DISCUSSION PAPER SERIES

DP17026

What is the optimal minimum wage?

Yujiang River Chen and Coen N Teulings

LABOUR ECONOMICS

CEPR

ISSN 0265-8003

What is the optimal minimum wage?

Yujiang River Chen and Coen N Teulings

Discussion Paper DP17026
Published 10 February 2022
Submitted 09 February 2022

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

What is the optimal minimum wage?

Abstract

The extensive literature on minimum wages has found evidence for the compression of relative wages and mixed results for employment. This literature has been plagued by a number of problems. First, the median-minimum wage ratio is used as the independent variable, where the median is endogenous. Second, it is difficult to disentangle compression of relative wages and truncation due to employment effects. Third, all effects are likely to depend on the initial level of the minimum. Fourth, employment effects are likely to differ between worker types. We offer solutions for these problems, by using instruments for the median, by using data on personal characteristics, and by using a flexible specification. We apply our method to US data starting from 1979, allowing for a wide variation in minimum wages. We find strong compression and positive employment effects for the lower half of the distribution, persisting for quite high levels of the minimum.

JEL Classification: J31, J58, R12

Keywords: Minimum Wages, US, employment effects, wage dispersion, Wage Share

Yujiang River Chen - yc349@cam.ac.uk
University of Cambridge

Coen N Teulings - c.n.teulings@outlook.com
Utrecht University and CEPR

What is the Optimal Minimum Wage?

Yujiang Chen and Coen Teulings*

February 2022

Abstract

The extensive literature on minimum wages has found evidence for the compression of relative wages and mixed results for employment. This literature has been plagued by a number of problems. First, the median-minimum wage ratio is used as the independent variable, where the median is endogenous. Second, it is difficult to disentangle compression of relative wages and truncation due to employment effects. Third, all effects are likely to depend on the initial level of the minimum. Fourth, employment effects are likely to differ between worker types. We offer solutions for these problems, by using instruments for the median, by using data on personal characteristics, and by using a flexible specification. We apply our method to US data starting from 1979, allowing for a wide variation in minimum wages. We find strong compression and positive employment effects for the lower half of the distribution, persisting for quite high levels of the minimum.

*Yujiang Chen, University of Cambridge, yc349@cam.ac.uk; Coen Teulings, Utrecht University, c.n.teulings@outlook.com.

1 Introduction

The extensive literature on the impact of minimum wages in the US typically reports strong compression of relative wages and a small employment effect. Many papers cannot rule out that an increase in the minimum wage will actually raise rather than reduce employment, see e.g. Card and Krueger (1994), Dube et al. (2010), Cengiz et al. (2019), Bailey et al. (2020) and Fishback and Seltzer (2020). Similar results have been reported for other countries, see Dolado et al. (1996) for Europe, Machin and Manning (1994) and Stewart (2012) for the UK, Ahlfeldt et al. (2018, 2022) for Germany and Engbom and Moser (2018) and Haanwinckel and Soares (2021) for Brazil.

Although this extensive literature has provided much insight, it has been plagued by a number of persistent problems, see e.g. Autor et al. (2016) and Neumark and Shirley (2021). First, most research uses the Kaitz index (the ratio of the minimum to the median wage), as a measure of the bindingness of the minimum wage. However, the median is endogenous, due to either truncation of workers with low human capital or compression of relative wages. Moreover, there might be reverse causality: some outside force may drive up both the median and wage dispersion in a region, e.g. the IT revolution in San Francisco. The rise in the median will then lead to a fall in the Kaitz index, which induces the researcher to conclude that a less binding minimum wage leads to higher wage dispersion. Second, disentangling truncation and compression effects is impossible when using data on wages only without making strong functional form assumptions. Third, all effects are likely to depend on the initial level of the minimum wage. For example, some papers found positive employment effects for low initial levels of the minimum wage, but these effects are unlikely to persist for higher levels. Finally, employment effects are likely to vary between workers with different levels of human capital. This paper addresses these problems.

A short review of the literature on minimum wages over the past 40 years is helpful for understanding our approach. Our review starts with Meyer and Wise (1983)'s analysis of the truncation effect of the minimum on the wage distribution, using data on wages only. They hypothesize that the minimum truncates an otherwise invariant wage distribution. The truncated lower tail can be split into three parts. For the first part, the wage is raised to the statutory minimum, yielding a spike in the wage distribution. For the second part, there is non-compliance: workers get paid below the statutory minimum. The third part measures the disemployment effect. Meyer and Wise report a substantial loss of employment.

Subsequent research by Card and Krueger (1994), based on time series evidence and a difference-in-difference approach between New Jersey and Pennsylvania found much smaller or even positive employment effects. This finding initiated a flurry of papers, at one hand disputing the empirical validity of these claims and at the other hand explaining the combination of small or even negative truncation effects and strong wage compression.

A first strand of papers use monopsony models and models with search frictions, notably Bontemps et al. (2000), Machin et al. (2003), Flinn (2006) and Engbom and Moser (2018). These models can explain why an increase in the minimum wage might raise rather than reduce employment. Models with search frictions predict that job seekers under-invest in search when the Hosios condition is violated by job seekers capturing too small a share of the match surplus. An outside legal intervention in wage setting can alleviate this hold up problem.

A second strand followed up on an idea first expressed by Rosen (1974) to apply hedonic pricing models for the analysis of minimum quality standards. In a Walrasian market, a minimum wage is akin to a minimum quality standard for human capital. Contributions in this strand are Teulings (1995, 2000) and Haanwinckel and Soares (2021). The idea is that small disemployment effects of minimum wages can go hand in hand with strong wage-compression effects by a chain of substitution effects, driving up the wages of those workers that are the closest substitutes for workers whose human capital falls below the implicit minimum quality standard imposed by the minimum wage and who therefore lose their job.

Motivated by this research, DiNardo et al. (1996), Lee (1999), and Teulings (2003) seek to explain the negative effect of the minimum on wage dispersion, not from truncation of workers with low human capital, as in Meyer and Wise (1983), but from compression of relative wages. In particular Lee (1999) and Teulings (2003) found that a minimum wage generates strong compression of wage-differentials above the minimum. Lee (1999)'s paper was probably the first to use interstate minimum wage differentials as a source of variation, allowing to control for time and region fixed effects. However, similar to Meyer and Wise (1983), he uses data on wages only. Hence, Lee has essentially assumed that the fall in wage dispersion is due to compression and not truncation. He concludes that the full increase in wage inequality during the eighties can be attributed to the freeze of the nominal minimum wage during the Reagan presidency. However, the problem in his analysis is that minimum wages were found to compress wages differentials not only in the bottom but also in the upper half of the distribution. This is implausible. Teulings (2003) also found strong compression, but mainly for the lower half of the distribution. He uses also data on workers' human capital, allowing separate inference on truncation (i.e. changing the distribution of human capital) versus compression (changing the wage distribution for a constant distribution of human capital).

Both Lee (1999) and Teulings (2003) used the Kaitz index as their independent variable. Autor et al. (2016) have argued that this procedure is suspect when e.g. 50-10% log wage differential serves as the endogenous variable. The median enters both the explanatory and endogenous variables. Hence, measurement error in the median introduces an artificial correlation, biasing the estimation results. Moreover, cities tend to have both a higher median and a larger wage dispersion than rural areas. Again, this creates an artificial correlation biasing the estimation results. Correcting for these biases, Autor et al. (2016) did not find significant evidence for compression effects above the spike, thereby confirming the initial assumption of Meyer and Wise (1983). Though one might dispute the validity of their instruments, their analysis shows the problem of using the Kaitz index as the explanatory variable.

Neumark and Shirley (2021) question the general view that the disemployment effects are small or even negative. They argue that there is clear evidence for negative employment effects for subgroups of low human capital workers. Their argument suggests that there is strong heterogeneity in the impact of minimum wages on employment, not only by the level of the minimum wage, but also between subgroups of workers.

Cengiz et al. (2019) try to correct the Meyer and Wise (1983) model for the compression effects above the minimum reported by Lee (1999) and Teulings (2003). Using data on wages only, they start from the Meyer and Wise (1983) estimate of the disemployment effect. From this estimate they subtract the added probability mass for wage levels slightly above the minimum, arguing that these workers earned less than minimum wage before its increase and should therefore not be included in Meyer and Wise (1983) estimate of the disemployment effect. They found the disemployment effect to be small. We argue that the magnitude of the employment effect can be estimated from wage data only by making strong functional form assumptions.

From this short history of the research on minimum wages, the main elements of our design are easy to understand. We tackle the problem of the endogeneity of the median by using the spike in the wage distribution rather than the Kaitz index as our explanatory variable. We interpret the spike as the objective of the policymaker and the nominal minimum wage as the instrument for implementing this objective. The spike will be determined by the real minimum wage. Hence, it depends both on the nominal minimum wage and on the counterfactual evolution of nominal wages in general. Apart from region and time fixed effects, we use a Bartik instrument for agglomeration externalities derived from a companion paper, Chen and Teulings (2021), to instrument for the counterfactual wage. Moreover, we address the systematic differences in the wage distribution between cities and the countryside by treating 34 SMSA's as separate regions. We disentangle the truncation and the compression effect by aggregating a vector of personal characteristics into a single index for workers' human capital, following Teulings (2003), which we use for the analysis of both relative wages and employment. We solve the problem of heterogeneity in employment effects by analyzing the effect on employment for each percentile of the human capital distribution and by allowing this effect to be non-linear in the spike. Our specification does not make a priori choices regarding the sign of employment effects for each quantile of the distribution. Moreover, it allows for a sign reversal when the spike exceeds some critical threshold. We use data starting from 1979, when the spike accounted for 5% of total employment for the country as a whole and even 10% in some low wage regions, to allow for sufficient variation in the spike for reliably establishing this turning point. These non-linearities allow us to quantify trade-offs legislators face when setting the minimum wage.

We find strong evidence both for the compression of wage differentials above the spike and for heterogeneous employment effects. The return to human capital for the median worker is 11% lower when the spike is 10% compared to the case without a spike; it is even 30% lower for a worker earning a wage just above the minimum. Total employment is maximized by a spike of 10%. The additional employment of this spike relative to a situation without a minimum wage is 14% of total employment, with a strong shift of the distribution of employment towards low skilled workers. Employment for workers with higher levels of human capital is even negatively affected by an increase in the minimum. These effects are precisely measured. This strong positive employment effect demonstrates the relevance of monopsony models above the hedonic pricing models.

We provide counterfactual simulations for several turning points in the evolution of the minimum. We find that the changes in minimum wages have contributed substantially to the variation in the return to human capital and wage dispersion in the bottom half of the distribution in particular in 1980s and that an increase in the minimum wage might be an effective instrument for boosting the labor share in aggregate output.

The structure of the paper is as follows. Section 2 discusses the compression and the truncation effects in hedonic models. These concepts guide our empirical approach, which is set out in Section 3. The data are discussed in Section 4, while the empirical specification and the estimation results are presented in Section 5. Section 6 contains the counterfactual analysis. Section 7 offers some concluding reflections on our results.

2 Some theoretical considerations

2.1 Walrasian models

Modelling spillover effects of a minimum wage to the wages above the minimum is not straightforward. Rosen (1974) was probably the first to observe that hedonic pricing models with heterogeneity on both sides of the market are a prerequisite for assessing the impact of minimum quality standards; a minimum wage is akin to a minimum quality standard for labor. Sattinger (1975) and Teulings (1995, 2005) analyzed equilibrium assignment models of heterogeneous workers to heterogeneous jobs, where the heterogeneity on each side of the market is captured by a *single index*, say, the worker's human capital h on the supply side and job-complexity z on the demand side.[1] Gabaix and Landier (2008)'s model of CEO compensation has the same structure. This section shows why an increase in the minimum yields wage compression in these models.

Returns to scale are constant in these models. Let $x\left(h,z\right)$ be the log productivity of a worker with human capital h in a job with complexity z; $x\left(h,z\right)$ is twice differentiable in both arguments. Human capital is assumed to have both an absolute advantage in all job-types, implying $x_h\left(h,z\right)>0$ (more human capital yields more input, irrespective of the job type), and a comparative advantage in more complex jobs, implying $x_{hz}\left(h,z\right)>0$ (log supermodularity: more human capital yields relatively more additional output in more complex jobs). The distributions of the supply of human capital and the demand for product complexity are exogenous; both distribution functions are assumed to be twice differentiable. For the sake of the argument, we ignore other factors of production. Finally, there is perfect competition in all markets.

Under these assumptions, absolute advantage implies that the equilibrium log wage function $w\left(h\right)$ is differentiable and strictly increasing, $w'\left(h\right)>0$, while comparative advantage implies that the equilibrium assignment of worker- to job-types, $z\left(h\right)$, is also a differentiable and strictly increasing function, $z'\left(h\right)>0$, see Teulings (2005) for a proof.

