Minimum Wage and Employer Variety*

Priyaranjan Jha and Antonio Rodriguez-Lopez†
University of California, Irvine and CESifo

First version: September 2021
This version: December 2022

Abstract

Exploiting minimum wage variation within multi-state commuting zones, we document a negative relationship between minimum wages and establishment counts in the United States. To explain this finding, we construct a heterogeneous-firm model with a monopsonistic labor market and endogenous firm variety. The decentralized equilibrium underprovides the mass of firms compared to the outcome achieved by a welfare-maximizing planner. A binding minimum wage further reduces the mass of firms, exacerbating the distortion. Workers value employer variety, and thus, by reducing firm variety the minimum wage reduces workers' welfare even if the average wage increases. Based on estimated elasticities, our model predicts that a 10 percent minimum wage hike reduces workers' welfare by 1.87 percent. The long-term effects are even larger, with an estimated welfare loss of 4.14 percent after four years.

JEL Classification: J38, J42

Keywords: minimum wage, number of firms, love of employer variety

*We thank Jeffrey Clemens, David Neumark and seminar participants at the Kiel Institute for the World Economy and the University of Nottingham for helpful comments and suggestions.

†E-mail addresses: pranjan@uci.edu (Jha) and jantonio@uci.edu (Rodriguez-Lopez).

1 Introduction

There is a large body of empirical research on the labor market implications of minimum wages, with the main focus being on the impact of the minimum wage on employment and average wages (see Neumark and Shirley, 2021 and Manning, 2021 for recent surveys). Although the evidence is mixed, studies that do not find a negative employment response to minimum wages (see, *e.g.*, Card and Krueger, 1994) often mention that a potential explanation is that labor markets are monopsonistic: facing an upward sloping labor supply, firms offer a wage that is below their marginal labor cost, and thus a minimum wage could lead to an increase in employment. This analysis, however, is incomplete because it assumes a fixed number of firms, and this number can be affected by a minimum wage. Simply put, a binding minimum wage increases firms' costs, which may lead to the exit of firms and reduce entry incentives, causing a decline in the mass of firms. This paper empirically analyzes the relationship between minimum wages and the number of firms, and studies the welfare implications of minimum wages when employer variety matters.

In section 2, we document a negative relationship between minimum wages and establishment counts in the United States. Using yearly U.S. data at the commuting zone–state level for the 1990-2016 period, we build on the cross-border design of Dube, Lester, and Reich (2010)—DLR hereafter—and estimate specifications that exploit minimum wage variation within multi-state commuting zones while controlling for time-varying spatial heterogeneity at the local level.[1] For the overall U.S. economy, we find a significant minimum wage elasticity of establishment counts of -0.091. After splitting U.S. establishments into 16 industries, we obtain negative elasticities for 13 of them, with significant estimates for the four lowest-earnings industries (including restaurants and retail trade). Moreover, we find statistically significant evidence of long-term effects of minimum wages on the number of U.S. establishments, with an estimated four-year elasticity of -0.224.

Previous literature has found evidence that minimum wages affect establishments' entry and exit dynamics. Luca and Luca (2019), for example, use Yelp ratings data from San Francisco restaurants and find that minimum wage hikes increase the exit of 3.5-star restaurants. Relatedly, using a cross-

[1] DLR study the effects of minimum wages on restaurant employment by exploiting minimum wage variation within county pairs sharing a state border, and find a positive—but small and insignificant—elasticity. Jha, Neumark, and Rodriguez-Lopez (2022) show that DLR's results crucially depend on their definition of local economic area. Using the same cross-border design, but defining local economic areas as multi-state commuting zones, Jha, Neumark, and Rodriguez-Lopez (2022) obtain a negative and significant minimum wage elasticity of restaurant employment of -0.141 when using DLR's QCEW 1990-2006 data, and an elasticity of -0.242 when using the same CBP 1990-2016 data as in this paper. Jha, Neumark, and Rodriguez-Lopez (2022) also show that multi-state commuting zones are better in capturing local economic shocks than cross-border county pairs. Allegretto, Dube, and Reich (2009) also use multi-state commuting zones in their analysis of minimum wages and teen employment, mentioning that: "these areas are not only contiguous; they are also demonstrably linked with each other by an economically meaningful criterion".

border county-pair approach, Aaronson, French, Sorkin, and To (2018) find that minimum wage hikes increase both the entry and exit of limited service restaurants, but the effect is larger on exit. As well, using U.S. establishment-level data at the county-industry level, Chava, Oettl, and Singh (2019) find that increases in the federal minimum wage increase exit rates and reduce entry rates. On the other hand, Rohlin (2011) finds that minimum wage hikes reduce entry in low-education U.S. industries but does not increase exit of existing businesses. In spite of their differences, all these studies imply a net reduction in the number of establishments after a minimum wage increase, which is consistent with our empirical results. Compared to these studies, our identification strategy is based on exploiting minimum wage variation within multi-state commuting zones.

To understand the mechanisms through which minimum wages affect firm variety, in section 3 we expand the monopsonistic labor market framework to endogenize the number of firms in a setting with heterogeneous firms. In our model, changes in the number of firms affect welfare because workers love variety of employers: the larger the number of firms a worker could choose to work for, the higher the worker's welfare is. Whereas the role of the number of firms is well understood for consumers' welfare—more firms imply a larger variety of goods—and the Dixit-Stiglitz model with love-of-variety preferences is the workhorse framework in industrial organization and international trade, the role of the number of firms in workers' welfare has received less attention. Card, Cardoso, Heining, and Kline (2018) discuss a model where workers have idiosyncratic preferences for employers due to factors such as location, work hours, and work culture. As first pointed out by Thisse and Toulemonde (2010), in such a setting the larger the mass of employers the greater the maximized expected utility of workers.[2]

Using a Melitz-type structure, we find that the decentralized equilibrium and the social planner's problem yield similar firm sizes, so that in spite of a monopsonistic labor market, there is no misallocation of resources across firms.[3] However, the mass of firms and total employment are smaller in the decentralized equilibrium than in the planner's solution. While the goods market is competitive in our set up, imperfect competition in the labor market leads firms to mark down wages. This results in leisure being cheaper than goods compared to what a planner would desire, resulting in lower labor supply and a smaller number of firms compared to the planner's problem. In this setting, a

[2]In related work, in Jha and Rodriguez-Lopez (2021) we study the effects of international trade on inequality and welfare when the labor market is monopsonistic and employer variety matters. Dustmann, Lindner, Schönberg, Umkehrer, and vom Berge (2021) study the reallocation effects of minimum wages following the introduction of a nationwide minimum wage in Germany that affected 15 percent of all employees. They find that the minimum wage raised wages and did not lower employment, but induced low-wage workers to move from small, low-paying firms to larger, higher-paying firms at the expense of increased commuting time. The latter finding highlights the importance of idiosyncratic, non-pecuniary factors in workers' preferences for employers.

[3]This parallels the result in Dhingra and Morrow (2019) where with CES preferences, monopolistic competition is efficient.

binding minimum wage wipes out the least productive firms, affecting the allocation of labor across firms, and reducing the total mass of firms and labor supply by workers. Since the mass of firms and labor supply by workers was suboptimal in the decentralized equilibrium, a binding minimum wage exacerbates the existing distortion and reduces total employment and welfare, even if the average wage increases.

2 Minimum Wages and the Number of Establishments: Empirical Assessment

This section documents a negative and statistically significant relationship between minimum wages and the number of establishments in the United States.

2.1 Data

From the Census's County Business Patterns (CBP) we obtain county-industry yearly establishment counts from 1990 to 2016, as well as annual payroll and employment.[4] We follow Jha, Neumark, and Rodriguez-Lopez (2022), who use the detailed programs of Acemoglu, Autor, Dorn, Hanson, and Price (2016), to process the CBP data into 479 industries and 722 commuting zones, but also splitting commuting zones by state. This yields 866 commuting zone-state entities, with 585 coming from single-state commuting zones, 130×2 from two-state commuting zones, and 7×3 from three-state commuting zones.[5] For our estimation strategy we focus on the 281 entities from the 137 multi-state commuting zones. These 137 zones—which account for 29.8% of U.S. employment and for 29% of U.S. establishments—contain 151 unique pairs: 130 pairs from the two-state commuting zones, and 21 pairs from the 7 three-state commuting zones.[6]

Whereas most minimum wage studies focus on the restaurant industry (including DLR and Jha, Neumark, and Rodriguez-Lopez, 2022), here we look at the effects of minimum wages on the overall number of establishments, as well as for 16 industries that encompass the entire U.S. private sector. In the classification of Acemoglu, Autor, Dorn, Hanson, and Price (2016), we use industries 5812

[4] Although the CBP data can be obtained up to 2019, we do not use 2017-2019 data because the CBP stopped reporting full establishment counts since the 2017 release. In particular, since 2017 the CBP omits establishment counts from county-industry cells with less than three establishments. This is an important limitation for CBP post-2016 data: whereas the 2016 data includes 784,474 county-industry cells with a positive establishment count, the 2017 data only includes 350,097.

[5] The District of Columbia, which belongs to the three-state commuting zone DC-VA-MD, is not included in our analysis because it appears in the CBP data starting in 2004. Therefore, we treat the DC-VA-MD area as a two-state commuting zone.

[6] Jha, Neumark, and Rodriguez-Lopez (2022) show that the 137 multi-state commuting zones are fairly distributed within the continental U.S., and that they follow similar patterns to the rest of the country for the number of establishments, total employment, employment-to-population ratios, earnings per worker, and average minimum wages.

(restaurants) and 5210 (retail trade), and aggregate the remaining 477 industries into 14 industries according to the classification used by the U.S. Bureau of Labor Statistics (BLS) in its yearly reports on minimum wage workers. Table 1 shows nominal earnings per worker, earnings rankings, and establishment and employment shares in U.S. totals for each of the 16 industries in 1990 and 2016. Table 1 sorts industries based on earnings per worker in 1990 (from lowest to highest), and below we use this ranking to identify each industry (*e.g.*, industry 1 refers to restaurants and industry 3 refers to retail trade).

In Table 1 we see that earnings-per-worker rankings by industry are very similar in both years. The eight lowest-earnings industries accounted for 52.5% of establishments and 50.6% of employment in 1990, and for 50.8% of establishments and 54% of employment in 2016. This masks, however, substantial heterogeneity across industries. For example, whereas restaurants, education and health services, and professional and business services had large increases in their shares from 1990 to 2016, industries such as retail trade, nondurable goods manufacturing, and durable goods manufacturing had large contractions.

Restaurants have by far the lowest earnings, which is the reason why most minimum wage studies focus on that industry.[7] Nevertheless, all industries are affected by minimum wages, as different firms within an industry pay different wages, and firms hire workers in different occupations and skill levels. To see this, for each of our 16 industries, Table 2 shows the fraction of hourly-paid workers that are paid at or below the federal minimum wage, as reported by the BLS for 2000, 2010, and 2016.[8] Importantly, the numbers reported in Table 2 understate the actual fractions of hourly-paid workers bound by the minimum wage, as a higher state-level minimum wage prevails in many states: whereas in 2000 only 10 states plus DC had a minimum wage above the federal level (being 16.5% higher on average), by 2016 there were 29 states plus DC (being 21% higher on average).

Table 2 shows that all industries have workers paid at or below the federal minimum wage. Indeed, restaurants is the industry where minimum wages bind at a wide scale, but also other low-earnings industries have non-negligible fractions. For example, in 2010 (the year after the last federal minimum wage increase), the fraction of bound workers was 6.5% overall, with the five lowest-earnings industries showing fractions between 7.3% for retail trade and 27.9% for restaurants.

In addition to CBP data, we obtain yearly working-age population at the commuting zone–state

[7]Compared to retail trade (the second or third lowest-earnings industry), earnings in restaurants were 43% lower in 1990 and 36% lower in 2016.

[8]The yearly report of the BLS on characteristics of minimum wage workers started in 2000, and is part of the labor force statistics constructed from the Current Population Survey. Data for all available years can be downloaded from FRED (Federal Reserve Economic Data). During the 2000-2016 period, the federal minimum wage was $5.15 up to July 2007, $5.85 up to July 2008, $6.55 up to July 2009, and $7.25 since then.

Table 1: Earnings per worker, establishment shares, and employment shares of 16 U.S. industries, 1990 and 2016

Industry	**1990** *Earnings ranking*	**1990** *Yearly earnings*	**1990** *Estab. share*	**1990** *Emp. share*	**2016** *Earnings ranking*	**2016** *Yearly earnings*	**2016** *Estab. share*	**2016** *Emp. share*
Restaurants	1	7.68	6.86%	7.21%	1	17.36	8.06%	9.37%
Accommodation sector	2	12.47	0.86%	1.71%	3	28.83	0.87%	1.69%
Retail trade	3	13.47	18.91%	13.94%	2	27.08	13.57%	12.81%
Other services	4	13.85	10.70%	4.90%	5	32.40	9.95%	4.65%
Arts and entertainment	5	15.62	1.28%	1.15%	4	32.15	1.77%	1.87%
Agriculture, forestry, fishing, hunting	6	16.71	1.68%	0.73%	6	37.25	2.33%	1.06%
Education and health services	7	21.45	10.03%	13.74%	7	46.40	12.90%	19.00%
Nondurable goods manufacturing	8	24.36	2.24%	7.26%	8	54.03	1.31%	3.56%
Professional and business services	9	24.97	11.45%	9.63%	12	63.10	14.64%	15.79%
Construction	10	25.22	10.12%	5.94%	10	58.69	9.06%	5.13%
Financial activities	11	27.66	9.83%	7.84%	16	85.04	11.45%	7.06%
Wholesale trade	12	27.97	8.02%	6.69%	13	66.68	6.23%	5.50%
Transportation and utilities	13	28.31	2.84%	4.22%	9	55.69	3.60%	4.29%
Durable goods manufacturing	14	29.03	3.46%	11.74%	11	59.84	2.33%	5.77%
Information industries	15	29.16	1.21%	2.64%	14	69.90	1.58%	1.93%
Mining, quarrying, oil and gas	16	33.53	0.49%	0.66%	15	83.76	0.37%	0.50%

Notes: Yearly earnings per worker are in thousands of U.S. dollars.

