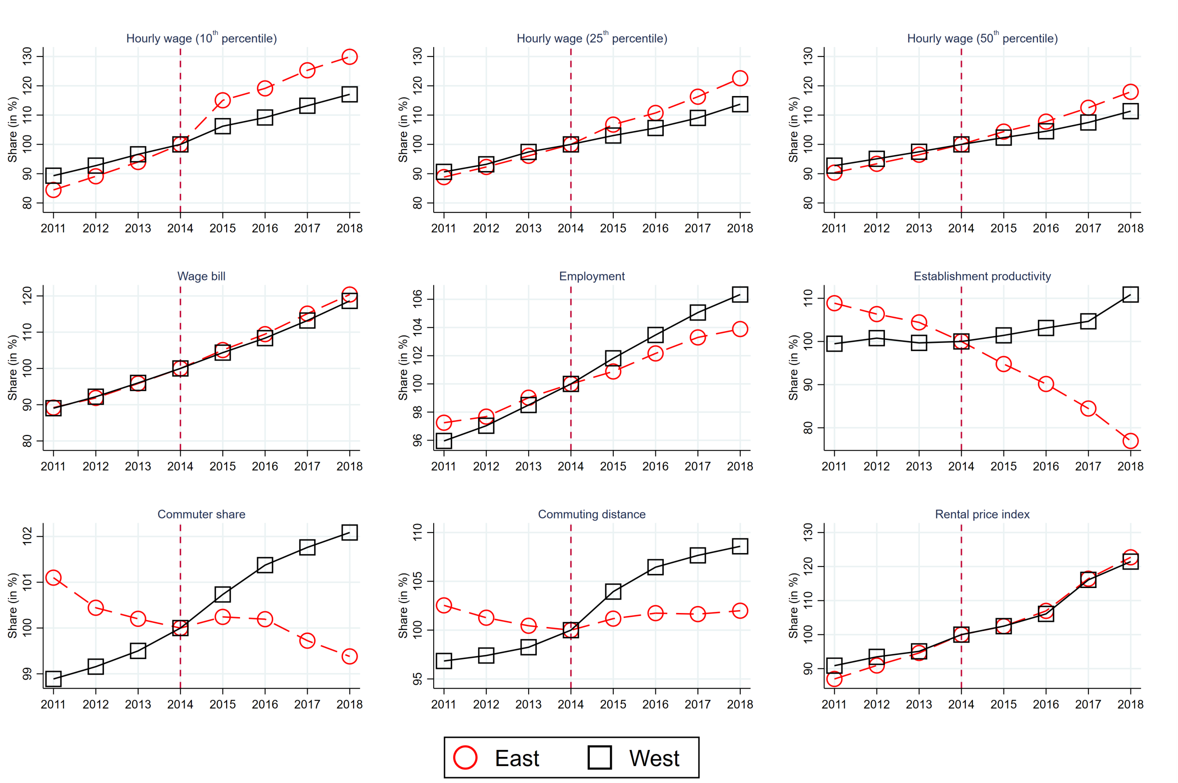
(a) Minimum wage bite in 2014

(b) 2014-2016 wage growth at 10th pct.

Note: Unit of observation is 4,421 municipality groups. The 10^{th} percentile wage refers to the 10^{th} percentile in the distribution of individuals within a workplace municipality, re-weighted to the residence using commuting flows. Wage and employment data based on the universe of full-time workers from the *BeH*.

comparison in Figure 2. Confirming Figure 1, a jump at the 10^{th} percentile of the wage distribution in the east is immediately apparent. A more moderate increase is also visible for the west. For higher percentiles, it is possible to eyeball some increase in the east, but not in the west. A first-order question from a policy-perspective is whether the policy-induced wage increase came at the cost of job loss as predicted by the competitive labour market model. While we argue that—without a general equilibrium model—it is difficult to establish a counterfactual for aggregate employment trends, the absence of an immediately apparent employment effect in these time series is still informative. It is worth noticing that, while employment continues to grow in both parts of the country after the minimum wage introduction, the rate of growth appears to slow down in the east compared to the west. However, even if one is willing to interpret this as suggestive evidence of a negative employment effect, it will be difficult to argue that negative employment effects turned out to be as severe as in some pessimistic scenarios circulated ahead of the implementation (Ragnitz and Thum, 2008). Since, following the minimum-wage introduction, the aggregate wage bill increases in the east, relative to the west, it seems fair to conclude

Figure 2: Outcome trends in western and eastern states



Note: All time series are normalized to 100% in 2014, the year before the minimum wage introduction. The establishment wage premium is the employment-weighted average across firm-year fixed effects from a decomposition of wages into worker and firm fixed effects following Abowd et al. (1999) (see Appendix B.4 for details).

that a positive wage effect has dominated a possibly negative employment effect, pointing to positive welfare effects. Figure 2 also illustrates the reallocation of workers to more productive establishments at greater commuting distance documented by Dustmann et al. (2021). Indeed, it appears that the effect has gained momentum subsequent to 2016, when their analysis ends. Finally, there appears to be a slight increase in the rate of property price appreciation after the minimum wage which could be reflective of increased demand.

These stylized facts echo a growing empirical literature on the effects of the German minimum wage which we summarize in Appendix A. They motivate us to develop a quantitative spatial model that nests a monopsonistic labour market in which a statutory minimum wage triggers different employment responses by firms of distinct productivity and workers decide where to live and work trading higher wages against greater commuting costs.

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 employment response to a uniform minimum wage differs qualitatively across firms, depending on their productivity. We then derive the novel prediction that the regional employment response is a hump-shaped function of regional productivity. Finally, we provide novel area-specific estimates of the employment effect of the minimum-wage that confirm this prediction.

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 QSM 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 firm's hiring probability. Unless otherwise indicated, we assume $\psi_j(\omega_j) = 1$ to ease notations. 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 Section 4). Under profit maximization and goods market clearing, we can express the revenue function as

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

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 C.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 C.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

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

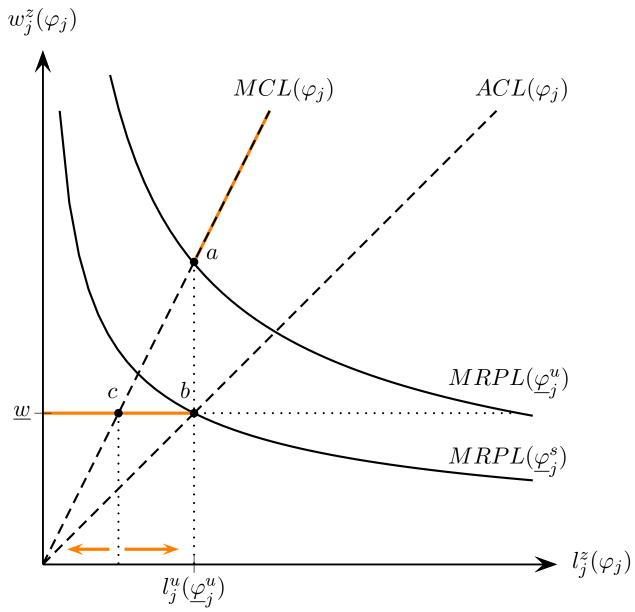
Facing an upward-sloping labour supply function, firms can only increase their employment by offering higher wages. Hence, the average cost of labour $ACL(\varphi_j) = c_j^u(\varphi_j)/l_j^u(\varphi_j)$ is upward-sloping as illustrated in Figure 3. The marginal cost of labour $MCL(\varphi_j) = \partial c_j^u(\varphi_j)/\partial l_j^u(\varphi_j) = \frac{\varepsilon+1}{\varepsilon} ACL(\varphi_j)$ 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 3. 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 $w_j(\underline{\varphi}_j^u) = \underline{w}$, 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

Figure 3: Optimal firm employment



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 3. 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 3).

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 3, 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

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

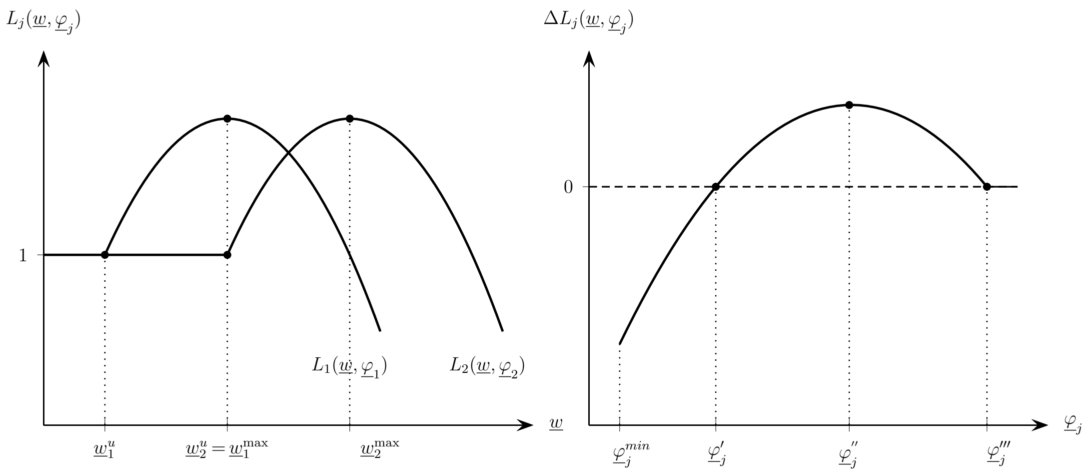
implying $\underline{w}_j^s/\underline{w}_j^u = (\rho/\eta)^{\frac{\sigma}{\sigma+\varepsilon}} > 1$. Using these critical minimum wages we derive the following proposition:

Proposition 1. *Assuming Pareto-distributed firm productivities, 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 C.4.

To develop the intuition, let's first consider the region indexed by $j = 1$ in panel a) of Figure 4. 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

Figure 4: 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.

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 absolute employment effect will turn negative. The generalizable insight is that for given fundamentals $\{S_j^r,\, S_j^h\}$ and regional productivity summarized by $\underline{\varphi}_j$, aggregate employment $L_j(\underline{w}_j, \underline{\varphi}_j)$ is hump-shaped in the minimum wage level $\underline{w}_j$.

To clear the regional labour market, the hump-shaped pattern must carry through to labour supply. Since the hiring probability is $\psi_j^z = 1 \;\forall\; z \in \{s, u\}$ for unconstrained and supply-constrained firms, labour supply defined in Eq. (1) increases for low but binding minimum wage levels $\underline{w}_j^u \leq \underline{w}_j < \underline{w}_j^s$. At a higher minimum wage level, the expected hiring probability adjusts to the hiring rate to account for the job rationing of demand-constrained firms ($\psi_j^d(\varphi_j) = l_j^d(\varphi_j)/h_j^d(\varphi_j)$). In other words, workers who are unlikely to get a job withdraw from the labour force. In the spatial general equilibrium introduced in Section 4.1, workers will adjust labour supply by choosing *if* to work and *where* to live (migration) and work (commuting). For a discussion of the effect of the minimum wage on firm profits and revenues, we refer to Appendix C.4.

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 4, 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 4, 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 Germany and many other countries. We refer to Appendix C.5 for a detailed and formal discussion of the comparative statics. Briefly summarized, we can distinguish between three types of regions. The minimum wage has no effect in regions where even the least productive firm is unconstrained ($\underline{\varphi_j} \geq \underline{\varphi_j^{'''}}$). 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 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 difference Eq. (9) twice to get

$$[\ln L_{j,t} - \ln L_{j,t-n}] - [\ln L_{j,t-n} - \ln L_{j,t-m}] = \Delta^2 \ln L_j = \overline{f} + f(\underline{\varphi}_j) + \tilde{\epsilon}_{i,t}, \tag{10}$$

where $t - n < \mathcal{J}$, $t - m < \mathcal{J}$ and $\tilde{\epsilon}_{i,t}$ is the twice differenced error term. Guided by the theoretical predictions summarized in Figure 4, we define the relative (up to the constant $\overline{f}$) before-after minimum wage effect as a polynomial spline function

$$\begin{aligned} f(\underline{\varphi}_j) &= \mathbb{E}\left[\Delta^2 \ln L_j | w_j^{\text{mean}}, (w^{\text{mean}_j} \leq \alpha_0)\right] - \mathbb{E}\left[\Delta^2 \ln L_j | w_j^{\text{mean}}, (w^{\text{mean}_j} > \alpha_0)\right] \\ &= \mathbb{1}\left(w_j^{\text{mean}} \leq \alpha_0\right) \times \left[\sum_{g=1}^{2} \alpha_g \left(w_j^{\text{mean}} - \alpha_0\right)^g\right], \end{aligned} \tag{11}$$

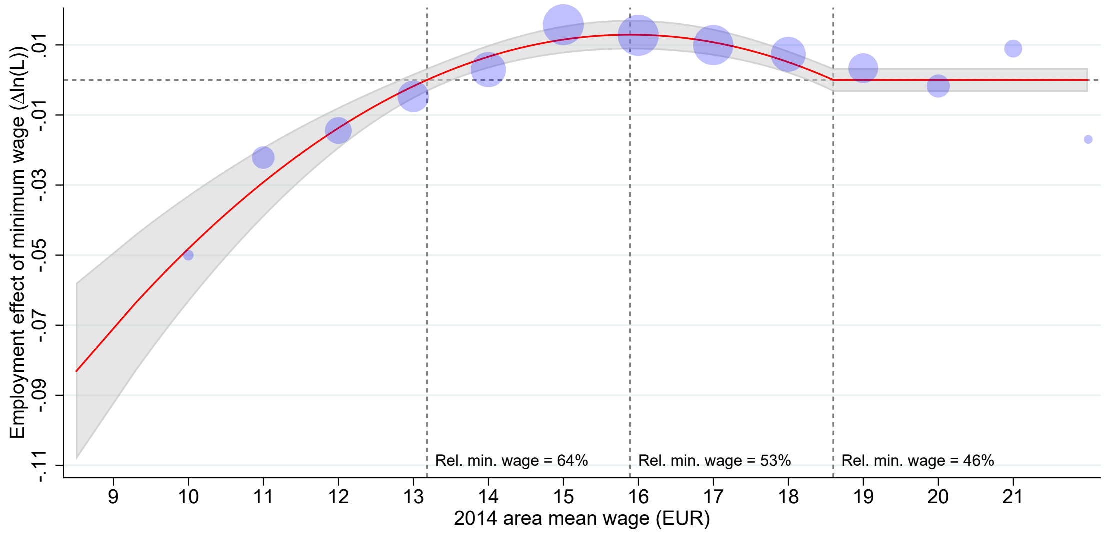
with the theory-consistent parameter restrictions $\{\alpha_0 > \frac{\alpha_1}{2\alpha_2}, \alpha_1 < 0, \alpha_2 < 0\}$. Since higher fundamental productivity maps to higher wages in our model, we use the 2014 mean wage w_j^{mean} as a proxy for regional productivity. Notice that the interpretation 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. (11), we are ready to estimate Eq. (10) for given years $\{t, t-n, t-m\}$. To obtain parameter values $\{\alpha_0, \alpha_1, \alpha_2\}$, 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 4 (see Appendix C.6 for details). Note that consistent with the partial-equilibrium nature of the analysis, Eq. (11) 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.

In the DGP laid out in Eq. (9), we assume that the linear area-specific trends extend from the $[t - m, t - n]$ to the $[t - n, t]$ period. This assumption is more likely to be true over shorter study periods. Hence, we set $\{t = 2016, m = 4, n = 2\}$ in Figure 5, 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 (see Appendix C.6).

Consistent with theory, we find an employment effect that is hump-shaped in the 2014

Figure 5: 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).

mean wage. 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 4. 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%.

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.

4 General equilibrium analysis

We develop the model in Section 4.1 and discuss the quantification in Section 4.2 before lay 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.

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 $Q_{i\nu}$, residential land use $T_{i\nu}$, an idiosyncratic amenity shock $\exp[b_{ij\nu}(\varphi_j)]$, and commuting costs $\kappa_{ij} > 1$, according to the Cobb-Douglas form

$$U_{ij\nu}(\varphi_j) = \frac{\exp[b_{ij\nu}(\varphi_j)]}{\kappa_{ij}} \left(\frac{Q_{i\nu}}{\alpha}\right)^{\alpha} \left(\frac{T_{i\nu}}{1-\alpha}\right)^{1-\alpha}. \tag{12}$$

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., 2021). We assume that $b_{ij\nu}(\varphi_j)$ is drawn from an independent Type I extreme value (Gumbel) distribution

$$F_{ij}(b) = \exp(-B_{ij}\exp\{-[\varepsilon b + \Gamma'(1)]\}), \quad \text{with} \quad B_{ij} > 0 \quad \text{and} \quad \varepsilon > 0, \tag{13}$$

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

$$Q_i = \left[\sum_j \int_{\varphi_j} q_{ij}(\varphi_j)^{\frac{\sigma-1}{\sigma}} d\varphi_j \right]^{\frac{\sigma}{\sigma-1}} \tag{14}$$

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. (14), 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:

$$P_i^T = \left(\frac{E_i^T}{\bar{T}_i} \right). \tag{15}$$

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

$$\tilde{\pi}_j = \frac{\Pi_j}{M_j} = f_j^e P_j^T. \tag{16}$$

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. (16) to get

$$M_j = \frac{\Phi_j^\Pi(\underline{w})(1-\eta)}{\Phi_j^R(\underline{w}) - \Phi_j^\Pi(\underline{w})(1-\eta)} \frac{\tilde{w}_j L_j}{P_j^T f_j^e}, \tag{17}$$

where

$$\tilde{w}_j = \frac{R_j - \Pi_j}{L_j} = \frac{1-(1-\eta)\Phi_j^\Pi(\underline{w})/\Phi_j^R(\underline{w})}{\eta} \frac{\chi_R \Phi_j^R(\underline{w})}{\chi_L \Phi_j^L(\underline{w})} w_j^u(\underline{\varphi}_j) \tag{18}$$

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

$$E_i^Q = \alpha \tilde{v}_i N_i + E_i^T = \tilde{v}_i N_i + \Pi_i, \tag{19}$$

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

$$\begin{aligned}\theta_{ij} &= \frac{M_j \int_{\varphi_j} p_{ij}(\varphi_j)^{1-\sigma} dG(\varphi_j)}{\sum_{k\in J} M_k \int_{\varphi_k} p_{ik}(\varphi_k)^{1-\sigma} dG(\varphi_k)}, \\ &= \frac{M_j \Phi_j^P(\underline{w}) \left(\left\{\Phi_j^L(\underline{w})/[\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^{\Pi}(\underline{w})]\right\} \tau_{ij}\tilde{w}_j/\underline{\varphi}_j\right)^{1-\sigma}}{\sum_{k\in J} M_k \Phi_k^P(\underline{w}) \left(\left\{\Phi_k^L(\underline{w})/[\Phi_k^R(\underline{w}) - (1-\eta)\Phi_k^{\Pi}(\underline{w})]\right\} \tau_{ik}\tilde{w}_k/\underline{\varphi}_k\right)^{1-\sigma}}.\end{aligned} \tag{20}$$

To derive Eq. (20) 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 C, it can be computed as

$$P_{ij} = \chi_P^{\frac{1}{1-\sigma}} \Phi_j^P(\underline{w})^{\frac{1}{1-\sigma}} M_j^{\frac{1}{1-\sigma}} p_{ij}^u(\underline{\varphi}_j), \tag{21}$$

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. (18) and (21) 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. (14) we obtain

$$\begin{aligned}P_i^Q &= \frac{\chi_L}{\chi_R}\chi_P^{\frac{1}{1-\sigma}} \left\{\sum_j M_j \Phi_j^P(\underline{w}) \left[\frac{\Phi_j^L(\underline{w})}{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^{\Pi}(\underline{w})} \frac{\tau_{ij}\tilde{w}_j}{\underline{\varphi}_j}\right]^{1-\sigma}\right\}^{\frac{1}{1-\sigma}}, \\ &= \frac{\chi_L}{\chi_R}\chi_P^{\frac{1}{1-\sigma}} \left[\frac{M_i \Phi_i^P(\underline{w})}{\theta_{ii}}\right]^{\frac{1}{1-\sigma}} \frac{\Phi_i^L(\underline{w})}{\Phi_i^R(\underline{w}) - (1-\eta)\Phi_i^{\Pi}(\underline{w})} \frac{\tau_{ii}\tilde{w}_i}{\underline{\varphi}_i},\end{aligned} \tag{22}$$