Since types of labor are the only factors of product and since there is perfect competition, wages account for the full value of output and profits are zero. Hence, profit maximization is equivalent to cost minimization. Since production is characterized by constant return to scale, cost minimization for a given quantity is equivalent to cost minimization per unit of output. Let $h\left(z\right)$ be the type of

[1] A single index is not the same as a single factor of production. In fact, single-index models have an infinite number of factors of production, since each value of the index corresponds to a different factor of production. The elasticity of substitution between two factors is a decreasing function of the distance between these factors measured along the index: *DIDES: Distance Dependent Elasticity of Substitution*, see Teulings (2005).

worker hired by an employer offering a job of complexity z; this function is therefore the inverse of $z(h)$, which exists since $z'(h) > 0$. The employer chooses the optimal level of human capital $h(z)$ as to minimize cost per unit of output

$$h(z) = \arg\min_h \left[\mathrm{e}^{w(h)-x(h,z)} \right],$$
$$w'(h) = x_h[h, z(h)], \tag{1}$$

where the second line is the first-order condition of the program in the first line, substituting z for $z(h)$ and hence $h(z)$ for $h[z(h)] = h$. This is a fundamental insight in this class of models: keeping constant the level of human capital h, the slope of the log wage function $w'(h)$(or equivalently: the Mincerian return to human capital) is an increasing function of the complexity $z(h)$ of the job which an h-type worker holds in equilibrium. The same result applies in Gabaix and Landier (2008)'s model of CEO pay, where the return to a CEO's talent is proportional to the size of his firm. Keeping constant his managerial talent, the larger are firms, the steeper is the CEO compensation curve.

The impact of a minimum wage is conveniently demonstrated by a simple parameterization of this model: human capital and jobs are uniformly distributed at the unit interval, $h \in [0, 1]$ and $z \in [0, 1]$, and the productivity function satisfies

$$x(h, z) = -\frac{1}{\gamma}\mathrm{e}^{\gamma(z-h)},$$

with $\gamma > 0$, satisfying the previous assumptions of absolute and comparative advantage. By equation (1) the first-order condition reads

$$w'(h) = \mathrm{e}^{\gamma[z(h)-h]}. \tag{2}$$

For $\gamma = 0$, relative wages are independent of the minimum wage. This is the case where worker types are perfect substitutes.

Since both h and z are distributed uniformly, the equilibrium assignment is $z(h) = h$. By equation (2), this implies $w'(h) = 1$. The situation is portrayed in Figure 1. The lower panel shows the assignment $z(h)$ of worker- to job-types, while upper panel shows the wage function $w(h)$; the red continuous lines depict the case without a minimum wage. All job-types are done and all worker-types are employed, where the worker with the least human capital is assigned to the simplest job, $z(0) = 0$, and the worker with the highest human capital to the most difficult job, $z(1) = 1$. We choose the numeraire of the model such that $w(0) = 0$, so that $w(h) = h$.

Consider the effect of log minimum wage m; since $w(0) = 0$, $m > 0$ for this minimum to be binding. The minimum wage causes the least skilled workers to lose their job. Let $\widetilde{h}$ be the worker with the lowest human capital who remains employed. Since h is distributed uniformly on the unit interval, $\widetilde{h}$ is the disemployment effect of the minimum wage as a fraction of total employment. $\widetilde{h}$ measures the *truncation effect* of a minimum wage on the wage distribution: wage dispersion is reduced since the workers with the least human capital lose their job.

Let $\widetilde{w}(h)$ and $\widetilde{z}(h)$ denote the wage function and the equilibrium allocation for this minimum

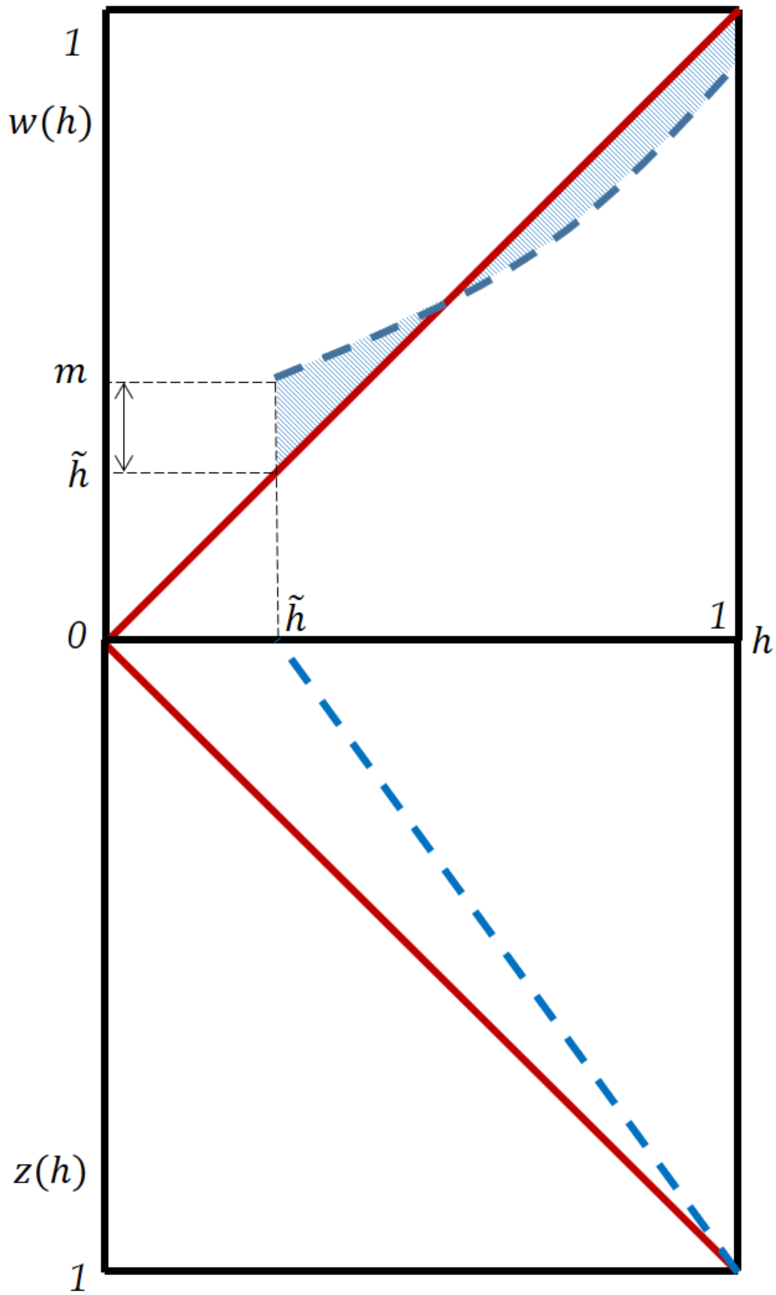

Figure 1: Equilibrium Assignment $z(h)$ and Wages $w(h)$

wage, the blue dashed lines in Figure 1. Both functions are also portrayed in Figure 1 as the dotted blue lines. Since all jobs have to be done and since $\widetilde{z}'(h) > 0$, the equilibrium assignment $\widetilde{z}(h)$ starts from $z\left(\widetilde{h}\right) = 0$ and ends at $z(1) = 1$. All worker-types who remain employed hold jobs that are less complex after rather than before the introduction of the minimum wage, $\widetilde{z}(h) \leq z(h)$, where equality holds only for the highest type, $h = 1$; the lower h, the larger $z(h) - \widetilde{z}(h)$. Summarizing our results till sofar: a minimum wage causes the least skilled workers ($h < \widetilde{h}$) to lose their job. Since these jobs have to be done anyway, all other workers move to less complex jobs than before the introduction of the minimum wage (except the worker with $h = 1$, who remains doing the most complex job $\widetilde{z}(1) = z(1) = 1$).

Since $\widetilde{w}(h)$ is differentiable and increasing, the wage of the least skilled worker who remains employed is equal to the minimum wage:

$$\widetilde{w}\left(\widetilde{h}\right) = m.$$

Due to equation (2), the decrease in job complexity for all workers (except for $h = 1$) implies that the slope of the wage function declines: $\widetilde{w}'\left(\widetilde{h}\right) \leq w'(h) = 1$, where equality holds for $h = 1$ only. This flattening is the strongest for least skilled worker who remains employed, $\widetilde{h}$. It gradually declines for higher levels of human capital. At the upper support of the human capital distribution, $h = 1$, the wage function before and after the introduction of the minimum wage run parallel, since $\widetilde{z}(1) = z(1)$, and hence $\widetilde{w}'(1) = w'(1)$. Since our example is constructed such that $w(h)$ is linear in the absence of a minimum wage (and hence $w''(h) = 0$), $\widetilde{w}''(h) \geq 0$ in the presence of a minimum wage (with equality holding only for $h = 1$).

These arguments establish the slope of $\widetilde{w}(h)$, but not its level. The latter follows from a Walrasian argument, see Teulings (2005) for the proof. Consider a marginal increase in the minimum wage, forcing workers with $h < \widetilde{h}$ out of employment. Since wages are equal to the value of the marginal product, the fall in aggregate output due to this increase in the minimum must be equal to the wage sum of the workers who lose their job. Hence, up to a term of order $O\left(m^{-2}\right)$, the sum of wages for all workers who remain employed must be equal before and after the introduction of the minimum wage:

$$\int_{\widetilde{h}}^{1} \mathrm{e}^{\widetilde{w}(h)} - \mathrm{e}^{w(h)} \mathrm{d}h = 0, \tag{3}$$

Since the wage function flattens, equation (3) implies that the wages at the bottom go up, while these at the top go down. This is the *compression effect* of a minimum wage. Roughly stated, the two blue shaded areas in Figure 1 between the functions $\widetilde{w}(h)$ and $w(h)$ must have equal surface.[2] We shall use equation (3) to calculate the aggregate shift in value-added between labor and other factors of production due to an increase in the minimum wage.

As equation (2) shows, this compression effect is proportional to γ. For $\gamma = 0$ (perfect substitu-

[2] The equality of the surfaces of the two blue shaded areas requires

$$\int_{\widetilde{h}}^{1} \widetilde{w}(h) - w(h) \,\mathrm{d}h = 0.$$

This condition is not exactly the same as equation (3), which uses the wage level $e^{w(h)}$ than its log $w(h)$.

tion of worker types), the return to human capital is independent of workers' assignment to jobs $z(h)$. Hence, there is no compression effect in that case.

Summarizing the conclusions of this analysis. For $\gamma > 0$, the introduction of a minimum wage has three effects:

1. *truncation*: workers with the least human capital, $h < \widetilde{h}$, lose their job;
2. *compression*: the return to human capital $w'(h)$ falls for all levels of human capital, except for the highest, the more so the lower h; hence, $w''(h)$ rises;
3. the sum of wages for workers who remain employed remains constant; hence, workers with h slightly above the disemployment threshold $\widetilde{h}$ gain by the introduction of a minimum wage, while workers at the top of the human capital distribution lose.

Note, however, that the compression effect in this model is driven by the truncation effect: workers take less complex jobs after the introduction of a minimum wage since the least skilled workers are truncated from the employment distribution. Without a truncation, this model does not generate compression. Note furthermore that this model does not predict a spike in the wage distribution.

2.2 Identification from only the wage distribution

Cengiz et al. (2019) seek to establish the disemployment effect of minimum wages just from the shape of the wage distribution. They set out to establish the additional probability mass in the wage distribution just above the minimum due to the compression effect and compare that additional mass to the truncated mass below the minimum. The disemployment effect is measured as the difference between this truncated mass minus the additional mass just above the minimum (workers previously employed below the minimum who found employment in jobs that pay above the minimum after its increase). This section applies the framework developed in the previous section to argue that this method works only under strong functional form assumptions.