Table 2: Fraction of hourly-paid workers paid at or below federal minimum wage by industry

Industry	**2000**	**2010**	**2016**
Restaurants	22.8	27.9	15.5
Accommodation sector	5.5	10.0	3.2
Retail trade	3.2	7.3	1.9
Other services	4.6	8.8	2.7
Arts and entertainment	8.5	10.3	5.5
Agriculture, forestry, fishing, hunting	5.8	4.2	1.4
Education and health services	2.3	3.4	1.4
Nondurable goods manufacturing	1.5	3.0	1.1
Professional and business services	1.8	3.1	1.0
Construction	0.5	1.1	0.4
Financial activities	1.0	2.3	0.7
Wholesale trade	1.1	2.9	0.5
Transportation and utilities	1.1	1.7	1.2
Durable goods manufacturing	0.7	1.2	0.5
Information industries	1.7	3.9	1.6
Mining, quarrying, oil and gas	0.1	0.0	0.4
All private-sector employment	*3.8*	*6.5*	*2.9*

Source: FRED (Federal Reserve Economic Data) with data from the U.S. Bureau of Labor Statistics. Industries are sorted according to the 1990 earnings ranking of Table 1.

level from the Census Bureau's Population Estimates Program, while from Vaghul and Zipperer (2016) we obtain yearly minimum wage data at the state level, with the minimum wage defined as the largest of the federal minimum wage and the state minimum wage.[9]

In the end, our dataset includes industry-level and overall establishment counts (as well as employment and earnings per worker), working-age population, and minimum wage for each of the 281 commuting zone–state entities that make the 151 pairs from multi-state commuting zones. Similar to DLR, we arrange the data so that for each industry in each period there are 302 observations (two for each pair), with each observation identified by its commuting zone, state, and pair: 260 commuting zone–state entities appear once (from the 130 two-state commuting zones), whereas 21 appear twice (from the 7 three-state commuting zones). If an industry appears in all commuting zone–state entities in all years, the maximum number of observations would be $302 \times 27 = 8,154$. In our regressions below for different industries, the number of observations ranges between $7,022$ and $8,134$.

[9]The minimum wage data is available at this link.

2.2 Econometric Specification and Results

DLR argue that conventional estimates of the minimum wage elasticity of employment—from two-way fixed effects specifications—are biased because they fail to account for spatial heterogeneity at the local level. To control for local economic conditions, they introduce a cross-state county-pair approach that exploits minimum wage variation within local economic areas.

Along these lines, but defining a local economic area as a pair within a multi-state commuting zone rather than as a pair of contiguous counties sharing a state border (Jha, Neumark, and Rodriguez-Lopez, 2022), our econometric specification for the relationship between the number of establishments and minimum wages is

$$\ln e_{ipt} = \alpha + \beta \ln MW_{it} + \gamma \ln P_{it} + \eta_i + \tau_{pt} + \nu_{it}, \tag{1}$$

where for commuting zone–state i from pair p in year t, e_{ipt} is the number of establishments, MW_{it} is the minimum wage, P_{it} is the working-age population, η_i is a commuting zone–state i fixed effect, and τ_{pt} denotes pair–year fixed effects (which control for spatial heterogeneity at the local level), and ν_{it} is the error term.

From the previous econometric specification, our coefficient of interest is β, which denotes the elasticity of the number of establishments to the minimum wage. Using the multi-way fixed-effect estimator of Correia (2016), which allows us to easily control for commuting zone–state fixed effects and pair–year fixed effects, Table 3 presents our results from the estimation of (1) for each of our 16 industries and overall. As in DLR, standard errors are two-way clustered at the state and border-segment levels using the procedure of Cameron, Gelbach, and Miller (2011), with a border segment defined as a pair of states sharing a multi-state commuting zone.

Table 3 shows an overall minimum wage elasticity of establishment counts of −0.091, which is significant at a 10% level. Moreover, Table 3 shows a negative relationship between minimum wages and establishment counts for 13 out of 16 industries, with the elasticity being statistically significant for five of them. From the numbers in Table 1, the industries with negative elasticities account for 94.7% of establishments in 1990, and for 93.1% in 2016, whereas the four lowest-earnings industries—all of which have significant elasticities—account for 37.3% of establishments in 1990 and for 32.5% in 2016. The point estimates for the statistically significant elasticities for the low-earnings industries are −0.208 for restaurants (industry 1), −0.194 for the accommodation sector (industry 2), −0.137 for retail trade (industry 3), and −0.09 for other services (industry 4).

Although the negative and significant elasticities for the four lowest-earnings industries are not surprising, the size and significance of the elasticity for the highest earnings industry (mining, quar-

Table 3: Pair-approach estimation of minimum wage responses of establishment counts in the U.S., 1990–2016

	Sixteen industries (sorted by 1990 earnings per worker)					
Industry →	**Overall**	**1**	**2**	**3**	**4**	**5**
ln(minimum wage)	-0.091*	-0.208*	-0.194*	-0.137**	-0.090*	0.016
	(0.048)	(0.119)	(0.110)	(0.053)	(0.052)	(0.149)
ln(population)	0.866***	1.063***	0.615***	0.641***	0.725***	1.072***
	(0.083)	(0.118)	(0.204)	(0.092)	(0.132)	(0.247)
Observations	8,134	8,134	7,798	8,134	8,132	7,920
Industry →	**6**	**7**	**8**	**9**	**10**	**11**
ln(minimum wage)	-0.001	-0.158	-0.234	-0.118	-0.105	-0.047
	(0.157)	(0.102)	(0.148)	(0.113)	(0.102)	(0.071)
ln(population)	1.206***	0.971***	1.259**	1.122***	0.853***	0.858***
	(0.199)	(0.132)	(0.513)	(0.229)	(0.192)	(0.123)
Observations	7,962	8,128	7,592	8,090	8,128	8,128
Industry →	**12**	**13**	**14**	**15**	**16**	
ln(minimum wage)	-0.048	0.028	-0.153	0.117	-1.585***	
	(0.093)	(0.101)	(0.109)	(0.182)	(0.423)	
ln(population)	0.753***	0.623***	0.362**	0.633*	1.400*	
	(0.131)	(0.182)	(0.174)	(0.361)	(0.777)	
Observations	8,094	8,112	7,808	8,014	7,022	

Notes: This table reports $\hat{\beta}$ and $\hat{\gamma}$ from the estimation of equation (1) for 16 industries and overall using yearly data from 1990 to 2016. All regressions include commuting zone–state fixed effects and pair–year fixed effects. Industries are sorted according to the 1990 earnings ranking of Table 1. Standard errors (in parentheses) are two-way clustered at the state and border segment levels. The coefficients are statistically significant at the *10%, **5%, or ***1% level.

rying, oil, and gas) is puzzling. This industry not only has the highest earnings per worker, but also has the smallest fraction of workers bound by the federal minimum wage (see Table 2). The following section shows that this result is a consequence of a very strong negative pre-existing trend in establishment counts for this industry. Importantly, this is the smallest industry—accounting for only 0.49% of establishments in 1990—and has a negligible impact on the overall estimated elasticity.[10]

Table 3 shows that the population control is positive and significant overall, so that a larger population is associated with more establishments. As well, the population coefficient is positive and significant in all individual industries. To account for cyclical changes in the number of establishments, Table A-1 in the Appendix re-estimates specification (1) for each of the 16 industries, but

[10]If we remove the mining, quarrying, oil and gas industry from our sample, the overall estimated elasticity changes slightly from −0.091 to −0.087 and retains its significance.

adds as control the log of establishment counts in all the other industries.[11] Table A-1 shows that 11 out of 16 estimated minimum wage elasticities are negative, but they become slightly smaller in size and lose significance when compared to those in Table 3. However, using establishment counts in the rest of the industries as control leads to an attenuation bias of the minimum wage effect, as minimum wages are negatively related to the number of establishments in most industries. That is, even though the regressor is intended to account for cyclical effects on establishment counts, it is a "bad control" or overcontrol.

2.3 Long-Term Effects of Minimum Wages

To assess the long-term effects of minimum wages on U.S. establishment counts, we follow DLR and estimate a distributed-lag version of (1) that includes two years of leads and four years of lags (see also Jha, Neumark, and Rodriguez-Lopez, 2022). The distributed-lag specification is

$$\ln e_{ipt} = \alpha + \sum_{k=-2}^{3} \beta_{-k} \Delta \ln MW_{i,t-k} + \beta_{-4} \ln MW_{i,t-4} + \gamma \ln P_{it} + \eta_i + \tau_{pt} + \nu_{it}, \tag{2}$$

where Δ is a one-year difference operator. From (2), the seven β parameters give the cumulative minimum wage elasticity of establishment counts, starting with β_2, which represents the lead effect two years before the minimum wage change, and up to β_{-4}, which denotes the cumulative elasticity four years after the minimum wage change. We estimate specification (2) for the 16 industries and overall, presenting all the results in Table A-2 in the Appendix.

For the overall regression, Figure 1 shows the estimates of the β parameters along with 90% confidence intervals. We assume that the minimum wage change occurs at time s, and hence the plot starts at $s-2$ with $\hat{\beta}_2$, and continues through $s+4$ with $\hat{\beta}_{-4}$. The plot shows statistically significant medium- and long-term negative effects of minimum wages on U.S. establishment counts, with a clear change in the slope occurring around time s. The cumulative elasticity becomes significant at a 10% level at $s+2$ with a point estimate of -0.121, and reaches -0.224 after four years of the minimum wage change. Notice that the four-year elasticity is about 2.5 times larger than the -0.091 contemporaneous elasticity of the previous section. Thus, there is evidence of persistent negative effects of minimum wages on the number of U.S. establishments.

For individual industries, Table A-2 shows that 15 out of 16 industries have negative four-year estimated elasticities, with five of them being statistically significant. Of those that are significant, the estimated four-year elasticity is -0.662 for restaurants, -0.298 for retail trade, -0.402 for education

[11] The Appendix is available at http://www.socsci.uci.edu/~jantonio/Papers/minwage_variety_app.pdf.

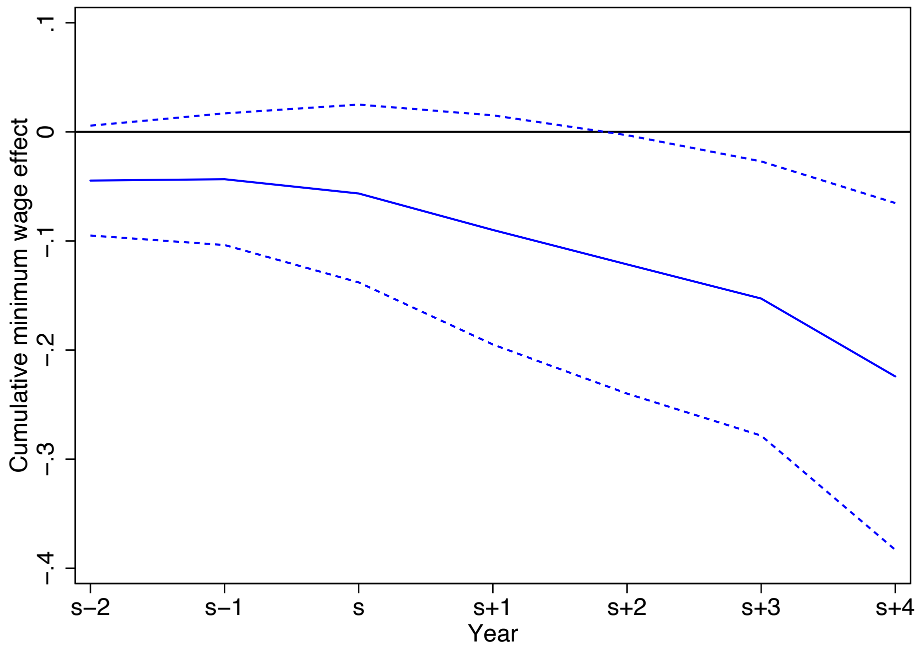

Figure 1: Time path of minimum wage effects on U.S. establishment counts with 90% confidence intervals

and health services. The other two industries with significant four-year elasticities are the mid-wage nondurable goods manufacturing industry (with a point estimate of −0.614), and the high-wage industry of mining, quarrying, oil and gas industry (with a point estimate of −2.593). These industries, however, have large, negative, and significant estimated elasticities two years before the minimum wage change (−0.32 and −1.145), which indicates important pre-existing trends. Given the well-documented secular decline in U.S. manufacturing during the last decades, the pre-trend for nondurable goods manufacturing is not surprising—Table 1 shows that this industry accounted for 2.24% of establishments and 7.26% of employment in 1990, whereas these shares decline to 1.31% and 3.56% by 2016. Although much smaller, the industry of mining, quarrying, oil, and gas also declined in importance during the 1990-2016 period, reducing its shares from 0.49% to 0.37% for establishments, and from 0.66% to 0.5% for employment.

3 The Model

This section presents a model with a monopsonistic labor market, heterogeneous firms, and endogenous firm variety. The model shows that a binding minimum wage reduces the mass of firms and employment, which then reduces welfare. Here we present the relevant parts of the model, and leave the most technical details for section B in the Appendix.

3.1 Household Preferences and Production

The economy produces a single final good that is chosen as the numéraire. The final good is traded in a perfectly competitive market, and is produced by a finite mass, M, of heterogeneous firms. Firms produce the final good using labor, which is procured from a unit measure of households participating in a monopsonistic labor market.

Similar to Berger, Herkenhoff, and Mongey (2022), the utility function of the representative household is

$$\mathbb{U} \equiv C - \frac{N^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}}, \tag{3}$$

where C is the consumption of the final good, N is the labor-supply index, and ψ is the Frisch elasticity of labor supply. The labor-supply index is defined as

$$N = \left(\int_{\omega \in \Omega} l(\omega)^{\frac{1+\theta}{\theta}} d\omega \right)^{\frac{\theta}{1+\theta}}, \tag{4}$$

where $l(\omega)$ is the amount of labor supplied to firm ω, Ω is a set of measure M of firms/employers offering jobs, and $\theta > \psi$ (this ensures that the objective function is concave). The parameter $\theta > 0$ is the elasticity of substitution across jobs from different firms, and accounts for workers' cost of mobility: the lower the value of θ, the higher the mobility costs, and thus the greater the monopsony power of firms.

The above specification captures the love-of-variety for employers. To see this clearly, suppose there are M identical firms, so that $l(\omega) = l$ is the same across firms and $N = M^{\frac{\theta}{1+\theta}} l$. For a constant amount of supplied labor, $L = Ml$, note that the labor-supply index can be written as $N = \frac{L}{M^{1/(1+\theta)}}$. Hence, the disutility from supplying the same amount of labor is lower the larger M is.