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:

$$\tilde{w}_j L_j = \frac{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^\Pi(\underline{w})}{\Phi_j^R(\underline{w})} \sum_i \theta_{ij} \left(\tilde{v}_i N_i + \Pi_i\right). \tag{23}$$

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. (12) given by

$$V_{ij\nu}(\varphi_j) = \frac{\exp[b_{ij\nu}(\varphi_j)]}{\kappa_{ij}} \frac{\psi_j(\varphi_j) w_j(\varphi_j)}{\left(P_i^Q\right)^\alpha \left(P_i^T\right)^{1-\alpha}}, \tag{24}$$

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

$$\lambda_{ij}(\varphi_j) = \frac{B_{ij}\left[\kappa_{ij}\left(P_i^Q\right)^\alpha \left(P_i^T\right)^{1-\alpha}\right]^{-\varepsilon} \left[\psi_j(\varphi_j) w_j(\varphi_j)\right]^\varepsilon}{\sum_r \sum_s B_{rs}\left[\kappa_{rs}\left(P_r^Q\right)^\alpha \left(P_r^T\right)^{1-\alpha}\right]^{-\varepsilon} \int_{\varphi_s} \left[\psi_s(\varphi_s) w_s(\varphi_s)\right]^\varepsilon d\varphi_s}. \tag{25}$$

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

$$S_j^h \equiv \frac{\sum_i B_{ij}\left[\kappa_{ij}\left(P_i^Q\right)^\alpha \left(P_i^T\right)^{1-\alpha}\right]^{-\varepsilon}}{\sum_r \sum_s B_{rs}\left[\kappa_{rs}\left(P_r^Q\right)^\alpha \left(P_r^T\right)^{1-\alpha}\right]^{-\varepsilon} W_s^\varepsilon} N, \tag{26}$$

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 C.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. (18) is proportional to the average wage $\tilde{w}_j$ in location j

$$\begin{aligned} W_j &= \chi_W^{\frac{1}{\varepsilon}} \Phi_j^W(\underline{w})^{\frac{1}{\varepsilon}} M_j^{\frac{1}{\varepsilon}} w_j^u(\underline{\varphi}_j) \\ &= \Omega_j(\underline{w}) \tilde{w}_j \frac{\chi_L}{\chi_R} \chi_W^{\frac{1}{\varepsilon}} M_j^{\frac{1}{\varepsilon}}, \text{where} \\ \Omega_j(\underline{w}) &\equiv \frac{\eta \Phi_j^W(\underline{w})^{\frac{1}{\varepsilon}} \Phi_j^L(\underline{w})}{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^\Pi(\underline{w})} \end{aligned} \tag{27}$$

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 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.

$$\lambda_{ij} = \int_{\varphi_j} \lambda_{ij}(\varphi_j)d\varphi_j = \frac{B_{ij}M_j\left[\frac{\Omega_j(\underline{w})\tilde{w}_j}{\kappa_{ij}\left(P_i^Q\right)^{\alpha}\left(P_i^T\right)^{1-\alpha}}\right]^{\varepsilon}}{\sum_r\sum_s B_{rs}M_j\left[\frac{\Omega_s(\underline{w})\tilde{w}_s}{\kappa_{rs}\left(P_r^Q\right)^{\alpha}\left(P_r^T\right)^{1-\alpha}}\right]^{\varepsilon}}, \tag{28}$$

From Eq. (28), we obtain the residential choice probability λ_i^N and the workplace choice probability λ_j^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 number of workers *working* at j, L_j to the 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):

$$L_j = \frac{L_j}{H_j}\lambda_j^H N = \frac{\Phi_j^L(\underline{w})}{\Phi_j^H(\underline{w})}\lambda_j^H N, \tag{29}$$

with the second equality following from results derived in Appendix C.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:

$$\lambda_{ij|i}^N \equiv \frac{\int_{\varphi_j}\lambda_{ij}(\varphi_j)d\varphi_j}{\lambda_i^N} = \frac{B_{ij}M_j\left[\Omega_j(\underline{w})\frac{\tilde{w}_j}{\kappa_{ij}}\right]^{\varepsilon}}{\sum_s B_{is}M_s\left[\Omega_s(\underline{w})\frac{\tilde{w}_s}{\kappa_{is}}\right]^{\varepsilon}}, \tag{30}$$

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 the measure of workers L_j employed in location j with the measure of workers that choose to commute to that location and that are successful in finding a job, namely,

$$L_j = \frac{L_j}{H_j} \sum_i \lambda^N_{ij|i} N_i = \frac{\Phi^L_j(\underline{w})}{\Phi^H_j(\underline{w})} \sum_i \lambda^N_{ij|i} N_i. \tag{31}$$

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

$$\tilde{v}_i = \sum_j \lambda^N_{ij|i} \frac{L_j}{H_j} \tilde{w}_j = \sum_j \lambda^N_{ij|i} \frac{\Phi^L_j(\underline{w})}{\Phi^H_j(\underline{w})} \tilde{w}_j. \tag{32}$$

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

$$\overline{V} = \left\{ \sum_i \sum_j B_{ij} M_j \left[\frac{\Omega_j(\underline{\omega}) \tilde{w}_j}{\kappa_{ij} \left(P^Q_i\right)^\alpha \left(P^T_i\right)^{1-\alpha}} \right]^\varepsilon \right\}^{\frac{1}{\varepsilon}}. \tag{33}$$

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. (33) 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 D.2, we can express the labour force participation rate as

$$\mu = \frac{\overline{V}^\zeta}{\overline{V}^\zeta + A}, \tag{34}$$

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

$$\sum_j H_j = \mu \overline{N}, \tag{35}$$

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:

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

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. (23); average residential income from Eq. (32); firm entry from Eq. (17); housing market clearing from Eq. (15); aggregate local employment from Eq. (31); $N_i = \lambda_i^N N$ based on Eq. (28) and the price index from Eq. (22). The conditions needed to determine the scalars $\{\mu, \overline{V}\}$ are labour force participation from Eq. (34) and the labour market clearing condition from Eq. (35).

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 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 Section D.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 in 2014, we get $\bar{N} = N/\mu$.

Reservation utility heterogeneity (ζ). As we show in Appendix D.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 use a novel estimation strategy that leverages on our firm-level data. We exploit the firm-level wage and firm size scale in firm productivity at elasticities that differ by multiplicative factor ε (see Table A3). 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 area fixed effects. Our estimate of $\varepsilon = 5.2$ is in between (Monte et al., 2018), who use larger spatial units, and (Ahlfeldt et al., 2015), who use smaller spatial units. We refer to Appendix D.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 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 $\tilde{w}_j$. The conventional reduced-form approach in the literature emphasizes the dispersion. However, in the context of the minimum wage evaluation, the left tail of the distribution is of particular relevance. Moreover, the aggregation of firm level outcomes requires several parameter constraints to be met. Therefore, we take our estimate of $\varepsilon = 5.2$ as given and nest the estimation of k using a GMM estimator into a grid search over σ values. We choose $\sigma = 1.5$ 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 0.53. 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 D.3.2 for details.

Minimum wage ($\underline{w}$). 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 minimum 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 an 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} at the bilateral area level in a theory-consistent way. We refer to Appendix Section D.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. (23) (substituting in Eq. (20)) using a conventional fixed-point solver. We refer to Appendix Section D.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. (28) 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. (16).

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. (15).

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 and set parameter values for $\{\alpha, \varepsilon, \sigma, \zeta, k\}$, we invert fundamental utility A using Eqs. (22), (33) and (34).

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 D.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 $\mathbf{X^0}$. We use these, instead of the observed values in the data, 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. 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. Using 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 $\Delta\mathbf{X} = \widehat{\mathbf{X}} \cdot \mathbf{X^D}$, where $\mathbf{X^D}$ indicates values observed in data. Whenever we refer to welfare, we use the expected utility $\overline{\mathcal{V}}$ from Eq. (36) 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 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 D.4.2.

4.4 The German minimum wage

We now use the model to quantitatively evaluate the effects of the German minimum wage in 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 as of now, 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 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. The high-level conclusion is that the minimum wage increases welfare at the cost of reducing employment. Given a workforce of approx. 30M, the 0.3%-reduction in employment translates into about 100k jobs lost, which is less than extant ex-ante predictions based on competitive labour market models (Knabe et al., 2014). Applying the relative welfare effect 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 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%. Intuitively, these workers become active

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)	-0.270	-23.37	5.300	-0.350	-32.88	5.440
Labour supply at residence (N)	0.590	0.110	1.500	0.590	-6.800	14.05
Employment probability (L/H)	-0.830	-19.78	0	-0.890	-20.82	0
Panel b: Wage and prices						
(Normalized) wage ($\tilde{w}$)	0.330	-1.350	25.27	0.400	-1.090	24.35
Real tradables price index (P^Q)	-3.010	-4.470	-2.150	-2.900	-5.490	-1.510
Real housing rent (P^T)	-1.100	-7.050	1.080	-1.130	-5.370	2.370
Panel c: Welfare components						
Exp. real wage $\tilde{v}\left[(P^Q)^{\alpha}(P^T)^{(1-\alpha)}\right]$	1.610	-0.310	4.870	1.620	0.390	4.240
# establishments (M)	-0.110	-11.42	1.010	-0.120	-24.66	2.800
Ease of commuting ($B\kappa^{-\epsilon}$)	1.070	-8.260	7.100	0.840	-21.02	8.490
Panel d: Welfare						
Worker welfare | working (V)	2.920	0.530	7.580	2.860	2.860	2.860
Worker welfare, all ($\mathcal{V}$)	2.160	0.390	5.590	2.120	2.120	2.120

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.

on the labour market because they prefer well-paid jobs with a lower hiring probability over badly paid jobs with a higher hiring probability. Thus, although the number of jobs decreases, the number of workers *working* or *searching* increases by about 180K. 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, implying that workers find jobs in more convenient reach. Since Figure 2 does not point to a significant reduction in commuting distance, it is likely that the increase in ease of commuting is driven by workers finding jobs at places that are well connected by transport infrastructure. 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 2.2%, ($\mathcal{V}$) is smaller that 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 job loss whereas employment increases in others. This mirrors a highly heterogeneous increase in real wages. 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.

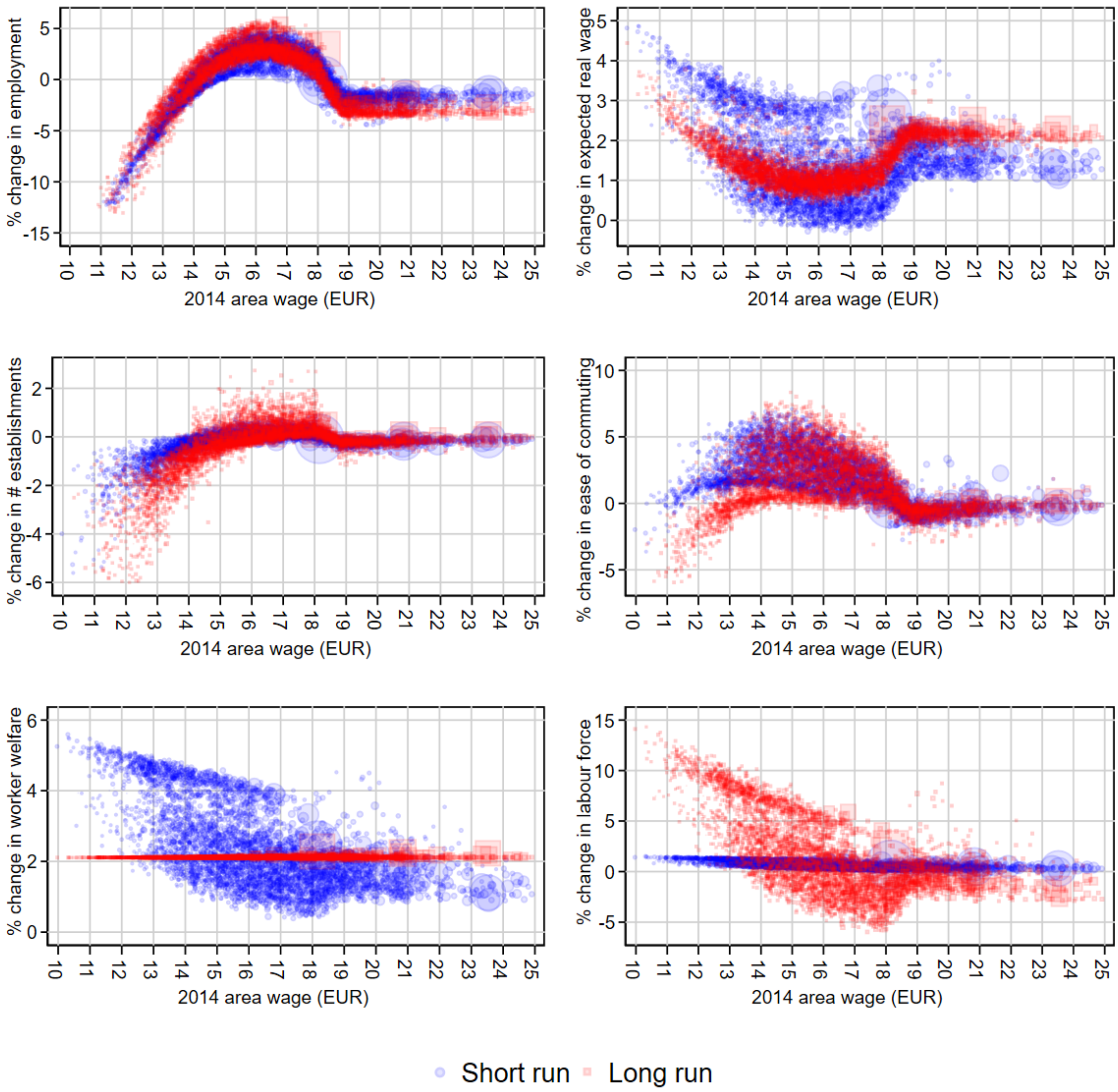
We dig deeper into the spatial heterogeneity in minimum wage effects in Figure 6, where we correlate our simulated relative changes in selected outcomes to the 2014 mean wage observed in our data. Expectedly, the employment effect follows the hump-shaped pattern that we have derived theoretically and substantiated empirically in partial equilibrium in Section 3. It is straightforward to eyeball the critical points introduced in Figure 4. 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 5, 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 5. Intuitively, the expansion of employment and production in municipalities of intermediate productivity comes at the expense of the most productive municipalities since aggregate demand is, albeit endogenous, finite. In keeping with intuition, this displacement effect is reinforced by worker mobility in the long-run.

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 reallocation effect of workers towards more productive establishments further away from their residences in low-productivity regions (Dustmann et al., 2021). 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.

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 7 shows that municipalities experiencing real wage growth and a reduction in employment probability are over-represented in the east (resembling the minimum wage bite in Figure 1). Because the former dominates the latter, expected real wages increase (see top-middle panel in Figure 6). The bottom panels of Figure 6 and 7 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

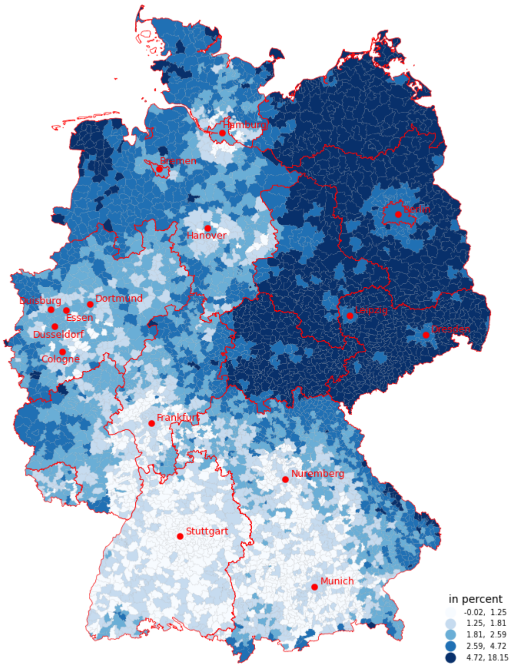
Figure 6: Short-run and long-run effects by regional productivity



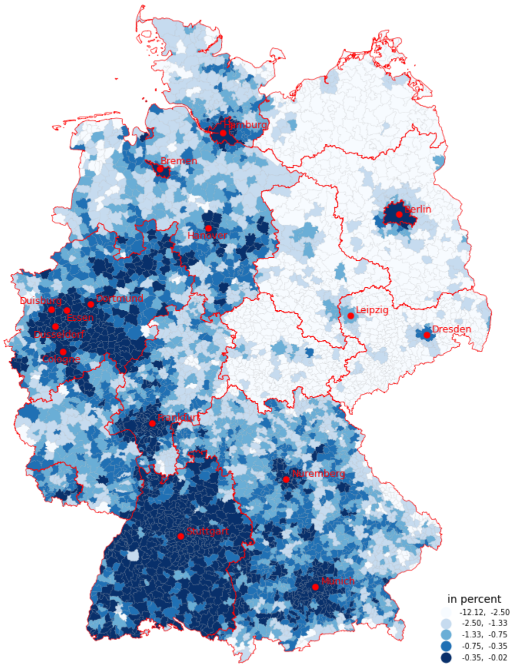
Note: Each icon represents one outcome for one area (*Verbandsgemeinde*). Results of model-based counterfactuals comparing the equilibrium under a federal minimum of 48% (the value observed in data) of national mean wage to the equilibrium with a zero minimum wage. *Blue circles* show outcomes when workers 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).

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.

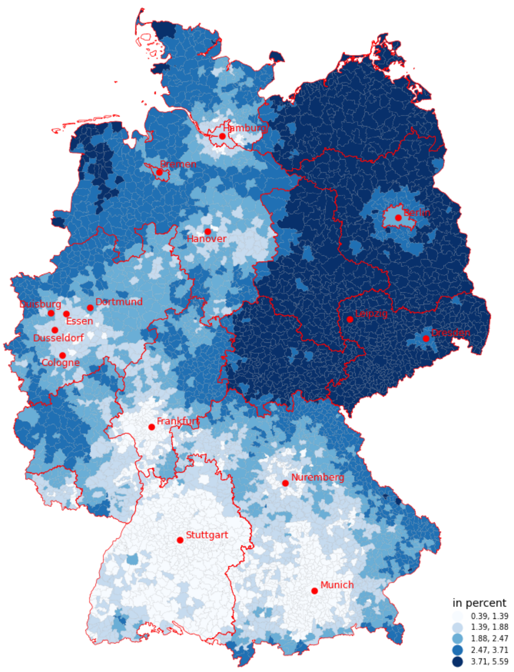
Figure 7: Regional effects of the German minimum wage



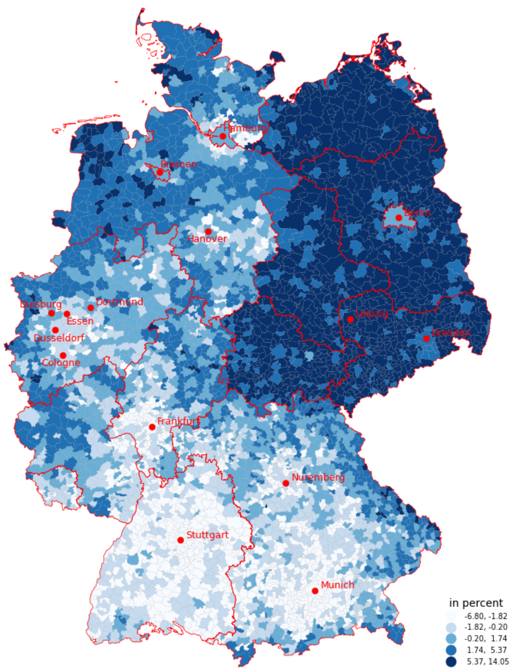
(a) Real wage, short run



(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. (32) 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.

