Minimum Wage and Firm Variety

Appendix — For Online Publication

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

September 2021

A Empirical Appendix
A.1 Supporting Figures and Tables

Figure A-1: The 137 multi-state commuting zones (dark blue)
Figure A-2: Comparison between commuting-zone groups for restaurant industry: 137 multi-state commuting zones (solid blue) and rest of the country (dashed red)

Notes: This figure shows that in both groups, the restaurant industry accounted for 6.1 percent of all establishments in 1990 and that this share increased to about 7.9 percent by 2011. From Figure A-2d, note that the average nominal wage increased from about $3.7 in 1990 to about $7.2 in 2011. These numbers are smaller than those reported for the average minimum wage, but this does not mean that restaurant workers are paid (on average) below the minimum—our measure of worker earnings underestimates the hourly wage because it assumes that every worker is employed for 40 hours per week during the entire year.
Figure A-3: Comparison between commuting-zone groups for retail-trade industry: 137 multi-state commuting zones (solid blue) and rest of the country (dashed red)

Notes: For the retail-trade industry, this figure shows that the multi-state commuting zone group is similar to the rest of the country, additionally documenting a steady decline in the share of retail establishments (as a fraction of total establishments) from about 19.2-19.9 percent to about 13.7-14.2 percent. The fact that the employment share of this industry only declined from 13.94 percent to 13.08 percent indicates that this industry has become more concentrated relative to the rest of the economy.
Table A-1: Employment shares and wage ranking of industries, 1990 and 2011

<table>
<thead>
<tr>
<th>Industry</th>
<th>1990</th>
<th></th>
<th>2011</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Employment</td>
<td>Nominal Wage (US$)</td>
<td>Wage Ranking</td>
<td>Employment</td>
</tr>
<tr>
<td>Eating and drinking places</td>
<td>7.21%</td>
<td>3.68</td>
<td>1</td>
<td>8.72%</td>
</tr>
<tr>
<td>Retail trade</td>
<td>13.94%</td>
<td>6.46</td>
<td>2</td>
<td>13.08%</td>
</tr>
<tr>
<td>Services</td>
<td>16.40%</td>
<td>6.74</td>
<td>3</td>
<td>22.14%</td>
</tr>
<tr>
<td>Textiles/apparel</td>
<td>1.99%</td>
<td>7.63</td>
<td>4</td>
<td>0.36%</td>
</tr>
<tr>
<td>Wood/furniture</td>
<td>1.36%</td>
<td>9.46</td>
<td>5</td>
<td>0.68%</td>
</tr>
<tr>
<td>Other manufacturing</td>
<td>0.44%</td>
<td>10.23</td>
<td>6</td>
<td>0.22%</td>
</tr>
<tr>
<td>Food/tobacco</td>
<td>1.62%</td>
<td>11.49</td>
<td>7</td>
<td>1.35%</td>
</tr>
<tr>
<td>Health services</td>
<td>9.83%</td>
<td>11.63</td>
<td>8</td>
<td>13.25%</td>
</tr>
<tr>
<td>Plastics, clay, stone</td>
<td>1.57%</td>
<td>11.85</td>
<td>9</td>
<td>0.92%</td>
</tr>
<tr>
<td>Agriculture, forestry, fishing, and mining</td>
<td>1.29%</td>
<td>12.03</td>
<td>10</td>
<td>1.56%</td>
</tr>
<tr>
<td>Construction</td>
<td>5.94%</td>
<td>12.09</td>
<td>11</td>
<td>4.71%</td>
</tr>
<tr>
<td>Paper/printing</td>
<td>2.44%</td>
<td>12.91</td>
<td>12</td>
<td>1.28%</td>
</tr>
<tr>
<td>Wholesale trade</td>
<td>6.69%</td>
<td>13.40</td>
<td>13</td>
<td>5.69%</td>
</tr>
<tr>
<td>Finance, insurance, and real estate</td>
<td>7.60%</td>
<td>13.42</td>
<td>14</td>
<td>7.24%</td>
</tr>
<tr>
<td>Metals</td>
<td>2.46%</td>
<td>13.57</td>
<td>15</td>
<td>1.42%</td>
</tr>
<tr>
<td>Transp., comm., elec., gas, and sanitary</td>
<td>5.88%</td>
<td>13.88</td>
<td>16</td>
<td>5.69%</td>
</tr>
<tr>
<td>Equipment</td>
<td>4.94%</td>
<td>14.53</td>
<td>17</td>
<td>2.46%</td>
</tr>
<tr>
<td>Legal, consulting, and computing services</td>
<td>5.27%</td>
<td>16.50</td>
<td>18</td>
<td>7.51%</td>
</tr>
<tr>
<td>Transportation manufacturing</td>
<td>2.03%</td>
<td>16.78</td>
<td>19</td>
<td>0.98%</td>
</tr>
<tr>
<td>Chemicals/petroleum</td>
<td>1.11%</td>
<td>17.22</td>
<td>20</td>
<td>0.73%</td>
</tr>
</tbody>
</table>
### Table A-2: Estimation of Responses to Minimum Wages Using County-Pairs Approach