Empirically, the distribution of human capital is bell shaped (like the normal distribution) rather than uniform. Consider an increase in the minimum wage. Does an increase in the fatness of the lower tail of the wage distribution provide evidence in favour of either truncation or compression? The answer is: hard to tell. Let $w(h, m)$ denote the equilibrium log wage as a function of the human capital h of the worker and the applicable log minimum wage m. It is convenient to construct the index h such that for a particular level of the log minimum wage m^o, the function is linear in h: $w_{hh}(h, m^o) = 0$ for all h (the subscript refers to the relevant partial derivative). This normalization is not essential, but makes the subsequent argument more easy to follow.[3] Let $h(m)$ be lowest level of human capital that is employed for that level of the log minimum m (the equivalent of $\widetilde{h}$ from the previous section). Like in the previous section, we assume that, apart from the truncation at $h(m)$, the supply of human capital $f(h, m)$ is exogenously fixed (labor supply

[3] Following the argument in the previous section, $w(h, m)$ is strictly increasing in h for any m. Hence, an index h for which $w_{hh}(h, m^o) = 0$ does always exist, by replacing the original human capital index h^* by a transformed index $h = w(h^*, m^o)$. This transformation can be applied without changing the empirical content of the model.

is fully price-inelastic). Except for the renormalisation due to this truncation, the density of the human capital distribution among employment is therefore invariant to changes in the minimum wage

$$f(h,m)=\frac{f(h)}{1-F[h(m)]} \text{ for } h\geq h(m)$$

where $f(h)$ denotes the untruncated distribution of human capital that applies in the absence of a minimum wage. Suppose that we start from a log minimum wage m^o. The density function of the log wage distribution conditional on a log minimum wage m, denoted $g(w,m)$, is equal to the density of human capital distribution $f(h,m)$ times the Jacobian $\mathrm{d}h/\mathrm{d}w=1/w_h(h,m)$:

$$g[w(h),m]=f(h,m)/w_h(h,m).$$

The relative change in the density function of log wages at the minimum is therefore equal to

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}m}[\log g(m,m)]_{|m=m^o} &= -\frac{w_{hm}[h(m^o),m^o]}{w_h[h(m^o),m^o]} \\ &\quad+\left(\frac{f'[h(m^o)]}{f[h(m^o)]}+\frac{f[h(m^o)]}{1-F[h(m^o)]}\right)h'(m^o), \end{aligned} \tag{4}$$

using $w_{hh}(h,m^o)=0$. The first term is the compression effect: the flattening of the wage function. This term is positive, since $w_{hm}(h,m)<0$ for $\gamma>0$. The second term is the truncation effect: the change in the lower support of the human capital distribution among employment. Since $f'(h)$ is positive in the left tail of a bell shaped distribution, this term is positive, too. If γ is high, so that there is strong compression and little truncation, then the first term dominates. In the reverse case, the second term dominates. The change in the fatness of the left tail of the wage distribution is therefore an uninformative statistic regarding the relative size of compression versus truncation.

This argument sketches the outline of a solution to this problem in the approach of Cengiz et al. (2019). When information on workers' human capital is available, we can disentangle changes in the return to human capital on the one hand and changes in its distribution on the other hand. This strategy will be pursued.

3 Empirical specification

We consider a world that consists of multiple regions r which are observed at multiple points in time t; the index s refers to a combination of region r and time t; we refer to each s as an economy. Each worker i is located in a single economy s, so the index i uniquely identifies the economy s in which i lives.

Similar to the model in Section 2, the human capital of worker i can be summarized in a *single index* h_i. This index has infinite support on the real domain. The wage return to this index may vary across economies, but the way in which various components of workers' human capital (like experience and years of education) are aggregated into this single index h_i is invariant across economies. However, the index h_i is observed only partially:

$$h_i = g_i + \varepsilon_i, \tag{5}$$

where g_i is the observable part of the human capital index h_i and $\varepsilon_i \sim N\left(0, \sigma_s^2\right)$ is the unobservable component which is orthogonal to g_i. Note that we allow the variance σ_s^2 of ε_i to vary between economies to allow for changes in the role of unobservable components of h_i.

Let m_s be the log minimum wage in economy s and let $w_s(h, m)$ be the log nominal wage for a worker with human capital h in economy s when the log minimum wage is equal to m. Hence, $w_s(h) \equiv w_s(h, m_s)$ is the log wage function evaluated at the actual log minimum wage m_s in economy s, is the function that generates our data on log wages w_i. Note that we allow the wage function $w_s(h, m)$ to vary between economies for other reasons than the minimum wage. In line with the analysis Section 2, we assume that this function is twice differentiable and strictly increasing in h everywhere, except for the spike at the minimum wage, where the function is flat. Let $h_s(m)$ be the upper support of the spike in economy s as a function of the log minimum wage: the highest level of h that does not earn more than the minimum wage. At $h_s(m)$, $w_s[h_s(m), m]$ is continuous but non-differentiable; the function is flat to the left of $h_s(m)$; it is increasing to its right. $h_s \equiv h_s(m_s)$ is the upper support of the spike in our data. Hence:

$$\begin{aligned} w_s[h_s(m), m] &= m, \\ w_s(h, m) &> m, \qquad \forall h > h_s. \end{aligned}$$

The Meyer and Wise (1983) model is the special case of this model where $w_s(h, m)$ does not depend on m for all $h > h_s(m)$.

The research question we address is how $w_s(h, m)$ and the density function of human capital among employment depends on the minimum wage. The standard approach has been to use the Kaitz index (ratio of the median to the minimum wage) as an index for the bindingness of the minimum wage. The problem of this approach is that the median wage itself is endogenous, as it is potentially affected by truncation and compression. Even more problematic, there are systematic differences between economies in their wage distribution unrelated to the minimum wage. In particular, both the mean and dispersion of the wage distribution tend to be higher in cities, e.g. due to agglomeration externalities. This yields a negative correlation between the Kaitz index and wage dispersion, which is unrelated to the minimum wage. We, therefore, do not use the Kaitz index, but the spike as a measure of the bindingness of the minimum wage. We view the minimum wage as an instrument of the policymaker to manipulate the level of the spike. Hence, our first stage regression analyses the effect of the minimum wage on the spike, while our second stage regressions analyze the effect of this instrument for the spike on wages and employment.

Following this argument, we apply a 5-step estimation procedure:

1. The construction of the index g_i for observed human capital for each individual i;

2. The estimation of the upper support of the spike (in terms of the human capital index) for each economy s;

3. The first-stage regression for the spike;
4. The second-stage regression for the effect of the spike on the wage function $w_s(h, m)$;
5. The second-stage regression for the effect of the spike on the distribution of observed human capital among employment.

These steps will be elaborated in the subsequent subsections. Before doing so, we provide a discussion of the data.

4 Data

We draw data from the Current Population Survey, Merged Outgoing Rotation Groups (CPS-MORG) from 1979 till 2019. We use the hourly wage, years of education, occupation, industry and other demography information as gender, age, marital status, and race. Our sample includes all workers aged between 16 and 64.

For our classifications of regions, we first select 34 Metropolitan Statistical Areas (MSAs). We then take the remaining part of each state as one non-city region. The definition of MSAs changes over time. To make the samples consistent, we match different IDs of these areas over time. From 1979 to 1985, we use the 1970 Census ranking to identify MSAs. From 1986 to 1988, we use CMSA and PMSA identifiers. From 1989 to 2003, we use MSAFIPS and for the rest of the samples we use CBSAFIPS. Out of the total sample of 2,099,847 observations, 36.2% lives in MSAs. We have 47 Non-MSA state regions: as is common practice, we exclude Hawaii and Alaska. Furthermore, we split New Jersey from NY-NJ MSA, and exclude Washington DC, leaving us with 34 MSAs and 47 non-city regions, 81 regions in total. The full list of MSAs is in Appendix Table A1. We use the industry definition by Autor et al. (2003) and the crosswalk constructed by IPUMS.

Let q_s be the spike in the wage distribution at the minimum wage. We operationalize the definition of the spike by including all workers whose log wage is equal to m_s plus or minus $.01$ (that is 1% above or below the minimum). The details of the construction of the spike are in the Appendix, while the summary statistics table is in Table A2.

5 Estimation

5.1 Step 1: the human capital index

For the construction of the index g_i for observed human capital, we apply a second-order Taylor expansion to the function $w_i = w_s(h_i)$:[4]

$$\begin{aligned} w_i &= \omega_{0s} + \omega_{1s} h_i + \omega_{2s}\left(h_i^2 - \sigma_s^2\right) \\ &= \omega_{0s} + \omega_{1s} g_i + \omega_{2s} g_i^2 + \varepsilon_{wi}, \\ g_i &= \chi' x_i, \end{aligned} \tag{6}$$

[4] Strictly speaking, this specification violates our assumption that $w_s(h)$ is strictly increasing. In practice, the support of h is limited to the domain $[-2, 2]$, so that this will not be a problem as long as $|\omega_{2s}| < \omega_{1s}/4$.

where we substitute equation (5) for h_i in the second line and where x_i is the standard vector of observable personal characteristics like gender, marital status, age, race and education. This specification implies that we allow for a separate fixed effect ω_{0s}, a separate return to human capital ω_{1s} and a separate second-order effect ω_{2s} for each economy s, while the parameter vector χ is common to all economies. Using equation (5), we can work out the that the error term $\varepsilon_{w,i}$ satisfies

$$\varepsilon_{wi} = (\omega_{1s} + 2\omega_{2s} g_i)\, \varepsilon_i + \omega_{2s} \left(\varepsilon_i^2 - \sigma_s^2 \right)$$

and has zero mean. Empirically, the error term captures not only unobserved human capital, but also measurement error and the effect of search frictions. Angrist and Krueger (1991) show that the measurement error accounts for 30% of the variance in log wages, while Gottfries and Teulings (2021) show that search frictions account for another 10%. The actual interpretation of ε_i does not matter for our estimation results, since we are only interested in the parameter vector χ aggregating the components of x_i into a single index for observed human capital g_i.

Equation (6) allows full flexibility ω_{0s} and ω_{1s} across economies s. There are good reasons for this: the return to human capital has increased between 1979 and 2019, in particular for higher levels of human capital, see Autor and Dorn (2013). However, equation (6) allows too much flexibility. In fact, it is under-identified.[5] We therefore impose further structure:

$$\begin{aligned} \mathrm{E}(x_i) = 0 \quad &\Rightarrow \quad \mathrm{E}(g_i) = 0, \\ \mathrm{E}(\omega_{1s}) = 1, & \end{aligned} \tag{7}$$

where the expectations are taken over all individuals in the sample in the first line and over all economies in the second line.[6] These assumptions imply that human capital index h_i is scaled as such that the "*average*" worker in our sample (with $h_i = g_i = 0$) has a return to this index of unity "*on average*" across regions and over the time span of our sample. This choice is just a normalization facilitating the interpretation of our results. Furthermore, we impose additional structure on ω_{2s}

$$\omega_{2s} = \omega_x (t - \mathrm{E}t)\,, \tag{8}$$

implying that ω_{2s} is the same across regions, while the variation over time is restricted to a linear time trend with zero mean over the time span of our sample. Hence, we impose "*on average linearity*" of $w_s(h, m)$ in h for the construction of the human capital index g_i. Again, this is not really a restriction, but a convenient normalization, since we can include everything and its square in the vector x_i to capture any non-linearity in the relation between g_i and w_i to an arbitrary degree of precision. We do so in our empirical specification, by allowing for a number of well-known non-linearities, e.g. the experience profile and the interaction between years of education and experience and by using dummies for each value of years of education. Moreover, deviating slightly from our assumption that χ is constant across economies, we included cross effects of marital sta-

[5] We can apply a linear transformation to g_i, $g_i^* = \chi_0 + \chi_1 g_i$, that is observationally equivalent to equation (6) by an appropriate change in the parameters ω_{0s}, ω_{1s} and ω_{2s}:.$\omega_{0s}^* = \omega_{0s} - \omega_{1s}\chi_0 - \omega_{2s}\chi_0^2$, $\omega_{1s}^* = (\omega_{1s} - 2\omega_{2s}\chi_0)/\chi_1$ and $\omega_{2s}^* = \omega_{2s}/\chi_1^2$

[6] $\mathrm{E}(x_i) = 0$ implies that x_i cannot contain an intercept.

tus, gender, and a time trend to account for the changes in the attitude towards working women. Similarly, we account for the differential impact of being black in Southern states.

Equation (6)-(8) is a simple NLLS model. It can be estimated in an iterative way, by first estimating a standard OLS earnings with economy fixed effects

$$w_i = \omega_{0s} + \chi' x_i + \varepsilon_{wi}.$$

These first round estimate of χ can be used to construct g_i, which is then use to estimate ω_{0s}, ω_{1s} and ω_x by one OLS regression for all economies simultaneously

$$w_i = \omega_{0s} + \omega_{1s} g_i + \omega_x \left(t - \mathrm{E}t\right) g_i^2 + \varepsilon_{wi}, \tag{9}$$

These estimates for ω_{0s}, ω_{1s} and ω_x can be used to reestimate the vector χ, etc., until the procedure converges. The converged results are the maximum likelihood estimates of the NLLS model. In practice, we stop after the first two steps, since the value of χ obtained after the second iteration hardly differs from the first iteration.

We run this regression for a subsample of the economies with the lowest spikes. We have two reasons for doing so. First, a large spike implies that the wage function is flat for a substantial part of the wage distribution since all workers in the spike earn the same wage, disturbing the estimation of χ. Second, we have a mild preference for a human capital index h_i which is "*on average*" linear in the counterfactual economy, where the minimum wage is only mildly binding. We approximate this by omitting the economies with the highest spikes from the sample used for the estimation of χ. Note that this is only a matter of presentation and does not affect the validity of our method, since we will control for non-linearities in the relation between log wages and the human capital index later on. We have experimented a bit by omitting 10%, 20%, 30%, and even 40% of the economies with the highest spike. It does not matter much for our estimates of the parameters χ. Moving from 30% to 40% does not affect our estimates of χ in a significant way. We choose to use the estimation results for χ excluding 30% of the economies to calculate g_i for all individuals, also for those for 30% of the economies that have been excluded in the estimation of χ. Tables A3 and A4 shows the number states and years that are included in bottom 70% regarding spike for each year and state respectively. The coefficients χ are presented in Table A5 in Appendix.