4.4.2 Comparison to data

To evaluate whether the model successfully predicts observed changes in the data, we use the area-based minimum wage effects discussed in Section 4.4.1 as treatment variables in a conventional dynamic difference-in-differences estimation. Intuitively, this approach compares before-after changes in selected outcomes in the data to respective before-after changes predicted by the model. For further detail on the empirical specification and the interpretation of the estimated treatment effect, we refer to Appendix D.5.

Figure 8 summarizes the results. The first insight 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, 2021). While Germany is no exception (Mindestlohnkommission, 2020), evidence from labour force surveys suggests that compliance has increased over time (Weinkopf, 2020).

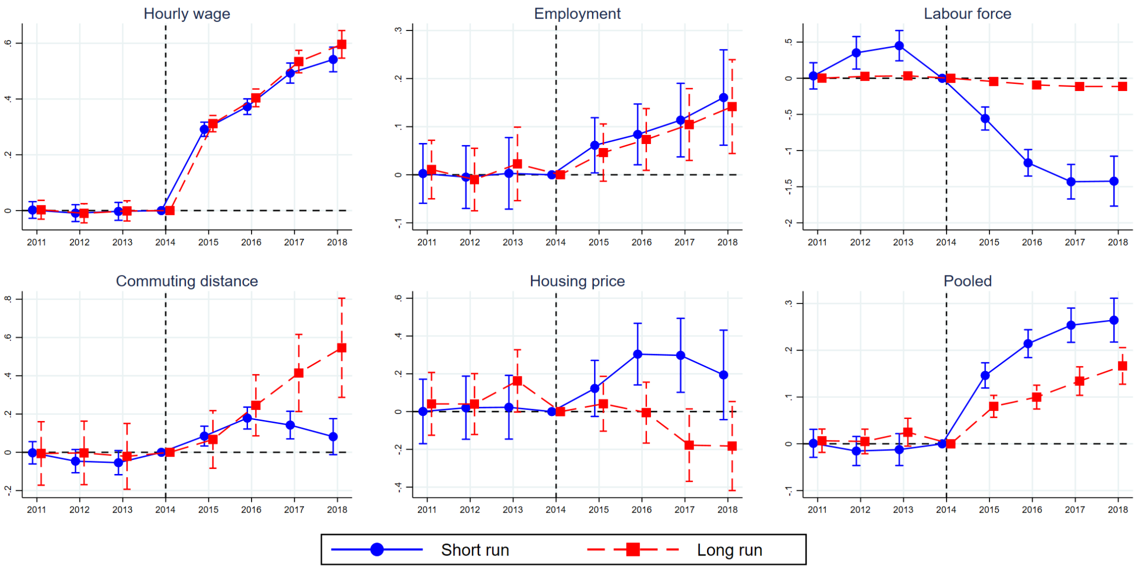
In contrast, workers do not seem to have started to relocate to regions with positive short-run welfare gains within the first four years of the policy. If anythings, those regions, which are mostly located in the East (see Figure 7), have continued to lose population, suggesting that migration decisions have been dominated by other forces. Notice that the correlations with the short-run predictions need to be taken with a grain of salt since—given fixed residential locations—the model predicts only small changes in the local labour force (see Table 1) owing to some workers starting searching for jobs, for which we do not have a good equivalent in the data.

For commuting distances, both short-run and long-run predictions are positively correlated with before-after changes in the data, but, as time goes by, the long-run predictions start outperforming the short-run predictions. Given that few workers appear to have re-optimized their residential locations with respect to the effects of the minimum wage, it is no surprise that we find positive correlations between predicted and observed changes for house prices based on the short-run predictions, but not for the long-run predictions.[10] Finally, a pooled regression in which we evaluate the predictive power for all outcomes simultaneously suggests that, overall, our short-run predictions provide a better description of the minimum-wage impact. We conclude that, four years after the introduction in Germany, the full effect of the minimum wage law has not yet materialized in the data.

This impression is confirmed by 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), which we can derive within our model as discussed in Appendix D.6.1. 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%

[10]This confirms Yamagishi (2021), who shows that desirable minimum wages increase housing rents.

Figure 8: 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). Icons denote point estimates. Error bars give 95% confidence bands. In the bottom-left panel, we pool over all outcomes, using area-by-outcome and year-by-zone-by-outcome fixed effects.

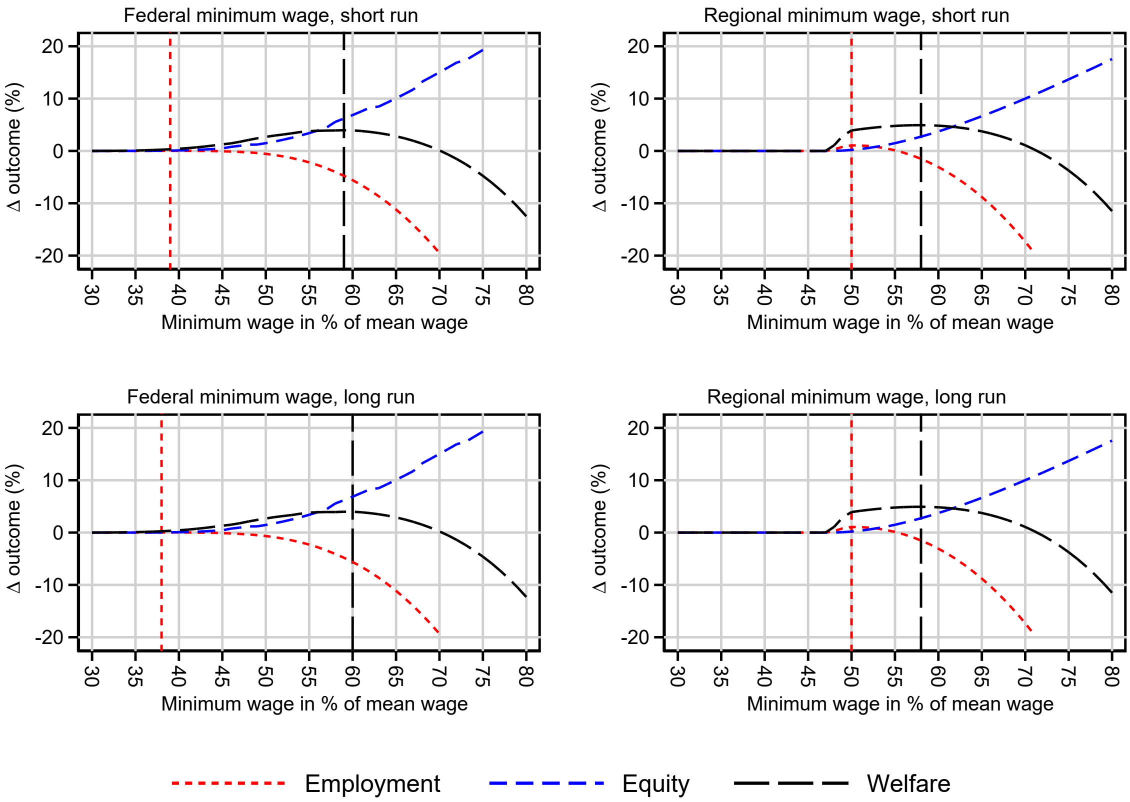
during the first three years of the minimum wage (see Appendix D.6.2 for details).

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 $\underline{\mathbf{w}}^{\mathbf{r}}_{\mathbf{j}} = \mathbf{w}^{\mathbf{m}}_{\mathbf{j}} \cdot \underline{\mathbf{w}}^{\mathbf{n}}$ that represents the regional minimum wage as a fraction of the $J \times 1$ vector of regional mean wages $\mathbf{w}^{\mathbf{m}}_{\mathbf{j}}$. As in Section 4.3, we evaluate the effects of both minimum wage schedules in a short-run scenario in which residential locations are fixed, and in a long-run scenario in which workers are fully mobile.

The most obvious optimality criterion for a successful minimum wage policy in the context of our model is expected worker welfare as defined in Eq. (36). 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 equity measure $1 - \mathcal{G}$, where $\mathcal{G}$ is the Gini coefficient of nominal wage inequality across all workers in all regions (see Appendix D.6 for details). Figure 9 summarizes how employment,

Figure 9: Minimum wage effects on employment, equity, 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. Welfare is the expected utility of as defined by Eq. (36). 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.

equity, and welfare effects vary in the level of a federal or regional minimum wage. We also compare the employment, equity, and welfare effects of employment-maximizing and welfare-maximizing federal and regional minimum wages to the effects of the actual German minimum wage 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 that trades aggregate welfare, equity and employment effects.

The first insight is that the welfare effect is hump-shaped in the minimum wage level, whether workers are mobile or not and whether the minimum wage is nationally uniform or regionally differentiated. The intuition is that up to the the welfare-maximizing minimum wage, the positive effect on real wages dominates the negative effect on employment probabilities, such that expected wages and welfare increase. With a federal minimum wage, this point is reached at a level of 60% 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 compresses 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.

Intuitively, the employment-maximizing minimum wage must 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 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 important takeaway is that, in setting federal minimum wages, policy makers trade positive aggregate welfare effects and progressive distributional effects (within employed workers) against negative employment effects.

This trade-off can be mitigated by setting minimum wages at the regional level. 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. 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 Appendix D.8. This confirms the intuition that a minimum wage needs to be sufficiently localized to account for productivity differentials *between* commuting zones since workers can relatively easily re-optimize to heterogeneous effects caused by productivity differentials *within* commuting zones. 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 9 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 6 and 7). 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

Table 2: Minimum wage schedules

		Level rel. to		Employment		Equity		Welfare	
Objective	Scheme	Mean	p50	SR	LR	SR	LR	SR	LR
Actual	Federal	48.00	52.81	-0.26	-0.35	1.20	1.16	2.08	2.11
Employment	Federal	38.00	41.81	0.02	0.02	0.05	0.03	0.24	0.25
Welfare	Federal	58.00	63.82	-3.94	-4.02	5.51	5.51	3.95	3.95
Employment	Regional	50.00	55.01	1.06	1.06	0.19	0.19	3.92	3.92
Welfare	Regional	58.00	63.82	-1.51	-1.51	2.75	2.76	4.93	4.92

Notes: All values are given in %. *Objective* describes if the minimum wage is employment-maximizing or welfare-maximizing. *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. Welfare is the expected utility of as defined by Eq. (36). 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.

benefits many low-productivity municipalities in the eastern states, the regional fortunes reverse under the 25% higher welfare-maximizing federal minimum wage. 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 labour force (see Appendix D.7).

In policy contexts, it is common to express minimum wages in terms of median wages. To convert the relative minimum wages discussed thus far into this metric, we just need to multiply them 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. For convenience, we also report the relative minimum wage in per-median terms in Table 2. Accordingly there is a range between the employment-maximizing and the welfare-maximizing minimum wage from 43-64% of the national median wage within which policy makers trade welfare gains against employment losses. Connecting to the current policy debate in many countries, our simulations suggest that ambitious minimum wages in the range of 60-70%, will likely increase aggregate welfare, but also put a sizable fraction of jobs at risk. In contrast a moderate regional minimum wage, set at about 55% of the municipal median wage, could deliver similar welfare effects *and* generate employment.

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. Some German political parties have recently proposed a minimum wage of €12 that would exceed 70% of the median wage. 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.[11] We inform this debate in a concrete, yet nuanced fashion.

Our simulations within a quantitative model calibrated to German micro-regional data reveal that such ambitious federal minimum wages may achieve a reduction in wage inequality without having a detrimental effect on welfare – compared to the counterfactual of no minimum wage. However, they will likely cause significant job loss. While employment effects remain small up until about 50% of the national mean wage, they build up at an increasing rate at higher levels. Therefore, we caution against extrapolating from encouraging reduced-form evidence on employment effects of moderate minimum wages to the likely effects of more ambitious levels. We recommend that ambitious minimum wages are implemented in small steps, under careful evaluation of short-run employment effects so that potential tipping points can be detected in time.

More generally, our results illustrate how the desirability of any minimum wage will depend on the considered relative level and the social welfare function. In setting minimum wages, policy makers trade employment, equity, and, welfare effects, and, depending on priorities, different minimum wage levels will be optimal. As an example, maximizing employment requires setting a relatively low level—in the case of Germany about 42% of the national median wage. This should generate a small positive employment effect, but also negligible equity and welfare effects. Maximizing welfare requires a more ambitious minimum wage of 64% of the national median wage, which will also lead to greater reduction in nominal wage inequality. Any increase in the minimum wage level within these bounds will trade positive equity and welfare effects against negative employment effects. Of course, even higher minimum wage levels can be advocated on the grounds of an equity objective. In fact, our simulations suggest that the minimum wage could be set as high as 77% of the national median wage before the welfare effect would turn negative. However, recommending such a high minimum wage level would imply that one strictly cared about the expected real wage—the product of the real wage earned conditional on being in employment and the employment probability—without any aversion to high unemployment rates.

While these trade-offs may appear frustrating from a policy perspective, our analysis also reveals some more encouraging news. Rather than 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 negative employment effects and positive welfare effects. To illustrate the potential, the employment-maximizing minimum wage, at 50% of the municipality mean wage, could

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

increase welfare by 4%—as much as the welfare-maximizing federal minimum wage—*and* generate a sizable positive employment effect of 1.1%.

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.*

_ , Stephan Heblich, and Tobias Seidel, "Micro-geographic property price and rent indices," *CEP Discussion Paper*, 2021, *1782.*

_ , 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.

Almagro, Milena and Tomas Domínguez-Iino, "Location sorting and endogenous amenities: Evidence from Amsterdam," *Working paper*, 2021.

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.

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.

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.

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.

Christl, Michael, Monika Köppl-Turyna, and Dénes Kucsera, "Revisiting the Employment Effects of Minimum Wages in Europe," *German Economic Review*, 2018, *19* (4), 426–465.

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*, 2021, p. forthcoming.

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*, 2021, p. forthcoming.

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.

Garnero, Andrea and Claudio Lucifora, "Turning a "Blind Eye"? Compliance with Minimum Wage Standards and Employment," 2021.

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

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.

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

Heblich, Stephan, Stephen J. Redding, and Daniel M. Sturm, "The Making of the Modern Metropolis: Evidence from London," *The Quarterly Journal of Economics*, 2020, *135* (4), 2059–2133.

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

Jha, Priyaranjan and Antonio Rodriguez-Lopez, "Trade, inequality, and welfare with Monopsonistic Labor Markets," *Working paper*, 2021.

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.

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.

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?," *National Bureau of Economic Research Working Paper Series*, 2021, *No. 28388.*

— **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," *unpublished manuscript, Banco de Mexico*, 2018, pp. 1–75.

Ragnitz, Joachim and Marcel Thum, "Beschäftigungswirkungen von Mindestlöhnen – eine Erläuterung zu den Berechnungen des ifo Instituts," *ifo Schnelldienst*, 2008, *1.*

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.

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.

Tsivanidis, Nick, "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's TransMilenio," *Working Paper*, 2019.

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.

APPENDIX FOR ONLINE PUBLICATION

This section presents an online appendix containing complementary material not intended for publication. It does not replace the reading of the main paper.

A Literature

This section complements Section 1 by providing a more complete discussion of the vast literature on the impact of the German statutory minimum wage (an overview of the extant literature can also be found in Caliendo et al. (2019), while Möller (2012) and Fitzenberger and Doerr (2016) discuss research on earlier sector-specific minimum wages in Germany).

The national minimum wage in Germany came into effect on 1 January 2015 (see Section B.1) and its introduction has been followed by a large amount of research on the effects that this policy has had on a variety of outcomes. A specificity of the German minimum wage is that it—with only a few exceptions—applies to all workers who earn less than the specified threshold. In contrast to the US literature in which the effects of the minimum wage are often identified from state-specific changes in minimum wage levels and where comparable workers from unaffected states can serve as a control group (e.g. Dube et al., 2010), such an approach is not feasible in Germany. Moreover, the possibility of spill-over effects makes it difficult to infer the effects of the minimum wage from a comparison of worker below and above the minimum wage threshold. Many empirical studies have therefore used a difference-in-differences approach in which the effects of the minimum wage are identified from the variation in the extent to which workers in given entities are directly affected by the introduction of the minimum wage—the regional minimum wage bite defined in Machin et al. (2003) being an example. Before turning to the evidence on the effects of the German statutory minimum wage, we provide a short description of the data sets that have been used in the empirical research will discuss.

Data sets. The evaluation of the effects of the German minimum wage is not restricted to a single data source. Most studies have, however, used one of the following data sets (a more detailed description can be found in Mindestlohnkommission (2020):

- The **German Socioeconomic Panel (SOEP)** is an annual survey currently consisting of a representative sample of about 15,000 households and 30,000 individuals, which was first conducted in 1984. Relevant for minimum wage research, participants provide information about their weekly working hours (actual and contractual) and monthly labour income which can be used to construct an estimate of hourly wages. Due to its comparatively small sample size, the potential for a regionally differentiated analysis are limited. Further information on SOEP can be found in Goebel et al. (2019).

- The **Structure of Earnings Survey (SES)** is mandatory establishment survey that is carried out by the German Statistical Offices. First carried out in 1951, it has been conducted every four years since 2006. The most recent survey refers to the year 2018 and contains information on approximately 60,000 establishments and 1,000,000 employees. As in the case of SOEP, the SES contains information about working hours and monthly earnings which can be used to estimate hourly wage rates and to determine whether a person earns more or less than a given minimum wage level. Evaluation of the effects of the minimum wage is facilitated by the availability of additional earnings surveys that have been conducted in years in which the SES was not carried out. Compared to the SES, these data sets are considerably smaller (between 6,000 and 8,000 establishments) and participation is not mandatory.

- The **Integrated Employment Biographies (IEB)** is prepared by the Institute for Employment Research (IAB) and covers episodes of employment, unemployment and participation in measures of active labour market policies for the majority of labour market participants in Germany (certain groups are, however, not covered: e.g. employment records do not contain information about civil servants or the self-employed). Employment records are based on mandatory notifications made by employers for the social security systems and, as such, are highly reliable. One advantage of the IEB is its size, which makes it possible to conduct analyses for specific groups or at a regionally differentiated level. A disadvantage in terms of minimum wage research is the fact, that the data set does not contain working hours which makes it necessary for this information to be provided by other data sources (see Section B.2.1).

- The **IAB Establishment Panel** is an annual establishment survey that is carried out by IAB. It covers a representative sample of about 15,000 establishments. The survey contains a unique establishment ID which can be used to link the survey with administrative data on the employees of the sampled establishments. Further information on the IAB Establishment Panel can be found in Ellguth et al. (2014).

- The **Federal Employment Agency** provides administrative statistics on various labour-market outcomes, such as employment levels (e.g. by year, region, sector of for various demographic groups).

Hourly wage outcomes. The extant literature has provided ample evidence that the introduction of the minimum wage has led to an increase in *hourly* wages at the lower end of the wage distribution. Burauel et al. (2020b) use SOEP data to estimate wage effect of the minimum wage introduction in a differential trend-adjusted difference-in-differences (DTADD) framework. Their results show that—conditional on their respective wage growth trends—workers, who initially earned less than the minimum wage, experienced an increase in hourly wage of 6.5% between 2014 and 2016 compared to workers above the minimum wage level. Evaluated at the mean hourly wage of workers in the treatment group, this suggests an increase of about €0.45 per hour. Qualitatively similar results are obtained by Caliendo et al. (2017) who also use SOEP data, but identify the effect of the minimum wage wage from the variation in the regional minimum wage bite, i.e. the share of workers who initially earned below the minimum wage threshold. Their findings show that a higher minimum wage bite is associated with faster hourly wage growth in the year 2015 (i.e. following the introduction of the minimum wage) for workers in the lowest quintile of the hourly wage distribution, while no significant effects are found for workers in higher quintiles. Dustmann et al. (2021) and Ahlfeldt et al. (2018) also use variation in the regional exposure to the minimum wage (in form of the Kaitz index and the minimum wage bite, respectively) to evaluate the impact on hourly wages in a difference-in-differences framework. Based on data from the IEB, their results suggest that regions with a higher degree of exposure experienced faster hourly wage growth at the lower end of the hourly wage distribution. Evidence by Fedorets and Shupe (2021) suggests that the introduction of the minimum wage not only affected realised hourly wages, but also led to an adjustment of reservation wages. Using SOEP data, the authors find that reservation wages increased considerably among non-employed job seekers. This adjustment, however, appears to have been temporary as reservation wages are found to return to their initial level. Even if only temporary, an increase in reservation wages represents a possible reason for why minimum wages may not lead to higher labour market participation.