Given the utility function in (3) and wages $w(\omega)$, for $\omega \in \Omega$, the representative household maximizes its utility by choosing, C, N, and allocating labor, $l(\omega)$, across firms. As shown in the Appendix, this maximization exercise yields the firm-level labor supply function

$$l(\omega) = \frac{w(\omega)^{\theta}}{W^{\theta-\psi}}, \tag{5}$$

for $\omega \in \Omega$, where W is an index of the wages available to the representative household, which is defined as

$$W = \left(\int_{\omega \in \Omega} w(\omega)^{1+\theta} d\omega \right)^{\frac{1}{1+\theta}}. \tag{6}$$

Firms take W as given, and hence, the function in (5) features a constant wage-elasticity of labor supply, θ.[12] The firm-level labor supply is increasing in $w(\omega)$, and thus, the firm has monopsony power in the labor market.

[12] Jha and Rodriguez-Lopez (2021) and Egger, Kreickemeier, Moser, and Wrona (2021) also derive a constant

The maximization problem also yields $N = W^{\psi}$ and hence, our assumption that $\theta > \psi$ simply means that the elasticity of the labor-supply index (also known as the Frisch elasticity) is smaller than the elasticity of firm-level labor supply. The wage bill of firms, which is given by $\int_{\omega\in\Omega} w(\omega)l(\omega)d\omega = WN = W^{1+\psi}$, equals the consumption of the household, C. Therefore, the welfare of the representative household in (3) can be rewritten as

$$\mathbb{U} = WN - \frac{N^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} = \frac{W^{1+\psi}}{1+\psi}, \tag{7}$$

Thus, welfare is increasing in the wage index, W.

As in Melitz (2003), after incurring an entry cost of f_E in terms of the final good, each firm draws its productivity φ from a distribution $G(\varphi)$. The firm produces for the market if and only if it can cover a fixed cost of operation, f, also in terms of the final good.[13] After meeting the fixed cost, the production function for a firm with productivity φ is $y(\varphi) = \varphi l$. Whereas the goods market is perfectly competitive, the labor market is monopsonistically competitive. Given (5) and using φ instead of ω to identify each firm, the profit maximization problem of a firm with productivity φ yields as solution

$$l(\varphi) = \left(\frac{\theta}{1+\theta}\right)^{\theta} \frac{\varphi^{\theta}}{W^{\theta-\psi}} \quad \text{and} \quad w(\varphi) = \left(\frac{\theta}{1+\theta}\right)\varphi. \tag{8}$$

The solution shows that employment and wages are increasing in φ: higher productivity firms employ more workers and pay higher wages. Note also that the solution for $w(\varphi)$ shows that a worker is paid a fraction $\frac{\theta}{1+\theta}$ of the value of the worker's marginal product, φ (*i.e.*, the proportional markdown on wages, $\frac{1}{1+\theta}$, is the same for every firm). It follows that the gross profit function of a firm with productivity φ, $\pi(\varphi) = \varphi l(\varphi) - w(\varphi)l(\varphi)$, is given by

$$\pi(\varphi) = \left[\frac{\theta^{\theta}}{(1+\theta)^{1+\theta}}\right] \frac{\varphi^{1+\theta}}{W^{\theta-\psi}}. \tag{9}$$

3.2 Decentralized Equilibrium and the Social Planner's Problem

As in Melitz (2003), there is a cutoff level of productivity, $\hat{\varphi}$, such that firms with productivity below $\hat{\varphi}$ cannot cover their fixed cost of operation and hence do not survive. The zero-cutoff-profit condition is $\pi(\hat{\varphi}) = f$, which from (9) implies that $W^{\theta-\psi} = \left[\frac{\theta^{\theta}}{(1+\theta)^{1+\theta}}\right] \frac{\hat{\varphi}^{1+\theta}}{f}$, so that we can rewrite firm-level employment in (8) as $l(\varphi) = \frac{(1+\theta)f\varphi^{\theta}}{\hat{\varphi}^{1+\theta}}$ and the gross profit function in (9) as $\pi(\varphi) = \left(\frac{\varphi}{\hat{\varphi}}\right)^{1+\theta} f$. Firms

elasticity labor supply function by using a random utility framework where workers have idiosyncratic preferences for employers. Close to the endogenous-labor-supply feature of our model, Bhaskar and To (1999) study the implications of minimum wages in a setting where firms' monopsony power emerges due to workers' commuting costs, so that firms must pay higher wages to attract workers from farther away.

[13] We model entry and fixed costs in terms of the final good so that it is the same for all firms. Modeling it in terms of labor will make it different for firms because each firm pays a different wage in equilibrium.

enter up to the point that the expected value of entry, $\int_{\hat{\varphi}}^{\infty} [\pi(\varphi) - f]\, g(\varphi) d\varphi$, is equal to the entry cost, f_E; therefore, the free entry condition is given by

$$\int_{\hat{\varphi}}^{\infty} \left[\left(\frac{\varphi}{\hat{\varphi}} \right)^{1+\theta} - 1 \right] f g(\varphi) d\varphi = f_E. \tag{10}$$

Once we solve for $\hat{\varphi}$ from (10), we obtain the equilibrium W from the zero-cutoff-profit condition, and thus, using (7) we obtain that welfare is $\mathbb{U} = \frac{1}{1+\psi} \left\{ \left[\frac{\theta^\theta}{(1+\theta)^{1+\theta}} \right] \frac{\hat{\varphi}^{1+\theta}}{f} \right\}^{\frac{1+\psi}{\theta-\psi}}$.

Letting M_E be the mass of entrants, the mass of producing firms, M, is simply the fraction of entrants that survive: $M = [1 - G(\hat{\varphi})] M_E$. The equilibrium M is determined from the definition of W in (6), which can be rewritten as

$$W = \left[M \int_{\hat{\varphi}}^{\infty} w(\varphi)^{1+\theta} g(\varphi | \varphi \geq \hat{\varphi}) d\varphi \right]^{\frac{1}{1+\theta}}. \tag{11}$$

Finally, total employment is obtained by aggregating firm-level employment across all producing firms, $L = M \int_{\hat{\varphi}}^{\infty} l(\varphi) g(\varphi | \varphi \geq \hat{\varphi}) d\varphi$, which can be rewritten as

$$L = \frac{(1+\theta) f M}{\hat{\varphi}^{1+\theta}} \int_{\hat{\varphi}}^{\infty} \varphi^\theta g(\varphi | \varphi \geq \hat{\varphi}) d\varphi. \tag{12}$$

In the following, we refer to the decentralized equilibrium values as $\hat{\varphi}_D$, $l_D(\varphi)$, $\mathbb{U}_D$, M_D, and L_D, where subscript D denotes 'dencentralized'.

A social planner chooses $\hat{\varphi}$, $l(\varphi)$, and the mass of entrants, M_E, so as to maximize (3) subject to (4) and final-good consumption (C) being equal to total final-good production minus final-good requirements to cover firms' entry and fixed costs. Section B.3 in the Appendix shows the details of this standard maximization problem. Letting $\hat{\varphi}_P$, $l_P(\varphi)$, $\mathbb{U}_P$, M_P, and L_P denote the solution values from the planner's problem, the Appendix shows that $\hat{\varphi}_P = \hat{\varphi}_D$, $l_P(\varphi) = l_D(\varphi)$ for every φ,

$$\frac{M_P}{M_D} = \frac{L_P}{L_D} = \left(\frac{1+\theta}{\theta} \right)^{\frac{(1+\theta)\psi}{\theta-\psi}} > 1 \quad \text{and} \quad \frac{\mathbb{U}_P}{\mathbb{U}_D} = \frac{\theta-\psi}{(1+\theta)} \left(\frac{1+\theta}{\theta} \right)^{\frac{\theta(1+\psi)}{\theta-\psi}} > 1.$$

Thus, the decentralized equilibrium has the same productivity cutoff and the same firm level employment as in the planner's problem, but the planner chooses a higher level of M, which results in a higher L and higher welfare. In other words, the decentralized equilibrium provides sub-optimal mass of firms and employment.

The distortion in the model is coming from the fact that wages are marked down while consumption goods are not. The marginal rate of transformation (MRT) between leisure and consumption is different in the market than in the planner's problem. In the planner's problem, one extra unit of labor is transformed into φ units of a consumption good at a firm with productivity φ. So, the MRT between consumption and leisure is φ. In the market, the price of one unit of leisure (same as the reward of one unit of labor) is $w(\varphi)$ while the price of a good is $p = 1$. Therefore, the opportunity cost of one unit of leisure is $w(\varphi)$ units of a consumption good. Since $w(\varphi) = \frac{\theta}{1+\theta}\varphi < \varphi$, the opportunity cost of leisure, and hence the MRT, is less in the market equilibrium than in the planner's problem. As a result, more leisure is consumed in the market equilibrium leading to a lower supply of labor and a lower mass of firms.

A natural question is whether there is a policy that can correct the distortion and restore optimality in the decentralized equilibrium. In section B.4 in the Appendix we show that this is indeed the case: a leisure tax (which effectively is a labor subsidy) of $\frac{1}{\theta}$ financed by a lump-sum tax on households restores optimality, so that the mass of firms, labor supply, and welfare in the decentralized case correspond to the planner's problem solution.[14]

3.3 Impact of a Minimum Wage Regulation

Denote a binding minimum wage by $\underline{w}$. Since the lowest productivity firm that survives in a decentralized equilibrium without any policy intervention offers a wage of $w(\hat{\varphi}_D)$, for the minimum wage to be binding it must be the case that $\underline{w} > w(\hat{\varphi}_D)$. Let $\underline{\varphi}$ denote lowest productivity level for which the desired wage offered by the firm, $\left(\frac{\theta}{1+\theta}\right)\underline{\varphi}$, equals $\underline{w}$. Therefore, $\underline{\varphi} = \frac{(1+\theta)\underline{w}}{\theta}$, and firms with $\varphi \geq \underline{\varphi}$ are unconstrained by the minimum wage, with their gross profits given by $\underline{\pi}(\varphi) = \left[\frac{\theta^\theta}{(1+\theta)^{1+\theta}}\right]\frac{\varphi^{1+\theta}}{\underline{W}^{\theta-\psi}}$, where $\underline{W}$ is the wage index.

Let $\hat{\underline{\varphi}}$ denote the cutoff level of productivity in an equilibrium with a binding minimum wage, so that firms with $\varphi \in [\hat{\underline{\varphi}}, \underline{\varphi})$ are minimum wage constrained. These firms' gross profits are directly proportional to their employed labor, $\underline{\pi}(\varphi) = (\varphi - \underline{w})l(\varphi)$, and therefore, each of them hires $\frac{\underline{w}^\theta}{\underline{W}^{\theta-\psi}}$, which from (5) we know is the amount of labor that workers supply to a firm paying wage $\underline{w}$. The zero-cutoff-profit condition, $\underline{\pi}(\hat{\underline{\varphi}}) = f$, is then $(\hat{\underline{\varphi}} - \underline{w})\frac{\underline{w}^\theta}{\underline{W}^{\theta-\psi}} = f$. Lastly, the free entry condition is $\int_{\hat{\underline{\varphi}}}^{\infty}[\underline{\pi}(\varphi) - f]g(\varphi)d\varphi = f_E$, which can be rewritten as

$$\int_{\hat{\underline{\varphi}}}^{\underline{\varphi}}\left[\frac{(\varphi - \underline{w})\underline{w}^\theta}{\underline{W}^{\theta-\psi}} - f\right]g(\varphi)d\varphi + \int_{\underline{\varphi}}^{\infty}\left\{\left[\frac{\theta^\theta}{(1+\theta)^{1+\theta}}\right]\frac{\varphi^{1+\theta}}{\underline{W}^{\theta-\psi}} - f\right\}g(\varphi)d\varphi = f_E. \quad (13)$$

Using the definition of $\underline{\varphi}$, we solve for the equilibrium values of $\hat{\underline{\varphi}}$ and $\underline{W}$ from the zero-cutoff-profit condition and the free entry condition. We then solve for the equilibrium mass of firms ($\underline{M}$), total

[14] Cahuc and Laroque (2014) study optimal policies in a monopsony model where workers are heterogeneous in productivity and in working opportunity costs—heterogeneous opportunity costs generate an upward-sloping labor supply curve, giving rise to monopsony power (as in Bhaskar and To, 1999). Similar to our findings, the monopsony distortion leads to suboptimal employment, which is corrected by a wage subsidy. The possibility of suboptimal employment also arises in the search model of Flinn (2006) if the bargaining power of workers is less than the elasticity of the matching function with respect to searchers. A minimum wage could potentially correct this distortion and increase employment.

employment ($\underline{L}$), and welfare ($\underline{\mathbb{U}}$) using similar expressions to those in section 3.2. As well, the average wage, $\underline{\bar{w}}$, can be calculated as the ratio of the total wage bill ($\underline{W}^{1+\psi}$) and total employment. The following proposition presents our model's main results.

Proposition 1. ***(The effects of a binding minimum wage)***

A binding minimum wage wipes out the least productive firms, and reduces the total mass of firms, the wage index, and workers' welfare. If productivity follows a Pareto distribution, then a binding minimum wage also reduces total employment and increases the average wage.