<table>
<thead>
<tr>
<th></th>
<th>Number of establishments</th>
<th>Employment</th>
<th>Average wage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) (2) (3)</td>
<td>(4) (5) (6)</td>
<td>(7) (8) (9)</td>
</tr>
<tr>
<td>A. Restaurant industry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(minimum wage)</td>
<td>-0.036 (-0.046)</td>
<td>-0.140**</td>
<td>0.259***</td>
</tr>
<tr>
<td></td>
<td>-0.034 (-0.052)</td>
<td>-0.123*</td>
<td>0.231***</td>
</tr>
<tr>
<td></td>
<td>-0.042 (-0.067)</td>
<td>-0.140</td>
<td>0.241***</td>
</tr>
<tr>
<td>ln(population)</td>
<td>0.573*** (0.098)</td>
<td>0.907***</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>0.625*** (0.096)</td>
<td>0.959***</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>0.649*** (0.104)</td>
<td>0.958***</td>
<td>0.117</td>
</tr>
<tr>
<td>ln(establishments−)</td>
<td>0.426*** (0.076)</td>
<td>0.045</td>
<td>(0.049)</td>
</tr>
<tr>
<td></td>
<td>0.399*** (0.072)</td>
<td>0.045</td>
<td>(0.058)</td>
</tr>
<tr>
<td></td>
<td>0.410*** (0.066)</td>
<td>0.045</td>
<td>(0.075)</td>
</tr>
<tr>
<td>ln(employment−)</td>
<td></td>
<td>0.165***</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.134***</td>
<td>-0.025</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.156**</td>
<td>0.036</td>
</tr>
<tr>
<td>ln(average wage−)</td>
<td></td>
<td>0.048</td>
<td>(0.043)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.049</td>
<td>(0.044)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.060</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Number of county pairs</td>
<td>1,165 770 410</td>
<td>1,165 770 410</td>
<td>1,165 770 410</td>
</tr>
<tr>
<td>Observations</td>
<td>25,497 16,893 9,009</td>
<td>25,495 16,893 9,009</td>
<td>25,336 16,815 8,972</td>
</tr>
<tr>
<td>B. Retail-trade industry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(minimum wage)</td>
<td>-0.060* (-0.033)</td>
<td>-0.002</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>-0.039 (-0.037)</td>
<td>0.017</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>-0.058 (-0.050)</td>
<td>-0.024</td>
<td>0.068</td>
</tr>
<tr>
<td>ln(population)</td>
<td>0.521*** (0.080)</td>
<td>0.792***</td>
<td>-0.068</td>
</tr>
<tr>
<td></td>
<td>0.531*** (0.102)</td>
<td>0.834***</td>
<td>-0.147**</td>
</tr>
<tr>
<td></td>
<td>0.333*** (0.096)</td>
<td>0.882***</td>
<td>-0.158*</td>
</tr>
<tr>
<td>ln(establishments−)</td>
<td>0.393*** (0.050)</td>
<td>0.082</td>
<td>(0.051)</td>
</tr>
<tr>
<td></td>
<td>0.407*** (0.054)</td>
<td>0.090</td>
<td>(0.057)</td>
</tr>
<tr>
<td></td>
<td>0.521*** (0.068)</td>
<td>0.106</td>
<td>(0.083)</td>
</tr>
<tr>
<td>ln(employment−)</td>
<td></td>
<td>0.127***</td>
<td>0.084***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.125***</td>
<td>0.081*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.097*</td>
<td>0.112**</td>
</tr>
<tr>
<td>ln(average wage−)</td>
<td></td>
<td>(0.035)</td>
<td>(0.032)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.040)</td>
<td>(0.043)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.057)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>Number of county pairs</td>
<td>1,166 770 410</td>
<td>1,166 770 410</td>
<td>1,166 770 410</td>
</tr>
<tr>
<td>Observations</td>
<td>25,581 16,911 9,016</td>
<td>25,581 16,911 9,016</td>
<td>25,423 16,833 8,979</td>
</tr>
</tbody>
</table>

Notes: This table reports $\hat{\beta}$, $\hat{\gamma}$, and $\hat{\delta}$ from the estimation of equation (3) for the restaurant and retail industries using DLR county pairs. For each county pair, the dependent variable is the difference in log number of establishments in columns 1-3, the difference in log employment in columns 4-6, and the difference in log average wage in columns 7-9. Columns 1, 4, and 7 use all DLR county pairs available in our dataset, columns 2, 5, and 8 use pairs where at least one of the counties is in our 137 multi-state commuting zones, and columns 3, 6, and 9 use only the pairs where both counties are in our 137 commuting zones. All regressions include entity fixed effects. Standard errors (in parentheses) are clustered at commuting-zone level. The coefficients are statistically significant at the *10%, **5%, or ***1% level.
A.2 Estimation Using the County Pairs of DLR

This section discusses the results of Table A-2, which shows the estimation of (3) for the number of establishments, employment, and average wages using the county pairs of DLR.