Hours worked and monthly wage outcomes. While evidence from different studies, using different data sources and identification strategies, have provided comparable evidence of a positive effect on hourly wages, it is ex ante unclear whether this finding also carries over to monthly labour earnings. The reason for this is that, faced with a higher cost per working hour, employers might choose to reduce the number of hours offered to minimum wage workers. In such a case, the impact of the minimum wage on monthly outcomes would be ambiguous and depend on whether the positive effect on hourly wages outweighed the potentially negative effect on the number of hours worked. An analysis by Burauel et al. (2020a) concludes that the number of contractual hours decreased by 5% in the year 2015 among workers who initially earned below the minimum wage level. No significant reduction is found, however, for the year 2016. This pattern corresponds with findings provided by Burauel et al. (2020b). According their these results, worker who initially earned below the minimum wage, did not experience a significant increase in monthly earnings (relative to workers from the control group) in 2015, but realised a 6.6% increase in the year 2016. Similar results are provided by Caliendo et al. (2017) for the year 2015. Slightly different results are provided by Bossler and Schank (2020). Based on IEB data and adopting a difference-in-differences framework based on the regional minimum wage bite, they find a statistically significant increase in monthly wage earnings in regions with a higher minimum wage bite from the year 2015 onward.

Wage spillovers. While minimum wages directly affect the wages of workers earning less than the specified threshold, there can also be effects on workers higher up the wage distribution. One reason for such spillover effects is that employers want to retain initial pay differences and therefore decide to also raise wages of workers above the threshold. Bossler and Gerner (2019) provide direct evidence on the extent of wage spillovers using information from the IAB Establishment Panel in which employers were asked whether they adjusted the remuneration of workers earning above the minimum wage threshold in response to the policy. Less than 5% of establishments in their sample report to have made such an adjustment. The analysis by Burauel et al. (2020b) relies on the assumption that the control group of workers above the minimum wage threshold is not affected by wage spillovers. To validate this assumption, they estimate the wage effects using a control group of workers further up the wage distribution, which yields comparable results. Based on the assumption that spillovers are likely to affect workers close to the minimum wage threshold, they conclude that spillover effects are limited. Dustmann et al. (2021) assess the existence of wage spillovers by comparing the change in two-year wage growth for the years following the introduction of the minimum wage between workers in different wage bins. As expected, excess wage growth (relative to the reference period 2011-13) is particularly pronounced for workers who initially earned less than the minimum wage. However, an increase in wage growth—though smaller—is also found up to the 12.50 € per hour bin, which suggests that the minimum wage also had an effect on workers above the threshold. Bossler and Schank (2020) find that the introduction of the minimum wage had an effect on monthly labour income up to the 50^{th} percentile.

Wage inequality, welfare receipt and in-work poverty. As described above, the introduction of the minimum wage led to an increase in wages at the lower end of the wage distribution. As such, it has been hypothesised that the minimum wage also contributed to a reduction in lower-tail wage inequality. It is, however, difficult to evaluate ex ante to what extent this is the case, as non-compliance or spillover effects might reduce the impact of the minimum wage. According to Bossler and Schank (2020) the minimum wage contributed considerably to the reduction in wage inequality. Based on counterfactual analyses, the authors conclude that between 40% and 60% of the observed decrease in wage inequality, as measured by the variance of log monthly wage earnings, can be ascribed to the introduction of the minimum wage. While wage income represents a worker-level outcome, poverty status and the eligibility of welfare benefits are determined on the basis of household-level income. In contrast to its effect on wages and wage inequality, existing evidence suggests that the minimum wage introduction only had a limited impact on welfare receipt and (in-work) poverty. According to results by Bruckmeier and Bruttel (2021), the minimum wage neither exerted downward pressure on the number of employees receiving top-up benefits nor did it alleviate poverty rates. Among other factors, the authors explain the absence of any sizeable effect by the fact that low household income is more often due to a low number of hours worked rather than a low hourly wage. Moreover, they argue that low-wage workers are not restricted to low-income households, but can rather be found throughout the household income distribution, so that a policy that increase the wages of low-wage workers does not necessarily improve the situation of low-income households.

Employment and unemployment. In a perfectly competitive labour market, a binding minimum wage will unambiguously lead to a lower equilibrium level of employment. As outlined in Section 3.1.2, this need not be the case in a monopsonistic labour market. From a theoretical perspective, the extent and sign of the employment effect of a minimum wage are, therefore, ex-ante unclear. A considerable amount of research has evaluated the impact that the introduction of the German minimum wage had on employment and unemployment. In contrast to the analysis presented in this paper, these studies are, however, based on partial equilibrium analysis. Caliendo et al. (2018) provide one of the earliest evaluations of the employment effects of the German minimum wage. Combining data from the SES and administrative statistics, their identification strategy rests on the regional variation in the extent to which the minimum wage "bites" into the wage distribution (measured by the minimum wage bite or the Kaitz index). Their findings suggest that the effect of the minimum wage differed substantially between regular and marginal employment. Specifically, they estimate that the introduction of the minimum wage reduced the number of marginal employment jobs by 180,000 in 2015, while the effect on regular employment is smaller and not statistically significant in all specifications. Similar results are obtained by two other studies: Schmitz (2019), who uses administrative statistics from the Federal Employment Agency, and Bonin et al. (2020), who combine SES data with administrative statistics, also find that there was a small negative effect on overall employment, which was driven mainly by a reduction in the number of marginal employment jobs. Schmitz (2019) estimates that the minimum wage led to a decrease of about 200,000 marginal employment jobs in 2015). Moreover, Bonin et al. (2020) do not find any evidence for a corresponding increase in unemployment. A possible explanation for the absence of such an effect is that workers, who were negatively affected by the introduction of the minimum wage, withdrew from the labour market. Slightly different results are reported by Holtemöller and Pohle (2020), who use variation in the exposure to the minimum wage across federal state-sector cells. Based on administrative statistics from the Federal Employment Agency, their results confirm previous findings that the introduction of the minimum wage led to a decrease in marginal employment (between 67,000 and 129,000 jobs, depending on the chosen specification). However, they also find a positive effect on regular employment in the range of 47,000 to 74,000 jobs. Interestingly, they do not find any evidence for a substitution of marginal for regular employment. Garloff (2019) also uses data from the Federal Employment Agency and exploits the variation in the minimum wage bite across regions and demographic groups or sectors. As in Holtemöller and Pohle (2020), his results show a negative relationship between the minimum wage bite and the development of marginal employment as well as a positive relationship with regular employment. With respect to overall employment, he finds a small positive association between the bite and the growth of total employment which amounts to approximately 11,000 additional jobs in the first year after the introduction of the minimum wage. Small positive effects of an increase in the minimum wage bite on total employment are also reported by Ahlfeldt et al. (2018) who use IEB data for their analysis. In contrast to the studies discussed above, which use the regional variation in the exposure to the minimum wage, Bossler and Gerner (2019) estimate the employment effects of the introduction of the minimum wage from the variation in establishment-level exposure. The authors use the IAB Establishment Panel to identify whether an establishment has at least one employee whose wage is directly affected by the policy. Comparing the development of employment among the treated establishments with a control group of unaffected establishments within a difference-in-differences framework, the authors find a reduction in employment in the post-treatment years among treated establishments of 1.7% as opposed to the control group. This result suggests that employment was lower by between 45,000 and 68,000 jobs at treated establishments as a result of the minimum wage introduction. The authors also provide evidence on the underlying mechanisms: according to their results, the negative employment effect is driven by a reduction in hires rather than by an increase in layoffs. Friedrich (2020) evaluates the impact that the minimum wage had on employment using the differential exposure to the policy between occupations. Consistent with the results from other contributions to the literature, he estimates that by the year 2017 the minimum wage (including its increase to a level of 8.84 € in 2017) led to a loss of approximately 50,000 jobs. This reduction is primarily driven by a decrease in marginal employment. Moreover, his findings suggest that there are considerable regional differences in the employment effects. Whereas, at least initially, the loss of marginal jobs was accompanied by an increase in regular employment in West Germany, such a compensating effect is not found for East Germany. While the employment effects that have been estimated by the extant literature differ in terms of size and sign, estimates of potential employment losses appear to be modest and considerably smaller than the large-scale job loss that was discussed before the introduction of the policy (e.g. Knabe et al., 2014).

Worker reallocation. Despite an absence of large-scale disemployment effects, the minimum wage introduction led to considerable changes in the structure of employment. Dustmann et al. (2021) provide evidence for a systematic reallocation of low-wage workers from lower-quality to higher-quality establishments. While the authors do not find that the minimum wage increased the share of workers who changed their employer, those workers who did so between 2014 and 2016 moved to establishments whose average daily wage was approximately 1.8% higher (relative to the corresponding change in establishment-level pay between 2011 and 2013). Evaluated for all workers who initially earned less than the minimum wage and who switched to a higher-paying establishment, this upgrade accounts for approximately 17% of the minimum wage-induced increase in daily wages. Receiving establishments are found to be significantly larger and to employ a higher share of full-time as well as university-educated workers. Moreover, the upgrade in establishment-level average daily wages can be almost exclusively ascribed to changes between establishments within in the same region, while about two thirds of the upgrade is associated with changes within the same three-digit industry, suggesting that worker reallocation is not driven by either regional or sectoral mobility.

Price pass-through and other establishment-level outcomes. Evidence on whether and to what extent firms in Germany adjusted their prices in response to the introduction of the minimum wage is limited. An exception is the study by Link (2019) whose results suggest that a substantial share of the increased costs induced by the minimum wage were passed on to consumers in the form of higher prices. Based on data from the ifo Business Survey–a monthly survey consisting of approximately 5,000 establishments from the manufacturing as well as the service sector in Germany—, he analyses how the extent of the sector-location-specific minimum wage bite is related to the probability of a firm planning to adjust prices. According to his results, there is a positive association around the time of the introduction of the minimum wage. Moreover, the results suggest that a minimum wage-induced increase in costs of 1% is associated with an increase in prices by 0.82%. No substantial difference is found between firms in the manufacturing and the service sector. However, the extent of price pass-through is estimated to be more pronounced when firms face less competition. Bossler et al. (2020) provide evidence on further channels through which establishments might have adjusted to the introduction of the minimum wage. Using data from the IAB Establishment Panel, they show that treated establishments, i.e. those employing at least one worker in the year 2014 earning less than 8.50 € per hour, experience an increase in labour costs in the years 2015 and 2016. In terms of investments, the results show a small and statistically insignificant reduction in the volume of investment in physical capital per employee following the introduction of the minimum wage. Likewise, the authors find no evidence that treated establishments adjusted investment in apprenticeship training — measured either as the share of apprentices per establishment or the number of apprenticeship offers per employee. However, the results point towards a small, but statistically significant reduction in the intensity of further training in the year 2015, measured by the share of employees receiving further training per establishment. This result is consistent with evidence by Bellmann et al. (2017) who also report a decrease in training intensity among treated establishments.

B Empirical context

B.1 The German minimum wage

This section complements Section 2 in the main paper. A statutory minimum wage, initially set at a level of €8.50 per hour, came into effect in Germany on 1 January 2015, having been ratified by Parliament on 3 July 2014. While the minimum wage, in principle, applies to all employees aged 18 years or older, certain groups are exempted: apprentices conducting vocational training, volunteers and internships as well as the long-term unemployed during the first six months of employment. Moreover, exemptions were made for existing sector-specific minimum wages that fell short of the level of the statutory minimum wage until 1 January 2017, when the value of €8.50 also applied in these cases. The number of employees covered by sector-specific minimum wages that were temporarily exempted from the new statutory minimum wage is comparatively small and has been estimated at approximately 115,000 by the Federal Statistical Office (Mindestlohnkommission, 2016).

The level of the statutory minimum wage is determined by the Minimum Wage Commission which consists of a chair person, three representatives each of employers and employees as well as two academic representatives (though, the latter two are not eligible to cast a vote). Following its introduction, the minimum wage has since been raised several times: to a level of €8.84 per hour from 1 January 2017 onward, €9.19 from 1 January 2018, €9.35 from 1 January 2021 and €9.60 from 1 July 2021. Further increases are scheduled for 1 January 2022 (€9.82) and 1 July 2022 (€10.45), while several political parties have campaigned for an increase of the minimum wage to a level of €12 per hour in the run-up to the 2021 Parliamentary elections. In deciding on adjustments to the level of the minimum wage, the Commission takes the development of collectively bargained wages into consideration. Further information on the statutory minimum wage in Germany can be found in Mindestlohnkommission (2016).

Table A1 shows the Kaitz index, the ratio of the minimum wage to the median wage, for the years 2015 to 2018. For full-time workers, the Kaitz index is fairly stable for the first three years, before rising slightly in 2018.

Table A1: Kaitz index

	2015	2016	2017	2018
All workers	52.85%	51.67%	52.14%	55.55%
Full-time workers	48.19%	47.35%	48.05%	51.59%

Notes: The Kaitz index is defined as the ratio of the minimum wage and the median hourly wage. See Section 2.2 for a description of how hourly wages are estimated.

B.2 Data

B.2.1 Hours worked

The wage information in the *BeH* dataset is defined as the average daily wage: the total wage earnings of an employment spell divided by the length of that spell. Since the German minimum wage is set at the hourly level, it is necessary to supplement the wage data in the *BeH* with an estimate of the number of hours worked per day. For this purpose, we use data from the 2021 version of the German *Mikrozensus*, which is a representative annual survey comprising 1% of households in Germany. Specifically, we use the information on the number of hours that an employed individual ω usually works per week and regress it on two sets of explanatory variables. In doing so, we differentiate between two worker groups g and estimate separate models for workers who are employed subject to social security contributions and marginally employed workers. The first set of control variables accounts for the fact that there are considerable differences in the working hours by gender, part-time status, sector and regions. The model therefore includes indicator variables for females (fem_ω), part-time workers ($part_\omega$) and the interaction of both variables as well as for 21 sector categories s (*Abschnitte* according to the 2008 version of the *Klassifikation der Wirtschaftszweige*) and the 16 federal states f (referring to a person's place of employment). Crucially, these variables are also available in the *BeH* dataset, so that we can compute predicted values for every combination and merge them into the *BeH*. The second set of control variables contains various worker- and household-level characteristics (age, German nationality, tertiary education, marital status, personal income, household size, number of children and household income). We mean-adjusted these variables (separately by sector s and worker group g), so that the predicted working hours refer to a worker with average characteristics in the corresponding sector.

$$\begin{aligned}\ln(hours^g_\omega) = {} & \alpha^g_0 + \alpha^g_1 fem^g_\omega + \alpha^g_2 part^g_\omega + \alpha^g_3 fem^g_\omega part^g_\omega \\ & + \sum_{s=1}^{21} \beta^g_s D^g_s(sector^g_\omega = s) + \sum_{f=1}^{16} \gamma^g_f D^g_f(state^g_\omega = f) \\ & + \boldsymbol{\delta^{g\prime}} \boldsymbol{x^g_\omega} + u^g_\omega,\end{aligned} \tag{37}$$

Table A2 provides an overview of the predicted weekly working hours. For compatibility with the average daily wage contained in the *BeH* dataset, we finally divide the predicted number of weekly hours by 7.

Table A2: Predicted weekly working hours

Gender	Part-time status	Hours (regular)	Hours (marginal)
Female	Full-time	39.43	-
Female	Part-time	21.24	9.98
Male	Full-time	41.22	-
Male	Part-time	20.71	10.43

Notes: Mean values are averaged across sectors and federal states of employment.

B.2.2 Trade

Throughout the paper, spatial variables are based on the delineation from 31 December 2018. The trade flow data, however, uses the delineation from the year 2010 which makes it necessary to apply a number of modifications to make it compatible with the 2018 delineation. Specifically, we merge counties *Göttingen (3152)* and *Osterode am Harz (3156)* into *Göttingen (3159)* and re-code the counties in Mecklenburg-Western Pomerania according to the 2011 reform. In doing so, we assign the former county *Demmin (13052)* completely to the new county *Mecklenburgische Seenplatte (13071)*.

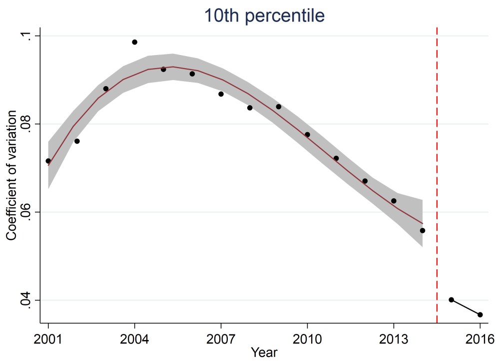
B.2.3 Spatial unit

The spatial units that are used in this paper are based on the delineation from 31 December 2018. The unit of analysis in the empirical analysis are municipality groups (*Verbandsgemeinden*), which contain one or more municipalities (*Gemeinden*). To arrive at the final set of 4,421 municipality groups, we perform the following steps. First, we remove 29 island municipalities that are not connected to the main land by either road or rail. Second, we merge all municipalities which are classified as being *gemeindefrei* and which typically do not contain any employees with the closest municipality in the same county (*Kreise und kreisfreie Städte*). This procedure leaves us with 10,987 municipalities which are then aggregated to the level of municipality groups. Third, for reasons of data anonymity six municipality groups cannot be included in the analysis. One such area is dropped (because it is an island) and the remaining five are merged with the closest municipality group in the same county.

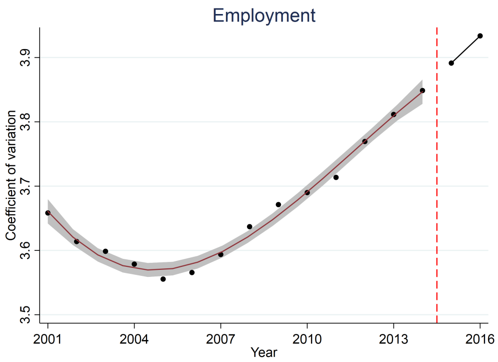
B.3 Spatial convergence

This section complements Section 2.3 in the main paper. Figure A1 substantiates that the introduction of the minimum wage is associated with a decrease in spatial inequalities in low wages. The dispersion in low wages across municipalities summarized by the coefficient of variation also decreases sharply in 2015, revealing spatial convergence. In contrast, there is little evidence for minimum wage effects on employment at the aggregate level. Dispersion in employment across municipalities, if anything, increases.

Figure A1: Convergence



(a) 10th pct. wage variation



(b) Employment variation

Note: Coefficient of variation in (c) and (d) computed across 4,460 municipalities. The 10th percentile wage refers to the 10th percentile in the distribution of individuals within a municipality. Wage and employment data based on the universe of full-time workers from the IAB.

B.4 Average establishment productivity by year-region

This section complements Section 2.3 in the main paper.

To estimate average establishment productivity within regions we perform an AKM-style wage decomposition (Abowd et al., 1999):

$$\ln(w_{\nu\omega jzt}) = \xi_\nu + \psi_\omega + \chi_{zt} + u_{\nu\omega jzt} \tag{38}$$

For this purpose, we regress the hourly wage of worker ν, who is employed at establishment ω in region j and zone z (East or West Germany) in year t, on worker (ξ_ν) and establishment (ψ_ω) fixed effects as well as on separate year fixed effects for East and West Germany. Restricting the sample to 2006-2014 ensures that the estimates are not contaminated by any effects that the introduction of the minimum wage in the year 2015 might have had on worker and establishment outcomes.

ψ_ω provides an estimate of the wage premium that establishment ω pays its workers. We interpret this quantity as a measure of establishment productivity. We then compute annual average regional productivity as the average of all establishment productivity estimates in a given region weighted by the number of workers in the corresponding establishment and year.[12]

C Partial equilibrium

This section complements Section 3 in the main paper.