The proof of this proposition is in section B.5 in the Appendix. Given that the minimum wage equilibrium approaches the decentralized equilibrium when $\underline{w} \to w(\hat{\varphi}_D)$, the proof simply shows that *(i)* for every productivity distribution, it holds that $\frac{d\underline{\hat{\varphi}}}{d\underline{w}} > 0$, $\frac{d\underline{M}}{d\underline{w}} < 0$, $\frac{d\underline{W}}{d\underline{w}} < 0$, $\frac{d\underline{\mathbb{U}}}{d\underline{w}} < 0$, and thus $\underline{\hat{\varphi}} > \hat{\varphi}_D$, $\underline{M} < M_D$, $\underline{W} < W_D$, and $\underline{\mathbb{U}} < \mathbb{U}_D$; and that *(ii)* if the productivity distribution is Pareto, then it also unambiguously holds that $\frac{d\underline{L}}{d\underline{w}} < 0$ and $\frac{d\underline{\bar{w}}}{d\underline{w}} > 0$, so that $\underline{L} < L_D$ and $\underline{\bar{w}} > \bar{w}_D$, where $\bar{w}_D$ is the average wage in the decentralized equilibrium.[15]

Supporting Proposition 1, our main empirical finding from section 2 is that an increase in the minimum wage is associated with reductions in U.S. establishment counts. In the model, a binding minimum wage makes the survival of low-productivity firms harder. This is captured by an increase in the cutoff productivity level, which then leads to a reduction in the mass of firms. The reduction in $\underline{M}$ causes a decline in the wage index, $\underline{W}$, which then translates into a reduction in welfare, $\underline{\mathbb{U}} = \frac{\underline{W}^{1+\psi}}{1+\psi}$. To see the role of workers' love-of-variety for employers in driving the welfare results it is useful to refer to the definition of the wage index in (11), which depends on the mass of firms and firm-level wages. Whereas firm-level wages increase to $\underline{w}$ for constrained firms, which pushes for an increase in $\underline{W}$, the reduction in $\underline{M}$ works in the opposite direction and is the dominant force.[16]

Notice that in spite of Melitz-type reallocation of resources from less productive firms to more productive firms, the minimum wage does not increase welfare. To understand why this is the case, recall from section 3.2 that the allocation of resources in the decentralized equilibrium is efficient

[15] Since Chaney (2008), the Pareto distribution has been used extensively in applications of the Melitz model, as it brings substantial tractability gains. This is also true for our case, where it allows us to obtain closed-form solutions for ${d\underline{L}}/{d\underline{w}}$ and ${d\underline{\bar{w}}}/{d\underline{w}}$. Although the second part of Proposition 1 cannot be proved for a general productivity distribution, we verified numerically that it holds for a lognormal distribution, $g(\varphi) = \frac{1}{\varphi\sqrt{2\pi\rho}} \exp\left(-{(\ln\varphi - \mu)^2}/{2\rho}\right)$, with parameter values $\mu = -0.02$ and $\rho = 0.35$. These values are from Combes, Duranton, Gobillon, Puga, and Roux (2012), who find that firm-level productivity of French firms is better approximated by a mix 95% lognormal and 5% Pareto, and that restricting the distribution to be 100% lognormal yields our assumed values.

[16] Note that firm-level wages for unconstrained firms do not change ($w(\varphi) = {\theta\varphi}/{(1+\theta)}$ for $\varphi \geq \underline{\varphi}$). If we assume instead a firm-level production function with decreasing returns to scale, $y(\varphi) = \varphi l^{\lambda}$, for $\lambda < 1$, the equilibrium firm-level wage for unconstrained firms is given by $w(\varphi) = \{[{\lambda\theta}/{(1+\theta)}]\, \underline{W}^{(\theta-\psi)(1-\lambda)}\varphi\}^{1/[1+\theta(1-\lambda)]}$, which is increasing in $\underline{W}$. The result ${d\underline{W}}/{d\underline{w}} < 0$ still holds with $\lambda < 1$, and therefore, a binding minimum wage reduces unconstrained firms' wages.

($\hat{\varphi}_P = \hat{\varphi}_D$ and $l_P(\varphi) = l_D(\varphi)$ for every φ). Therefore, any reallocation induced by a binding minimum wage cannot be a source of welfare gain. On the other hand, we also know from section 3.2 that the mass of firms is suboptimal in the decentralized equilibrium. By further reducing the mass of firms, a binding minimum exacerbates the existing distortion.[17]

3.4 Welfare Effects of Minimum Wages

Given that the wage bill, $\underline{L}\,\bar{\underline{w}}$, equals $\underline{W}^{1+\psi}$, from (7) we can rewrite the expression for welfare as $\underline{\mathbb{U}} = \frac{\underline{L}\,\bar{\underline{w}}}{1+\psi}$. It follows that

$$\zeta_{\underline{\mathbb{U}},\underline{w}} = \zeta_{\underline{L},\underline{w}} + \zeta_{\bar{\underline{w}},\underline{w}}, \tag{14}$$

where $\zeta_{\underline{\mathbb{U}},\underline{w}}$ is the minimum wage elasticity of welfare, and $\zeta_{\underline{L},\underline{w}}$ and $\zeta_{\bar{\underline{w}},\underline{w}}$ are minimum wage elasticities of total employment and the average wage. Hence, to quantify the welfare implications of our model, we use the data described in section 2.1 to estimate the elasticities on the right-hand side of equation (14).

Similar to Table 3, Tables A-3 and A-4 in the Appendix present the results from the estimation of equation (1), but using employment and earnings per worker instead of the number of establishments. The point estimate for the overall minimum wage elasticity of employment is significant and equal to -0.204, whereas the point estimate for the overall minimum wage elasticity of earnings per worker is positive but insignificant at just 0.017. Hence, in accordance with Proposition 1, our empirical results show that an increase in the minimum wage is associated with lower employment and higher earnings per worker in the United States, though the effect on worker earnings is small. Based on these values, from (14) we obtain that the estimated minimum wage elasticity of welfare is -0.187 so that, as predicted by Proposition 1, a 10 percent increase in the minimum wage is associated with a 1.87 percent loss in welfare.[18]

We also estimate equation (2) for employment and earnings per worker to calculate the long-term welfare effects of minimum wage changes. Tables A-5 and A-6 in the Appendix show the results. The point estimates for the overall four-year minimum wage elasticities are a statistically significant -0.418 for employment, and an insignificant 0.004 for worker earnings. Therefore, after four years

[17] Berger, Herkenhoff, and Mongey (2022) show that a binding minimum wage in their oligopsony model can alleviate the distortion induced by firms' markdowns on wages, but in contrast to our model, they abstract from the welfare effects of employer variety because they consider a fixed number of firms.

[18] Note that the estimated elasticity of employment (-0.204) is larger in magnitude than the estimated elasticity of establishment counts (-0.091), which implies that average firm size declines with the minimum wage. On the other hand, our model predicts that average firm size increases with the minimum wage, which is a consequence of our simplifying assumption of a production function with constant returns to scale. In the case with decreasing returns to scale (see footnote 16) it is possible that average firm size declines in response to the minimum wage. This possibility arises because in that case the minimum wage can reduce employment in some constrained firms.

the minimum wage elasticity of employment doubles, whereas the elasticity of earnings becomes essentially zero, yielding a long-term minimum wage elasticity of welfare of -0.414. Hence, a 10 percent increase in the minimum wage causes a 4.14 percent welfare loss after four years.

Tables A-3–A-6 also show estimated elasticities for the 16 industries. For restaurants—the industry with the lowest earnings per worker—our estimated contemporaneous minimum wage elasticities are -0.273 for employment and 0.164 for earnings, whereas the four-year elasticities are -0.719 for employment and 0.169 for earnings, with all of them being significant. Thus, the contemporaneous minimum wage elasticity of welfare in the restaurant industry is -0.109, but it reaches a value of -0.55 after four years. Nevertheless, and in contrast to the overall welfare elasticity, industry-level welfare elasticities do not take into account that workers that lose their job in one industry may find a job in another industry.

4 Conclusion

This paper showed that minimum wages are associated with reductions in establishment counts in the United States. Whereas we presented a simple model with competitive firms and monopsonistic labor markets to explain this fact, the model's key insights regarding the impact of a binding minimum wage on the mass of firms and employment will go through even if the product market were imperfectly competitive, as long as the number of firms is endogenous. The welfare effects would be different, however, depending on the precise specification of imperfect competition. For example, with standard CES preferences and monopolistic competition, a minimum wage-induced reduction in the mass of firms will reduce welfare through the consumption channel, in addition to the welfare loss through the employer variety channel.

References

Aaronson, D., E. French, I. Sorkin, and T. To (2018): "Industry Dynamics and the Minimum Wage: A Putty-Clay Approach," *International Economic Review*, 59(1), 51–84.

Acemoglu, D., D. Autor, D. Dorn, G. H. Hanson, and B. Price (2016): "Import Competition and the Great US Employment Sag of the 2000s," *Journal of Labor Economics*, 34(S1), S141–S198.

Allegretto, S., A. Dube, and M. Reich (2009): "Spatial Heterogeneity and Minimum Wages: Employment Estimates for Teens Using Cross-State Commuting Zones," Working paper, Institute for Research on Labor and Employment.

Berger, D., K. Herkenhoff, and S. Mongey (2022): "Labor Market Power," *American Economic Review*, 112(4), 1147–93.

Bhaskar, V., and T. To (1999): "Minimum Wages for Ronald McDonald Monopsonies: A Theory of Monopsonistic Competition," *The Economic Journal*, 109(455), 190–203.

Cahuc, P., and G. Laroque (2014): "Optimal Taxation and Monopsonistic Labor Market: Does Monopsony Justify the Minimum Wage?," *Journal of Public Economic Theory*, 16(2), 259–273.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011): "Robust Inference With Multiway Clustering," *Journal of Business & Economic Statistics*, 29(2), 238–249.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018): "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 36(S1), S13–S70.

Card, D., and A. B. Krueger (1994): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84(4), 772–793.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), 1707–21.

Chava, S., A. Oettl, and M. Singh (2019): "Does a One-Size-Fits-All Minimum Wage Cause Financial Stress for Small Businesses?," Working Paper 26523, National Bureau of Economic Research.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012): "The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection," *Econometrica*, 80(6), 2543–2594.

Correia, S. (2016): "A Feasible Estimator for Linear Models with Multi-Way Fixed Effects," Discussion paper, Duke University, Working Paper.

Dhingra, S., and J. Morrow (2019): "Monopolistic Competition and Optimum Product Diversity Under Firm Heterogeneity," *Journal of Political Economy*, 127(1), 196–232.

Dube, A., T. W. Lester, and M. Reich (2010): "Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties," *Review of Economics and Statistics*, 92(4), 945–964.

Dustmann, C., A. Lindner, U. Schönberg, M. Umkehrer, and P. vom Berge (2021): "Reallocation Effects of the Minimum Wage," *Quarterly Journal of Economics*, 137(1), 267–328.

Egger, H., U. Kreickemeier, C. Moser, and J. Wrona (2021): "Exporting and Offshoring with Monopsonistic Competition," *The Economic Journal*, ueab078.

Flinn, C. J. (2006): "Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates," *Econometrica*, 74(4), 1013–1062.

Jha, P., D. Neumark, and A. Rodriguez-Lopez (2022): "What's Across the Border? Re-Evaluating the Cross-Border Evidence on Minimum Wage Effects," Working paper, University of California, Irvine.

Jha, P., and A. Rodriguez-Lopez (2021): "Monopsonistic Labor Markets and International Trade," *European Economic Review*, 140, 103939.

Luca, D. L., and M. Luca (2019): "Survival of the Fittest: The Impact of the Minimum Wage on Firm Exit," Working Paper 25806, National Bureau of Economic Research.

Manning, A. (2021): "Monopsony in Labor Markets: A Review," *ILR Review*, 74(1), 3–26.

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

Neumark, D., and P. Shirley (2021): "Myth or Measurement: What Does the New Minimum Wage Research Say about Minimum Wages and Job Loss in the United States?," Working Paper 28388, National Bureau of Economic Research.

Rohlin, S. M. (2011): "State Minimum Wages and Business Location: Evidence from a Refined Border Approach," *Journal of Urban Economics*, 69(1), 103–117.

Thisse, J.-F., and E. Toulemonde (2010): "The Distribution of Earnings Under Monopsonistic/Polistic Competition," Discussion Paper 5136, IZA.

Vaghul, K., and B. Zipperer (2016): "Historical State and Sub-State Minimum Wage Data," *Washington Center for Equitable Growth*.

Minimum Wage and Employer Variety

Appendix — For Online Publication

Priyaranjan Jha and Antonio Rodriguez-Lopez
University of California, Irvine and CESifo

December 2022

A Supporting Tables

Table A-1: Pair-approach estimation of minimum wage responses of U.S. establishment counts with additional control, 1990–2016

	Sixteen industries (sorted by 1990 earnings per worker)					
Industry →	**Overall**	**1**	**2**	**3**	**4**	**5**
ln(minimum wage)	-0.091*	-0.183	-0.166	-0.097**	-0.044	0.063
	(0.048)	(0.112)	(0.111)	(0.040)	(0.046)	(0.148)
ln(population)	0.866***	0.789***	0.336	0.242**	0.279**	0.599*
	(0.083)	(0.155)	(0.249)	(0.110)	(0.122)	(0.340)
ln(establishments$^-$)		0.322***	0.314**	0.442***	0.504***	0.519***
		(0.109)	(0.150)	(0.072)	(0.076)	(0.174)
Observations	8,134	8,134	7,798	8,134	8,132	7,920
Industry →	**6**	**7**	**8**	**9**	**10**	**11**
ln(minimum wage)	0.080	-0.128	-0.192	-0.021	-0.049	0.028
	(0.139)	(0.093)	(0.141)	(0.093)	(0.095)	(0.078)
ln(population)	0.480*	0.676***	0.778	0.276	0.301	0.218
	(0.248)	(0.149)	(0.624)	(0.206)	(0.277)	(0.138)
ln(establishments$^-$)	0.847***	0.353**	0.528**	1.026***	0.647***	0.743***
	(0.183)	(0.139)	(0.250)	(0.135)	(0.231)	(0.121)
Observations	7,962	8,128	7,592	8,090	8,128	8,128
Industry →	**12**	**13**	**14**	**15**	**16**	
ln(minimum wage)	-0.001	0.075	-0.090	0.151	-1.497***	
	(0.084)	(0.104)	(0.102)	(0.178)	(0.415)	
ln(population)	0.316**	0.176	-0.269	0.309	0.534	
	(0.137)	(0.227)	(0.211)	(0.552)	(0.859)	
ln(establishments$^-$)	0.501***	0.513***	0.705***	0.373	1.092**	
	(0.142)	(0.152)	(0.147)	(0.280)	(0.421)	
Observations	8,094	8,112	7,808	8,014	7,022	

Notes: This table reports $\hat{\beta}$ and $\hat{\gamma}$ from the estimation of equation (1)—after also adding establishment counts in all other industries as control—for 16 industries and overall using yearly data from 1990 to 2016. All regressions include commuting zone–state fixed effects and pair–year fixed effects. Industries are sorted according to the 1990 earnings ranking of Table 1. Standard errors (in parentheses) are two-way clustered at the state and border segment levels. The coefficients are statistically significant at the *10%, **5%, or ***1% level.