The mean of the distribution of g_i in economy s evolves over time due to the rise in educational attainment and it differs between regions, due to the strong agglomeration of highly educated workers in metropolitan areas. For reasons discussed later on, it is convenient to define g_i relative to a proxy for its mean $\overline{g}_s$ for economy s. However, we do not want to demean by the sample mean for the same reasons as that we do not want to use the Kaitz index as a measure of the bindingness of the minimum wage: the sample mean is endogenous. Hence, we regress g_i on fixed time and region effects and an instrument, based on a companion paper on the size of regional agglomeration externalities, see Chen and Teulings (2021). In line with the evidence in Gennaioli et al. (2013), we show that higher educated agglomerate in particular regions. We use a Bartik instrument for $\overline{g}_s$, using the idea that its evolution is driven by its industry mix. We use nationwide changes in the level of human capital in each industry. We hypothesize that regions where industries with rising human capital are overrepresented experience a rise in $\overline{g}_s$.[7] We exclude the own region from the calculation of nationwide mean of these instruments, see Chen and Teulings (2021) for details. The coefficient of this Bartik instrument is 0.410 ($t = 12.62$), implying that when for the industry mix of a region, the nationwide mean of g increases by one unit, the regional mean is predicted to increase by 0.410. In the remainder of the paper, g_i refers to this regionally demeaned version of the index.

Figure 2 shows the overall distribution of g_i, controlled for the local mean in the way described above. This distribution has zero mean by construction, see equation (7), and its variance and standard deviation are 0.135 and 0.367 respectively. It is approximately normal, though there are some clear spikes, associated with spikes in the distribution of years of education at 12, 14, and 16 years.

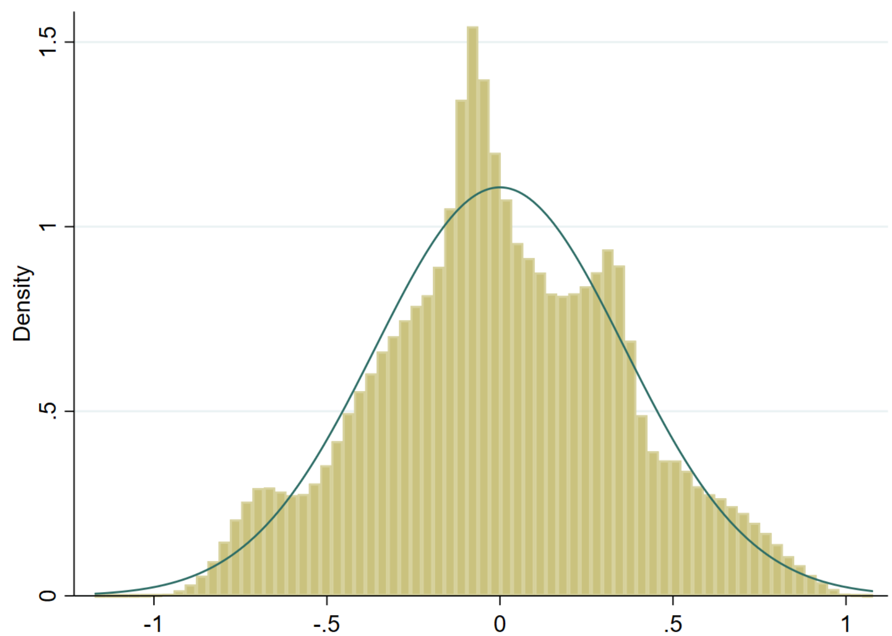

Figure 2: Histogram of g_i

5.2 Step 2: the upper support of the spike

The next step in our empirical strategy is to establish the uppersupport of the spike h_s for each economy. The challenge here is that we have to account for the fact that h_i is only partially observed. Equation (6) implies

$$w_i = w_s\left(h_i\right) = w_s\left(g_i + \varepsilon_i\right).$$

[7] Along the same lines, one can develop a second instrument, using nationwide changes in the industry mix rather than nationwide changes in $\overline{g}_s$ within industries. When entering both instruments in a first stage regression, this second instrument was insignificant, so we omit it from our estimation results.

Since $w_s(h_i) > m_s$ for $h_i > h_s$ and $w_s(h_i) \leq m_s$ for $h_i \leq h_s$, h_s can be estimated by means of a simple Probit model:

$$\Pr(w_i < m_s + .01|g_i) = \Pr(g_i + \varepsilon_i < h_s|g_i) = \Phi\left(\frac{h_s - g_i}{\sigma_s}\right), \quad (10)$$

where we add add $.01$ to m_s to be consistent with the definition of the spike. We can estimate this model separately for each economy. The Probit-parameter on g_i is an estimate of $-\sigma_s^{-1}$; the intercept is an estimate of h_s/σ_s.

For small values of q_s, the number of observations in the spike is low: an average economy has slightly less than 1500 observations on workers in our data. A spike of 1.5% is equivalent to roughly 25 workers. For a smaller number, the estimate of h_s becomes highly unreliable. We estimate the model for all economies with $q_s > 1.5\%$; 1414 economies meet this restriction. In all subsequent stages, we report both the results for the full sample of economies and for the restricted sample for which $q_s \geq 1.5\%$.

Since the distribution of g_i is approximately normal, see Figure 2, and using our assumption that the distribution of ε_i is normal, we can calculate the share of workers that earn less or equal to the minimum

$$\Pr(w_i < m_s + .01) = \Pr(h_i < h_s) = \Phi\left(\frac{h_s - \overline{g}_s}{\sigma_{hs}}\right), \quad (11)$$

$$\sigma_{hs}^2 = \sigma_s^2 + \sigma_{gs}^2,$$

where σ_{gs}^2 and σ_{hs}^2 are the variances of g_i and h_i respectively for economy s. The second line follows from equation (5). We use $\mathrm{E}(\varepsilon_i|s) = 0$ and hence $\mathrm{E}(h_i|s) \equiv \overline{g}_s$. The estimation results for equation (10) provide an estimate of σ_s, while $\overline{g}_s$ and σ_{gs} can be calculated from the data. Since σ_{hs}^2 is subject to substantial measurement error, we instrument σ_{hs} by region and time fixed effects; a hat on a variable denotes the instrument (the explained part of a first stage regression)

$$\sigma_{hs} = \widehat{\sigma}_{hs} + \varepsilon_{\sigma hs},$$

$$\widehat{\sigma}_{hs} = \sigma_{hr} + \sigma_{ht}.$$

where σ_{hr} and σ_{ht} denote vectors of region and time fixed effects.

There is convenient way to evaluate the estimates of h_s. Ignoring non-compliance (the part of the wage distribution below the minimum wage) for the sake of the argument, $\Pr(w_i < m_s + .01)$ is equal to the spike q_s. Define h_{qs} as the inverse of q_s with respect to the normal distribution:

$$h_{qs} = \widehat{\sigma}_{hs}\Phi^{-1}(q_s). \quad (12)$$

h_{qs} is the equivalent of h_s, but now derived form q_s rather than from the Probit model. Hence

$$h_{qs} \cong h_s - \overline{g}_s. \quad (13)$$

Equation (13) can be tested. Since $\mathrm{Var}(h_{qs}) = 0.128 > \mathrm{Var}(h_s) = 0.042$, the measurement error in h_{qs} is likely to be larger than in h_s. Hence, we run equation (13) as a regression taking h_{qs} as the endogenous variable. The estimation results for the restricted sample of economies are reported Table 1. Since h_{qs} is a non-linear transformation of q_s, we expect outliers among the error terms of this regression. We therefore estimate this equation by both OLS and a robust method Hamilton (1992), by performs a regression, calculates case weights from absolute residuals, and regresses again using those weights, allowing for non-normality of the error term.

Table 1: OLS and Robust Regression results for h_{qs}

	(1)	(2)
VARIABLES	h_{qs}	h_{qs}
h_s	1.014	1.019
	(59.45)	(56.89)
$\bar{g}_s$	0.320	0.322
	(2.04)	(1.96)
R-squared	0.721	0.703
RMSE	0.108	0.114
Regression	OLS	Robust
Observations	1,414	1,414
t-statistics in parentheses		

The coefficient for h_s is indeed equal to unity. The coefficient for $\overline{g}_s$ is positive, but much smaller unity. However, most of the true variation in $\overline{g}_s$ is already absorbed by the demeaning of g_i by means of region and time fixed effects and the Bartik instrument, see the discussion of Step 1. Hence, the share of measurement error in the remaining variance is substantial, compressing the regression coefficient. OLS or robust estimation makes little difference in this case. We conclude that our estimation results for h_s and the index h_{qs}, which derived from the data for q_s, are mutually consistent.

5.3 Step 3: the first stage regression for the spike

We use the previous analysis of the relation between the spike q_s and its upper support h_s as the benchmark for the construction of an instrument for the spike. Suppose that Meyer and Wise (1983)'s model had applied. Then, the wage function above the minimum wage would be independent of the minimum. Due to our normalization of the index g_i, it would be linear "*on average*" in w_i with a unit slope. The coefficient of a regression of h_{qs} on m_s would therefore be equal to unity in that case; h_{qs} is, therefore, a natural starting point for the construction of an instrument.

Three factors determine the evolution of the real regional minimum wage:

1. Increases in the federal nominal minimum wage; since the federal minimum is not binding in all states, changes in the federal minimum are not fully absorbed by the inclusion of time fixed effects;

2. Increases in the state nominal minimum wage;

3. The gradual increase in nominal wages reduces the real value of a fixed nominal minimum wage. Since the nominal minimum wage is adjusted at irregular intervals, this factor plays an important role in the variation of the minimum. In fact, the fall in the real minimum wage during the Reagan presidency was fully due to the nominal freeze of the federal minimum. To the extent that the evolution of real wages is the same across regions, it is fully absorbed by time dummies. However, we can expect interregional heterogeneity in the evolution of wages.

Our first stage regression allows for these three factors.

$$h_{qs} = \alpha_r + \alpha_t + \alpha_m m_s + \alpha_w w_{\text{IV}s} + \varepsilon_{qt}, \tag{14}$$

where α_r and α_t are region and time fixed effects, where m_s is the maximum of the federal and state minimum wage, and where $w_{\text{IV}s}$ is an instrument for the real wage. The first two factors listed above are captured by m_s. The third factor is captured by the region and time fixed effects and by the Bartik instrument discussed in Step 2.

Table 2: Instrumental Variable First Stage Regression

	(1)	(2)	(3)	(4)
VARIABLES	h_{qs}	h_{qs}	h_{qs}	h_{qs}
m_s	1.211	1.357	1.396	1.414
	(21.97)	(24.80)	(34.24)	(40.30)
Bartik IV	-4.040	-4.319	-2.576	-3.138
	(-7.01)	(-7.55)	(-3.60)	(-5.09)
Observations	1,414	1,414	3,321	3,321
R-squared	0.886	0.890	0.838	0.870
RMSE	0.0724	0.0719	0.147	0.126
Time Dummy	Y	Y	Y	Y
Region Dummy	Y	Y	Y	Y
Regression	OLS	Robust	OLS	Robust

t-statistics in parentheses

Table 2 presents the estimation results for equation (14). We present results for both the full and the restricted sample of economies [8]. Again, we present both OLS and robust estimation results. The instrument is strong. The regression results are very similar for all four regressions. The coefficient on m_s is between 1.2 and 1.4, where the estimation results for the full sample are in the upper part of that bracket. Had Meyer and Wise (1983)'s model had applied, the coefficient would have been equal to unity. The estimation results are in this order of magnitude. More remarkable is the coefficient on the Bartik instrument. The sign is in accordance with the theoretical predictions (higher nominal wages reduce the spike), but its magnitude is much higher. For demeaning of g_i, its coefficient was 0.410. Since the wage function has "*on average*" a unit slope in g_i, this effect is "*reinstalled*" by a coefficient of -0.410. The actual coefficient is 6 to 10 times larger. In our

[8] For some economies, the measured value of q_s is zero. Hence, h_{qs} cannot be calculated. For these economies, h_{qs} is set at the lowest value observed among other economies

companion paper Chen and Teulings (2021), we argue that this points to substantial agglomeration benefits. Since this result is discussed extensively in that paper, we do not discuss it here.