C.1 Derivation of Eq. (3)

Firm ω_j maximizes its profits

$$\max_{y_j(\omega_j),\, q_{ij}(\omega_j)} \quad \sum_i (S_i^q)^{\frac{1}{\sigma}} q_{ij}(\omega_j)^{\frac{\sigma-1}{\sigma}} - w_j(\omega_j)\frac{y_j(\omega_j)}{\varphi_j(\omega_j)} \quad \text{s.t.} \quad y_j(\omega_j) = \sum_i \tau_{ij} q_{ij}(\omega_j), \tag{39}$$

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

$$\frac{q_{ij}(\omega_j)}{q_{\ell j}(\omega_j)} = \frac{S_i^q}{S_\ell^q}\left(\frac{\tau_{ij}}{\tau_{\ell j}}\right)^{-\sigma} \quad \forall\ \ell \in J, \tag{40}$$

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

[12] The parameter ψ_ω cannot necessarily be estimated for every establishment in the sample. This is the case when an establishment is only observed in a single year. Another possible reason is that an establishment's workers never move to another establishment so that worker and establishment fixed effects cannot be identified separately. Whenever the parameter ψ_ω cannot be identified, we replace the missing value by the average establishment productivity in the corresponding 3-digit sector-year combination. We use the same procedure in the case of establishments that first appear after 2014.

$\sum_i p_{ij}(\omega_j)q_{ij}(\omega_j) = \sum_i (S_i^q)^{\frac{1}{\sigma}} q_{ij}(\omega_j)^{\frac{\sigma-1}{\sigma}}$ in order to obtain $r_j(\omega_j)$ in Eq. (3). $\blacksquare$

C.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 A3 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 $c_j^u(\varphi_j)$ 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 labor 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 labor 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 labor 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 $r_j^d(\varphi_j)$.

Table A3: Firm-level outcomes

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

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

$$\psi_j^d(\varphi_j) = \rho^{\frac{\sigma}{\varepsilon+1}} (S_j^r)^{\frac{1}{\varepsilon+1}} (S_j^h)^{-\frac{1}{\varepsilon+1}} \varphi_j^{\frac{\sigma-1}{\varepsilon+1}} \underline{w}^{-\frac{\sigma+\varepsilon}{\varepsilon+1}}, \tag{41}$$

$$h_j^d(\varphi_j) = \rho^{\frac{\sigma\varepsilon}{\varepsilon+1}} (S_j^r)^{\frac{\varepsilon}{\varepsilon+1}} (S_j^h)^{\frac{1}{\varepsilon+1}} \varphi_j^{\frac{(\sigma-1)\varepsilon}{\varepsilon+1}} \underline{w}^{-\frac{(\sigma-1)\varepsilon}{\varepsilon+1}}. \tag{42}$$

C.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

$$X_j = \chi_X \Phi_j^X(\underline{w}) M_j x_j^u(\underline{\varphi}_j), \tag{43}$$

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

$$\frac{\underline{\varphi}_j^z(\underline{w})}{\underline{\varphi}_j} = \left(\frac{\underline{w}}{\underline{w}_j^z}\right)^{\frac{\sigma+\varepsilon}{\sigma-1}} \quad \forall\ z \in \{s, u\}. \tag{44}$$

Eq. (44) 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

$$\begin{aligned} L_j = M_j \Bigg\{ & l_j^d(\underline{\varphi}_j) \int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}} \frac{l_j^d(\varphi_j)}{l_j^d(\underline{\varphi}_j)} \frac{dG(\varphi_j)}{1-G(\underline{\varphi}_j)} \\ & + l_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\ & \times \int_{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}} \frac{l_j^s(\varphi_j)}{l_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \\ & + l_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\ & \times \int_{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}}^{\infty} \frac{l_j^u(\varphi_j)}{l_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \Bigg\}. \end{aligned} \tag{45}$$

Using firm-level outcomes, Eq. (44), and Eq. (7) we can solve for L_j as defined by Eq. (43), 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

$$H_j = M_j \Bigg\{ h_j^d(\underline{\varphi}_j) \int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}} \frac{h_j^d(\varphi_j)}{h_j^d(\underline{\varphi}_j)} \frac{dG(\varphi_j)}{1-G(\underline{\varphi}_j)}$$
$$+ h_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)}$$
$$\times \int_{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}} \frac{h_j^s(\varphi_j)}{h_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \tag{46}$$
$$+ h_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)}$$
$$\times \int_{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}}^{\infty} \frac{h_j^u(\varphi_j)}{h_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \Bigg\}.$$

Using $h_j^u(\varphi_j) = h_j^s(\varphi_j) = 1$ and $h_j^d(\varphi_j)$ from Eq. (42) in combination with the Eqs. (44) and Eq. (7) allows us to solve for H_j as defined by Eq. (43), with $\chi_H \equiv k/\{k-[\varepsilon/(\varepsilon+1)]\gamma\}$ and

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

Aggregate revenues in location j are defined as

$$
\begin{aligned}
R_j = M_j \Bigg\{ & r_j^d(\underline{\varphi}_j) \int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\}} \frac{r_j^d(\varphi_j)}{r_j^d(\underline{\varphi}_j)} \frac{dG(\varphi_j)}{1-G(\underline{\varphi}_j)} \\
& + r_j^s(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\}} \frac{r_j^s(\varphi_j)}{r_j^s(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})} \\
& + r_j^u(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\}}^{\infty} \frac{r_j^u(\varphi_j)}{r_j^u(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})} \Bigg\}.
\end{aligned}
\tag{47}
$$

Using firm-level outcomes, Eq. (44), and Eq. (7) we can solve for aggregate revenues R_j as defined by Eq. (43), with $\chi_R \equiv k/(k-\gamma)$ and

$$
\begin{aligned}
\Phi_j^R(\underline{w}) \equiv & \frac{r_j^d(\underline{\varphi}_j)}{r_j^u(\underline{\varphi}_j)} \frac{k-\gamma}{k-(\sigma-1)} \left\{ 1 - \left(\frac{w_j^s}{\max\{w_j^s,\underline{w}\}} \right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}} \right\} \\
& + \frac{r_j^s(\underline{\varphi}_j)}{r_j^u(\underline{\varphi}_j)} \frac{k-\gamma}{k-(\sigma-1)/\sigma} \left\{ \left(\frac{w_j^s}{\max\{w_j^s,\underline{w}\}} \right)^{\frac{[k-(\sigma-1)/\sigma](\sigma+\varepsilon)}{\sigma-1}} \right. \\
& \left. - \left(\frac{w_j^u}{\max\{w_j^u,\underline{w}\}} \right)^{\frac{[k-(\sigma-1)/\sigma](\sigma+\varepsilon)}{\sigma-1}} \right\} + \left(\frac{w_j^u}{\max\{w_j^u,\underline{w}\}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}}, \\
= & \left(\frac{w_j^u}{\underline{w}} \right)^{\sigma-1} \left(\frac{\rho}{\eta} \right)^{\sigma-1} \frac{k-\gamma}{k-(\sigma-1)} \left\{ 1 - \left(\frac{w_j^s}{\max\{w_j^s,\underline{w}\}} \right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}} \right\} \\
& + \left(\frac{w_j^u}{\underline{w}} \right)^{-\frac{\sigma-1}{\sigma}\varepsilon} \frac{k-\gamma}{k-(\sigma-1)/\sigma} \left\{ \left(\frac{w_j^s}{\max\{w_j^s,\underline{w}\}} \right)^{\frac{[k-(\sigma-1)/\sigma](\sigma+\varepsilon)}{\sigma-1}} \right. \\
& \left. - \left(\frac{w_j^u}{\max\{w_j^u,\underline{w}\}} \right)^{\frac{[k-(\sigma-1)/\sigma](\sigma+\varepsilon)}{\sigma-1}} \right\} + \left(\frac{w_j^u}{\max\{w_j^u,\underline{w}\}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}}.
\end{aligned}
\tag{48}
$$

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

$$
\begin{aligned}
C_j = M_j \Bigg\{ & c_j^d(\underline{\varphi}_j) \int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}} \frac{c_j^d(\varphi_j)}{c_j^d(\underline{\varphi}_j)} \frac{dG(\varphi_j)}{1-G(\underline{\varphi}_j)} \\
& + c_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}} \frac{c_j^s(\varphi_j)}{c_j^s(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^s, \underline{\varphi}_j\})} \\
& + c_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}) \frac{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\}}^{\infty} \frac{c_j^u(\varphi_j)}{c_j^u(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^u, \underline{\varphi}_j\})} \Bigg\}.
\end{aligned}
\tag{49}
$$

Using firm-level outcomes, Eq. (44), 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

$$
\begin{aligned}
\Phi_j^C(\underline{w}) \equiv & \frac{c_j^d(\underline{\varphi}_j)}{c_j^u(\underline{\varphi}_j)} \frac{k-\gamma}{k-(\sigma-1)} \left\{ 1 - \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}} \right\} \\
& + \frac{c_j^s(\underline{\varphi}_j)}{c_j^u(\underline{\varphi}_j)} \frac{k-\gamma}{k} \left\{ \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} - \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right\} \\
& + \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}}, \\
= & \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{\sigma-1} \left(\frac{\rho}{\eta} \right)^{\sigma} \frac{k-\gamma}{k-(\sigma-1)} \left\{ 1 - \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{[k-(\sigma-1)](\sigma+\varepsilon)}{\sigma-1}} \right\} \\
& + \left(\frac{\underline{w}_j^u}{\underline{w}} \right)^{-(\varepsilon+1)} \frac{k-\gamma}{k} \left\{ \left(\frac{\underline{w}_j^s}{\max\{\underline{w}_j^s, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right. \\
& \left. - \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{k(\sigma+\varepsilon)}{\sigma-1}} \right\} + \left(\frac{\underline{w}_j^u}{\max\{\underline{w}_j^u, \underline{w}\}} \right)^{\frac{(k-\gamma)(\sigma+\varepsilon)}{\sigma-1}}.
\end{aligned}
\tag{50}
$$

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. (43).

In order to derive the **price index** in Eq. (21) we start out from the definition

$$
\begin{aligned}
P_{ij}^{1-\sigma} = M_j \Bigg\{ & [p_{ij}^d(\underline{\varphi}_j)]^{1-\sigma} \int_{\underline{\varphi}_j}^{\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\}} \left[\frac{p_{ij}^d(\varphi_j)}{p_{ij}^d(\underline{\varphi}_j)}\right]^{1-\sigma} \frac{dG(\varphi_j)}{1-G(\underline{\varphi}_j)} \\
& + [p_{ij}^s(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})]^{1-\sigma} \frac{1-G(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\}}^{\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\}} \left[\frac{p_{ij}^s(\varphi_j)}{p_{ij}^s(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})}\right]^{1-\sigma} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^s,\underline{\varphi}_j\})} \\
& + [p_{ij}^u(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})]^{1-\sigma} \frac{1-G(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})}{1-G(\underline{\varphi}_j)} \\
& \times \int_{\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\}}^{\infty} \left[\frac{p_{ij}^u(\varphi_j)}{p_{ij}^u(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})}\right]^{1-\sigma} \frac{dG(\varphi_j)}{1-G(\max\{\underline{\varphi}_j^u,\underline{\varphi}_j\})} \Bigg\}.
\end{aligned} \tag{51}
$$

Using firm-level outcomes, Eq. (44), and Eq. (7) we can solve for P_{ij} from Eq. (21), 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. (48). As a consequence, it follows that we have $\Phi_j^P(\underline{w})|_{\underline{w}<w_j^u} = 1$ and $d\Phi_j^P(\underline{w})|_{\underline{w}<w_j^u}/d\underline{w} = 0$ for $\underline{w} < w_j^u$, $d\Phi_j^P(\underline{w})|_{w_j^u\leq\underline{w}<w_j^s}/d\underline{w} > 0$ for $w_j^u \leq \underline{w} < w_j^s$, and $d\Phi_j^P(\underline{w})|_{w_j^s\leq\underline{w}}/d\underline{w} < 0$ for $w_j^s \leq \underline{w}$. Finally, it is easily verified that $\lim_{\underline{w}\to\infty} \Phi_j^P(\underline{w})|_{w_j^s\leq\underline{w}} = 0$.

In order to derive the (expected) **wage index** in Eq. (27) 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. (44), and Eq. (7) we can solve for W_j from Eq. (21), 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. (47) As a consequence, it follows that we have $\Phi_j^W(\underline{w})|_{\underline{w}<w_j^u} = 1$ and $d\Phi_j^W(\underline{w})|_{\underline{w}<w_j^u}/d\underline{w} = 0$ for $\underline{w} < w_j^u$ as well as $d\Phi_j^W(\underline{w})|_{w_j^u\leq\underline{w}<w_j^s}/d\underline{w} > 0$ for $w_j^u \leq \underline{w} < w_j^s$. For $w_j^s \leq \underline{w}$ we have $d\Phi_j^W(\underline{w})|_{w_j^s\leq\underline{w}}/d\underline{w} < 0$ for sufficiently large values of $\underline{w}$. If $\underline{w} > w_j^s$ is small, $d\Phi_j^W(\underline{w})|_{w_j^s\leq\underline{w}}/d\underline{w}$ can be positive or negative. Finally, it is easily verified that $\lim_{\underline{w}\to\infty} \Phi_j^W(\underline{w})|_{w_j^s\leq\underline{w}} = 0$.

C.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 $\lim_{\underline{w}\to\infty}\Phi_j^L(\underline{w})|_{\underline{w}_j^s \leq \underline{w}} = 0$. 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 $\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s$ we have

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

and it is easily verified that

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

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

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

and it is straightforward to show that

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

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 $\underline{w}_j^s$

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

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 $\lim_{\underline{w}\to\infty} \Phi_j^H(\underline{w})|_{\underline{w}_j^s \le \underline{w}} = 0$. 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 $\underline{w}_j^u \le \underline{w} < \underline{w}_j^s$ we have

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

and it is easily verified that

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

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

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

and it is straightforward to show that

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

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

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

and it is easily verified that

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

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

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

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

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

and it is straightforward to show that

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

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

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

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 $\Phi_j^{\Pi}(\underline{w})|_{\underline{w}_j^s \le \underline{w}}$ 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 $\underline{w} \ge \underline{w}_j^u$. At a low, but binding minimum wage $\underline{w}_j^u \le \underline{w} < \underline{w}_j^s$, supply-constrained firms increase labour input, which results in greater output at lower prices. Because prices decrease in quantity at an elasticity $-1/\sigma \ge -1$ (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.

C.5 Comparative statics

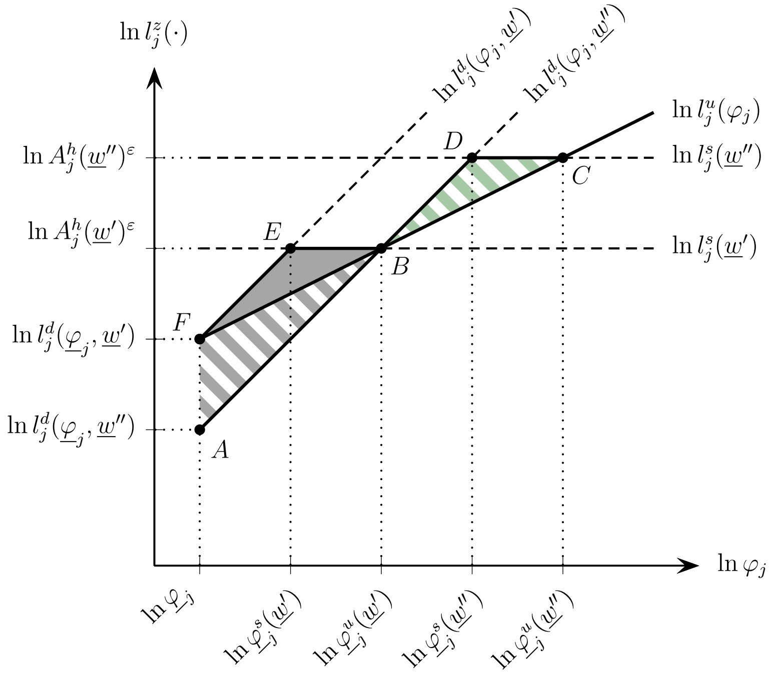
In the following, we discuss how exactly aggregate employment L_j in location j is affected by the introduction of a binding minimum wage $\underline{w}$. For this purpose, we plot in Figure A2 firm-level employment $l_j^z(\cdot)$ as a log-linear function of the firm-specific productivity level $\varphi_j \geq \underline{\varphi}_j$, with $\underline{\varphi}_j > 0$ as the lower bound of location j's productivity distribution.

Without a binding minimum wage location j only features unconstrained firms, whose (log) employment $\ln l_j^u(\varphi_j)$ is increasing in the (log) productivity $\ln \varphi_j$ with slope $[\varepsilon/(\varepsilon+1)]\gamma > 0$. Let us now introduce a low binding minimum wage $\underline{w}'$. Location j then features unconstrained firms (with productivities $\underline{\varphi}_j^u(\underline{w}') \leq \varphi_j < \infty$), supply-constrained firms (with productivities $\underline{\varphi}_j^s(\underline{w}') \leq \varphi_j < \underline{\varphi}_j^u(\underline{w}')$), and demand-constrained firms (with productivities $\underline{\varphi}_j \leq \varphi_j < \underline{\varphi}_j^s(\underline{w}')$).[13] Because the monopsony power of constrained firms is limited or even eliminated by a binding minimum wage $\underline{w}'$, these firms are restricted in their ability to depress their workers' wages by voluntarily reducing their employment level. Supply-constrained firms rather find it optimal to expand their workforce beyond the employment level of equally productive unconstrained firms. And although they are limited in their expansion by the exogenously given labour supply, their employment level $l_j^u(\varphi_j)$ exceeds the employment $l_j^u(\varphi_j)$ of comparable unconstrained firms. Due to a binding minimum wage $\underline{w}'$ the marginal cost of demand-constrained firms do not depend on the firms' underlying productivity level φ_j, which is why their employment $l_j^d(\varphi_j)$ is increasing in firm-level productivity φ_j with an elasticity $\sigma - 1 > 0$, that is larger than the respective employment elasticity $[\varepsilon/(\varepsilon+1)]\gamma$ of unconstrained firms. For productivity levels φ_j in the vicinity of the critical productivity level $\underline{\varphi}_j^s(\underline{w}')$ the employment gain from the elimination of monopsony power in the labor market is large enough to compensate for the employment drop in low-productivity firms, which – due to the binding minimum wage – are confronted with higher marginal costs. As a consequence, we find that at a low binding minimum wage $\underline{w}'$ all constrained firms in Figure A2 feature a higher employment level than in a situation without a binding minimum wage. In the special case of uniformly distributed productivities, the aggregate employment gain from introducing a low binding minimum wage $\underline{w}'$ would be represented by the blue-colored triangle $\overline{FBE}$.

Now suppose the minimum wage is raised from the level of a low binding minimum

[13]Note that not all firm-types have to exist. If the binding minimum wage $\underline{w}$ is sufficiently small, location j does not feature demand-constrained firms.

Figure A2: Comparative statics, minimum wage



wage $\underline{w}'$ to the level of a high binding minimum wage $\underline{w}''$. Raising the minimum wage results in a higher labour supply for supply-constrained firms, and, hence, in an upward shift in (log) employment $\ln l_j^s(\underline{w})$, as well as in a lower labour demand for demand-constrained firms, and, hence in a downward shift in log employment $\ln l_j^d(\varphi_j, \underline{w})$.[14] By the same logic as before, there is an employment gain from the elimination of monopsony power among supply-constrained firms and demand-constrained firms with relatively high productivity levels just below $\underline{\varphi}_j^s(\underline{w}'')$. Ignoring that firms are not necessarily equally distributed, we can indicate this employment gain through the green-colored triangle $\overline{BCD}$. In addition to this aggregate employment gain there also exists an aggregate employment loss (indicated through the red-colored trapezoid $\overline{ABEF}$), that emerges because all incumbent demand-constrained firms see their employment levels decline. For a sufficiently high binding minimum wage $\underline{w}$, this employment loss is not only large enough to offset the aforementioned employment gain, but also to push aggregate employment below the level in a situation without a binding minimum wage.