Table A-2: Long-term minimum wage responses of U.S. establishment counts, 1990–2016

Industry ↓	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_{-1}$	$\hat{\beta}_{-2}$	$\hat{\beta}_{-3}$	$\hat{\beta}_{-4}$
Overall	-0.045	-0.043	-0.056	-0.090	-0.121*	-0.153**	-0.224**
	(0.030)	(0.036)	(0.049)	(0.063)	(0.071)	(0.075)	(0.095)
1	-0.186**	-0.160	-0.215*	-0.321*	-0.403**	-0.502**	-0.662**
	(0.082)	(0.096)	(0.120)	(0.163)	(0.191)	(0.219)	(0.272)
2	-0.113	-0.028	-0.004	-0.116	-0.113	-0.087	-0.327
	(0.126)	(0.126)	(0.178)	(0.187)	(0.214)	(0.243)	(0.242)
3	-0.075*	-0.088**	-0.140**	-0.154**	-0.215***	-0.250***	-0.298***
	(0.039)	(0.044)	(0.058)	(0.075)	(0.078)	(0.078)	(0.092)
4	-0.029	-0.039	-0.051	-0.082	-0.063	-0.070	-0.120
	(0.043)	(0.052)	(0.050)	(0.058)	(0.070)	(0.076)	(0.107)
5	0.189	0.016	-0.177	-0.154	-0.270	-0.257	-0.420
	(0.114)	(0.139)	(0.191)	(0.210)	(0.238)	(0.286)	(0.364)
6	-0.003	-0.085	-0.099	-0.147	-0.024	-0.001	-0.237
	(0.115)	(0.129)	(0.150)	(0.182)	(0.204)	(0.239)	(0.328)
7	-0.006	-0.028	-0.061	-0.109	-0.166	-0.242	-0.402*
	(0.063)	(0.068)	(0.086)	(0.125)	(0.146)	(0.168)	(0.232)
8	-0.320**	-0.434***	-0.352**	-0.426**	-0.557**	-0.730**	-0.614**
	(0.145)	(0.147)	(0.163)	(0.196)	(0.263)	(0.286)	(0.280)
9	0.006	0.058	0.010	-0.078	-0.103	-0.139	-0.178
	(0.072)	(0.094)	(0.112)	(0.139)	(0.165)	(0.203)	(0.256)
10	-0.041	-0.022	-0.048	-0.105	-0.129	-0.191	-0.218
	(0.079)	(0.090)	(0.119)	(0.132)	(0.154)	(0.160)	(0.215)
11	0.003	0.007	0.034	0.030	-0.006	-0.014	-0.133
	(0.053)	(0.054)	(0.064)	(0.086)	(0.096)	(0.103)	(0.125)
12	-0.078	-0.175**	-0.162**	-0.134	-0.130	-0.123	-0.151
	(0.065)	(0.071)	(0.080)	(0.098)	(0.110)	(0.127)	(0.148)
13	0.020	0.074	0.067	0.123	0.143	0.119	0.123
	(0.072)	(0.093)	(0.098)	(0.108)	(0.123)	(0.134)	(0.173)
14	-0.062	-0.126	-0.193*	-0.165	-0.186	-0.219	-0.324
	(0.081)	(0.100)	(0.107)	(0.133)	(0.160)	(0.186)	(0.217)
15	0.007	0.206	0.204	0.178	0.230	0.258	-0.005
	(0.125)	(0.136)	(0.151)	(0.201)	(0.195)	(0.229)	(0.214)
16	-1.145***	-0.598	-0.503	-1.090***	-1.290***	-1.839***	-2.593***
	(0.379)	(0.422)	(0.418)	(0.372)	(0.441)	(0.532)	(0.631)

Notes: This table reports $\hat{\beta}_k$, for $k \in \{2, 1, 0, -1, -2, -3, -4\}$ from the estimation of equation (2) for 16 industries and overall using yearly data from 1990 to 2016. All regressions include the log of working-age population as control, as well as commuting zone–state fixed effects and pair–year fixed effects. Industries are sorted according to the 1990 earnings ranking of Table 1. Standard errors (in parentheses) are two-way clustered at the state and border segment levels. The coefficients are statistically significant at the *10%, **5%, or ***1% level.

Table A-3: Pair-approach estimation of minimum wage responses of U.S. employment, 1990–2016

	Sixteen industries (sorted by 1990 earnings per worker)					
Industry →	**Overall**	**1**	**2**	**3**	**4**	**5**
ln(minimum wage)	-0.204**	-0.273**	0.140	-0.119*	-0.048	-0.310
	(0.084)	(0.119)	(0.349)	(0.067)	(0.085)	(0.340)
ln(population)	0.999***	1.093***	1.054	0.753***	0.707***	0.720*
	(0.098)	(0.179)	(0.705)	(0.102)	(0.171)	(0.376)
Observations	8,134	8,134	7,798	8,134	8,132	7,920
Industry →	**6**	**7**	**8**	**9**	**10**	**11**
ln(minimum wage)	-0.024	-0.114	-0.036	-0.221	-0.215	-0.132
	(0.315)	(0.109)	(0.369)	(0.200)	(0.151)	(0.111)
ln(population)	0.801*	0.899***	2.201**	1.574***	0.959***	0.894***
	(0.399)	(0.233)	(0.824)	(0.425)	(0.295)	(0.220)
Observations	7,962	8,128	7,592	8,090	8,128	8,128
Industry →	**12**	**13**	**14**	**15**	**16**	
ln(minimum wage)	-0.057	-0.470*	-0.329	0.104	-2.491***	
	(0.144)	(0.267)	(0.319)	(0.238)	(0.587)	
ln(population)	1.422***	0.411	-0.167	0.999*	2.139**	
	(0.201)	(0.294)	(0.447)	(0.511)	(0.922)	
Observations	8,094	8,112	7,808	8,014	7,022	

Notes: This table reports $\hat{\beta}$ and $\hat{\gamma}$ from the estimation of equation (1) for 16 industries and overall using 1990-2016 yearly data, but using log employment instead of log establishment counts as the dependent variable. All regressions include commuting zone–state fixed effects and pair–year fixed effects. Industries are sorted according to the 1990 earnings ranking of Table 1. Standard errors (in parentheses) are two-way clustered at the state and border segment levels. The coefficients are statistically significant at the *10%, **5%, or ***1% level.

Table A-4: Pair-approach estimation of minimum wage responses of U.S. earnings per worker, 1990–2016

	Sixteen industries (sorted by 1990 earnings per worker)					
Industry →	**Overall**	**1**	**2**	**3**	**4**	**5**
ln(minimum wage)	0.017	0.164***	0.144	0.039	-0.012	0.063
	(0.055)	(0.055)	(0.153)	(0.043)	(0.089)	(0.170)
ln(population)	0.093	0.096	0.100	-0.169	0.095	-0.161
	(0.083)	(0.087)	(0.195)	(0.111)	(0.131)	(0.250)
Observations	8,134	8,134	7,798	8,134	8,132	7,920
Industry →	**6**	**7**	**8**	**9**	**10**	**11**
ln(minimum wage)	0.094	0.126*	-0.118	0.070	-0.037	-0.088
	(0.107)	(0.069)	(0.107)	(0.093)	(0.077)	(0.058)
ln(population)	-0.222*	0.107	-0.006	0.040	0.144	0.099
	(0.127)	(0.122)	(0.119)	(0.163)	(0.101)	(0.075)
Observations	7,962	8,128	7,592	8,090	8,128	8,128
Industry →	**12**	**13**	**14**	**15**	**16**	
ln(minimum wage)	0.043	0.094	-0.127	0.010	0.086	
	(0.089)	(0.077)	(0.098)	(0.126)	(0.166)	
ln(population)	-0.022	0.113	0.055	0.035	-0.434	
	(0.194)	(0.106)	(0.137)	(0.144)	(0.337)	
Observations	8,094	8,112	7,808	8,014	7,022	

Notes: This table reports $\hat{\beta}$ and $\hat{\gamma}$ from the estimation of equation (1) for 16 industries and overall using 1990-2016 yearly data, but using log earnings per worker instead of log establishment counts as the dependent variable. All regressions include commuting zone–state fixed effects and pair–year fixed effects. Industries are sorted according to the 1990 earnings ranking of Table 1. Standard errors (in parentheses) are two-way clustered at the state and border segment levels. The coefficients are statistically significant at the *10%, **5%, or ***1% level.

Table A-5: Long-term minimum wage responses of U.S. employment, 1990–2016

Industry ↓	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_{-1}$	$\hat{\beta}_{-2}$	$\hat{\beta}_{-3}$	$\hat{\beta}_{-4}$
Overall	-0.080*	-0.118*	-0.167**	-0.271**	-0.316**	-0.400***	-0.418**
	(0.045)	(0.060)	(0.079)	(0.107)	(0.131)	(0.148)	(0.188)
1	-0.093	-0.156	-0.185*	-0.354***	-0.547***	-0.577***	-0.719***
	(0.084)	(0.107)	(0.098)	(0.125)	(0.156)	(0.162)	(0.188)
2	-0.146	-0.351	-0.019	0.083	0.172	0.152	0.340
	(0.207)	(0.269)	(0.329)	(0.436)	(0.497)	(0.584)	(0.664)
3	-0.079	-0.075	-0.069	-0.117	-0.126	-0.222**	-0.200
	(0.062)	(0.063)	(0.071)	(0.092)	(0.100)	(0.108)	(0.137)
4	0.011	0.011	0.034	0.017	-0.002	0.031	-0.108
	(0.075)	(0.084)	(0.098)	(0.107)	(0.105)	(0.126)	(0.135)
5	0.045	-0.154	-0.299	-0.122	-0.444	-0.395	-0.924
	(0.329)	(0.394)	(0.495)	(0.606)	(0.668)	(0.645)	(0.889)
6	-0.340*	-0.096	-0.176	0.023	0.129	0.075	0.486
	(0.174)	(0.191)	(0.209)	(0.264)	(0.298)	(0.377)	(0.518)
7	-0.123	-0.097	-0.245**	-0.164	-0.213	-0.242	-0.308
	(0.079)	(0.083)	(0.118)	(0.125)	(0.138)	(0.150)	(0.205)
8	-0.261	-0.007	0.014	0.022	-0.023	0.378	0.360
	(0.344)	(0.348)	(0.329)	(0.415)	(0.499)	(0.509)	(0.647)
9	0.054	0.049	0.180	-0.137	-0.141	-0.342	-0.193
	(0.157)	(0.182)	(0.177)	(0.245)	(0.271)	(0.311)	(0.434)
10	0.005	-0.128	-0.046	-0.365*	-0.125	-0.440*	-0.312
	(0.094)	(0.110)	(0.151)	(0.199)	(0.239)	(0.228)	(0.275)
11	-0.037	-0.095	-0.091	-0.079	-0.209	-0.215	-0.384*
	(0.094)	(0.120)	(0.131)	(0.156)	(0.178)	(0.182)	(0.210)
12	0.049	0.052	-0.090	-0.114	-0.050	-0.169	-0.276
	(0.113)	(0.137)	(0.157)	(0.205)	(0.193)	(0.197)	(0.208)
13	-0.433**	-0.474**	-0.570*	-0.630*	-0.659	-0.824*	-0.842
	(0.176)	(0.219)	(0.292)	(0.352)	(0.403)	(0.451)	(0.553)
14	0.147	-0.048	-0.290	-0.367	-0.575	-0.663	-0.626
	(0.162)	(0.172)	(0.232)	(0.346)	(0.452)	(0.555)	(0.797)
15	0.223	0.276	0.116	-0.039	-0.140	-0.245	-0.506
	(0.168)	(0.216)	(0.247)	(0.273)	(0.295)	(0.348)	(0.411)
16	-1.945***	-1.580*	-1.522*	-2.472***	-2.740***	-3.253***	-4.338***
	(0.663)	(0.849)	(0.900)	(0.861)	(0.968)	(1.029)	(1.067)

Notes: This table reports $\hat{\beta}_k$, for $k \in \{2, 1, 0, -1, -2, -3, -4\}$ from the estimation of equation (2) for 16 industries and overall using yearly data from 1990 to 2016, but using log employment instead of log establishment counts as the dependent variable. All regressions include the log of working-age population as control, as well as commuting zone–state fixed effects and pair–year fixed effects. Industries are sorted according to the 1990 earnings ranking of Table 1. Standard errors (in parentheses) are two-way clustered at the state and border segment levels. The coefficients are statistically significant at the *10%, **5%, or ***1% level.

Table A-6: Long-term minimum wage responses of U.S. earnings per worker, 1990–2016

Industry ↓	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_{-1}$	$\hat{\beta}_{-2}$	$\hat{\beta}_{-3}$	$\hat{\beta}_{-4}$
Overall	-0.001	-0.027	-0.025	0.011	0.012	0.045	0.004
	(0.043)	(0.052)	(0.062)	(0.069)	(0.071)	(0.078)	(0.090)
1	0.165*	0.148	0.287**	0.345***	0.306***	0.257***	0.169*
	(0.096)	(0.110)	(0.112)	(0.121)	(0.086)	(0.094)	(0.086)
2	-0.017	0.238	0.367*	0.323	0.199	0.299	0.293
	(0.109)	(0.147)	(0.202)	(0.258)	(0.288)	(0.339)	(0.346)
3	-0.042	-0.067*	0.011	0.037	0.030	0.070	0.076
	(0.041)	(0.038)	(0.050)	(0.061)	(0.071)	(0.079)	(0.088)
4	-0.083	-0.182*	-0.050	-0.152	-0.060	-0.130	-0.107
	(0.074)	(0.095)	(0.107)	(0.122)	(0.120)	(0.136)	(0.173)
5	0.161	0.135	0.239	0.332	0.365	0.187	0.007
	(0.163)	(0.203)	(0.299)	(0.329)	(0.327)	(0.321)	(0.356)
6	-0.014	0.013	0.218	0.114	-0.004	0.201	-0.024
	(0.113)	(0.125)	(0.145)	(0.149)	(0.149)	(0.171)	(0.195)
7	0.045	-0.011	0.079	0.071	0.132	0.192*	0.209
	(0.065)	(0.062)	(0.075)	(0.080)	(0.089)	(0.105)	(0.129)
8	-0.002	-0.133	-0.225**	-0.229**	-0.243*	-0.297**	-0.325*
	(0.079)	(0.080)	(0.091)	(0.111)	(0.133)	(0.134)	(0.165)
9	-0.083	0.040	-0.019	0.171	0.063	0.346*	0.235
	(0.100)	(0.117)	(0.127)	(0.133)	(0.162)	(0.182)	(0.229)
10	0.036	-0.004	-0.076	-0.041	-0.142	-0.010	-0.124
	(0.082)	(0.088)	(0.116)	(0.118)	(0.120)	(0.130)	(0.158)
11	-0.081	-0.064	-0.115	-0.133	-0.099	-0.202**	-0.255**
	(0.062)	(0.064)	(0.073)	(0.093)	(0.092)	(0.095)	(0.111)
12	0.045	0.021	-0.043	0.091	-0.116	-0.018	0.039
	(0.075)	(0.078)	(0.091)	(0.103)	(0.125)	(0.111)	(0.137)
13	-0.059	-0.001	0.017	0.064	0.075	0.194	0.071
	(0.100)	(0.105)	(0.116)	(0.144)	(0.151)	(0.155)	(0.188)
14	-0.068	-0.163*	-0.134	-0.202	-0.195	-0.223	-0.246
	(0.066)	(0.082)	(0.104)	(0.143)	(0.182)	(0.190)	(0.234)
15	-0.187**	-0.084	-0.065	0.029	-0.110	0.033	-0.031
	(0.082)	(0.121)	(0.156)	(0.165)	(0.177)	(0.198)	(0.244)
16	-0.312	-0.077	-0.278	0.016	0.079	0.332	0.359
	(0.237)	(0.217)	(0.241)	(0.306)	(0.405)	(0.349)	(0.484)

Notes: This table reports $\hat{\beta}_k$, for $k \in \{2, 1, 0, -1, -2, -3, -4\}$ from the estimation of equation (2) for 16 industries and overall using yearly data from 1990 to 2016, but using log earnings per worker instead of log establishment counts as the dependent variable. All regressions include the log of working-age population as control, as well as commuting zone–state fixed effects and pair–year fixed effects. Industries are sorted according to the 1990 earnings ranking of Table 1. Standard errors (in parentheses) are two-way clustered at the state and border segment levels. The coefficients are statistically significant at the *10%, **5%, or ***1% level.