We use the results for the robust regression on the full sample and the inverse of equation (12) to calculate the instrument $\widehat{q}_s$

$$\widehat{q}_s = \Phi\left(\widehat{h}_{qs}/\widehat{\sigma}_s\right). \tag{15}$$

5.4 Step 4: the second stage regression for wages

A simple approach to the estimation of the effect of the spike on the wage function would be to regress individual log wages w_i on a second-order polynomial in g_i similar to equation (6) for all individuals earning more than the minimum wage, $w_i > m_s + 0.01$, and then to use simple regressions to analyze how the coefficients ω_{0s}, ω_{1s} and ω_{2s} depend on the instrument for the spike $\widehat{q}_s$. This approach fails, however, due to the selectivity at the lower bound: individuals in the sample are positively selected on earning a wage above the minimum wage. This problem could be resolved by estimating a non-linear Tobit model for all economies simultaneously. This approach is computationally infeasible, however, since there are several thousand economies in the full sample. Instead, we apply a 2-step procedure to correct for this selection bias similar to the classic Heckman 2-step method.

First, we regress log wages on a second-order polynomial in g_i as suggested above for each economy s, simply ignoring the selection bias problem:

$$w_i = \widetilde{w}_s(g_i) + \widetilde{\varepsilon}_{wi},$$
$$\widetilde{w}_s(g_i) = \widetilde{\omega}_{0s} + \widetilde{\omega}_{1s} g_i + \widetilde{\omega}_{2s} g_i^2.$$

$\widetilde{w}(g_i) >$E$(w_i|g_i)$ due to the selection bias. Let $\text{Bias}_s(g_i)$ be this selection bias in economy s as a function of the observed component of the human capital index g_i

$$\text{Bias}_s(g_i) \equiv \widetilde{w}_s(g_i) - w_s(g_i).$$

The bias at the upper support of the spike h_s can calculated using the fact that by construction $w_s(h_s) = m_s$. Hence

$$\text{Bias}_s(h_s) = \widetilde{w}_s(h_s) - w_s(h_s) = \widetilde{w}_s(h_s) - m_s. \tag{16}$$

Alternatively, $\text{Bias}_s(h_s)$ can be calculated from a first-order Taylor expansion of $w_s(h)$ around $h = h_s$:

$$\begin{aligned}\text{Bias}_s(h_s) &= \mathrm{E}\left[w_s(h_s+\varepsilon) - w_s(h_s)\,|\varepsilon > 0\right] \\ &= w_s'(h_s)\,\mathrm{E}\left[\varepsilon|\varepsilon > 0\right] + O\left(\mathrm{E}\left[\varepsilon^2|\varepsilon > h_s - g_i\right]\right) \\ &= \widehat{\sigma}_s w_s'(h_s)\frac{\phi(0)}{\Phi(0)} + O\left(\widehat{\sigma}_s^2\right),\end{aligned} \tag{17}$$

This Taylor expansion would hold exactly if $w_s(h)$ were linear in h. It is a reasonable approximation for small non-linearities in $w_s(h)$ and for small $\widehat{\sigma}_s$.

Equation (17) can be generalized to the calculation of $\mathrm{Bias}_s(g_i)$ for $g_i \neq h_s$:

$$\begin{aligned}\mathrm{Bias}_s(g_i) &= \mathrm{E}\left[w_s(g_i+\varepsilon) - w_s(g_i)\,|\varepsilon > h_s - g_i\right] \\ &\cong w_s'(h_s)\,\mathrm{E}\left[\varepsilon|\varepsilon > h_s - g_i\right] = \widehat{\sigma}_s w_s'(h_s)\,\frac{\phi\left[(g_i-h_s)/\widehat{\sigma}_s\right]}{\Phi\left[(g_i-h_s)/\widehat{\sigma}_s\right]} \\ &= \left[\widetilde{w}(h_s) - m_s\right]\frac{\Phi(0)}{\phi(0)}\frac{\phi\left[(g_i-h_s)/\widehat{\sigma}_s\right]}{\Phi\left[(g_i-h_s)/\widehat{\sigma}_s\right]},\end{aligned}$$

where we drop higher order terms and where the last step follows from combining equation (16) and (17). Similar to the second step of the Heckman 2-step method, we use the last expression to calculate

$$\widehat{w}_i = w_i - \mathrm{Bias}_s(g_i),$$

and then run the wage regression

$$\widehat{w}_i = \omega_{0s} + \omega_{1s} g_i + \omega_{2s} g_i^2 + \varepsilon_{wi}.$$

The estimates for ω_{0s}, ω_{1s} and ω_{2s} do not suffer from selection bias.

Next, we run a regression of the parameters ω_{0s} for each economy s on a polynomial in $\widehat{q}_s$, the Bartik instrument and fixed region and time effects, and the same for ω_{1s} and ω_{2s}. Again, we run this regression for both the full and the restricted sample of economies and we use both OLS and robust regression techniques. The estimation results are in Table 3.

The results for the full and the restricted sample are qualitatively similar, but the coefficients are larger when using the full sample. This is to be expected since the variance in the explanatory variable $\widehat{q}_s$ is reduced in the restricted sample. This yields a lower signal-noise ratio. Hence, we focus on the results for the full sample.

A comparison of Panel A and Panel B shows that the estimated coefficients are very similar in OLS and robust regression, but that they are estimated more precisely when accounting for outliers by using robust regression. We, therefore, focus on the latter results.

The difference between Panel B and C is that the former uses a second-order polynomial in $\widehat{q}_s$, while the latter uses $\widehat{q}_s$ and $\widehat{q}_s \times \ln \widehat{q}_s$ as explanatory variables. Since the results in Panel C provide a better fit, we use the latter for our counterfactual simulations in Section 6. However, the results for the second-order polynomial in Panel B are easier to interpret. The subsequent discussion, therefore, focuses on these results.

Recall that $\overline{g}_s$ (the mean value of g_i for economy s) is normalized to zero for each economy by means of region and time fixed effects and by the Bartik instrument (see Step 1). Hence, ω_{0s} is the log wage $w_s(g_i, m_s)$ for the worker with mean level of human capital in economy s (since $\overline{g}_s \cong 0$), ω_{1s} is the first derivative of this function for $g_i = 0$ (i.e. her return to human capital), while $2\omega_{2s}$ is the second derivative. The regression results imply therefore that for a low spike $\widehat{q}_s$, an increase in the spike raises the wage of the median worker, reduces the return to additional human capital for this worker, while it raises the second derivative. Due to the positive second derivative, the negative effect on the first derivative is lower for higher levels of g_i; it has disappeared for

Table 3: Regression with Bias_s corrected $\widehat{\omega}_{0s}$ $\widehat{\omega}_{1s}$ and $\widehat{\omega}_{2s}$

VARIABLES	(1) $\widehat{\omega}_{0s}$	(2) $\widehat{\omega}_{1s}$	(3) $\widehat{\omega}_{2s}$	(4) $\widehat{\omega}_{0s}$	(5) $\widehat{\omega}_{1s}$	(6) $\widehat{\omega}_{2s}$
Panel A: with 2SLS Regression						
$\widehat{q}_s$	2.333	-2.403	2.864	2.720	-4.834	5.268
	(4.16)	(-3.79)	(2.99)	(9.87)	(-13.82)	(10.99)
$\widehat{q}_s^2$	-4.753	28.30	-43.94	-7.524	46.35	-60.22
	(-0.96)	(5.06)	(-5.20)	(-2.73)	(13.23)	(-12.55)
Bartik IV	5.945	0.134	-6.216	4.071	0.224	-3.659
	(15.49)	(0.31)	(-9.48)	(17.49)	(0.76)	(-9.04)
R-squared	0.988	0.701	0.624	0.985	0.579	0.497
RMSE	0.0454	0.0514	0.0776	0.0464	0.0589	0.0807
Panel B: with Robust Regression						
$\widehat{q}_s$	2.506	-0.829	0.785	3.322	-4.914	5.297
	(5.75)	(-1.36)	(0.86)	(13.62)	(-14.21)	(10.94)
$\widehat{q}_s^2$	-6.710	16.18	-29.02	-16.68	53.46	-68.89
	(-1.84)	(3.17)	(-3.78)	(-6.89)	(15.57)	(-14.34)
Bartik IV	6.933	0.00568	-6.333	4.329	0.496	-3.910
	(23.09)	(0.01)	(-10.03)	(23.35)	(1.89)	(-10.63)
R-squared	0.993	0.731	0.663	0.991	0.676	0.575
RMSE	0.0366	0.0512	0.0769	0.0369	0.0524	0.0733
Panel C: with Robust Regression						
$\widehat{q}_s$	0.427	3.484	-6.510	-1.262	7.527	-9.283
	(0.85)	(4.96)	(-6.19)	(-4.29)	(18.22)	(-15.75)
$\widehat{q}_s \times \ln\widehat{q}_s$	-0.739	1.410	-2.284	-1.488	3.731	-4.111
	(-2.79)	(3.81)	(-4.11)	(-11.06)	(19.75)	(-15.25)
Bartik IV	6.987	-0.323	-5.784	4.889	-1.008	-2.255
	(22.86)	(-0.76)	(-9.03)	(26.21)	(-3.85)	(-6.03)
R-squared	0.993	0.731	0.665	0.991	0.689	0.579
RMSE	0.0366	0.0511	0.0767	0.0364	0.0511	0.0729
Time Dummy	Y	Y	Y	Y	Y	Y
Region Dummy	Y	Y	Y	Y	Y	Y
Observations	1,414	1,414	1,414	3,321	3,321	3,321

t-statistics in parentheses

$g_i = 0.46$,[9] which is one and a quarter standard deviations of g_i above its mean. This roughly fits the theoretical notion developed in Section 2.1 that the effect should be zero at the uppersupport of the distribution of g_i, where the actual and counterfactual wage functions $w_s(h, m)$ run parallel.

For a higher spike, the marginal effect of a further increase in the spike switches signs for all three variables; for the level ω_{0s}, this occurs at a spike of 10%; for the return to human capital at 5%; and for the second derivative at 4%. The wage of the median worker is 16% higher[10] for a spike of 10% rather than 0%, while the return to the human capital index g_i is 11% lower[11] for a spike of 5% rather than 0%. At the upper support of the spike, the compression of the return to human capital is even 30%. Summarizing: we find strong compression effects for wage levels above the minimum, in accordance with the theoretical model in Section 2, which persist even for quite high levels of the spike.

5.5 Step 5: the second stage regression for employment

The final step is to estimate the employment effect. We deal with this issue in two steps: first, we analyze the effect on aggregate employment, and second on the distribution of observed human capital among employed workers.

The results on the first step are reported in Table 4. We apply two specifications, one with a second-order polynomial in $\widehat{q}_s$ as explanatory variables, and another with $\widehat{q}_s$ and $\widehat{q}_s \times \ln \widehat{q}_s$. In both cases, we add the Bartik instrument as a control. Both specifications are estimated using both OLS and Robust methods. Since the second-order polynomial in $\widehat{q}_s$ and the specification with $\widehat{q}_s$ and $\widehat{q}_s \times \ln \widehat{q}_s$ yield an equal fit for both estimation methods and since the former is easier to interpret, we focus on the estimation results for the second-order polynomial. Since the Robust method outperforms OLS, we focus on the former. Hence, we take column (3) as a benchmark for the subsequent discussion. Aggregate employment is increasing in the spike, though at a declining rate. The turning point is at a spike of 10%, where employment is 14% higher than at a spike of zero;[12] 10% is about the highest value of the spike in our sample of economies, observed at the beginning of our sample period around 1980 in some Southern states.

The second step uses the log density function of human capital as the endogenous variable. Let g_{sp} be the p-quantile of the distribution of g_i among employment in economy s. We calculate g_{sp} for 100 percentiles ($p = .01, .02, .03...., 1.00$). The log density is calculated as $\ln(.01) - \ln(g_{sp} - g_{s,p-.01})$. As explanatory variables, we use a second-order polynomial in $\widehat{q}_s$ multiplied by a fourth-order polynomial in $p - 0.50$, where use a full set of fixed effects for each percentiles p to account for the general shape of the distribution. We present estimation results using both OLS and Robust methods. Again, the Robust method outperforms OLS.

As shown in Figure 2, the actual distribution of g_i is approximately normal, but with mass points associated with spikes in the distribution of years of education 12, 14, and 16 years of education. Since we have demeaned the values of g_i for each economy, using region and time fixed

[9] $4.9/(2 \times 5.3)$

[10] $\frac{1}{4}\ 3.3^2/16.5$

[11] $-\frac{1}{4}\ 4.9^2/53.3$. Note that the return to the human capital index g_i is normalized to unity "*on average*". Hence, this effect can be interpreted as a relative change in the return to human capital.