Out of the 1,181 DLR county pairs, we have 1,165 for the restaurant industry, and 1,166 for the retail-trade industry.\(^1\) Column 4 presents the results for the specification that is the closest to DLR, who obtain a non-significant minimum-wage elasticity of employment of 0.016 for the restaurant industry. In contrast, our county-pair results show an elasticity of $-0.14$, which is statistically significant at a 5 percent level. In comparison with the $-0.325$ elasticity of column 5 in Table 2, the county-pair specification coefficient declines in size by more than half.\(^2\)

For a more direct comparison with our 137 multi-state commuting zones, columns 5 and 6 in Table A-2 show the employment estimation using county pairs for which at least one of the counties is in our 137 commuting zones (column 5), and using county pairs for which both counties are in our 137 zones.\(^3\) The point estimate for the minimum-wage elasticity of employment in the restaurant industry barely changes, but becomes non-significant in the regression where county pairs are fully contained in our 137 zones. For the retail-trade industry, the minimum-wage elasticity of employment is close to zero and non-significant in our three county-pair specifications.

For the number of establishments, which is our main interest and is not analyzed by DLR, the first three columns of Table A-2 show negative but small and non-significant minimum-wage elasticities for the restaurant industry, which again highlights how the county-pair approach underestimates the effect of minimum wages. Something similar happens for the retail-trade industry, with the exception of a mildly significant elasticity in the specification using all available county pairs. Finally, the point estimates for the minimum-wage elasticities of average wages in the restaurant industry (between 0.231 and 0.259) are similar to those obtained in Table 3 and retain their statistical significance.\(^4\) The fact that the elasticities for average wages in the restaurant industry are not really affected when using the county-pair approach simply reflects that no matter the location, there is always a sizable fraction of

---

\(^1\)We lose four pairs that involve DC, six pairs for which both counties are within the same state (DLR use sub-state minimum wage data, whereas we use state-level minimum wage data), and the rest are county pairs for which there is no data for our industries of interest.

\(^2\)As a robustness check, DLR also use CBP data and obtain a non-significant elasticity of -0.034. However, DLR are skeptical about their CBP data because of the changes in industry classification (from SIC to NAICS) and the fact that due to confidentiality reasons, the CBP reports many county-industry cells as an employment range. As a consequence, DLR’s CBP data presents some important deficiencies. These problems are minimal in our CBP data, which was processed using the sophisticated fixed-point imputation and industry-classification method of Autor, Dorn, and Hanson (2013) and Acemoglu, Autor, Dorn, Hanson, and Price (2016). Additionally, for the main objective of this empirical section, the number of establishments is always reported in the CBP.

\(^3\)There are 770 county pairs for which at least one county is in our 137 zones, and 410 for which both counties are in our 137 zones. That is, out of 770 there are at least 360 county pairs that involve two different commuting zones (there may be more, as within the 410 pairs there may be some that involve two of our 137 zones), indicating a potentially large number of county pairs that share little economic activity.

\(^4\)They are also close to the point estimates obtained by DLR in their county-pair earnings regressions (0.188 when using the QCEW and 0.22 when using the CBP).
restaurant employees that are paid the minimum wage. Hence, an increase in the minimum wage in one county will be met by an increase in the average wage gap between this county and all counties in other states (irrespective of whether a county pair has any joint economic activity).

B Theoretical Appendix

B.1 Household Maximization Problem

Given the utility function in (4), 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

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

The solution to this problem yields that the labor supply to firm $\omega$ is

$$l(\omega) = N^{\frac{\theta}{\psi}} w(\omega)^{\theta}.$$  \hspace{1cm} (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

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

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

$$l(\omega) = \frac{w(\omega)^{\theta}}{W^{\theta-\psi}}.$$  \hspace{1cm} (B-2)

Using (B-2), the aggregate labor supply is

$$L = \int_{\omega \in \Omega} l(\omega) d\omega = \frac{1}{1+\theta} \int_{\omega \in \Omega} w(\omega)^{\theta} d\omega.$$  \hspace{1cm} (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 (11), the left-hand side is strictly decreasing with $\hat{\varphi}$, approaching zero as $\hat{\varphi} \rightarrow \infty$. Thus, as long as $f_E$ is sufficiently small, there is unique value for $\hat{\varphi}$ that solves (11). From the zero-cutoff-profit condition, the solution for the wage index is then

$$W_D = \left\{ \left[ \frac{\theta}{(1+\theta)^{1+\theta}} \right] \frac{\varphi_D^{1+\theta}}{f} \right\}^{\frac{1}{1-\psi}},$$  \hspace{1cm} (B-4)
and from (9) and (B-4) we know that
\[ l_D(\varphi) = \left( \frac{\theta}{1 + \theta} \right)^{\theta} \frac{\varphi^\theta}{W_D^{\theta - \psi}} = \frac{(1 + \theta) f \varphi^\theta}{\varphi^{1 + \theta}}. \] (B-5)