These results imply two important takeaways: First, location j's aggregate employment L_j is hump-shaped in $\underline{w}$ for all $\underline{w} \geq \underline{w}_j^u$, with $\underline{w}_j^u \equiv w_j^u(\underline{\varphi}_j)$ as the critical minimum wage level below which a minimum wage $\underline{w}$ is non-binding in location j and $\underline{w}_j^s$ as the critical wage level at which aggregate employment L_j in location j is maximized (see

[14]Note that the relative positions of $\ln l_j^s(\underline{w}')$ versus $\ln l_j^s(\underline{w}'')$ and $\ln l_j^d(\varphi_j, \underline{w}')$ versus $\ln l_j^d(\varphi_j \underline{w}'')$ are determined by the requirement that according to Eq. (6) the difference between the critical (log) productivity thresholds $\ln \underline{\varphi}_j^u$ and $\ln \underline{\varphi}_j^u$ is constant and equal to $[\sigma/(\sigma-1)]\ln(\rho/\eta) > 0$.

Appendix C.3).[15] A rather low binding minimum wage $\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s$ is associated with an aggregate employment increase, because all supply-constrained firms optimally expand their employment in response to the minimum wage that limits their monopsony power in the labour market. On the contrary, a rather high binding minimum wage $\underline{w} \geq \underline{w}_j^s$ is associated with an aggregate employment loss, that is the result of falling employment levels among demand-constrained firms, which scale down their production in response to a cost shock associated with the introduction of a high binding minimum wage $\underline{w}$. In Figure 4, we plot the hump-shaped aggregate employment patterns for two locations $j \in \{1, 2\}$ with notionally fixed location-specific fundamentals S_j^r and S_j^h that are assumed to be the same across both locations and varying lower-bound productivities that are ranked $\underline{\varphi}_1 < \underline{\varphi}_2$.

Second, our results imply that the absolute and marginal employment effects of introducing a minimum wage are location-specific. According to Figure 4, the marginal effect of the minimum wage $\underline{w}$ on location j's aggregate employment is positive for $\underline{w}_j^u \leq \underline{w} < \underline{w}_j^s$ and negative for $\underline{w}_j^s \leq \underline{w}$. The critical thresholds $\underline{w}_j^u$ and $\underline{w}_j^s$ thereby inherit their ranking from the productivity ranking $\underline{\varphi}_1 < \underline{\varphi}_2$ (for identical fundamentals S_j^r and S_j^h). For minimum wages in the range $\underline{w} \in (\max\{\underline{w}_1^s, \underline{w}_2^u\}.\underline{w}_2^s)$ it therefore is possible that the marginal effect on location j's aggregate employment in Figure 4 is positive for the high-productivity location $j = 2$, and negative for the low-productivity location $j = 1$. Taking stock, we can conclude that the marginal effect of an increasing minimum wage on aggregate employment is hump-shaped in the location's (lower-bound) productivity $\underline{\varphi}_j$.

C.6 Reduced-form evidence

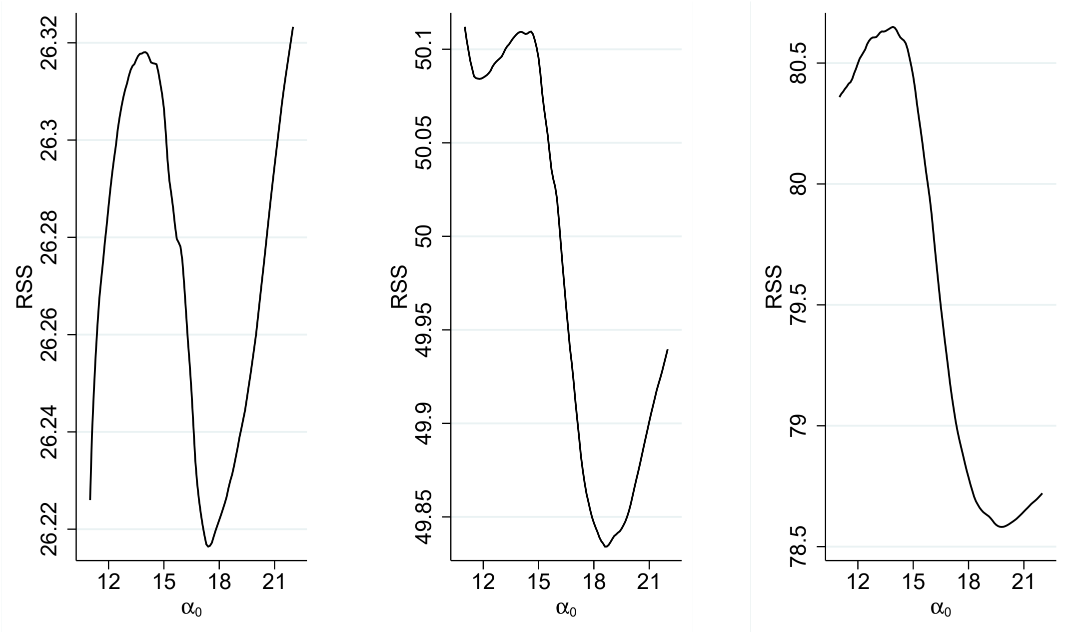
This appendix complements Section 3.2 in the main paper. We provide additional background on the critical points estimated in Figure 5. We also provide the results from robustness tests in which we select alternative temporal windows.

Objective functions for α_0. To identify α_0 introduced in Eq. (11), we estimate Eq. (10) using OLS for set values of α_0 over the parameter space $[\underline{\alpha}_o, \bar{\alpha}_o] = [10, 10.1, ..., 22]$. For each set value α_0 and corresponding estimates of α_1, α_2, we predict $f(\underline{\varphi}_j$ and compute the sum of squared residuals $RSS = \sum_j^J \tilde{\epsilon}_j$. We pick the parameter combination that minimizes the value of this objective function. Figure A3 shows that the objective function is well-behaved in the parameter space around the global minimum for any of the spatial windows in the outcome trends we consider.

Mapping to critical productivity values. The following mapping from the reduced-form parameters $\{\alpha_0, \alpha_1, \alpha_2\}$ to the mean wage levels $\{w^{\text{mean}'}, w^{\text{mean}''}, w^{\text{mean}'''}\}$, which in

[15] The critical minimum wage level $\underline{w}_j^s = (\eta/\rho)^{\sigma/(\sigma-1)}\underline{w}_j^u > \underline{w}_j^u$ also separates a scenario with $\underline{w} < \underline{w}_j^s$, in which location j features unconstrained firms and supply-constrained firms, from a scenario with $\underline{w} \geq \underline{w}_j^s$, in which location j features unconstrained, supply- and demand-constrained firms. Intuitively, $\underline{w}_j^s$ is implicitly defined through $\varphi_j^s(\underline{w}_j^s) = \underline{\varphi}_j$.

Figure A3: Value in objective function of identification of α_0



(a) One-year window (b) Two-year window (c) Three-year window

Note: Each panel shows the sum of squared residuals resulting from the estimation of Eq (10) for varying values of α_0 (introduced in Eq. (11). A one-year spatial window implies that we take second differences over two one-year periods centered on 2014, the year of the minimum wage introduction, i.e. we difference periods 2015-2014 and 2014-2013 when computing the outcome trend in Eq. (10).

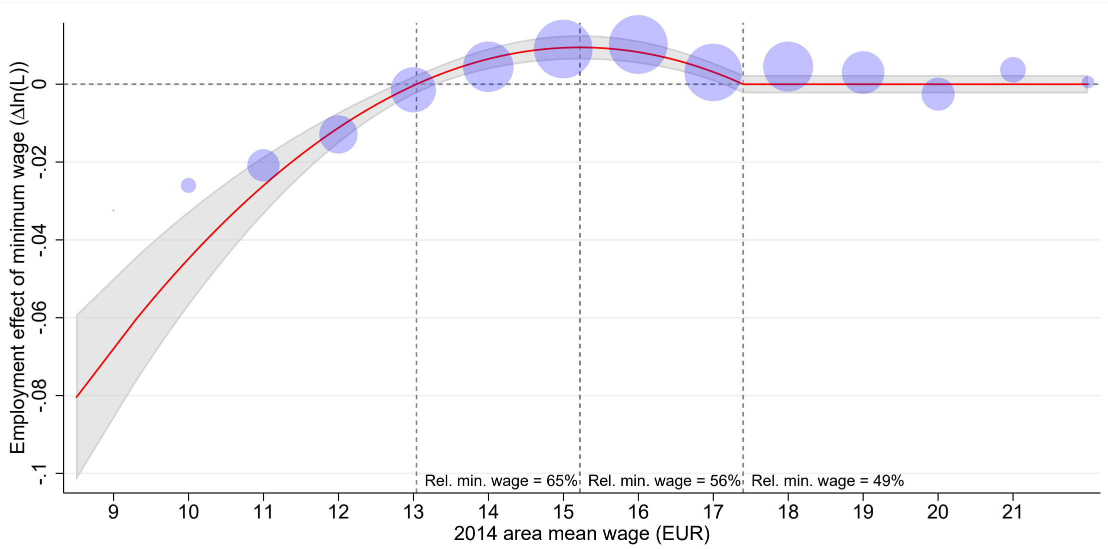
turn correspond to the productivity levels $\{\underline{\varphi}', \underline{\varphi}'', \underline{\varphi}'''\}$, follows directly from the second-order polynomial function in Eq. (11).

$$\begin{aligned} w^{\text{mean}'} &== \alpha_0 - \frac{\alpha_1}{\alpha_2} \\ w^{\text{mean}''} &= \alpha_0 - \frac{\alpha_1}{2\alpha_2} \\ w^{\text{mean}'''} &= \alpha_0 \end{aligned}$$

Alternative temporal windows. To control for unobserved trends at the area level, we take second-differences in Eq. (10). In Figure 5, we have set $\{t = 2016, m = 4, n = 2\}$, which implies that we take differences over the two two-year periods 2012-2014 and 2014-2016, i.e. we have used a two-year spatial window. As robustness tests, we replicate the procedure using a one-year and a three-year window in Figures A4 and A5. Reassuringly, the critical values for the relative minimum wages remain in the same ballpark.

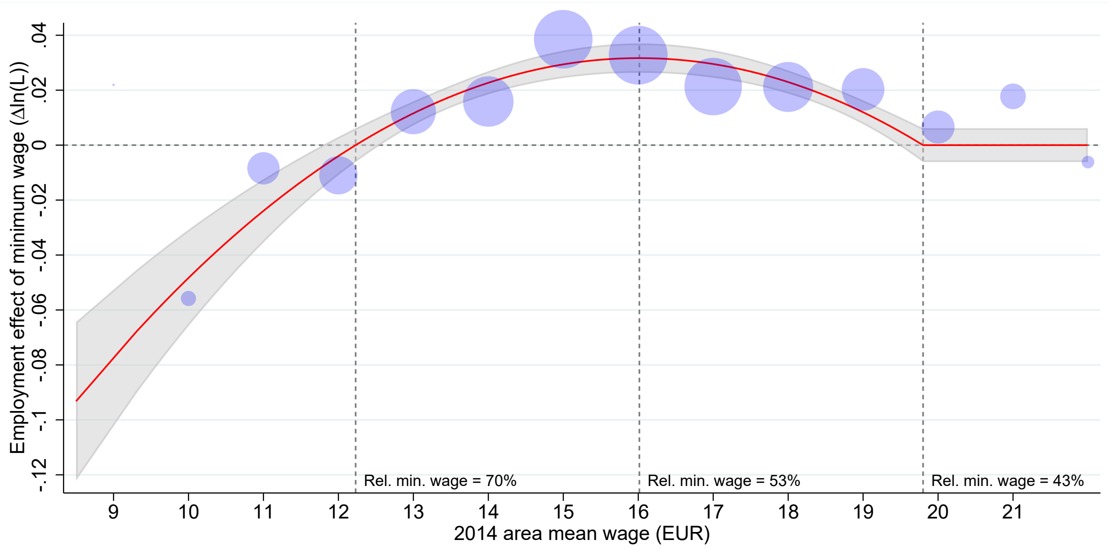
A one-year spatial window implies that we take second differences over two one year periods centered on 2014, the year of the minimum wage introduction, i.e. we difference periods 2015-2014 and 2014-2013 when computing the outcome trend in Figures A4 and A5.

Figure A4: Reduced-form evidence with one-year window



Note: Dependent variable is the second difference in log employment over the 2013-14 and 2014-15 periods. Markers give averages within one-euro bins, with the marker size representing the number of municipalities within a bin. The last bin (22.5) includes all municipalities with higher wages because observation are sparse. The red solid line is the quadratic fit, weighted by bin size. Two outlier bin effects 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).

Figure A5: Reduced-form evidence with three-year window



Note: Dependent variable is the second difference in log employment over the 2011-14 and 2014-17 periods. Markers give averages within one-euro bins, with the marker size representing the number of municipalities within a bin. The last bin (22.5) includes all municipalities with higher wages because observation are sparse. Red solid line is the quadratic fit, weighted by bin size. Two outliers bin effects are are excluded to improve readability, but they 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).

D General equilibrium

This section complements Section 4 in the main paper.

D.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.

$$\lambda_i^N = \frac{N_i}{L} = \sum_j \int_{\varphi_j} \lambda_{ij}(\varphi_j) d\varphi_j,$$
$$= \frac{\sum_j B_{ij} \left[\kappa_{ij} \left(P_i^Q\right)^{\alpha} \left(P_i^T\right)^{1-\alpha}\right]^{-\varepsilon} \left[\frac{\eta \Phi_j^W(\underline{w})^{\frac{1}{\varepsilon}} \Phi_j^L(\underline{w})}{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^{\Pi}(\underline{w})}\right]^{\varepsilon} M_j \tilde{w}_j^{\varepsilon}}{\sum_r \sum_s B_{rs} \left[\kappa_{rs} \left(P_r^Q\right)^{\alpha} \left(P_r^T\right)^{1-\alpha}\right]^{-\varepsilon} \left[\frac{\eta \Phi_s^W(\underline{w})^{\frac{1}{\varepsilon}} \Phi_s^L(\underline{w})}{\Phi_s^R(\underline{w}) - (1-\eta)\Phi_s^{\Pi}(\underline{w})}\right]^{\varepsilon} M_s \tilde{w}_s^{\varepsilon}}. \quad (52)$$

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

$$\lambda_j^H = \frac{H_j}{L} = \sum_i \int_{\varphi_j} \lambda_{ij}(\varphi_j) d\varphi_j,$$
$$= \frac{\sum_i B_{ij} \left[\kappa_{ij} \left(P_i^Q\right)^{\alpha} \left(P_i^T\right)^{1-\alpha}\right]^{-\varepsilon} \left[\frac{\eta \Phi_j^W(\underline{w})^{\frac{1}{\varepsilon}} \Phi_j^L(\underline{w})}{\Phi_j^R(\underline{w}) - (1-\eta)\Phi_j^{\Pi}(\underline{w})}\right]^{\varepsilon} M_j \tilde{w}_j^{\varepsilon}}{\sum_r \sum_s B_{rs} \left[\kappa_{rs} \left(P_r^Q\right)^{\alpha} \left(P_r^T\right)^{1-\alpha}\right]^{-\varepsilon} \left[\frac{\eta \Phi_s^W(\underline{w})^{\frac{1}{\varepsilon}} \Phi_s^L(\underline{w})}{\Phi_s^R(\underline{w}) - (1-\eta)\Phi_s^{\Pi}(\underline{w})}\right]^{\varepsilon} M_s \tilde{w}_s^{\varepsilon}}. \quad (53)$$

D.2 Labour market entry

This section complements Section 4.1.4 in the main paper.

D.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

$$V_{i\nu}^o = V_i^o \exp(a_{i\nu}^o) \quad (54)$$

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

$$G_i^o(a) = \exp(-A_i^o \exp[-\zeta a - \Gamma]), \quad (55)$$

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

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

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

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

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

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

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

$$\mu = \frac{(V_i^{emp})^{\zeta}}{(V_i^{emp})^{\zeta} + A_i^{non}} \tag{56}$$

The labor supply elasticity can be computed as

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

D.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. (33), 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:

$$V^o = \bar{V}^o \exp(a^o), \tag{57}$$

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

$$
\begin{aligned}
E(V) =& \int_0^\infty V dG(V) \\
=& \int_0^\infty -\zeta \sum_o A^o (V^o)^\zeta \exp(-\Gamma) V^{-\zeta} \\
& \times \exp\left\{-\sum_o A^o (V^o)^\zeta \exp(-\Gamma) V^{-\zeta}\right\} dV
\end{aligned} \tag{58}
$$

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

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

D.3 Quantification

This section complements Section 4.2 in the main paper.

D.3.1 Preference heterogeneity (ε)

As discussed in detail in Section C.2, ε governs how, at the firm level, greater productivity φ translates into higher wages $w(\varphi)$ and larger employment $l(\varphi)$. Importantly, ε also monitors the relationship between the two endogenous variables $w(\varphi)$ and $l(\varphi)$.

Using the conditions in Table A3, we can solve the employment of the unconstrained firm for productivity, which we can substitute into the wage equation. Taking logs, we obtain a reduced-form equation:

$$
\ln w_{\omega,j} = \mu_j^e + \tilde{\varepsilon} \ln l_{\omega,j} + \epsilon_{\omega,j}^e, \tag{59}
$$

where $\tilde{\varepsilon} \equiv 1/\varepsilon$, the area fixed effect μ_j^e absorbs the effects of the general equilibrium terms $\{S_j^r, S_j^h\}$ as well as the constant $\eta^{\frac{\sigma}{\sigma+\epsilon}}$, and $\epsilon_{\omega,j}^e$ is an error term that accounts for measurement error in hours worked.

We present our estimates of Eq. (59) in Table A4. Column (1) contains our preferred theory-consistent baseline. The remaining columns are robustness tests. We find that wages scale in firm-level employment at an elasticity of slightly below 0.2, with the estimate being insensitive to controls for firm characteristics. The implied value of $\varepsilon = 5.2$ is in 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 is in between those in Monte et al. (2018) (counties) and and Ahlfeldt et al. (2015) (housing blocks). It is intuitive, that the dispersion of tastes for places increases in the size of the considered spatial units.

Table A4: Preference heterogeneity

	(1)	(2)	(3)	(4)	(5)
Ln establishment employment	0.1923***	0.1890***	0.1772***	0.1873***	0.1753***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Constant	2.0461***	2.0545***	2.0702***	2.0567***	2.0719***
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Region FE	Yes	Yes	Yes	No	No
Sector FE (1-digit)	No	Yes	No	No	No
Sector FE (2-digit)	No	No	Yes	No	No
Region-by-sector FE (1 digit)	No	No	No	Yes	No
Region-by-sector FE (2 digit)	No	No	No	No	Yes
Observations	2390350	2291881	2291881	2290834	2232591
R^2	.192	.255	.331	.276	.393
$\hat{\varepsilon}$	5.2	5.3	5.6	5.3	5.7
	(0.0072)	(0.0078)	(0.0090)	(0.0080)	(0.0097)

Notes: Unit of observation is establishment-level. The estimate of ε is defined as $1/\hat{\hat{\varepsilon}}$. Robust standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

D.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 A3, $l_j^u(\varphi_j)$, delivers:

$$\mathcal{S}_j^b = 1 - \frac{\int_{\varphi_j^b}^{\infty} l_j^u(\varphi_j) dG(\varphi_j)}{\int_{\underline{\varphi}_j}^{\infty} l_j^u(\varphi_j) dG(\varphi_j)} = 1 - \frac{\int_{\varphi_j^b}^{\infty} \varphi_j^{\frac{(\sigma-1)\varepsilon}{\sigma+\varepsilon}} k_j \varphi_j^{-k_j-1} \underline{\varphi}_j^{k_j} d\varphi_j}{\int_{\underline{\varphi}_j}^{\infty} \varphi_j^{\frac{(\sigma-1)\varepsilon}{\sigma+\varepsilon}} k_j \varphi_j^{-k_j-1} \underline{\varphi}_j^{k_j} d\varphi_j} \tag{60}$$

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

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

Substituting $\underline{\varphi}_j$ and φ_j^b using Eq. (5) and the formular for average wages, Eq. (18), we obtain:

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

$$= 1 - \left(\frac{k_j(\sigma+\varepsilon) - (\varepsilon+1)(\sigma-1)}{k_j(\sigma+\varepsilon) - \varepsilon(\sigma-1)} \frac{\tilde{w}_j}{w_j^b(\varphi_j^b)} \right)^{\frac{k_j(\sigma+\varepsilon)-(\sigma-1)\varepsilon}{\sigma-1}} \tag{61}$$

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.

$$\tilde{S}_j^b = S_j^b - e_j^b \tag{62}$$

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

$$\mathbb{E}\left(w^b e_j^b \right) = 0, \tag{63}$$

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

$$\mathbb{E}\left(w^b \left[1 - \tilde{S}_j^b - \left(\frac{k_j(\sigma+\varepsilon) - (\varepsilon+1)(\sigma-1)}{k_j(\sigma+\varepsilon) - \varepsilon(\sigma-1)} \frac{\tilde{w}_j}{w_j^b(\varphi_j^b)} \right)^{\frac{k_j(\sigma+\varepsilon)-(\sigma-1)\varepsilon}{\sigma-1}} \right] \right) = 0 \tag{64}$$

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 C.3:

$$\begin{aligned} k &> \sigma - 1 \\ k &> \frac{(-1)\varepsilon}{\sigma + \varepsilon} \\ k &> \frac{(\sigma - 1)\varepsilon}{\varepsilon + 1} \\ k &> \frac{\sigma - 1}{\sigma} \\ k &> \frac{(\sigma - 1))(\varepsilon + 1)}{\varepsilon + \sigma} \end{aligned} \tag{65}$$

Therefore, we set $\varepsilon = 5.2$ to the value estimated in Section D.3.1 and nest a GMM estimation of k using the moment condition in Eq. (64) 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. (64) 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 $\bar{b} = 15$ as these appear like generous bounds of minimum wages to be considered by policy.