[12] The turning point is at $\frac{1}{2}2.6/12.6$ while the maximum employment effect is $\frac{1}{4}2.6^2/12.6$.

effects and the Bartik instrument, see Step 1, the location of these spikes differs between economies. Hence, fixed effects for each p do not fully absorb these spikes. To correct for this, we use full sets of fixed effects for each p both the original version of g_i (dummy p in the table) and its demeaned equivalent (dummy p_g in the table). We present estimation results both with and without the latter set of dummies. The additional set of dummies significantly improves the fit. Hence, we focus on the results including this additional set of dummies and using the Robust estimation method, see column 4 of Table A6. Standard errors are clustered at the economy level. The marginal significance of the fourth-order terms $\widehat{q}_s \times (p - 0.5)^4$ and $\widehat{q}_s^2 \times (p - 0.5)^4$ suggests that the fourth-order polynomial provides sufficient flexibility to cover the non-linearities in the relation between the spike and the log density for various point in the distribution.

Table 4: Regression with weighted total employment $\ln(N_s)$

	(1)	(2)	(3)	(4)
VARIABLES	$\ln(N_s)$	$\ln(N_s)$	$\ln(N_s)$	$\ln(N_s)$
$\widehat{q}_s$	1.153	4.593	2.647	0.156
	(1.14)	(3.43)	(4.30)	(0.19)
$\widehat{q}_s^2$	13.08		-12.64	
	(1.28)		(-2.03)	
$\widehat{q}_s \times \ln \widehat{q}_s$		0.897		-0.580
		(1.71)		(-1.83)
Bartik IV	6.790	6.825	3.889	4.267
	(7.78)	(7.69)	(7.35)	(7.95)
Time Dummy	Y	Y	Y	Y
Region Dummy	Y	Y	Y	Y
Regression	OLS	OLS	Robust	Robust
Observations	3,321	3,281	3,321	3,281
R-squared	0.948	0.948	0.979	0.979
RMSE	0.174	0.174	0.105	0.105

t-statistics in parentheses

The results are presented in Table A6 in the Appendix. They can are hard to assess. However, a graphical representation is easy to interpret, see Figure 3. They show the effect of $\widehat{q}_s$ (black), $\widehat{q}_s^2/10$ (red) and the Bartik instrument (green) as a function of p on the log density and the effect on log aggregate employment, see Table 4 (clearly, the latter does not depend on p), joint with their one standard deviation bands. The effect on log employment at percentile p is the sum of distribution and aggregate effect.

The first-order effect of $\widehat{q}_s$ is monotonically declining in p. The combined aggregate and distribution effect is positive for low initial levels of $\widehat{q}_s$ roughly until the median of the distribution, while it is declining for the upper half of the distribution. The second-order effect has the opposite sign, such that the maximum employment gain for low values of p is achieved for a spike of 10%. The switch of the sign of first- and second-order term of $\widehat{q}_s$ is almost at the same percentile, adding to the credibility of the estimation results. If there were a substantial difference in the point of crossing, the effect of $\widehat{q}_s$ would be convex rather than concave for the part of the support where the

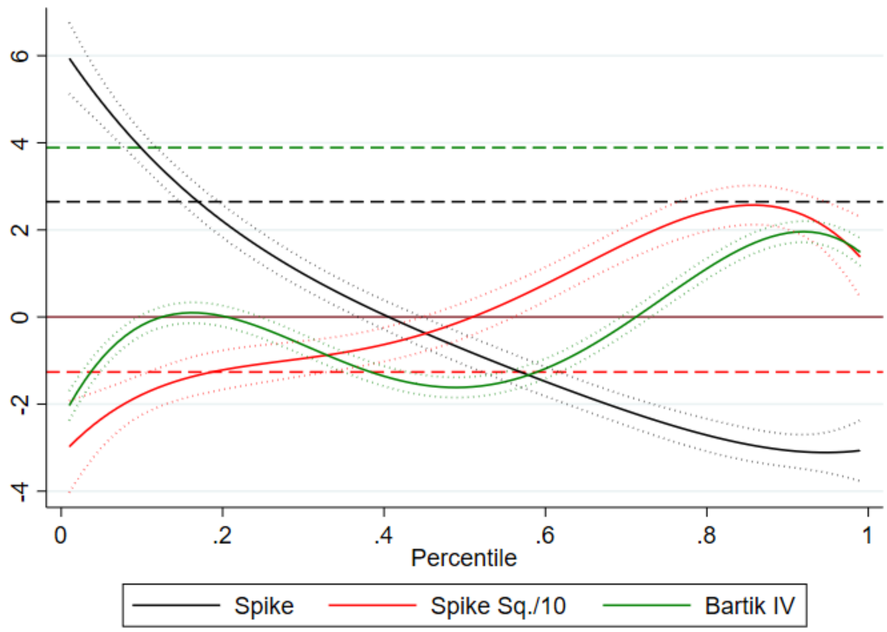

Figure 3: Estimated Marginal Employment Effect of q_s

first- and second-order term had the same sign.

The reported employment effects are large. One would like to have a plausibility test of this methodology. The results on the Bartik instrument provide this. The pattern for the Bartik instrument in Figure 3 shows two effects. First, there is polarization in the human capital distribution, raising employment in both tails of the distribution, as reported by Autor and Dorn (2013). This polarization can interpreted as a rising demand for personal services in economies with high employment of high paid high human capital workers who can afford to outsource these services. Second, the mean of the human capital distribution goes up by 0.662 for a unit increase in the instrument.[13] This number can be checked against the direct estimate of the effect of the instrument

[13]Let $g(p)$ be the inverse of the distribution function of g. Assuming this distribution to be approximately normal, we obtain $p = \Phi[g(p)/\sigma_g]$ or $g(p) = \sigma_g \Phi^{-1}(p)$. Since this relation applies identically for all p, its first derivative applies: $1 = g'(p)\,\phi[g(p)/\sigma_g]/\sigma_g$. Hence

$$\mathrm{E}[g] = \int_{-\infty}^{\infty} \frac{g}{\sigma_g}\phi\left(\frac{g}{\sigma_g}\right)\mathrm{d}g = \int_0^1 g(p)\,\frac{g'(p)}{\sigma_g}\phi\left(\frac{g(p)}{\sigma_g}\right)\mathrm{d}p = \int_0^1 \sigma_g \Phi^{-1}(p)\,\mathrm{d}g$$

Let $P(p)$ be the derivative of the polynomial reported in Table A6. Hence, using the approximation by the normal distribution, we obtain

$$P(p) = \frac{\mathrm{d}\log\phi[g(p)/\sigma_g]}{\mathrm{dBartik}}$$

Therefore

$$\begin{aligned}
\frac{\mathrm{dE}[g]}{\mathrm{dBartik}} &= \int_0^1 \sigma_g \Phi^{-1}(p)\,P(p)\,\mathrm{d}p \\
&= 0.367\int_0^1 \Phi^{-1}(p)\left(-4.4 + 2.1\left(p-\frac{1}{2}\right) + 96.5\left(p-\frac{1}{2}\right)^2 + 32.3\left(p-\frac{1}{2}\right)^3 - 338.3\left(p-\frac{1}{2}\right)^4\right)\mathrm{d}p \\
&= 0.662
\end{aligned}$$

on the mean 0.410, see the demeaning g_i for each economy s discussed in Step 1. These numbers are roughly similar, lending credibility to our methodology and hence to the estimated positive employment effects of an increase in the spike.

One might worry about the interaction between the effects of the Bartik instrument and the spike. The Bartik instrument increases both the mean and the dispersion of the human capital distribution, the former due to agglomeration externalities of high skilled workers and the latter due to the induced polarization of the distribution. The increase in the mean makes the minimum wage less binding and reduces therefore the spike, see Table 2. The estimation results in Figure 3 suggest that a lower spike reduces employment at the bottom, partly undoing the polarization-effect of an increase in the Bartik instrument. Might the inclusion of the Bartik instrument in some way artificially generate positive employment effect of an increase in the spike? To check this, we rerun all regression without the Bartik instrument, see Appendix. The results are basically the same, except for the turning point for the sign of the employment effect of an increase in the spike from positive to negative, which is now not at a spike of 10%, but even of 20%. Our results are therefore not driven by the inclusion of the Bartik instrument in our regressions.

The observed pattern is consistent with a monopsony model, where higher minimum wages attract a larger supply of labor at the lower end of the human capital distribution, but which reduces employment at the upper end of the distribution. For higher levels of the spike ($> 10\%$), the adverse demand effect of higher wages stops the increase in employment in the lower tail of the human capital distribution, while the negative effect on demand and supply reduces employment in the upper tail.

6 Counterfactuals

Our empirical results can be used for an analysis of the impact of the changes in the spike on the distribution of wages and employment over the past forty years. We consider the average impact across all regions in 1980, 1991, 1998, 2004, 2010 and 2019, taking 1980 as the point of reference. [14] These years are chosen since they mark turning point in the policy regarding minimum wages, see Table A2 in the Appendix. The counterfactuals raise the spike in all regions by the difference between nation-wide mean with that in 1980, so that the counterfactual nation-wide mean is equal to the actual mean in 1980.[15] The details of the procedure for the calculation of the counterfactuals

[14]We can calculate counterfactuals for a higher, but not for a lower spike, because for a lower spike there is no one-to-one correspondence of the actual and counterfactual wage for workers in the spike, since part of these workers may remain in the spike, while others will earn more than minimum wage after its reduction.

In the Appendix we present counterfactuals taking 2010 (the year with the second highest spike) as the point of reference, excluding the year with the highest spike.

[15]We adjust the share q_s^- of workers earning less than the minimum using a simple regression

$$q_s^- = \beta_0 + \beta q_s + \varepsilon_q.$$

We find $\beta = 0.52$ ($t = 34.8$). Hence

$$\begin{aligned} q_{s,\text{counterfactual}} &= q_s + \overline{q}_{1980} - \overline{q}_t, \\ q^-_{s,\text{counterfactual}} &= q_s^- + \beta\left(\overline{q}_{1980} - \overline{q}_t\right), \end{aligned}$$

where $\overline{q}_t$ denotes the mean of the spike for year t.

are relegate to Appendix A.

The results are summarized in Table 5. The spike in the benchmark year 1980 is 5.4%. A decade of keeping the minimum nominally constant reduced the spike to 0.6% in 1991. Raising the spike back to the level applying 1980 would require an increase in the average nominal minimum wage of 0.335 log points or 40%. Employment in the bottom quintile of the distribution would increase by 0.029 log points. During that decade, the 50-10 log wage differential increased by 0.073 log points. The fall in the minimum wage fully explains this increase ($0.010 + 0.065 = 0.075$); 85% of the explanation comes from the compression effect, the remaining 15% from the truncation effect. Following the logic of Meyer and Wise (1983) that truncation due to a higher minimum wage leads to compression of the wage distribution, one would expect that the reversal of the sign of the employment effect would lead to a decompression. However, the positive employment effect in the lower half of the human capital distribution reduces the human capital of the median worker and therefore 50-10 log wage differential. Obviously, this effect is offset by larger wage differentials due to truncation in the upper tail. Since the compression and the truncation effect have opposite signs, the net effect of an increase in the spike on the 90-50 log wage differential is small. Altogether, the fall in the minimum wage explains 76% of the increase in wage dispersion during the eighties. The increase in the minimum raises the total wage bill by 7%, largely due to the higher wages paid to workers in the bottom half of the wage distribution. In 2004, when the spike is even lower than in 1991, the fall in the minimum wage compared to 1980 explains only 45% of the increase in wage dispersion.

We add a counterfactual for 2019 where the federal minimum wage is increased by 40%. Since the federal minimum is not binding in all states, the actual increase is smaller, only 0.181 log points. This experiment is predicted to raise the spike from 1% to 2.9%. The truncation effect reduces the wage bill by 1.5% due to an increase in employment at lower skill levels, while the compression effect would increase the wage bill by 4.5%. Employment in the lowest quintile would increase with 0.029 log points.