To obtain the expression for the equilibrium mass of firms, \( M_D \), we use equations (12) and (9) to obtain
\[ M_D = \left[ \int_{\varphi_D}^{\infty} \varphi^{1 + \theta} g(\varphi \geq \varphi_D) d\varphi \right]^{-1} \left( \frac{1 + \theta}{\theta} \right)^{1 + \theta} W_D^{1 + \theta} \] (B-6)

From equation (13), the equilibrium total employment is
\[ L_D = \frac{(1 + \theta) f M_D}{\varphi_D^{1 + \theta}} \int_{\varphi_D}^{\infty} \varphi^\theta g(\varphi \geq \varphi_D) d\varphi. \] (B-7)

Finally, from (8) we know that \( U_D = \left( \frac{1}{1 + \psi} \right) W_D^{1 + \psi} = \left( \frac{1}{1 + \psi} \right) N_D^{1 + \frac{1}{\psi}} \), 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 (4) subject to equation (5) and \( C = M \int_{\hat{\varphi}}^{\infty} \varphi l(\varphi) g(\varphi \geq \hat{\varphi}) d\varphi - M_E f_E - M f \), 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 \( g(\varphi \geq \hat{\varphi}) = \frac{g(\varphi)}{1 - G(\hat{\varphi})} \), we can write the planner’s problem as
\[
\max_{l(\varphi), M_E, \hat{\varphi}} \left\{ M_E \int_{\hat{\varphi}}^{\infty} \varphi l(\varphi) g(\varphi) d\varphi - M_E f_E - [1 - G(\hat{\varphi})] M f - \left[ \int_{\hat{\varphi}}^{\infty} \varphi l(\varphi) g(\varphi) d\varphi \right]^{\frac{1}{1 + \frac{1}{\psi}}} \right\}.
\] (B-8)

The first order conditions with respect to \( l(\varphi) \), \( M_E \), and \( \varphi \) are respectively
\[
\varphi - N^{\frac{\theta - \psi}{\theta}} l(\varphi) \frac{\theta}{1 + \theta} = 0, \quad \text{(B-9)}
\]
\[
\int_{\hat{\varphi}}^{\infty} \varphi l(\varphi) g(\varphi) d\varphi - \left( \frac{\theta}{1 + \theta} \right) N^{\frac{\theta - \psi}{\theta}} \int_{\hat{\varphi}}^{\infty} l(\varphi)^{1 + \theta} g(\varphi) d\varphi - f_E - [1 - G(\hat{\varphi})] f = 0, \quad \text{(B-10)}
\]
\[
\hat{\varphi} l(\hat{\varphi}) - \left( \frac{\theta}{1 + \theta} \right) N^{\frac{\theta - \psi}{\theta}} l(\hat{\varphi})^{1 + \theta} - f = 0. \quad \text{(B-11)}
\]

Note that (B-9) implies that \( N^{\frac{\theta - \psi}{\theta}} = \frac{\hat{\varphi}}{l(\hat{\varphi})^\theta} \), 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
\[
N = \left[ \frac{\hat{\varphi}^{1 + \theta}}{(1 + \theta) f} \right]^\frac{\psi}{\psi - \theta}. \quad \text{(B-12)}
\]
Note also from (B-9) that \( \frac{\bar{l}(\phi)}{l(\hat{\phi})} = \left(\frac{\bar{\phi}}{\hat{\phi}}\right)^\theta \), which from (9) 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

\[
l(\phi) = \frac{(1 + \theta) f \phi^\theta}{\hat{\phi}^{1+\theta}}.
\]

(B-13)

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

\[
\int_{\hat{\phi}}^\infty \phi l(\phi) g(\phi) d\phi = (1 + \theta) f \int_{\hat{\phi}}^\infty \left(\frac{\phi}{\hat{\phi}}\right)^{1+\theta} g(\phi) d\phi,
\]

(B-14)

\[
\int_{\hat{\phi}}^\infty l(\phi)^{1+\theta} g(\phi) d\phi = \left[\frac{(1 + \theta) f}{\phi}\right]^{\frac{1+\theta}{\theta}} \int_{\hat{\phi}}^\infty \left(\frac{\phi}{\hat{\phi}}\right)^{1+\theta} g(\phi) d\phi,
\]

(B-15)

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

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

(B-16)

Notice that (B-16) is exactly the same as (11), which is the equation determining \( \hat{\phi}_D \) in the decentralized case. Using \( \hat{\phi}_P \) to denote the equilibrium cutoff productivity level in the planner’s problem, it follows that \( \hat{\phi}_P = \hat{\phi}_D \). Once we obtain \( \hat{\phi}_P \), we plug it into (B-12) and (B-13) to obtain \( N_P \) and \( l_P(\phi) \).