This procedure identifies a Pareto firm productivity shape parameter of $k = 0.5$ and an elasticity of substitution of $\sigma = 1.5$. 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 A5. 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.

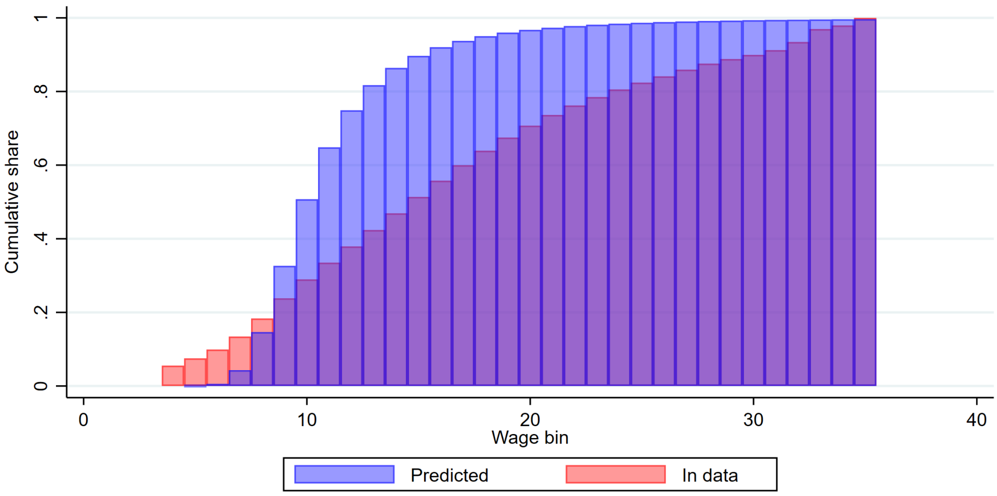
Table A5: Estimation of firm productivity distribution parameter k

	$\sigma = 1.5$	$\sigma = 3$	$\sigma = 4.5$	$\sigma = 6$
k	0.5313***	1.7366***	1.7837***	3.1785***
	(0.0001)	(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 15 Euro per hour. Estimation by GMM. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In Figure A6, 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 A6: 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. (61).

D.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:

$$\tau_{ij} = \exp\left(b_i^T DIST_{ij} + d^{T,EW} D_{ij}^{T,EW} + d^{T,WE} D_{ij}^{T,WE}\right), \tag{66}$$

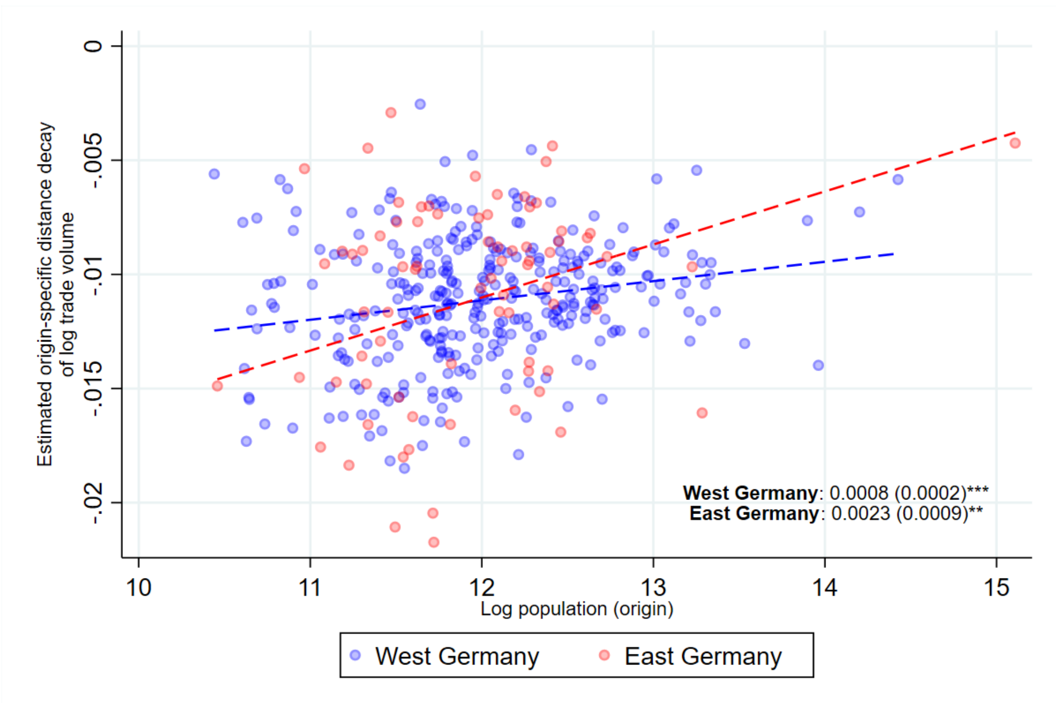
where $D^{T,EW}$ takes the value of one if i refers to a region in East Germany and j to a region in West Germany, while $D^{T,WE}$ takes the value of one for routes starting in West and ending in East Germany. Note that the distance effect on trade cost b_i^T is origin-specific. This allows some regions 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. (66) in Eq. (20), we can derive a gravity equation of trade:

$$\ln(F_{lk}) = c^T + O_l^T + D_k^T + \tilde{b_l}^T DIST_{lk} + \tilde{d}^{T,EW} D_{lk}^{T,EW} + \tilde{d}^{T,WE} D_{lk}^{T,WE} + e_{lk}^T, \tag{67}$$

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. (67) 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 A6). Figure A7 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 A7: Estimated distance elasticity of trade volumes



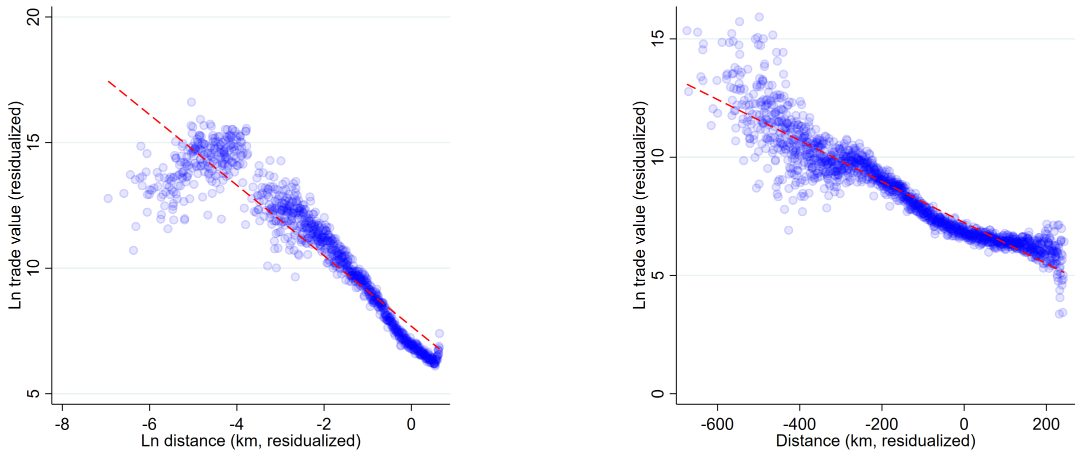
Note: Unit of observation is the county level. The figure plots the estimated origin-specific distance elasticities of trade volumes (given by $\tilde{b}_l^T$ in Eq. (67)) against log origin population size at the county level separately for East and West Germany. The dashed lines show the linear fit 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

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

Figure A8 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 A8: 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.

D.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. (20) in (23) to obtain:

$$\tilde{w}_j L_j = \sum_i \left[N_i \frac{M_j (\tau_{ij} \tilde{w}_j / \underline{\varphi}_j)^{1-\sigma}}{\sum_{k \in J} M_k (\tau_{ik} \tilde{w}_k / \underline{\varphi}_k)^{1-\sigma}} \sum_j^J \lambda_{ij|i}^N \tilde{w}_j \right] \tag{68}$$

Since we observe $\{\tilde{w}_j, L_j, M_j, \lambda_{ij|i}^N\}$ and have parametrized τ_{ij} in Section D.3.3, Eq. (68) 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 A6: 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	114951	114951	114951
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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D.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.

D.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. (32).
   (b) Compute L_j based on Eq. (31).
   (c) Compute residents according to $N_i = \lambda_i^N L$.
   (d) Compute house price index P_i^T according to Eq. (15).
   (e) Using the free-entry condition Eq. (17), compute expenditure shares θ_{ij} according to Eq. (20).
   (f) Compute the goods price index P_i^Q according to Eq. (22).
2. Derive new values of initially guessed variables:
   (a) Compute new value of $\tilde{w}_j$ according to Eq. (23) and normalize values with employment-weighted average wage.
   (b) Compute new value of λ_{ij} according to Eq. (28).

(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 C.3.

(d) Compute new value of labor force participation rate μ according to Eq. (34) 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.

D.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. (33). 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. (33) 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. (34), 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

$$\mu = \frac{\tilde{V}_i^{\zeta}}{\tilde{V}_i^{\zeta} + \tilde{A}_i} \tag{69}$$

Consequentially, conditional welfare becomes location-specific:

$$\tilde{\mathcal{V}}_i = \left(\tilde{A}_i + \tilde{V}_i^{\zeta}\right)^{\frac{1}{\zeta}} \tag{70}$$

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. (69) and spatially varying changes in conditional welfare as Eq. (70). While the working-age population is a fixed endowment in the short run, the labour force remains an endogenous variable as per

$$N_i = \mu_i \overline{N}_i. \tag{71}$$

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. (30). Third, under 2.(d), we compute new values for local labor force participation rates based on Eq. (71) and thus new N_i. All other steps remain the same.

D.5 The German minimum wage

D.5.1 Comparison to data

This section complements Section 4.4.2 in the main paper. To investigate the model's out-of-sample predictive power, we use the minimum wage effect predicted by model $\hat{\mathbf{X}} = \frac{\mathbf{X^C}}{\mathbf{X^0}}$ as an input into a dynamic difference-in-difference model with time-varying treatment effects:

$$\ln \mathbf{X^D_{i,t}} = \sum_{z \neq 2014} a_z \ln \hat{\mathbf{X}}_{\mathbf{i}} + a^T_{z(i),t} + a^I_i + e^D_{it}, \tag{72}$$

where $\mathbf{X^D_{i,t}}$ is an outcome observed for region i in year t, $z(i), t$ is a year-by-zone (former East and West Germany) fixed effect, a^I_i region fixed effect and e^D_{it} is an error term. This specification generates the following intensive-margin difference-in-difference treatment effects:

$$\alpha^z = \frac{\partial \ln \mathbf{X^D_{i,t=z}}}{\partial \ln \mathbf{\hat{X}_i}} - \frac{\partial \ln \mathbf{X^D_{i,t=2014}}}{\partial \ln \mathbf{\hat{X}_i}} = \frac{\partial \left(\ln \mathbf{X^D_{i,t=z}} - \ln \mathbf{X^D_{i,t=2014}}\right)}{\partial \left(\ln \mathbf{X^C_i} - \ln \mathbf{X^0_i}\right)} \tag{73}$$

Thus, if changes in the data scale proportionately in to changes predicted by models, we will observe treatment effects α^z close to one.

D.6 Gini coefficient

D.6.1 The Gini coefficient in the model

We derive the GINI-coefficient according to the following steps.

1. We derive the CDF of firm-level employment for each location.
2. We aggregate the CDFs to the national level by taking the sum over the employment-weighted location-specific CDFs.
3. We define wage bins and compute PDFs from differentiating CDFs across adjacent bins.
4. Multiplying the employment densities with the wage level in each bin and computing the cumulative sum delivers the CDF of labor income.
5. Plotting the CDF for employment and labor income against each other delivers the Lorenz curve. The Gini coefficient is defined as $\mathcal{G} = 1 - 2B$, where B is the area under the Lorenz curve.

Cumulative distribution function of firm-level employment. First, we derive the number of workers in location j who are employed at firms with productivities between $\underline{\varphi}_j$ and φ^b:

$$\begin{aligned}
L_j(\varphi^b_j) = M_j \Bigg\{ & l^d_j(\underline{\varphi}_j) \int_{\underline{\varphi}_j}^{\min\{\varphi^b_j, \max\{\underline{\varphi}^s_j, \underline{\varphi}_j\}\}} \frac{l^d_j(\varphi_j)}{l^d_j(\underline{\varphi}_j)} \frac{dG(\varphi_j)}{1 - G(\underline{\varphi}_j)} \\
& + l^s_j(\min\{\varphi^b_j, \max\{\underline{\varphi}^s_j, \underline{\varphi}_j\}\}) \frac{1 - G(\min\{\varphi^b_j, \max\{\underline{\varphi}^s_j, \underline{\varphi}_j\}\})}{1 - G(\underline{\varphi}_j)} \\
& \times \int_{\min\{\varphi^b_j, \max\{\underline{\varphi}^s_j, \underline{\varphi}_j\}\}}^{\min\{\varphi^b_j, \max\{\underline{\varphi}^u_j, \underline{\varphi}_j\}\}} \frac{l^s_j(\varphi_j)}{l^s_j(\min\{\varphi^b_j, \max\{\underline{\varphi}^s_j, \underline{\varphi}_j\}\})} \frac{dG(\varphi_j)}{1 - G(\min\{\varphi^b_j, \max\{\underline{\varphi}^s_j, \underline{\varphi}_j\}\})} \\
& + l^u_j(\min\{\varphi^b_j, \max\{\underline{\varphi}^u_j, \underline{\varphi}_j\}\}) \frac{1 - G(\min\{\varphi^b_j, \max\{\underline{\varphi}^s_j, \underline{\varphi}_j\}\})}{1 - G(\underline{\varphi}_j)} \\
& \times \int_{\min\{\varphi^b_j, \max\{\underline{\varphi}^u_j, \underline{\varphi}_j\}\}}^{\varphi^b_j} \frac{l^u_j(\varphi_j)}{l^u_j(\min\{\varphi^b_j, \max\{\underline{\varphi}^u_j, \underline{\varphi}_j\}\})} \frac{dG(\varphi_j)}{1 - G(\min\{\varphi^b_j, \max\{\underline{\varphi}^u_j, \underline{\varphi}_j\}\})} \Bigg\}.
\end{aligned}$$

We now substitute productivity thresholds with critical minimum wage levels according to Eq. (44) and use

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

To compute the share of workers that earn less than w^b, we use the facts that all constrained firms pay the minimum wage and that $w^b \geq w_j^s(\underline{\varphi}_j) > w_j^u(\underline{\varphi}_j)$. Following the same procedure as in Appendix C.3, we get

$$L_j(\underline{w}, w^b) = \chi_L \Phi_j^L(\underline{w}, w^b) M_j l_j^u(\underline{\varphi}_j)$$

where $\chi_L \equiv k/\{k - [\varepsilon/(\varepsilon+1)]\gamma\}$ and

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

Notice that for a given wage level w^b the density for any of the three firm types must not be negative. We ensure this in the code by manually assigning appropriate values to w^b for the respective firm types. To give an example, for demand-constrained firms, if $\underline{w} > \underline{w}_j^s$ and $w^b < \underline{w}_j^s$, we set $w^b = \underline{w}_j^s$. This ensures that then density of demand-constrained firms for wage bins smaller than the mandatory minimum wage is zero. We apply this logic to all cases and firm types.

Relating $L(\underline{w}, w^b)$ to L_j delivers the cumulative density of workers as a function of wages:

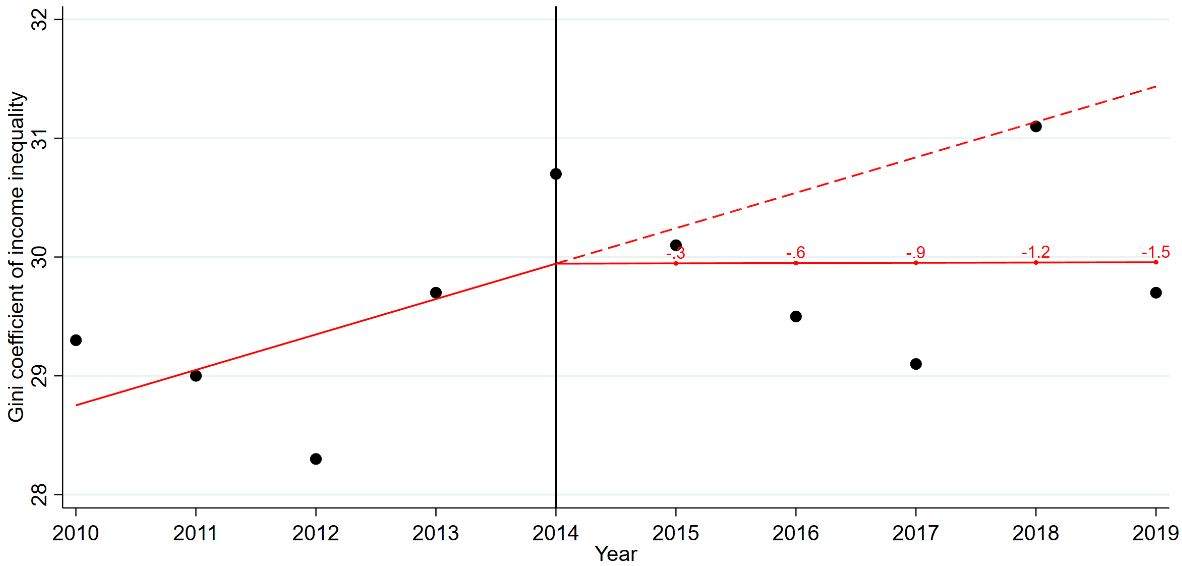
$$Z_j(w \leq w^b) = \Phi_j^L(\underline{w}, w^b)/\Phi_j^L \equiv \Phi_j^L(\underline{w}, w^b),$$

where we take Φ_j^L from Appendix C.3.

The remaining steps as introduced above can be executed straightforwardly.

D.6.2 The Gini coefficient in data

Figure A9: Gini coefficient in data



Note: Own illustration using Gini coefficients from the German Statistical Office. Each dot represents a Gini coefficient of the income distribution across all workers in all regions measured in data. The red solid line is the fit of a linear spline function with a knot in 2014. The dashed red line is the linear extrapolation of the pre-policy trend.