Table 5: Counterfactual Estimation with $\hat{q}$

						Δlog employment[(3)]					log wage differential[(4)]				
year		$\overline{q}_t{}^{(1)}$	$\overline{q}_t^-$	$\Delta\overline{m}_t{}^{(2)}$		1%	5%	10%	20%		50-10	50-15	50-20	90-50	$\Delta\ln\,\Sigma W_i^{(5)}$
1980	actual	0.054	0.061							actual	0.580	0.526	0.446	0.683	
1991	actual	0.006	0.024		dens	0.221	0.185	0.143	0.073	actual	0.073	0.036	0.024	0.041	
	c.fact.	0.054	0.051	0.344	distr	0.002	0.010	0.018	0.029	trunc.	-0.010	-0.011	0.003	0.028	-0.036
										compr.	-0.065	-0.041	-0.019	-0.035	0.107
1998	actual	0.013	0.041		dens	0.178	0.150	0.117	0.059	actual	0.079	0.045	0.024	0.051	
	c.fact.	0.054	0.064	0.271	distr	0.002	0.008	0.015	0.024	trunc.	-0.006	-0.012	0.007	0.013	-0.030
										compr.	-0.062	-0.019	-0.014	-0.013	0.091
2004	actual	0.006	0.025		dens	0.223	0.186	0.145	0.074	actual	0.094	0.044	0.027	0.080	
	c.fact.	0.054	0.053	0.437	distr	0.002	0.010	0.019	0.029	trunc.	-0.007	-0.009	0.005	0.010	-0.038
										compr.	-0.055	-0.020	-0.019	-0.027	0.117
2010	actual	0.018	0.040		dens	0.148	0.126	0.098	0.049	actual	0.098	0.055	0.045	0.105	
	c.fact.	0.054	0.061	0.210	distr	0.001	0.007	0.012	0.020	trunc.	-0.014	-0.006	-0.008	0.010	-0.026
										compr.	-0.040	-0.009	-0.009	-0.011	0.076
2019	actual	0.010	0.055		dens	0.192	0.161	0.126	0.064	actual	0.080	0.029	0.011	0.131	
	c.fact.	0.054	0.080	0.338	distr	0.002	0.009	0.016	0.025	trunc.	-0.003	0.004	0.007	0.015	-0.032
										compr.	-0.065	-0.048	-0.018	-0.025	0.099
Counterfactual for a 40% increase in the federal minimum wage															
2019	actual	0.010	0.055		dens	0.085	0.072	0.056	0.029	actual	0.664	0.554	0.456	0.814	
	c.fact.	0.029	0.065	0.181	distr	0.001	0.004	0.007	0.011	trunc.	-0.003	0.001	0.006	0.004	-0.015
										compr.	-0.039	-0.025	-0.007	-0.009	0.045

Note: 1. The unweighed average of the spike q_s across regions for a particular year. 2. The unweighed average of the log minimum wage m_s across regions for a particular year. 3. Δ ln employment: (c.fact.) see Appendix A. 4. The actual and counter-factual log wage differentials among the workers. 5. The difference between the actual and the counter-factual log of the sum of wage for all workers earning more than the minimum wage.

7 Conclusion

We reexamined the evidence on the effect of minimum wages on wage spillovers and employment, by addressing the problems of (i) the endogeneity of the median wage, (ii) disentangling the truncation and compression effect, (iii) its non-linearity/dependence on the initial level of the minimum wage, and (iv) the heterogeneity of the employment effect for different points in the human capital distribution. We find strong evidence for an increase of the minimum wage to raise median of the wage distribution and to compress wage differentials above the minimum wage, largely confirming Lee (1999) and Teulings (2003), though unlike the former reference, the compression is heavily concentrated in the bottom half of the distribution. Moreover, we find strong evidence for ”reverse truncation”: employment in the bottom half of the distribution responds positively to an increase in the minimum wage. There is a turning point where these effects switch signs, but for most metrics, this turning point is at a surprisingly high level of the spike of about 10%; this is about the highest spike observed in our sample of 3000 economies, occurring in some Southern states in 1980. We find that 75% of the increase in wage dispersion during the eighties of the previous century was due to the erosion of the real minimum wage by inflation. The contribution of the minimum wage to raising wage inequality in the next twenty years was much smaller.

From a theory point of view, our results are inconsistent with the hedonic pricing model discussed in Section 2. The compression of the wage differentials is fully consistent with the theoretical predictions of this model, but the increase in the average wage and the combination of wage compression and positive employment effects in the bottom tail is not: compression can only be rationalized in this model by negative employment effects in the bottom tail, such that workers who remain employed have to take the low complex left vacant by the disemployment of the workers with lower human capital. Only a monopsony model seems to be able to explain these results, suggesting that the Hosios condition is violated to the detriment of workers. This conclusion is consistent with an increasing body of evidence in favour of the monopsony model, see Ashenfelter et al. (2021).

From a policy point of view, we have not attempted to provide a welfare evaluation of the cost and benefits for various parts of the human capital distribution and of the rewards for other factors of production than labor. However, for those with a strong preference for an equal wage distribution and a higher labor share this paper provides arguments for a much higher spike than we currently observe.

References

AHLFELDT, G., D. ROTH, AND T. SEIDEL (2018): "The Regional Effects of a National Minimum Wage," CESifo Working Paper Series 6924, CESifo.

——— (2022): "Optimal minimum wages," Working paper series, CEPR.

ANGRIST, J. D. AND A. B. KRUEGER (1991): "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics*, 106, 979–1014.

ASHENFELTER, O. C., D. CARD, H. S. FARBER, AND M. RANSOM (2021): "Monopsony in the Labor Market: New Empirical Results and New Public Policies," Working Paper 29522, National Bureau of Economic Research.

AUTOR, D. H. AND D. DORN (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103, 1553–97.

AUTOR, D. H., L. F. KATZ, AND M. S. KEARNEY (2006): "The Polarization of the U.S. Labor Market," *American Economic Review*, 96, 189–194.

——— (2008): "Trends in US wage inequality: Revising the revisionists." *The Review of economics and statistics*, 90, 300–323.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 118, 1279–1333.

AUTOR, D. H., A. MANNING, AND C. L. SMITH (2016): "The Contribution of the Minimum Wage to US Wage Inequality over Three Decades: A Reassessment," *American Economic Journal: Applied Economics*, 8, 58–99.

AZAR, J., E. HUET-VAUGHN, I. MARINESCU, B. TASKA, AND T. VON WACHTER (2019): "Minimum Wage Employment Effects and Labor Market Concentration," Working Paper 26101, National Bureau of Economic Research.

BAILEY, M. J., J. DINARDO, AND B. A. STUART (2020): "The Economic Impact of a High National Minimum Wage: Evidence from the 1966 Fair Labor Standards Act," Working Paper 26926, National Bureau of Economic Research.

BONTEMPS, C., J.-M. ROBIN, AND G. J. VAN DEN BERG (2000): "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation," *International Economic Review*, 41, 305–358.

BOUND, J. AND A. B. KRUEGER (1991): "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics*, 9, 1–24.

BROCHU, P., D. A. GREEN, T. LEMIEUX, AND J. TOWNSEND (2015): "The Minimum Wage, Turnover, and the Shape of the Wage Distribution," .

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, 772–93.

——— (1995): "Time-Series Minimum-Wage Studies: A Meta-analysis," *American Economic Review*, 85, 238–43.

——— (2000): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Reply," *American Economic Review*, 90, 1397–1420.

——— (2016): *Myth and Measurement: The New Economics of the Minimum Wage Twentieth-Anniversary Edition*, Princeton University Press, 2 ed.

CENGIZ, D., A. DUBE, A. LINDNER, AND B. ZIPPERER (2019): "The Effect of Minimum Wages on Low-Wage Jobs*," *The Quarterly Journal of Economics*, 134, 1405–1454.

CHEN, Y. AND C. N. TEULINGS (2021): "Agglomeration, Sorting and the Return to Human Capital," Working paper series, University of Cambridge.

DICKENS, R., S. MACHIN, AND A. MANNING (1998): "Estimating the effect of minimum wages on employment from the distribution of wages: A critical view," *Labour Economics*, 5, 109–134.

DICKENS, R. AND A. MANNING (2004): "Has the National Minimum Wage Reduced UK Wage Inequality?" *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 167, 613–626.

DICKENS, R., A. MANNING, AND T. BUTCHER (2012): "Minimum Wages and Wage Inequality: Some Theory and an Application to the UK," Working paper series, Department of Economics, University of Sussex Business School.

DINARDO, J., N. M. FORTIN, AND T. LEMIEUX (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64, 1001–1044.

DOLADO, J., F. KRAMARZ, S. MACHIN, A. MANNING, D. MARGOLIS, C. TEULINGS, G. SAINT-PAUL, AND M. KEEN (1996): "The Economic Impact of Minimum Wages in Europe," *Economic Policy*, 11, 319–372.

DUBE, A., T. W. LESTER, AND M. REICH (2010): "Minimum Wage Effects across State Borders: Estimates using Contiguous Counties," *The Review of Economics and Statistics*, 92, 945–964.

——— (2016): "Minimum Wage Shocks, Employment Flows, and Labor Market Frictions," *Journal of Labor Economics*, 34, 663–704.

ENGBOM, N. AND C. MOSER (2018): "Earnings Inequality and the Minimum Wage: Evidence from Brazil," Working paper series, CESifo.

FISHBACK, P. V. AND A. SELTZER (2020): "The Rise of American Minimum Wages, 1912-1968," Working Paper 26937, National Bureau of Economic Research.

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

GABAIX, X. AND A. LANDIER (2008): "Why has CEO Pay Increased So Much?*," *The Quarterly Journal of Economics*, 123, 49–100.

GENNAIOLI, N., R. LA PORTA, F. LOPEZ-DE SILANES, AND A. SHLEIFER (2013): "Human Capital and Regional Development," *The Quarterly Journal of Economics*, 128, 105–164.

GOLDSMITH-PINKHAM, P., I. SORKIN, AND H. SWIFT (2020): "Bartik Instruments: What, When, Why, and How," *American Economic Review*, 110, 2586–2624.

GOOS, M., A. MANNING, AND A. SALOMONS (2010): "Explaining Job Polarization in Europe: The Roles of Technology, Globalization and Institutions," Cep discussion papers, Centre for Economic Performance, LSE.

GOTTFRIES, A. AND C. N. TEULINGS (2021): "Returns to on-the-job search and the dispersion of wages," Working paper series, CEPR.

HAANWINCKEL, D. (2017): "Supply, Demand, Institutions, and Firms: A Theory of Labor Market Sorting and the Wage Distribution," Working paper series, CEPR.

HAANWINCKEL, D. AND R. R. SOARES (2021): "Workforce Composition, Productivity, and Labor Regulations in a Compensating Differentials Theory of Informality," *The Review of Economic Studies*, rdab017.

HAGEDORN, M., I. MANOVSKII, AND Y. WANG (2018): "Search Frictions and Wage Dispersion," Working paper series, University of Oslo.

HAMILTON, L. C. (1992): "How Robust is Robust Regression?" *Stata Technical Bulletin*, 1.

JUHN, C., K. M. MURPHY, AND B. PIERCE (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101, 410–442.

KATZ, L. F. AND K. M. MURPHY (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107, 35–78.

LEE, D. S. (1999): "Wage Inequality in the United States during the 1980s: Rising Dispersion or Falling Minimum Wage?" *The Quarterly Journal of Economics*, 114, 977–1023.

MACHIN, S. AND A. MANNING (1994): "The Effects of Minimum Wages on Wage Dispersion and Employment: Evidence from the U.K. Wages Councils," *Industrial and Labor Relations Review*, 47, 319–329.

MACHIN, S., A. MANNING, AND L. RAHMAN (2003): "Where the Minimum Wage Bites Hard: Introduction of Minimum Wages to a Low Wage Sector," *Journal of the European Economic Association*, 1, 154–180.

MEYER, R. H. AND D. A. WISE (1983): "Discontinuous Distributions and Missing Persons: The Minimum Wage and Unemployed Youth," *Econometrica*, 51, pp. 1677–1698.

NEUMARK, D. AND L. F. M. CORELLA (2019): "Do Minimum Wages Reduce Employment in Developing Countries? A Survey and Exploration of Conflicting Evidence," NBER Working Papers 26462, National Bureau of Economic Research, Inc.

NEUMARK, D., J. I. SALAS, AND W. WASCHER (2014): "More on Recent Evidence on the Effects of Minimum Wages in the United States," *IZA J Labor Policy*, 3.

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?" Tech. rep., National Bureau of Economic Research.

NEUMARK, D. AND W. WASCHER (2000): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Comment," *American Economic Review*, 90, 1362–1396.

ROSEN, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82, 34–55.

RUGGLES, S., K. GENADEK, R. GOEKEN, J. GROVER, AND M. SOBEK (2017): "Integrated Public Use Microdata Series: Version 7.0 [dataset]," Tech. rep., Minneapolis: University of Minnesota.

SATTINGER, M. (1975): "Comparative Advantage and the Distributions of Earnings and Abilities," *Econometrica*, 43, 455–68.

STEWART, M. B. (2012): "Quantile Estimates of Counterfactual Distribution Shifts and the Effect of Minimum Wage Increases on the Wage Distribution," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175, 263–287.

TEULINGS, C. N. (1995): "The Wage Distribution in a Model of the Assignment of Skills to Jobs," *Journal of Political Economy*, 103, 280–315.

——— (2000): "Aggregation Bias in Elasticities of Substitution and the Minimum Wage Paradox," *International Economic Review*, 41, 359–398.

——— (2003): "The Contribution of Minimum Wages to Increasing Wage Inequality," *The Economic Journal*, 113, 801–833.

——— (2005): "Comparative Advantage, Relative Wages, and the Accumulation of Human Capital," *Journal of Political Economy*, 113, 425–461.