From (B-5), it follows that \( l_P(\phi) = l_D(\phi) \), so that firm size is the same in both cases. Moreover, using \( N_D = W_{\psi}^\psi \), (B-4), (B-12), and \( \hat{\phi}_P = \hat{\phi}_D \) we get that

\[
\frac{N_P}{N_D} = \left(\frac{1 + \theta}{\theta}\right)^{\frac{\#_{\psi}}{W_{\psi}^\psi}} > 1.
\]

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 \( M = [1 - G(\hat{\phi})] M_E \) to get

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

(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

\[
\frac{M_P}{M_D} = \left[\frac{\theta N_P^{1/\psi}}{(1+\theta) W_{\psi}^\psi}\right]^{1+\theta},
\]

which using (B-4), (B-12), and \( \hat{\phi}_P = \hat{\phi}_D \), can be rewritten as

\[
\frac{M_P}{M_D} = \left(\frac{1 + \theta}{\theta}\right)^{\frac{1+\theta}{\#_{\psi}}} > 1.
\]

(B-18)

Thus, the decentralized outcome yields a suboptimal mass of firms. Total employment in the planner’s case is

\[
L_P = M_P \int_{\hat{\phi}_P}^\infty l_P(\phi) g(\phi|\phi \geq \hat{\phi}_P) d\phi,
\]

which using (B-13) can be rewritten as

\[
L_P = \frac{(1 + \theta) f M_P}{\hat{\phi}_P^{1+\theta}} \int_{\hat{\phi}_P}^\infty \phi^{1+\theta} g(\phi|\phi \geq \hat{\phi}_D) d\phi.
\]

(B-19)
From (B-19), (B-7), and $\hat{\varphi}_p = \hat{\varphi}_D$, it follows that
\[
\frac{L_p}{L_D} = \frac{M_p}{M_D} = \left(\frac{1 + \theta}{\theta}\right)^{\frac{1+\theta}{\psi-\psi}} > 1. \tag{B-20}
\]

Regarding welfare, note first from (B-16) that
\[
M_E f_E + [1 - G(\hat{\varphi})]M_E f = M_E \int_{\hat{\varphi}}^{\infty} \varphi^{1+\theta} f(\varphi) d\varphi,
\]
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
\[
U_P = \frac{M_P}{M_D} \frac{\theta f(\hat{\varphi})}{1+\psi} - N_P \frac{1+\frac{\varphi}{\psi}}{1+\frac{1+\theta}{\psi}} \left(1 + \frac{\theta}{1+\theta}\right) \ln \left(\frac{1+\theta}{\theta}\right) < 0,
\]
where the second equality uses (B-17). From (B-12), we obtain that
\[
\frac{\theta f(\hat{\varphi}_p)}{1+\psi} = \frac{\theta}{(1+\theta)N_p^{\theta(1+\psi)}} \left(\frac{1}{1+\theta}\right)^{\theta(1+\psi)} > 1.
\tag{B-21}
\]

Using the expression for welfare in the decentralized case given in the end of section B.2, we get that
\[
\frac{U_P}{U_D} = \frac{\theta - \psi}{1+\theta} \left(\frac{N_D}{N_P}\right)^{1+\psi}.
\]
Using (B-12), $N_D = W_D^{\psi}$, (B-4), and $\hat{\varphi}_p = \hat{\varphi}_D$, this ratio can be rewritten as
\[
\frac{U_P}{U_D} = \frac{(\theta - \psi)N_p^{1+\frac{\psi}{\theta}}}{(1+\theta)(1+\psi)}. \tag{B-22}
\]
We know that $\frac{U_P}{U_D} > 1$ because $\frac{U_P}{U_D}$ is strictly decreasing in $\theta$,
\[
\frac{d (U_P/U_D)}{d \theta} = - \psi \left(1 + \frac{\psi}{\theta - \psi} \left(1 + \theta \right) \frac{\theta(1+\psi)}{\psi - \psi} \ln \left(\frac{1+\theta}{\theta}\right) < 0,
\]
and $\lim_{\theta \to \infty} \frac{U_P}{U_D} = 1$.

**B.4 Optimality of Wage Subsidy**

Here we show that a wage subsidy can restore optimality. With a proportional wage subsidy of $s = \frac{1}{1+\theta}$, the effective wage bill of a firm becomes $(1-s)w(\varphi)l(\varphi) = (1-s)W_s^{\theta - \psi} l(\varphi)^{\frac{\theta}{\theta - \psi}}$. Therefore, the firm’s profit maximization problem yields as solution
\[
l_s(\varphi) = \frac{\varphi^{\theta-\psi}}{W_s^{\theta - \psi}}, \tag{B-24}
\]
so that $w_s(\varphi) = \varphi$; that is, including the subsidy, each workers is paid the value of its marginal product. We use subscript $S$ to indicate that this is the model with a wage subsidy. The gross profit function of a firm with productivity $\varphi$ is then given by $\pi_s(\varphi) = \left(\frac{1}{1+\theta}\right) \frac{\varphi^{1+\theta}}{W_s^{\theta - \psi}}$. 