B Theoretical Appendix

B.1 Household Maximization Problem

Given the utility function in (3), the representative household maximizes its utility by choosing N and allocating labor, $l(\omega)$, across firms. With the final good being the numéraire, consumption of the representative household, C, equals the wage income. Therefore, we can write the household's problem as

$$\max_{l(\omega)} \int_{\omega\in\Omega} w(\omega)l(\omega) - \frac{N^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} \quad \text{subject to} \quad N = \left(\int_{\omega\in\Omega} l(\omega)^{\frac{1+\theta}{\theta}} d\omega\right)^{\frac{\theta}{1+\theta}}.$$

The solution to this problem yields that the labor supply to firm ω is

$$l(\omega) = N^{\frac{\psi-\theta}{\psi}} w(\omega)^{\theta}. \tag{B-1}$$

Using the definition of the wage index, $W \equiv \left(\int w(\omega)^{1+\theta}\right)^{\frac{1}{1+\theta}}$, it follows that

$$N^{\frac{1}{\theta}} \equiv \left(\int_{\omega\in\Omega} l(\omega)^{\frac{1+\theta}{\theta}}\right)^{\frac{1}{1+\theta}} = N^{\frac{\psi-\theta}{\psi\theta}} \left(\int_{\omega\in\Omega} w(\omega)^{1+\theta}\right)^{\frac{1}{1+\theta}} = N^{\frac{\psi-\theta}{\psi\theta}} W,$$

and hence $N = W^{\psi}$, so that we can rewrite the firm-level labor supply in (B-1) as

$$l(\omega) = \frac{w(\omega)^{\theta}}{W^{\theta-\psi}}. \tag{B-2}$$

Using (B-2), the aggregate labor supply is $L = \int_{\omega\in\Omega} l(\omega)d\omega = \frac{1}{W^{\theta-\psi}} \int_{\omega\in\Omega} w(\omega)^{\theta} d\omega$. Lastly, the aggregate wage bill is given by

$$\int_{\omega\in\Omega} w(\omega)l(\omega)d\omega = N^{\frac{\psi-\theta}{\psi}} \int_{\omega\in\Omega} w(\omega)^{1+\theta} d\omega = N^{\frac{\psi-\theta}{\psi}} W^{1+\theta} = N^{1-\frac{\theta}{\psi}} W^{1+\theta} = NW = W^{1+\psi}, \tag{B-3}$$

where first equality follows from (B-1), the second equality follows from the definition of W, and the third and fourth equalities follow from $N = W^{\psi}$.

B.2 Decentralized Equilibrium

In the free-entry condition in (10), the left-hand side is strictly decreasing with $\hat{\varphi}$, approaching zero as $\hat{\varphi} \to \infty$. Thus, as long as f_E is sufficiently small, there is unique value for $\hat{\varphi}$ that solves (10). From the zero-cutoff-profit condition, the solution for the wage index is then

$$W_D = \left\{ \left[\frac{\theta^{\theta}}{(1+\theta)^{1+\theta}}\right] \frac{\hat{\varphi}_D^{1+\theta}}{f} \right\}^{\frac{1}{\theta-\psi}}, \tag{B-4}$$

and from (8) and (B-4) we know that

$$l_D(\varphi) = \left(\frac{\theta}{1+\theta}\right)^{\theta} \frac{\varphi^{\theta}}{W_D^{\theta-\psi}} = \frac{(1+\theta) f \varphi^{\theta}}{\hat{\varphi}_D^{1+\theta}}. \tag{B-5}$$

To obtain the expression for the equilibrium mass of firms, M_D, we use equations (11) and (8) to obtain

$$M_D = \left[\int_{\hat{\varphi}_D}^{\infty} \varphi^{1+\theta} g(\varphi|\varphi \geq \hat{\varphi}_D) d\varphi\right]^{-1} \left(\frac{1+\theta}{\theta}\right)^{1+\theta} W_D^{1+\theta} \tag{B-6}$$

From equation (12), the equilibrium total employment is

$$L_D = \frac{(1+\theta) f M_D}{\hat{\varphi}_D^{1+\theta}} \int_{\hat{\varphi}_D}^{\infty} \varphi^{\theta} g(\varphi|\varphi \geq \hat{\varphi}) d\varphi. \tag{B-7}$$

Finally, from (7) we know that $\mathbb{U}_D = \left(\frac{1}{1+\psi}\right) W_D^{1+\psi} = \left(\frac{1}{1+\psi}\right) N_D^{1+\frac{1}{\psi}}$, where the second equality follows from $N_D = W_D^{\psi}$.

B.3 Social Planner's Problem

The social planner chooses the cutoff productivity level ($\hat{\varphi}$), the mass of entrants (M_E), and firm-level labor supplies ($l(\varphi)$ for every $\varphi \geq \hat{\varphi}$) that maximize the household utility function in (3) subject to equation (4) and $C = M \int_{\hat{\varphi}}^{\infty} \varphi l(\varphi) g(\varphi|\varphi \geq \hat{\varphi}) d\varphi - M_E f_E - M f$, for $M = [1 - G(\hat{\varphi})] M_E$. Notice that household consumption equals total output minus entry costs and fixed costs of operation (recall that f and f_E are in terms of the final good). Using $g(\varphi|\varphi \geq \hat{\varphi}) = \frac{g(\varphi)}{1-G(\hat{\varphi})}$, we can write the planner's problem as

$$\max_{l(\varphi), M_E, \hat{\varphi}} \left\{ M_E \int_{\hat{\varphi}}^{\infty} \varphi l(\varphi) g(\varphi) d\varphi - M_E f_E - [1 - G(\hat{\varphi})] M_E f - \frac{\left[\overbrace{\left[\left(M_E \int_{\hat{\varphi}}^{\infty} l(\varphi)^{\frac{1+\theta}{\theta}} g(\varphi) d\varphi\right)^{\frac{\theta}{1+\theta}}\right]}^{N}\right]^{1+\frac{1}{\psi}}}{1 + 1/\psi} \right\}. \tag{B-8}$$

The first order conditions with respect to $l(\varphi)$, M_E, and $\hat{\varphi}$ are respectively

$$\varphi - N^{\frac{\theta-\psi}{\theta\psi}} l(\varphi)^{\frac{1}{\theta}} = 0, \tag{B-9}$$

$$\int_{\hat{\varphi}}^{\infty} \varphi l(\varphi) g(\varphi) d\varphi - \left(\frac{\theta}{1+\theta}\right) N^{\frac{\theta-\psi}{\theta\psi}} \int_{\hat{\varphi}}^{\infty} l(\varphi)^{\frac{1+\theta}{\theta}} g(\varphi) d\varphi - f_E - [1 - G(\hat{\varphi})] f = 0, \tag{B-10}$$

$$\hat{\varphi} l(\hat{\varphi}) - \left(\frac{\theta}{1+\theta}\right) N^{\frac{\theta-\psi}{\theta\psi}} l(\hat{\varphi})^{\frac{1+\theta}{\theta}} - f = 0. \tag{B-11}$$

Note that (B-9) implies that $N^{\frac{\theta-\psi}{\theta\psi}} = \frac{\hat{\varphi}}{l(\hat{\varphi})^{\frac{1}{\theta}}}$, which plugged into (B-11) yields $l(\hat{\varphi}) = \frac{(1+\theta)f}{\hat{\varphi}}$, so that we can solve for N as a function of $\hat{\varphi}$ as

$$N = \left[\frac{\hat{\varphi}^{1+\theta}}{(1+\theta) f}\right]^{\frac{\psi}{\theta-\psi}}. \tag{B-12}$$

Note also from (B-9) that $\frac{l(\varphi)}{l(\hat{\varphi})} = \left(\frac{\varphi}{\hat{\varphi}}\right)^{\theta}$, which from (8) we know that it is the same allocation of resources across firms as in the decentralized case. This proves that the allocation of resources across firms in the decentralized case is efficient. Plugging in (B-12) into (B-9), we can write the planner's optimal firm-level labor supply as

$$l(\varphi) = \frac{(1+\theta) f \varphi^{\theta}}{\hat{\varphi}^{1+\theta}}. \tag{B-13}$$

Using (B-13), we obtain the following expressions

$$\int_{\hat{\varphi}}^{\infty} \varphi l\left(\varphi\right) g(\varphi) d\varphi = (1+\theta) f \int_{\hat{\varphi}}^{\infty} \left(\frac{\varphi}{\hat{\varphi}}\right)^{1+\theta} g(\varphi) d\varphi, \tag{B-14}$$

$$\int_{\hat{\varphi}}^{\infty} l\left(\varphi\right)^{\frac{1+\theta}{\theta}} g(\varphi) d\varphi = \left[\frac{(1+\theta) f}{\hat{\varphi}}\right]^{\frac{1+\theta}{\theta}} \int_{\hat{\varphi}}^{\infty} \left(\frac{\varphi}{\hat{\varphi}}\right)^{1+\theta} g(\varphi) d\varphi, \tag{B-15}$$

which along with (B-12) can be plugged into (B-10) to obtain

$$\int_{\hat{\varphi}}^{\infty} \left[\left(\frac{\varphi}{\hat{\varphi}}\right)^{1+\theta} - 1\right] f g(\varphi) d\varphi = f_E. \tag{B-16}$$

Notice that (B-16) is exactly the same as (10), which is the equation determining $\hat{\varphi}_D$ in the decentralized case. Using $\hat{\varphi}_P$ to denote the equilibrium cutoff productivity level in the planner's problem, it follows that $\hat{\varphi}_P = \hat{\varphi}_D$. Once we obtain $\hat{\varphi}_P$, we plug it into (B-12) and (B-13) to obtain N_P and $l_P(\varphi)$. From (B-5), it follows that $l_P(\varphi) = l_D(\varphi)$, so that firm size is the same in both cases. Moreover, using $N_D = W_D^{\psi}$, (B-4), (B-12), and $\hat{\varphi}_P = \hat{\varphi}_D$ we get that

$$\frac{N_P}{N_D} = \left(\frac{1+\theta}{\theta}\right)^{\frac{\theta\psi}{\theta-\psi}} > 1.$$

To obtain the mass of firms, we use the definition of N as written in the planner's maximization problem in (B-8), along with (B-13) and $M = [1 - G(\hat{\varphi})] M_E$ to get

$$M_P = \left\{\left[\frac{(1+\theta) f}{\hat{\varphi}_P^{1+\theta}}\right]^{\frac{1+\theta}{\theta}} \int_{\hat{\varphi}_P}^{\infty} \varphi^{1+\theta} g(\varphi|\varphi \geq \hat{\varphi}_P) d\varphi\right\}^{-1} N_P^{\frac{1+\theta}{\theta}} = \left[\int_{\hat{\varphi}_P}^{\infty} \varphi^{1+\theta} g(\varphi|\varphi \geq \hat{\varphi}_P) d\varphi\right]^{-1} N_P^{\frac{1+\theta}{\psi}}, \tag{B-17}$$

where the second equality follows from (B-12). From (B-17) and (B-6), note that the ratio between M_P and M_D is $\frac{M_P}{M_D} = \left[\frac{\theta N_P^{1/\psi}}{(1+\theta) W_D}\right]^{1+\theta}$, which using (B-4), (B-12), and $\hat{\varphi}_P = \hat{\varphi}_D$, can be rewritten as

$$\frac{M_P}{M_D} = \left(\frac{1+\theta}{\theta}\right)^{\frac{(1+\theta)\psi}{\theta-\psi}} > 1. \tag{B-18}$$

Thus, the decentralized outcome yields a suboptimal mass of firms. Total employment in the planner's case is $L_P = M_P \int_{\hat{\varphi}_P}^{\infty} l_P(\varphi) g(\varphi|\varphi \geq \hat{\varphi}_P) d\varphi$, which using (B-13) can be rewritten as

$$L_P = \frac{(1+\theta) f M_P}{\hat{\varphi}_P^{1+\theta}} \int_{\hat{\varphi}_D}^{\infty} \varphi^{\theta} g(\varphi|\varphi \geq \hat{\varphi}_D) d\varphi. \tag{B-19}$$

From (B-19), (B-7), and $\hat{\varphi}_P = \hat{\varphi}_D$, it follows that

$$\frac{L_P}{L_D} = \frac{M_P}{M_D} = \left(\frac{1+\theta}{\theta}\right)^{\frac{(1+\theta)\psi}{\theta-\psi}} > 1. \tag{B-20}$$