In Figure A9, we plot Gini coefficients of wage inequality across German workers by year. They are generally around 30% in Germany, which is a typical value for a European country and within close range of the wage inequality we generate within our model. While there is some volatility across years, there are clear trends within the three years preceding and succeeding the minimum wage inequality: Inequality increased before the introduction and decreased afterwards, consistent with the intended policy objective. If we expand the temporal window, there is more noise, but the perception of a reduction in wage inequality persists. 2018—a suspicious outlier—aside, Gini coefficients are lower during the post-policy period than in 2014 and certainly lower than predicted by an extrapolation of previously observed trends. Comparing a linear trend interpolation within the post-policy period to a linear trend extrapolation from the pre-policy period, we estimate a reduction in the Gini coefficient of 1.5 percentage points which is close to the 2-percentage reduction predicted by our model.

D.7 Optimal minimum wages

This section complements Section 4.5 by providing additional detail on the causes, effects, and the regional distribution of the effects of the optimal minimum wages discussed in Table 2.

To this end, we provide descriptive statistics on the four outcomes illustrated in Figure 7 derived under alternative minimum wages in Table A7. We present the results for employment-maximizing and welfare-maximizing federal and regional minimum wages introduced in Table 2. We further distinguish between short-run and long-run effects. The, perhaps, most striking insight from Table 2 is that *regional* minimum wages—because they "bite" similarly in all regions—have effects that hardly vary by region. In contrast, *federal* minimum wages lead to great spatial heterogeneity in the welfare incidence in the short-run. The long-run, spatial arbitrage results in large reallocation of the labour force towards those regions experiencing short-run welfare gains. Because the welfare-maximizing federal minimum wage is more ambitious than the employment-maximizing minimum wage, the spatial heterogeneity in the effects is particularly striking. A comparison of Figure A10 to Figure 7 reveals that the way this heterogeneity plays out depends on the level of the minimum wage.

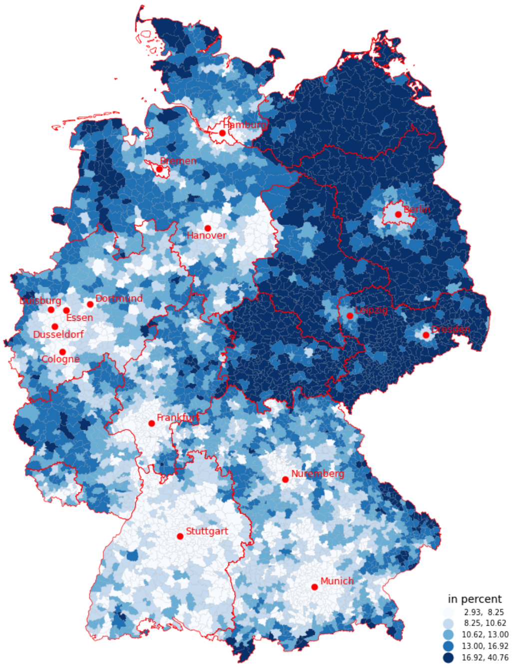
As the welfare-maximizing minimum wage is higher than the implemented German minimum wage (60% vs 48%), we observe more pronounced increases in real wages from panel (a) of Figure A10. These wage increases that are particularly prevalent in East Germany are associated with reductions in employment probabilities that can be quite substantial in certain municipalities (up to -25%). While the former effect dominates the latter, it is remarkable that the net effect is now larger in West Germany. This is almost exactly the opposite of the actual German minimum wage that is about 20% lower. As an immediate consequence, the higher welfare gains in the west cause a different migration responses. In contrast to Figure 7, we now observe emigration from East Germany. Increasing the minimum wage from moderate levels therefore affects different regions at different stages. For medium to high minimum wages (relative to the mean wage), it is the medium to high productivity locations that experience short-run welfare gains and long-run immigration.

Table A7: Optimal minimum wages

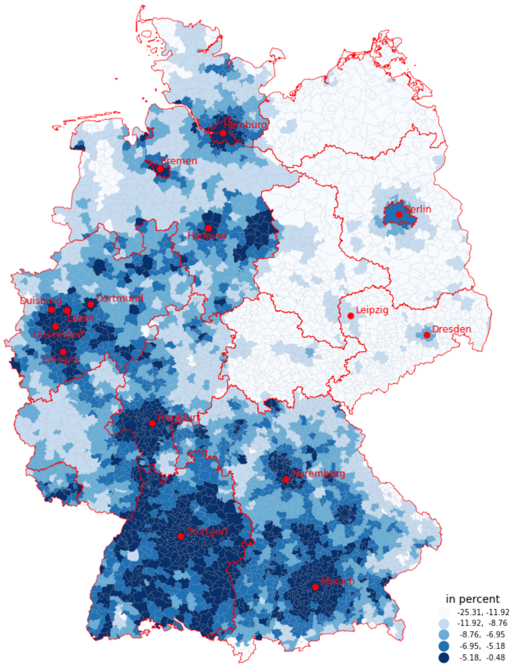
Objective	Scheme	Case	Outcome	Mean	S.D.	Min.	Max.
Employment	Federal	SR	Real wage	0.070	0.450	-1.260	3.730
Employment	Federal	SR	Employment prob.	-0.110	0.280	-4.330	0.000
Employment	Federal	SR	Welfare	0.290	0.710	-0.230	9.560
Employment	Federal	SR	Labour force	0.080	0.100	-0.150	0.720
Employment	Federal	LR	Real wage	0.030	0.430	-1.150	3.340
Employment	Federal	LR	Employment prob.	-0.110	0.300	-4.590	0.000
Employment	Federal	LR	Welfare	0.270	0.600	-0.230	9.150
Employment	Federal	LR	Labour force	0.050	1.470	-4.260	8.200
Employment	Regional	SR	Real wage	4.610	0.010	4.590	4.640
Employment	Regional	SR	Employment prob.	-0.010	0.000	-0.010	-0.010
Employment	Regional	SR	Welfare	-0.010	0.000	-0.020	0.020
Employment	Regional	SR	Labour force	1.070	0.000	1.060	1.070
Employment	Regional	LR	Real wage	4.610	0.000	4.590	4.630
Employment	Regional	LR	Employment prob.	-0.010	0.000	-0.010	-0.010
Employment	Regional	LR	Welfare	0.000	0.000	-0.020	0.020
Employment	Regional	LR	Labour force	1.060	0.020	0.930	1.120
Welfare	Federal	SR	Real wage	12.660	5.250	2.930	40.760
Welfare	Federal	SR	Employment prob.	-8.470	3.910	-25.310	-0.480
Welfare	Federal	SR	Welfare	7.590	8.690	-6.310	45.120
Welfare	Federal	SR	Labour force	1.160	0.150	0.570	1.670
Welfare	Federal	LR	Real wage	12.710	5.270	4.010	41.050
Welfare	Federal	LR	Employment prob.	-8.460	3.850	-25.030	-0.450
Welfare	Federal	LR	Welfare	7.490	8.700	-6.090	44.960
Welfare	Federal	LR	Labour force	2.220	2.540	-7.990	14.220
Welfare	Regional	SR	Real wage	6.610	0.030	6.540	6.840
Welfare	Regional	SR	Employment prob.	-2.800	0.000	-2.800	-2.800
Welfare	Regional	SR	Welfare	-0.030	0.020	-0.150	0.040
Welfare	Regional	SR	Labour force	1.330	0.010	1.290	1.350
Welfare	Regional	LR	Real wage	6.610	0.030	6.550	6.810
Welfare	Regional	LR	Employment prob.	-2.800	0.000	-2.800	-2.800
Welfare	Regional	LR	Welfare	-0.020	0.030	-0.060	0.040
Welfare	Regional	LR	Labour force	1.250	0.100	0.570	1.600

Notes: This table provides a additional outputs of the simulated minimum wage effects summarised in Table 2. *Objective* describes if the minimum wage is employment-maximizing or welfare-maximizing. SR = short run; LR = long run. *Mean* is the unweighted average across municipalities. It does not correspond to the national average.

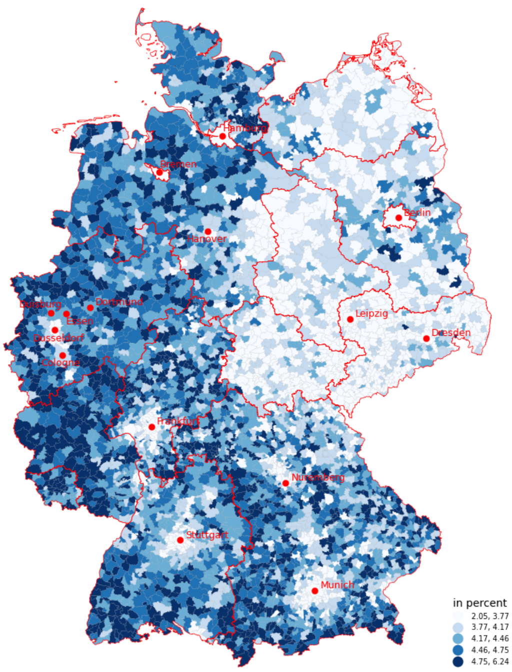
Figure A10: Effects of the welfare-maximizing federal minimum wage



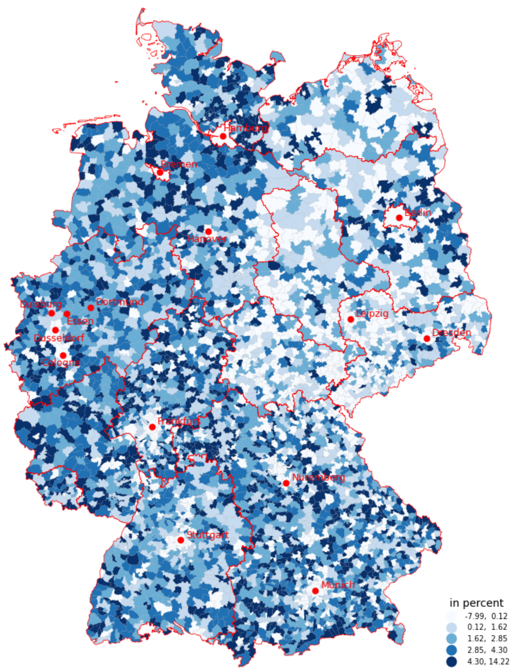
(a) Real wage, short run



(b) Employment probability, short run



(c) Welfare, short run



(d) Labour force, long run

Note: Unit of observation are 4,421 municipality groups. The welfare-maximizing federal minimum wage is set at 60% of the national employment-weighted mean wage. 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. (32) into two components. The first is the residential wage conditional on working $\sum_j \lambda^N_{ij|i} \tilde{w}_j$, 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 $\sum_j \lambda^N_{ij|i} L_j / H_j$, which captures the probability that a worker finds a job within the area-specific commuting zone.

Table A8: Minimum wage schedules

		Level rel. to		Employment		Equity		Welfare	
Objective	Scheme	Mean	p50	SR	LR	SR	LR	SR	LR
Employment	State	42.00	46.21	0.04	0.04	0.12	0.12	0.47	0.50
Welfare	State	58.00	63.82	-3.19	-3.26	4.23	4.22	4.35	4.38
Employment	County	50.00	55.01	0.42	0.42	0.56	0.56	3.14	3.16
Welfare	County	58.00	63.82	-2.22	-2.23	3.35	3.36	4.71	4.71

Notes: All values are given in %. *Objective* describes if the minimum wage is employment-maximizing or welfare-maximizing. *State* indicates a minimum wage that is set the respective *level* of the state (*Bundesland*) mean. *County* indicates a minimum wage that is set the respective *level* of the county (*Kreis*) 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. Welfare is the expected utility of as defined by Eq. (36). 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.

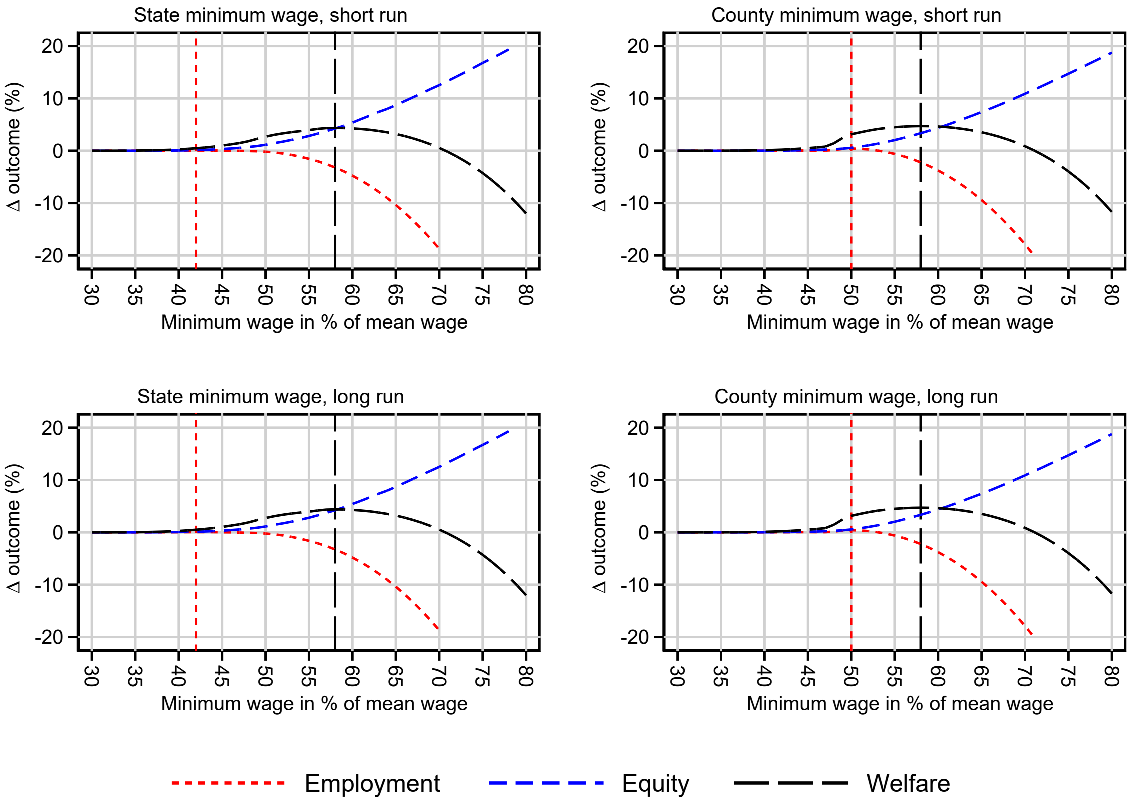
D.8 Regional minimum wages for alternative spatial units

This section complements Section 4.5 in which we quantitatively evaluate the effect of a regional minimum wage set at the municipality level. Here, we consider regional minimum wages set at the level of federal states (*Bundesländer*) and counties (*Kreise and Kreisfreie Städte*) as alternatives. To this end, we compute the worker-weighted wage across all municipalities in a region (county or state) and set the regional minimum wage such that it corresponds to a given fraction of the regional mean wage. Otherwise, the procedure is identical to the one used in Section 4.5.

The main insight from A11, which is the analog to Figure 9 in the main paper, is that the state minimum wage resembles the federal minimum wage, whereas the county minimum wage resembles the municipality minimum wage. This impression is reinforced by Table A8, which is the analog to Table 2 in the main paper. The employment-maximizing and welfare-maximizing levels of the *state* minimum wage are close to those of the *federal* minimum wage, and so are the employment, equity and welfare effects. Similarly, the levels of employment-maximizing and welfare-maximizing *county* minimum wage are close to those of the *municipality* minimum wage, and so are the employment, equity and welfare effects.

We conclude that for regional minimum wages to play out their strengths—mitigating the trade-off of positive welfare and negative employment effects—they need to be set for relatively small spatial units, at least in countries where productivity varies strongly between cities and towns within broader regions.

Figure A11: Regional minimum wages at state and county levels



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. Welfare is the expected utility of as defined by Eq. (36). 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.

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

_ , 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.

Bellmann, Lutz, Mario Bossler, Hans-Dieter Gerner, and Olaf Hübler, "Training and minimum wages: first evidence from the introduction of the minimum wage in Germany," *IZA Journal of Labor Economics*, 2017, *6* (1), 8.

Bonin, Holger, Ingo E Isphording, Annabelle Krause-Pilatus, Andreas Lichter, Nico Pestel, and Ulf Rinne, "The German Statutory Minimum Wage and Its Effects on Regional Employment and Unemployment," *Jahrbücher für Nationalökonomie und Statistik*, 2020, *240* (2-3), 295–319.

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.

_ and Thorsten Schank, "Wage Inequality in Germany after the Minimum Wage Introduction," 2020.

_ , Nicole Gürtzgen, Benjamin Lochner, Ute Betzl, and Lisa Feist, "The German Minimum Wage: Effects on Productivity, Profitability, and Investments," *Jahrbücher für Nationalökonomie und Statistik*, 2020, *240* (2-3), 321–350.

Bruckmeier, Kerstin and Oliver Bruttel, "Minimum Wage as a Social Policy Instrument: Evidence from Germany," *Journal of Social Policy*, 2021, *50* (2), 247–266.

Burauel, Patrick, Marco Caliendo, Markus M Grabka, Cosima Obst, Malte Preuss, and Carsten Schröder, "The Impact of the Minimum Wage on Working Hours," *Jahrbücher für Nationalökonomie und Statistik*, 2020, *240* (2-3), 233–267.

_ , _ , _ , _ , _ , _ , and Cortnie Shupe, "The Impact of the German Minimum Wage on Individual Wages and Monthly Earnings," *Jahrbücher für Nationalökonomie und Statistik*, 2020, *240* (2-3), 201–231.

Caliendo, Marco, Alexandra Fedorets, Malte Preuß, Carsten Schröder, and Linda Wittbrodt, "The Short-Term Distributional Effects of the German Minimum Wage Reform," 2017.

_ , _ , 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.

_ , Carsten Schröder, and Linda Wittbrodt, "The Causal Effects of the Minimum Wage Introduction in Germany – An Overview," *German Economic Review*, 8 2019, *20*

(3), 257–292.

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.

Dube, Arindrajit, 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*, 2021, p. forthcoming.

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

Ellguth, Peter, Susanne Kohaut, and Iris Möller, "The IAB Establishment Panel—methodological essentials and data quality," *Journal for Labour Market Research*, 2014, *47* (1), 27–41.

Fedorets, Alexandra and Cortnie Shupe, "Great expectations: Reservation wages and minimum wage reform," *Journal of Economic Behavior & Organization*, 2021, *183*, 397–419.

Fitzenberger, Bernd and Annabelle Doerr, "Konzeptionelle Lehren aus der ersten Evaluationsrunde der Branchenmindestlöhne in Deutschland," *Journal for Labour Market Research*, 2016, *49* (4), 329–347.

Friedrich, Martin, "Using Occupations to Evaluate the Employment Effects of the German Minimum Wage," *Jahrbücher für Nationalökonomie und Statistik*, 2020, *240* (2-3), 269–294.

Garloff, Alfred, "Did the German Minimum Wage Reform Influence (Un)employment Growth in 2015? Evidence from Regional Data," *German Economic Review*, 8 2019, *20* (3), 356–381.

Goebel, Jan, Markus M Grabka, Stefan Liebig, Martin Kroh, David Richter, Carsten Schröder, and Jürgen Schupp, "The German Socio-Economic Panel (SOEP)," *Jahrbücher für Nationalökonomie und Statistik*, 2019, *239* (2), 345–360.

Holtemöller, Oliver and Felix Pohle, "Employment effects of introducing a minimum wage: The case of Germany," *Economic Modelling*, 2020, *89*, 108–121.

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

Link, Sebastian, "The Price and Employment Response of Firms to the Introduction of Minimum Wages," 2019, (March), 1–57.

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.

Mindestlohnkommission, "Erster Bericht zu den Auswirkungen des gesetzlichen Mindestlohns. Bericht an die Bundesregierung nach §9 Abs. 4 Mindestlohngesetz," Technical

Report, Berlin 2016.

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

Möller, Joachim, "Minimum wages in German industries—what does the evidence tell us so far?," *Journal for Labour Market Research*, 2012, *45* (3), 187–199.

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).

Schmitz, Sebastian, "The Effects of Germany's Statutory Minimum Wage on Employment and Welfare Dependency," *German Economic Review*, 8 2019, *20* (3), 330–355.

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