The zero-cutoff-profit condition, \( \pi_S(\hat{\varphi}_S) = f \), implies that \( W_{S}^{\theta-\psi} = \frac{\hat{\varphi}_S^{1+\theta}}{(1+\theta)f} \), which can be used to rewrite firm-level employment and gross profits as \( l_S(\varphi) = \frac{(1+\theta)\hat{\varphi}_S^{\theta}}{(1+\theta)f} \) and \( \pi_S(\varphi) = \left( \frac{\varphi}{\hat{\varphi}_S} \right)^{1+\theta} f \). It follows that the free-entry condition is exactly the same as in equation (11) for the decentralized case, and therefore, \( \hat{\varphi}_S = \hat{\varphi}_D = \hat{\varphi}_P \). This result along with the zero-cutoff-profit condition and (B-4) yields that \( \frac{W_S}{W_D} = \left( \frac{1+\theta}{\theta} \right)^{\psi/(\theta-\psi)} > 1 \). That is, the wage subsidy increases the wage index. As well, from (B-5), (B-13), and our expression for \( l_S(\varphi) \) above, it also follows that \( l_S(\varphi) = l_D(\varphi) = l_P(\varphi) \).

Given that \( N_S = W_S^{\psi} \), it follows from the zero-cutoff-profit condition that

\[
N_S = \left[ \frac{\hat{\varphi}_S^{1+\theta}}{(1+\theta)f} \right]^{\frac{\psi}{1+\theta}} = N_P,
\]

where the second equality follows (B-12) and \( \hat{\varphi}_S = \hat{\varphi}_P \). From \( N_S = N_P \), \( l_S(\varphi) = l_P(\varphi) \), \( \hat{\varphi}_S = \hat{\varphi}_P \), and the definition of \( N_S \), \( N_S \equiv \left[ M_S \int_{\hat{\varphi}_S}^{\infty} l(\varphi) \frac{1+\theta}{\theta} g(\varphi | \varphi \geq \hat{\varphi}_S) d\varphi \right]^{\frac{1}{1+\theta}} \), it follows that that \( M_S = M_P \) as given by (B-17). As well, from the definition of \( L_S \), \( L_S \equiv M_S \int_{\hat{\varphi}_S}^{\infty} l_S(\varphi) g(\varphi | \varphi \geq \hat{\varphi}_S) d\varphi \), it also follows that \( L_S = L_P \) as given by (B-19).

Assuming that the wage subsidy is financed by a lump-sum tax on households, \( T \), total household consumption is then given by \( W_S N_S - T \), where \( W_S N_S \) is the wage bill. The total amount of the subsidy is \( sW_S N_S \), and thus, \( C_S \equiv (1-s)W_S N_S = \left( \frac{\theta}{1+\theta} \right) N_S^{1+\frac{\psi}{\theta}} \). Therefore, from equation (4), we can write welfare as

\[
U_S = \left( \frac{\theta}{1+\theta} \right) N^{1+\frac{\psi}{\theta}} - \left( \frac{\psi}{1+\psi} \right) N^{1+\frac{\psi}{\theta}} \left( \frac{\theta - \psi}{1+\theta} \right) N^{1+\frac{\psi}{\theta}} = \frac{\theta - \psi}{1+\theta} N^{1+\frac{\psi}{\theta}} = U_P, \tag{B-25}
\]

where the last equality follows from \( N_S = N_P \) and equation (B-22). Thus, a proportional wage subsidy of \( s = \frac{1}{1+\theta} \) restores optimality.

**B.5 The Effects of a Binding Minimum Wage**

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

From the zero-cutoff-profit condition in section 3.3, we know that \( W_{\theta-\psi} = \left( \frac{\theta}{\psi} - w \right) \frac{\psi}{\theta} \), which allows us to rewrite the free-entry condition in (14) as

\[
\int_{\hat{\varphi}}^{\varphi} \left( \frac{\varphi-w}{\hat{\varphi}-w} - 1 \right) f g(\varphi) d\varphi + \int_{\hat{\varphi}}^{\infty} \left\{ \left[ \frac{\theta}{(1+\theta)^{1+\theta}} \right] \frac{\varphi^{1+\theta}}{(\hat{\varphi}-w)^{\psi}} - 1 \right\} f g(\varphi) d\varphi = f_E. \tag{B-26}
\]