Regarding welfare, note first from (B-16) that $M_E f_E + [1-G(\hat{\varphi})]M_E f = M_E \int_{\hat{\varphi}}^{\infty} \left(\frac{\varphi}{\hat{\varphi}}\right)^{1+\theta} f g(\varphi) d\varphi$, which along with (B-14) and $M_E = \frac{M}{1-G(\hat{\varphi})}$ can be plugged into the maximized value of welfare in (B-8) to get

$$\mathbb{U}_P = \frac{M_P \theta f}{\hat{\varphi}_P^{1+\theta}} \int_{\hat{\varphi}_P}^{\infty} \varphi^{1+\theta} g(\varphi|\varphi \geq \hat{\varphi}_P) d\varphi - \frac{N_P^{1+\frac{1}{\psi}}}{1+1/\psi} = \left(\frac{\theta f}{\hat{\varphi}_P^{1+\theta}}\right) N_P^{\frac{1+\theta}{\psi}} - \frac{N_P^{1+\frac{1}{\psi}}}{1+1/\psi}, \tag{B-21}$$

where the second equality uses (B-17). From (B-12), we obtain that $\frac{\theta f}{\hat{\varphi}_P^{1+\theta}} = \frac{\theta}{(1+\theta)N_P^{(\theta-\psi)/\psi}}$, which plugged into (B-21) yields that welfare in the planner's case is

$$\mathbb{U}_P = \frac{(\theta-\psi)N_P^{1+\frac{1}{\psi}}}{(1+\theta)(1+\psi)}. \tag{B-22}$$

Using the expression for welfare in the decentralized case given in the end of section B.2, we get that $\frac{\mathbb{U}_P}{\mathbb{U}_D} = \frac{\theta-\psi}{1+\theta}\left(\frac{N_P}{N_D}\right)^{1+1/\psi}$. Using (B-12), $N_D = W_D^{\psi}$, (B-4), and $\hat{\varphi}_P = \hat{\varphi}_D$, this ratio can be rewritten as

$$\frac{\mathbb{U}_P}{\mathbb{U}_D} = \left(\frac{\theta-\psi}{1+\theta}\right)\left(\frac{1+\theta}{\theta}\right)^{\frac{\theta(1+\psi)}{\theta-\psi}} > 1. \tag{B-23}$$

We know that $\frac{\mathbb{U}_P}{\mathbb{U}_D} > 1$ because $\frac{\mathbb{U}_P}{\mathbb{U}_D}$ is strictly decreasing in θ,

$$\frac{d\left(\mathbb{U}_P/\mathbb{U}_D\right)}{d\theta} = -\frac{\psi(1+\psi)}{(\theta-\psi)(1+\theta)}\left(\frac{1+\theta}{\theta}\right)^{\frac{\theta(1+\psi)}{\theta-\psi}} \ln\left(\frac{1+\theta}{\theta}\right) < 0,$$

and $\lim_{\theta\to\infty} \frac{\mathbb{U}_P}{\mathbb{U}_D} = 1$.

B.4 Optimality of Labor Subsidy/Leisure Tax

Suppose the government subsidizes labor at rate s and finances it using a lump sum tax on workers, T. The utility maximization exercise is then

$$\max_{l(\omega)} \int_{\omega\in\Omega} (1+s)w(\omega)l(\omega) - T - \frac{N^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}} \quad \text{subject to} \quad N = \left(\int_{\omega\in\Omega} l(\omega)^{\frac{1+\theta}{\theta}} d\omega\right)^{\frac{\theta}{1+\theta}},$$

which yields the following labor supply curve to firm ω:

$$l_S(\omega) = N_S^{\frac{\psi-\theta}{\psi}} (1+s)^{\theta} w(\omega)^{\theta}. \tag{B-24}$$

We use subscript S to indicate that this is the model with a wage subsidy.

Using the definition of the wage index, W, it follows that

$$N_S^{\frac{1}{\theta}} \equiv \left(\int_{\omega\in\Omega} l(\omega)^{\frac{1+\theta}{\theta}}\right)^{\frac{1}{1+\theta}} = N_S^{\frac{\psi-\theta}{\psi\theta}}(1+s)\left(\int_{\omega\in\Omega} w(\omega)^{1+\theta}\right)^{\frac{1}{1+\theta}} = N_S^{\frac{\psi-\theta}{\psi\theta}}(1+s)W_S,$$

and hence $N_S = (1+s)^{\psi}W_S^{\psi}$, so that we can rewrite the firm-level labor supply in (B-24) as

$$l_S(\omega) = (1+s)^{\psi}W_S^{\psi-\theta}w(\omega)^{\theta} = \frac{(1+s)^{\psi}w(\omega)^{\theta}}{W_S^{\theta-\psi}}. \tag{B-25}$$

We can also verify that the total wage bill is still given by $W_S N_S$:

$$\int_{\omega\in\Omega} w(\omega)l(\omega)d\omega = N_S^{\frac{\psi-\theta}{\psi}}(1+s)^{\theta}\int_{\omega\in\Omega} w(\omega)^{1+\theta}d\omega = (1+s)^{\theta}N_S^{1-\frac{\theta}{\psi}}W_S^{1+\theta} = W_S N_S.$$

From the firm optimization problem we obtain

$$w_S(\varphi) = \left(\frac{\theta}{1+\theta}\right)\varphi \quad \text{and} \quad l_S(\varphi) = \left(\frac{\theta}{1+\theta}\right)^{\theta}\frac{\varphi^{\theta}}{W_S^{\theta-\psi}}(1+s)^{\psi}, \tag{B-26}$$

and hence the gross profit is

$$\pi_S(\varphi) = \left(\frac{\theta^{\theta}}{(1+\theta)^{1+\theta}}\right)\frac{(1+s)^{\psi}\varphi^{1+\theta}}{W_S^{\theta-\psi}}.$$

The zero cutoff profit condition is $\pi_S(\hat{\varphi}_S) = f$. Since the free entry condition given in (10) is unchanged, $\hat{\varphi}$ remains unchanged: $\hat{\varphi}_S = \hat{\varphi}_D = \hat{\varphi}_P$. Therefore, W_S is determined by the zero cutoff profit condition and is given by

$$\left(\frac{\theta^{\theta}}{(1+\theta)^{1+\theta}}\right)\frac{(1+s)^{\psi}\hat{\varphi}_S^{1+\theta}}{W_S^{\theta-\psi}} = f, \tag{B-27}$$

which implies that

$$W_S = \left\{\left(\frac{\theta^{\theta}}{(1+\theta)^{1+\theta}}\right)\frac{\hat{\varphi}^{1+\theta}(1+s)^{\psi}}{f}\right\}^{\frac{1}{\theta-\psi}}, \tag{B-28}$$

From $N_S = (1+s)^{\psi}W_S^{\psi}$, it follows that

$$N_S = (1+s)^{\frac{\theta\psi}{\theta-\psi}}\left\{\left(\frac{\theta^{\theta}}{(1+\theta)^{1+\theta}}\right)\frac{\hat{\varphi}^{1+\theta}}{f}\right\}^{\frac{\psi}{\theta-\psi}}. \tag{B-29}$$

Recall that in the planner's problem $N_P = \left[\frac{\hat{\varphi}^{1+\theta}}{(1+\theta)f}\right]^{\frac{\psi}{\theta-\psi}}$. Therefore, $N_S = N_P$ if and only if $s = \frac{1}{\theta}$. Thus, a labor subsidy or a leisure tax in the amount of $s = \frac{1}{\theta}$ yields the same N as in the planner's problem.

Next, note that using (B-28), the expression for firm-level employment given in (B-26) becomes

$$l_S(\varphi) = \frac{(1+\theta)f\varphi^{\theta}}{\hat{\varphi}_S^{1+\theta}}. \tag{B-30}$$

That is, this expression remains the same as in the decentralized equilibrium without policy intervention and in the planner's problem. From the definition of N_S,

$$N_S \equiv \left[M_S \int_{\hat{\varphi}_S}^{\infty} l(\varphi)^{\frac{1+\theta}{\theta}} g(\varphi | \varphi \geq \hat{\varphi}_S) d\varphi \right]^{\frac{\theta}{1+\theta}},$$

it follows that $M_S = M_P$ as given by (B-17).

We have seen that the total wage bill is still given by $W_S N_S$. Therefore, the consumption of households is $(1+s)W_S N_S - T = W_S N_S$ because of balanced budget. We have shown that $N_S = (1+s)^{\psi} W_S^{\psi}$, and hence $W_S N_S = \frac{N_S^{1+\frac{1}{\psi}}}{1+s} = \left(\frac{\theta}{1+\theta}\right) N_S^{1+\frac{1}{\psi}}$. This exactly equals the expression for consumption in the planner's problem because $N_S = N_P$. Hence, $\mathbb{U}_S = \mathbb{U}_P$. Thus, a proportional labor subsidy (or leisure tax) of $s = \frac{1}{\theta}$ restores optimality.

B.5 The Effects of a Binding Minimum Wage

This section presents the proof of Proposition 1. For a binding minimum wage $\underline{w}$, so that $\underline{w} > w(\hat{\varphi}_D)$, we show that for every productivity distribution, $\frac{d\hat{\varphi}}{d\underline{w}} > 0$, $\frac{dM}{d\underline{w}} < 0$, $\frac{dW}{d\underline{w}} < 0$, $\frac{d\mathbb{U}}{d\underline{w}} < 0$, and that if the productivity distribution is Pareto, then it also holds that $\frac{dL}{d\underline{w}} < 0$ and $\frac{d\overline{w}}{d\underline{w}} > 0$.

From the zero-cutoff-profit condition in section 3.3, we know that $\underline{W}^{\theta-\psi} = (\hat{\underline{\varphi}} - \underline{w})\frac{\underline{w}^{\theta}}{f}$, which allows us to rewrite the free-entry condition in (13) as

$$\int_{\hat{\underline{\varphi}}}^{\underline{\varphi}} \left(\frac{\varphi - \underline{w}}{\hat{\underline{\varphi}} - \underline{w}} - 1 \right) f g(\varphi) d\varphi + \int_{\underline{\varphi}}^{\infty} \left\{ \left[\frac{\theta^{\theta}}{(1+\theta)^{1+\theta}} \right] \frac{\varphi^{1+\theta}}{(\hat{\underline{\varphi}} - \underline{w})\underline{w}^{\theta}} - 1 \right\} f g(\varphi) d\varphi = f_E. \tag{B-31}$$

Taking the derivative of (B-31) with respect to $\underline{w}$, and given that $\frac{d\underline{\varphi}}{d\underline{w}} = \frac{1+\theta}{\theta}$, we obtain that

$$\frac{d\hat{\underline{\varphi}}}{d\underline{w}} = \left[\frac{\int_{\hat{\underline{\varphi}}}^{\underline{\varphi}} \left(\frac{\varphi - \hat{\underline{\varphi}}}{\underline{\varphi} - \hat{\underline{\varphi}}} \right) g(\varphi) d\varphi + \int_{\underline{\varphi}}^{\infty} \left(\frac{\varphi}{\underline{\varphi}} \right)^{1+\theta} g(\varphi) d\varphi}{\int_{\hat{\underline{\varphi}}}^{\underline{\varphi}} \left(\frac{\varphi - \underline{w}}{\underline{\varphi} - \underline{w}} \right) g(\varphi) d\varphi + \int_{\underline{\varphi}}^{\infty} \left(\frac{\varphi}{\underline{\varphi}} \right)^{1+\theta} g(\varphi) d\varphi} \right] \left(\frac{\underline{\varphi} - \hat{\underline{\varphi}}}{\underline{\varphi} - \underline{w}} \right) > 0. \tag{B-32}$$

All the terms in (B-32) are positive because $\underline{w} < \hat{\underline{\varphi}} < \underline{\varphi}$. We know that $\underline{w} < \hat{\underline{\varphi}}$ from the zero-cutoff-profit condition, $(\hat{\underline{\varphi}} - \underline{w})\frac{\underline{w}^{\theta}}{\underline{W}^{\theta-\psi}} = f$, and we know that $\hat{\underline{\varphi}} < \underline{\varphi}$ because firms with $\varphi \in [\hat{\underline{\varphi}}, \underline{\varphi})$ are constrained by the minimum wage. Moreover, the term in brackets is less than 1 because $\int_{\hat{\underline{\varphi}}}^{\underline{\varphi}} \left(\frac{\varphi - \hat{\underline{\varphi}}}{\underline{\varphi} - \hat{\underline{\varphi}}} \right) g(\varphi) d\varphi < \int_{\hat{\underline{\varphi}}}^{\underline{\varphi}} \left(\frac{\varphi - \underline{w}}{\underline{\varphi} - \underline{w}} \right) g(\varphi) d\varphi$, which follows from $\frac{(\hat{\underline{\varphi}} - \underline{w})(\underline{\varphi} - \varphi)}{(\underline{\varphi} - \underline{w})(\underline{\varphi} - \hat{\underline{\varphi}})} > 0$ for every $\varphi < \underline{\varphi}$. Therefore,

$$\frac{d\hat{\underline{\varphi}}}{d\underline{w}} \in \left(0, \frac{\underline{\varphi} - \hat{\underline{\varphi}}}{\underline{\varphi} - \underline{w}} \right). \tag{B-33}$$

From the zero-cutoff-profit condition we get that $\underline{W} = \left[\frac{(\hat{\underline{\varphi}} - \underline{w})\underline{w}^{\theta}}{f} \right]^{1/(\theta-\psi)}$, and thus

$$\frac{d\underline{W}}{d\underline{w}} = \frac{\underline{w}^{\theta}}{(\theta - \psi)\underline{W}^{\theta-\psi-1} f} \left[\frac{\theta(\hat{\underline{\varphi}} - \underline{w})}{\underline{w}} + \frac{d\hat{\underline{\varphi}}}{d\underline{w}} - 1 \right].$$

The sign of $\frac{d\underline{W}}{d\underline{w}}$ is determined by the term within the brackets, which is negative if and only if

$$\frac{d\hat{\underline{\varphi}}}{d\underline{w}} < 1 - \frac{\theta(\hat{\underline{\varphi}} - \underline{w})}{\underline{w}} = 1 - \frac{\hat{\underline{\varphi}} - \underline{w}}{\underline{\varphi} - \underline{w}} = \frac{\underline{\varphi} - \hat{\underline{\varphi}}}{\underline{\varphi} - \underline{w}}, \tag{B-34}$$

where we use that $\underline{\varphi} - \underline{w} = \frac{\underline{w}}{\theta}$. From (B-33) we know that (B-34) holds, and therefore, $\frac{d\underline{W}}{d\underline{w}} < 0$. Similar to (7), welfare with a binding minimum wage is given by $\underline{\mathbb{U}} = \frac{\underline{W}^{1+\psi}}{1+\psi}$, and thus, $\frac{d\underline{\mathbb{U}}}{d\underline{w}} = \underline{W}^{\psi}\frac{d\underline{W}}{d\underline{w}} < 0$.