Taking the derivative of (B-26) with respect to \( w \), and given that \( \frac{d\varphi}{dw} = \frac{1+\theta}{\theta} \), we obtain that

\[
\frac{d\hat{\varphi}}{dw} = \left[ \int_{\hat{\varphi}}^{\varphi} \left( \frac{\varphi-w}{\hat{\varphi}-w} \right) g(\varphi) d\varphi + \int_{\hat{\varphi}}^{\infty} \left( \frac{\varphi}{\hat{\varphi}} \right)^{1+\theta} g(\varphi) d\varphi \right] \left( \frac{\varphi - \hat{\varphi}}{\hat{\varphi} - w} \right) > 0. \tag{B-27}
\]
All the terms in (B-27) are positive because $w < \hat{\varphi} < \varphi$. We know that $w < \hat{\varphi}$ from the zero-cutoff-profit condition, $(\hat{\varphi} - w) \frac{w^\theta}{W^\theta - w} = f$, and we know that $\hat{\varphi} < \varphi$ because firms with $\varphi \in [\hat{\varphi}, \varphi]$ are constrained by the minimum wage. Moreover, the term in brackets is less than 1 because $\int_{\hat{\varphi}}^{\varphi} \left( \frac{\varphi - w}{\varphi - \hat{\varphi}} \right) g(\varphi) d\varphi < \int_{\hat{\varphi}}^{\varphi} \left( \frac{\varphi - w}{\varphi - \hat{\varphi}} \right) g(\varphi) d\varphi$, which follows from $(\hat{\varphi} - w)/(\varphi - \hat{\varphi}) > 0$ for every $\varphi < \varphi$. Therefore,

$$\frac{d\hat{\varphi}}{dw} \in \left( 0, \frac{\varphi - \hat{\varphi}}{\varphi - w} \right).$$

(B-28)

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

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

The sign of $\frac{dW}{dw}$ is determined by the term within the brackets, which is negative if and only if

$$\frac{d\hat{\varphi}}{dw} < 1 - \frac{\theta(\hat{\varphi} - w)}{w} = 1 - \frac{\varphi - w}{\varphi - \hat{\varphi}} = \frac{\varphi - \hat{\varphi}}{\varphi - w},$$

where we use that $\varphi - w = \frac{w}{\theta}$. From (B-28) we know that (B-29) holds, and therefore, $\frac{dW}{dw} < 0$. Similar to (8), welfare with a binding minimum wage is given by $U = W^{1+\psi}$, and thus, $\frac{dU}{dw} = W^\psi \frac{dW}{dw} < 0$.

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

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

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

$$\frac{d\ln M}{dw} = (1 + \theta) \frac{d\ln W}{dw} - \left[ \frac{g(\varphi)}{1 - G(\varphi)} \right] \frac{d\hat{\varphi}}{dw} + \frac{w^{1+\theta}g(\varphi)^{\frac{d\hat{\varphi}}{dw}} - (1 + \theta)w^\theta[G(\varphi) - G(\hat{\varphi})]}{w^{1+\theta}[G(\varphi) - G(\hat{\varphi})] + \int_{\hat{\varphi}}^{\varphi} w(\varphi)^{1+\theta} g(\varphi) d\varphi}
$$

$$= (1 + \theta) \frac{d\ln W}{dw} - g(\varphi) \frac{d\hat{\varphi}}{dw} \left[ \frac{1}{1 - G(\varphi)} - \frac{G(\varphi) - G(\hat{\varphi})}{G(\varphi) - G(\hat{\varphi}) + \int_{\varphi}^{\varphi} [w(\varphi)/w]^{1+\theta} g(\varphi) d\varphi} \right]
$$

$$- \left( \frac{1 + \theta}{w} \right) \frac{G(\varphi) - G(\hat{\varphi})}{G(\varphi) - G(\hat{\varphi}) + \int_{\varphi}^{\varphi} [w(\varphi)/w]^{1+\theta} g(\varphi) d\varphi}.\quad (B-30)$$

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

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

12
which is true if \( \int_{\frac{\varphi}{w}}^{\infty} \left[ \frac{w(\varphi)}{w} \right]^{1+\theta} g(\varphi) d\varphi > 1 - G(\varphi) \), which implies \( \int_{\frac{\varphi}{w}}^{\infty} \left[ \frac{w(\varphi)}{w} \right]^{1+\theta} g(\varphi|\varphi \geq \varphi) d\varphi > 1 \). This condition holds because \( \frac{w(\varphi)}{w} \geq 1 \) for \( \varphi \geq \varphi \) (with equality if and only if \( \varphi = \varphi \)). Therefore, \( \frac{dM}{dw} < 0 \).