The wage index is defined as $\underline{W} = \left[\underline{M}\int_{\hat{\underline{\varphi}}}^{\infty} w(\varphi)^{1+\theta} g(\varphi|\varphi \geq \hat{\underline{\varphi}})d\varphi\right]^{1/(1+\theta)}$ where $w(\varphi) = \underline{w}$ for $\varphi \in [\hat{\underline{\varphi}}, \underline{\varphi})$ and $w(\varphi) = \left(\frac{\theta}{1+\theta}\right)\varphi$ for $\varphi \geq \underline{\varphi}$. Thus, we can solve for the mass of firms as

$$\underline{M} = \frac{[1 - G(\hat{\underline{\varphi}})]\underline{W}^{1+\theta}}{\underline{w}^{1+\theta}[G(\underline{\varphi}) - G(\hat{\underline{\varphi}})] + \int_{\underline{\varphi}}^{\infty} w(\varphi)^{1+\theta} g(\varphi)d\varphi}.$$

Taking the derivative of $\ln \underline{M}$ with respect to $\underline{w}$ yields

$$\begin{aligned}\frac{d\ln \underline{M}}{d\underline{w}} =& (1+\theta)\frac{d\ln \underline{W}}{d\underline{w}} - \left[\frac{g(\hat{\underline{\varphi}})}{1 - G(\hat{\underline{\varphi}})}\right]\frac{d\hat{\underline{\varphi}}}{d\underline{w}} + \frac{\underline{w}^{1+\theta} g(\hat{\underline{\varphi}})\frac{d\hat{\underline{\varphi}}}{d\underline{w}} - (1+\theta)\underline{w}^{\theta}[G(\underline{\varphi}) - G(\hat{\underline{\varphi}})]}{\underline{w}^{1+\theta}[G(\underline{\varphi}) - G(\hat{\underline{\varphi}})] + \int_{\underline{\varphi}}^{\infty} w(\varphi)^{1+\theta} g(\varphi)d\varphi} \\ =& \underbrace{(1+\theta)\frac{d\ln \underline{W}}{d\underline{w}}}_{<0} - g(\hat{\underline{\varphi}})\frac{d\hat{\underline{\varphi}}}{d\underline{w}}\left[\frac{1}{1 - G(\hat{\underline{\varphi}})} - \frac{1}{G(\underline{\varphi}) - G(\hat{\underline{\varphi}}) + \int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{1+\theta} g(\varphi)d\varphi}\right] \\ & \underbrace{- \left(\frac{1+\theta}{\underline{w}}\right)\frac{G(\underline{\varphi}) - G(\hat{\underline{\varphi}})}{G(\underline{\varphi}) - G(\hat{\underline{\varphi}}) + \int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{1+\theta} g(\varphi)d\varphi}}_{<0}.\end{aligned} \tag{B-35}$$

Thus, a sufficient condition for $\frac{d\ln \underline{M}}{d\underline{w}} < 0$ is that

$$-g(\hat{\underline{\varphi}})\frac{d\hat{\underline{\varphi}}}{d\underline{w}}\left[\frac{1}{1 - G(\hat{\underline{\varphi}})} - \frac{1}{G(\underline{\varphi}) - G(\hat{\underline{\varphi}}) + \int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{1+\theta} g(\varphi)d\varphi}\right] < 0,$$

which is true if $\int_{\underline{\varphi}}^{\infty}\left[\frac{w(\varphi)}{\underline{w}}\right]^{1+\theta} g(\varphi)d\varphi > 1 - G(\underline{\varphi})$, which implies $\int_{\underline{\varphi}}^{\infty}\left[\frac{w(\varphi)}{\underline{w}}\right]^{1+\theta} g(\varphi|\varphi \geq \underline{\varphi})d\varphi > 1$. This condition holds because $\frac{w(\varphi)}{\underline{w}} \geq 1$ for $\varphi \geq \underline{\varphi}$ (with equality if and only if $\varphi = \underline{\varphi}$). Therefore, $\frac{d\underline{M}}{d\underline{w}} < 0$.

Total employment is defined as $\underline{L} = \underline{M}\int_{\hat{\underline{\varphi}}}^{\infty} l(\varphi)g(\varphi|\varphi \geq \hat{\underline{\varphi}})d\varphi$ where $l(\varphi) = \frac{\underline{w}^{\theta}}{\underline{W}^{\theta-\psi}}$ for $\varphi \in [\hat{\underline{\varphi}}, \underline{\varphi})$ and $l(\varphi) = \frac{w(\varphi)^{\theta}}{\underline{W}^{\theta-\psi}}$ for $\varphi \geq \underline{\varphi}$. Hence, we can rewrite $\underline{L}$ as

$$\underline{L} = \frac{\underline{M}\underline{w}^{\theta}}{[1 - G(\hat{\underline{\varphi}})]\underline{W}^{\theta-\psi}}\left\{G(\underline{\varphi}) - G(\hat{\underline{\varphi}}) + \int_{\underline{\varphi}}^{\infty}\left[\frac{w(\varphi)}{\underline{w}}\right]^{\theta} g(\varphi)d\varphi\right\}. \tag{B-36}$$

The derivative of $\ln \underline{L}$ with respect to $\underline{w}$ is then given by

$$\begin{aligned}
\frac{d\ln \underline{L}}{d\underline{w}} =& \frac{d\ln \underline{M}}{d\underline{w}} - (\theta-\psi)\frac{d\ln \underline{W}}{d\underline{w}} + g(\underline{\hat{\varphi}})\frac{d\underline{\hat{\varphi}}}{d\underline{w}}\left[\frac{1}{1-G(\underline{\hat{\varphi}})} - \frac{1}{G(\underline{\varphi})-G(\underline{\hat{\varphi}})+\int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{\theta} g(\varphi)d\varphi}\right] \\
& + \left(\frac{\theta}{\underline{w}}\right)\frac{G(\underline{\varphi})-G(\underline{\hat{\varphi}})}{G(\underline{\varphi})-G(\underline{\hat{\varphi}})+\int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{\theta} g(\varphi)d\varphi} \\
=& -g(\underline{\hat{\varphi}})\frac{d\underline{\hat{\varphi}}}{d\underline{w}}\left[\frac{1}{G(\underline{\varphi})-G(\underline{\hat{\varphi}})+\int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{\theta} g(\varphi)d\varphi} - \frac{1}{G(\underline{\varphi})-G(\underline{\hat{\varphi}})+\int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{1+\theta} g(\varphi)d\varphi}\right] \\
& +(1+\psi)\frac{d\ln \underline{W}}{d\underline{w}} + \frac{G(\underline{\varphi})-G(\underline{\hat{\varphi}})}{\underline{w}}\times \\
& \left[\frac{\theta}{G(\underline{\varphi})-G(\underline{\hat{\varphi}})+\int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{\theta} g(\varphi)d\varphi} - \frac{1+\theta}{G(\underline{\varphi})-G(\underline{\hat{\varphi}})+\int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{1+\theta} g(\varphi)d\varphi}\right] \\
=& -\frac{1}{D}\Bigg\{\underbrace{\left[g(\underline{\hat{\varphi}})\frac{d\underline{\hat{\varphi}}}{d\underline{w}} - \frac{\theta[G(\underline{\varphi})-G(\underline{\hat{\varphi}})]}{\underline{w}}\right]\left[\int_{\underline{\varphi}}^{\infty}\left[\left(\frac{w(\varphi)}{\underline{w}}\right)^{1+\theta} - \left(\frac{w(\varphi)}{\underline{w}}\right)^{\theta}\right]g(\varphi)d\varphi\right]}_{\text{Term 1}} \\
& + \underbrace{\left[\frac{G(\underline{\varphi})-G(\underline{\hat{\varphi}})}{\underline{w}}\right]\left[G(\underline{\varphi})-G(\underline{\hat{\varphi}})+\int_{\underline{\varphi}}^{\infty}\left[\frac{w(\varphi)}{\underline{w}}\right]^{\theta} g(\varphi)d\varphi\right]}_{\text{Term 2}}\Bigg\} \underbrace{+(1+\psi)\frac{d\ln \underline{W}}{d\underline{w}}}_{<0},
\end{aligned} \tag{B-37}$$

where $D = \{G(\underline{\varphi})-G(\underline{\hat{\varphi}})+\int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{\theta} g(\varphi)d\varphi\}\{G(\underline{\varphi})-G(\underline{\hat{\varphi}})+\int_{\underline{\varphi}}^{\infty}\left[w(\varphi)/\underline{w}\right]^{1+\theta} g(\varphi)d\varphi\} > 0$. The second equality above uses (B-35) to substitute for $\frac{d\ln \underline{M}}{d\underline{w}}$. Without further assumptions on the productivity distribution, we cannot pin down the sign of Term 1 + Term 2 in (B-37). However, if we assume a Pareto distribution for productivity, $g(\varphi) = \frac{k}{\varphi^{k+1}}$ and $G(\varphi) = 1 - \frac{1}{\varphi^k}$ for $k > 1+\theta$, we can show that Term 1 + Term 2 > 0, so that $\frac{d\ln \underline{L}}{d\underline{w}} < 0$.

Using $\frac{w(\varphi)}{\underline{w}} = \frac{\varphi}{\underline{\varphi}}$ for $\varphi \geq \underline{\varphi}$, $g(\underline{\hat{\varphi}}) = \frac{k}{\underline{\hat{\varphi}}^{k+1}}$, $G(\underline{\varphi}) - G(\underline{\hat{\varphi}}) = \frac{1}{\underline{\hat{\varphi}}^k} - \frac{1}{\underline{\varphi}^k}$, $\int_{\underline{\varphi}}^{\infty}\left(\frac{\varphi}{\underline{\varphi}}\right)^{\theta} g(\varphi)d\varphi = \left(\frac{k}{k-\theta}\right)\frac{1}{\underline{\varphi}^k}$, $\int_{\underline{\varphi}}^{\infty}\left(\frac{\varphi}{\underline{\varphi}}\right)^{1+\theta} g(\varphi)d\varphi = \left(\frac{k}{k-\theta-1}\right)\frac{1}{\underline{\varphi}^k}$, and defining $u \equiv \frac{\underline{\varphi}}{\underline{\hat{\varphi}}} \in \left(1, \frac{1+\theta}{\theta}\right)$, we obtain that Term 1+Term 2 $>$ 0 if and only if

$$\underbrace{\frac{u^k}{u^k-1}\left\{\frac{\frac{u^k}{k-1} + \frac{k\theta u}{(k-1)(k-\theta-1)} - \frac{1+\theta}{k-\theta-1}}{\frac{(1+\theta)u^{k-1}}{k-1} - \frac{\theta u^k}{k} + \frac{\theta(1+\theta)}{k(k-1)(k-\theta-1)}}\right\}}_{\text{Term 3}} + \underbrace{\left(\frac{k-\theta-1}{\theta}\right)\left[1+\left(\frac{k-\theta}{k}\right)(u^k-1)\right]}_{>0} > 1. \tag{B-38}$$

Thus, a sufficient condition for (B-38) to hold is that Term 3 > 1. In the term in braces within Term 3, both the numerator and denominator are positive and increasing in u, with the numerator approaching 0 and the denominator approaching $\frac{1}{k-\theta-1}$ as $u \to 1$. Rearranging terms, we get that

Term 3 > 1 if and only if

$$\frac{\theta(1+\theta)}{(k-\theta-1)u^k}+\frac{(1+\theta)k}{u}-(k-1)\theta > (1+\theta)ku^{k-1}-[(k-1)\theta+k]u^k-\frac{k^2\theta u}{k-\theta-1}+\frac{(1+\theta)[k(k-1)+\theta]}{k-\theta-1}. \tag{B-39}$$

Both the left-hand side (LHS) and the right-hand side (RHS) approach $\frac{k(k-1)}{k-\theta-1}$ as $u \to 1$, and they are both decreasing in u, with

$$\begin{aligned}\frac{d\text{LHS}}{du} &= -k(1+\theta)\left[\frac{1}{u^2}+\frac{\theta}{(k-\theta-1)u^{k+1}}\right] < 0,\\ \frac{d\text{RHS}}{du} &= -k\left[[k(1+\theta)-\theta]u^{k-1}\left(1-\frac{1}{u}\right)+u^{k-2}+\frac{k\theta}{k-\theta-1}\right] < 0.\end{aligned}$$

Hence, if the LHS declines faster with u than the RHS, then it must be that condition (B-39) holds. We obtain that $\left|\frac{d\text{LHS}}{du}\right| > \left|\frac{d\text{RHS}}{du}\right|$ if

$$\underbrace{u^k\left[\left(k-\frac{\theta}{1+\theta}\right)u-k+1\right]+\frac{k\theta u^2}{(1+\theta)(k-\theta-1)}}_{\text{Term 4}} > \underbrace{1+\frac{\theta}{(k-\theta-1)u^{k-1}}}_{\text{Term 5}},$$

which is always true because both Term 4 and Term 5 approach $\frac{k-1}{k-\theta-1}$ as $u \to 1$, and Term 4 is strictly increasing with u, whereas Term 5 is strictly decreasing with u. Therefore, Term 3 > 1, and thus, Term 1+Term 2 > 0 and $\frac{d\ln \underline{L}}{d\underline{w}} < 0$.

From (B-3) we know that the wage bill is equal to $\underline{W}^{1+\psi}$. It is also the case that the wage bill equals the product of the average wage and total employment, $\underline{\bar{w}}\underline{L}$. It follows that $\underline{\bar{w}} = \frac{\underline{W}^{1+\psi}}{\underline{L}}$, and therefore, $\frac{d\ln \underline{\bar{w}}}{d\underline{w}} = (1+\psi)\frac{d\ln \underline{W}}{d\underline{w}} - \frac{d\ln \underline{L}}{d\underline{w}}$. Using (B-37), it follows that

$$\frac{d\ln \underline{\bar{w}}}{d\underline{w}} = \frac{1}{D}\left[\text{Term } 1 + \text{Term } 2\right] > 0.$$

Therefore, if firm productivity has a Pareto distribution, an increase in the minimum wage increases the average wage but reduces total employment.