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

\[
L = \frac{M w^\theta}{[1 - G(\varphi)]w^\theta - \varphi} \left\{ G(\varphi) - G(\varphi) + \int_{\frac{\varphi}{w}}^{\infty} \left[ \frac{w(\varphi)}{w} \right]^{1+\theta} g(\varphi) d\varphi \right\}. \tag{B-31}
\]

The derivative of \( \ln L \) with respect to \( w \) is then given by

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

\[
+ \left( \frac{\theta}{w} \right) \left[ \frac{G(\varphi) - G(\varphi) + \int_{\frac{\varphi}{w}}^{\infty} \left[ \frac{w(\varphi)}{w} \right]^{1+\theta} g(\varphi) d\varphi}{G(\varphi) - G(\varphi) + \int_{\frac{\varphi}{w}}^{\infty} \left[ \frac{w(\varphi)}{w} \right]^{1+\theta} g(\varphi) d\varphi} \right] \frac{g(\varphi)}{w}
\]

\[
+ (1 + \psi) \frac{d \ln W}{d w} + \frac{G(\varphi) - G(\varphi)}{w} \times
\]

\[
\left[ \frac{1}{G(\varphi) - G(\varphi) + \int_{\frac{\varphi}{w}}^{\infty} \left[ \frac{w(\varphi)}{w} \right]^{1+\theta} g(\varphi) d\varphi} - \frac{1 + \theta}{G(\varphi) - G(\varphi) + \int_{\frac{\varphi}{w}}^{\infty} \left[ \frac{w(\varphi)}{w} \right]^{1+\theta} g(\varphi) d\varphi} \right]
\]

\[
= - \frac{1}{D} \left\{ \left[ g(\varphi) \frac{d \hat{\varphi}}{d w} - \theta [G(\varphi) - G(\varphi)] \right] \left[ \int_{\frac{\varphi}{w}}^{\infty} \left[ \left( \frac{w(\varphi)}{w} \right)^{1+\theta} - \left( \frac{w(\varphi)}{w} \right)^{\theta} \right] g(\varphi) d\varphi \right] \right\}
\]

\[
+ \left[ \frac{G(\varphi) - G(\varphi)}{w} \right] \left[ G(\varphi) - G(\varphi) + \int_{\frac{\varphi}{w}}^{\infty} \left[ \frac{w(\varphi)}{w} \right]^{1+\theta} g(\varphi) d\varphi \right]
\]

\[
+ (1 + \psi) \frac{d \ln W}{d w} < 0, \tag{B-32}
\]

where \( D = \{G(\varphi) - G(\varphi) + \int_{\frac{\varphi}{w}}^{\infty} \left[ \frac{w(\varphi)}{w} \right]^{1+\theta} g(\varphi) d\varphi\} \{G(\varphi) - G(\varphi) + \int_{\frac{\varphi}{w}}^{\infty} \left[ \frac{w(\varphi)}{w} \right]^{1+\theta} g(\varphi) d\varphi\} > 0 \). The second equality above uses (B-30) to substitute for \( \frac{d \ln M}{d w} \). Without further assumptions on the productivity distribution, we cannot pin down the sign of Term 1 + Term 2 in (B-32). However, if we assume a Pareto distribution for productivity, \( g(\varphi) = \frac{k}{\varphi^{k+\theta+1}} \) for \( k > 1 + \theta \), we can show that Term 1 + Term 2 > 0, so that \( \frac{d \ln L}{d w} < 0 \).

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

\[
\frac{u^k}{u^k - 1} \left\{ \frac{k \theta u}{k-1} + \frac{\theta (k-1)(k-\theta-1)}{k-\theta} \right\} + \left( \frac{k-\theta - 1}{\theta} \right) \left[ 1 + \left( \frac{k-\theta}{k} \right) (u^k - 1) \right] > 1. \tag{B-33}
\]
Thus, a sufficient condition for (B-33) 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

\[
\frac{\theta(1 + \theta)}{(k - \theta - 1)u^k} + \frac{(1 + \theta)k}{u} - (k-1)\theta > (1+\theta)k u^{k-1} - [(k-1)\theta + k]u^k - \frac{k^2\theta u}{k-\theta-1} + \frac{(1 + \theta)[k(k - 1) + \theta]}{k - \theta - 1}.
\]

(B-34)

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

\[
\frac{d\text{LHS}}{du} = -k(1 + \theta) \left[ \frac{1}{u^2} + \frac{\theta}{(k - \theta - 1)u^k + 1} \right] < 0,
\]

\[
\frac{d\text{RHS}}{du} = -k \left[ (k(1 + \theta) - \theta)u^{k-1} \left( 1 - \frac{1}{u} \right) + u^{k-2} + \frac{k\theta}{k - \theta - 1} \right] < 0.
\]

Hence, if the LHS declines faster with \( u \) than the RHS, then it must be that condition (B-34) holds. We obtain that \( |\frac{d\text{LHS}}{du}| > |\frac{d\text{RHS}}{du}| \) if

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

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 \( \frac{d\ln L}{dw} < 0 \).

From (B-3) we know that the wage bill is equal to \( W^{1+\psi} \). It is also the case that the wage bill equals the product of the average wage and total employment, \( \bar{w}L \). It follows that \( \bar{w} = \frac{W^{1+\psi}}{L} \), and therefore, \( \frac{d\ln \bar{w}}{dw} = (1 + \psi)\frac{d\ln W}{dw} - \frac{d\ln L}{dw} \). Using (B-32), it follows that

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

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

References