TEULINGS, C. N. AND T. VAN RENS (2008): "Education, Growth, and Income Inequality," *The Review of Economics and Statistics*, 90, 89–104.

A Appendix

Our numerical procedure is as follows. The superscript c denotes a counterfactual. For each economy s we do:

1. Sort observations by w_i ($i = 1$ is the lowest wage).
2. Calculate the quantile for all i in the actual distribution: $p_i = i/N_s$.
3. Calculate the employment change for observation i

$$f_i = \exp\left[\sigma_g^{-1}\left[P\left(p, q_s^c\right) - P\left(p, q_s\right)\right]\right]$$

where $P\left(p, q_s\right)$ is the polynomial in p and q_s, using the estimation results from Table A6, column (4).

4. Calculate the counterfactual quantile for all i: $p_i^c = \Sigma_{j=1}^{i} f_j / \Sigma_{i=1}^{N_s} f_i$.
5. Define i_s^- and i_s such that $p_{i_s^-}^c = q_s^{-c}$ and $p_{i_s}^c = q_s^{-c} + q_s^c$ respectively (the value of i at the lower- and uppersupport respectively of the spike in the counterfactual human capital distribution).
6. Calculate the counterfactual wage for w_i^c for all i
   (a) For all $i \geq i_s$ (those earning more than minimum wage in the counterfactual)

   $$w_i^c = w_i + w_s\left(g_i, q_s^c\right) - w_s\left(g_i, q_s\right)$$

   using the estimation results for $w_s\left(g_i, q_s\right)$ from Table 4, column (3).

   (b) We use this step to calculate to the counterfactual minimum wage $m_s^c = w_{i_s}^c$.

   (c) For $i_s^- < i \leq i$ (the counterfactual spike), $w_i^c = m_s^c$.

   (d) For all $i \leq i_s^-$ (those earning less than the minimum wage in the counterfactual), $w_i^c = \mathrm{E}(w_i | p_i < q_s^-)$.
7. Define i_{sp} such that $p_{i_{sp}} = p$ and i_{sp}^c such that $p_{i_{sp}}^c = p$ for $p = .01, .05, .10, .20, .50, .90$ (the factual and counterfactual values of i corresponding with these quantiles).
8. Calculate the cumulative relative change in employment from the lowest quantile 0 up till quantile p: $p^{-1}\Sigma_{i=1}^{i_{sp}}\left(f_i - 1\right)$.
9. Calculate the actual log wage differentials: $\Delta w_{sp} = w_{i_{sp}} - w_{i_{s.50}}$ (reversing sign for $\Delta w_{s.90}$)
10. Calculate the change in the log wage differential: $\Delta^2 w_{sp} = \Delta w_{sp} - \Delta w_{sp}^b$ (where the superscript b denotes the benchmark year).
11. Calculate the change in the log wage differential due to truncation: $\Delta^2 w_{sp\text{trunc}} = w_{i_{sp}^c} - w_{i_{sp}} - w_{i_{s.50}^c} + w_{i_{s.50}}$
12. Calculate log wage differential due to compression: $\Delta^2 w_{sp\text{comp}} = w_{i_{sp}^c}^c - w_{i_{sp}^c} - w_{i_{s.50}^c}^c + w_{i_{s.50}}$

13. Calculate the relative change in the total wage sum due to truncation: $\Delta \log W_{i\text{trunc}} = \log \Sigma_{i \in t} f_i \mathrm{e}^{w_i} - \log \Sigma_{i \in t} \mathrm{e}^{w_i}$

14. Calculate the relative change in the total wage sum due to compression: $\Delta \log W_{i\text{comp}} = \log \Sigma_{i \in t} f_i \mathrm{e}^{w_i^c} - \log \Sigma_{i \in t} f_i \mathrm{e}^{w_i}$

Table A1: CBSA Observations Distribution Among States

CBSA	State I	State II	State III	State IV	Pct SI	Pct SII	Pct SIII	Pct SIV	NAME
31100	CA				100.00%				Los Angeles-Long Beach-Anaheim, CA
40140	CA				100.00%				Riverside-San Bernardino-Ontario, CA
41740	CA				100.00%				San Diego-Carlsbad, CA
41860	CA				100.00%				San Francisco-Oakland-Hayward, CA
41940	CA				100.00%				San Jose-Sunnyvale-Santa Clara, CA
19740	CO				100.00%				Denver-Aurora-Lakewood, CO
47900	DC	VA	MD	WV	45.91%	25.90%	28.19%	0.00%	Washington-Arlington-Alexandria, DC-VA-MD-WV
33100	FL				100.00%				Miami-Fort Lauderdale-West Palm Beach, FL
45300	FL				100.00%				Tampa-St. Petersburg-Clearwater, FL
12060	GA				100.00%				Atlanta-Sandy Springs-Roswell, GA
16980	IL	IN	WI		98.23%	1.77%	0.00%		Chicago-Naperville-Elgin, IL-IN-WI
26900	IN				100.00%				Indianapolis-Carmel-Anderson, IN
35380	LA				100.00%				New Orleans-Metairie, LA
14460	MA	NH			86.75%	13.25%			Boston-Cambridge-Newton, MA-NH
12580	MD				100.00%				Baltimore-Columbia-Towson, MD
19820	MI				100.00%				Detroit-Warren-Dearborn, MI
33460	MN	WI			99.99%	0.01%			Minneapolis-St. Paul-Bloomington, MN-WI
28140	MO	KS			45.36%	54.64%			Kansas City, MO-KS
41180	MO	IL			80.98%	19.02%			St. Louis, MO-IL
24660	NC				100.00%				Greensboro-High Point, NC
15380	NY				100.00%				Buffalo-Cheektowaga-Niagara Falls, NY
35620	NY	NJ			69.24%	30.76%			New York-Newark-Jersey City, NY-NJ
40380	NY				100.00%				Rochester, NY
17140	OH	KY	IN		77.70%	22.30%	0.00%		Cincinnati, OH-KY-IN
17460	OH				100.00%				Cleveland-Elyria, OH
18140	OH				100.00%				Columbus, OH
38900	OR	WA			91.57%	8.43%			Portland-Vancouver-Hillsboro, OR-WA
37980	PA	NJ	DE	MD	62.06%	23.32%	14.62%	0.00%	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
38300	PA				100.00%				Pittsburgh, PA
19100	TX				100.00%				Dallas-Fort Worth-Arlington, TX
26420	TX				100.00%				Houston-The Woodlands-Sugar Land, TX
47260	VA	NC			100.00%	0.00%			Virginia Beach-Norfolk-Newport News, VA-NC
42660	WA				100.00%				Seattle-Tacoma-Bellevue, WA
33340	WI				100.00%				Milwaukee-Waukesha-West Allis, WI

Note: Information for 34 city areas: CBSA code in 2013, city belong to which state(s) and the percentage of sample observations in the CPS 1979-2015, name of cities. *Data sources*: the Current Population Survey MORG and the US Census Bureau.

Table A5: Individual Level Mincer Regression

VARIABLES β	(1) $\tilde{w}$ Full Obs.	(2) $\tilde{w}$ excl. 10%	(3) $\tilde{w}$ excl. 20%	(4) $\tilde{w}$ excl. 30%	(5) $\tilde{w}$ excl. 40%
Male	0.310	0.303	0.285	0.269	0.260
	(179.08)	(155.45)	(129.15)	(106.57)	(92.36)
Male $\times$ Trend	0.00548	0.00499	0.00419	0.00354	0.00321
	(131.11)	(104.32)	(75.88)	(54.84)	(43.75)
Single	0.0126	0.0121	0.00997	0.00863	0.00975
	(9.11)	(7.51)	(5.31)	(3.84)	(3.82)
Single $\times$ Trend	-0.00242	-0.00223	-0.00206	-0.00197	-0.00197
	(-44.50)	(-36.34)	(-29.48)	(-24.20)	(-21.32)
Divorced	0.0257	0.0296	0.0267	0.0269	0.0268
	(16.78)	(16.36)	(12.59)	(10.61)	(9.29)
Divorced $\times$ Trend	-0.00164	-0.00175	-0.00164	-0.00162	-0.00160
	(-25.43)	(-24.00)	(-19.67)	(-16.86)	(-14.67)
Male $\times$ Single	-0.211	-0.216	-0.213	-0.207	-0.206
	(-112.79)	(-99.02)	(-83.67)	(-67.93)	(-59.17)
Male $\times$ Single $\times$ Trend	0.00398	0.00399	0.00380	0.00353	0.00345
	(54.11)	(48.10)	(39.97)	(31.84)	(27.43)
Male $\times$ Divorced	-0.113	-0.117	-0.116	-0.115	-0.117
	(-45.36)	(-40.38)	(-34.22)	(-28.59)	(-25.48)
Male $\times$ Divorced $\times$ Trend	0.00164	0.00182	0.00176	0.00173	0.00182
	(15.94)	(15.70)	(13.41)	(11.42)	(10.61)
South	0.00609	0.00583	0.00716	0.00736	0.00714
	(1.76)	(1.67)	(1.98)	(2.01)	(1.89)
Black	-0.100	-0.104	-0.110	-0.117	-0.120
	(-103.19)	(-103.42)	(-102.89)	(-100.16)	(-96.59)
Other Race	-0.0764	-0.0790	-0.0820	-0.0815	-0.0826
	(-73.75)	(-73.54)	(-71.08)	(-64.42)	(-60.07)
South $\times$ Black	-0.0349	-0.0314	-0.0246	-0.0158	-0.00996
	(-25.71)	(-21.57)	(-16.05)	(-9.59)	(-5.63)
South $\times$Others	0.00133	0.00405	0.00889	0.00739	0.0104
	(0.58)	(1.71)	(3.60)	(2.84)	(3.70)
Edu = 0	-0.637	-0.637	-0.622	-0.602	-0.594
	(-107.31)	(-98.90)	(-88.62)	(-77.16)	(-69.69)
Edu = 1	-0.532	-0.518	-0.500	-0.465	-0.423
	(-36.52)	(-27.67)	(-22.09)	(-15.16)	(-12.28)
Edu = 2	-0.530	-0.531	-0.509	-0.467	-0.456

Continued on next page

Table A5 – continued from previous page

VARIABLES β	(1) $\tilde{w}$ Full Obs.	(2) $\tilde{w}$ excl. 10%	(3) $\tilde{w}$ excl. 20%	(4) $\tilde{w}$ excl. 30%	(5) $\tilde{w}$ excl. 40%
	(-72.20)	(-60.87)	(-50.08)	(-35.15)	(-30.42)
Edu = 3	-0.525	-0.523	-0.520	-0.513	-0.509
	(-80.23)	(-65.61)	(-55.10)	(-40.58)	(-35.31)
Edu = 4	-0.453	-0.459	-0.449	-0.419	-0.429
	(-68.01)	(-55.38)	(-45.67)	(-32.23)	(-29.15)
Edu = 5	-0.465	-0.468	-0.451	-0.449	-0.436
	(-103.70)	(-88.79)	(-75.07)	(-59.31)	(-50.99)
Edu = 6	-0.427	-0.436	-0.437	-0.412	-0.418
	(-118.86)	(-101.40)	(-87.31)	(-62.13)	(-55.21)
Edu = 7	-0.357	-0.371	-0.365	-0.362	-0.363
	(-109.67)	(-93.89)	(-80.11)	(-66.97)	(-60.03)
Edu = 8	-0.262	-0.268	-0.278	-0.277	-0.282
	(-119.28)	(-98.37)	(-84.43)	(-67.22)	(-60.17)
Edu = 9	-0.259	-0.270	-0.272	-0.270	-0.272
	(-166.68)	(-155.34)	(-142.54)	(-127.25)	(-117.00)
Edu = 10	-0.194	-0.200	-0.204	-0.206	-0.208
	(-167.84)	(-155.40)	(-143.73)	(-131.43)	(-121.51)
Edu = 11	-0.157	-0.163	-0.166	-0.171	-0.172
	(-148.07)	(-138.94)	(-129.94)	(-120.72)	(-111.43)
Edu = 13	0.0574	0.0549	0.0508	0.0439	0.0423
	(60.89)	(52.55)	(43.56)	(33.33)	(29.18)
Edu = 14	0.165	0.170	0.174	0.176	0.178
	(187.37)	(177.70)	(166.44)	(151.76)	(139.95)
Edu = 15	0.209	0.217	0.223	0.224	0.226
	(135.36)	(126.84)	(116.73)	(104.21)	(95.54)
Edu = 16	0.420	0.429	0.433	0.435	0.437
	(369.70)	(347.58)	(324.91)	(298.35)	(275.43)
Edu = 17	0.399	0.409	0.421	0.425	0.432
	(167.36)	(148.71)	(129.95)	(107.31)	(95.26)
Edu = 18	0.574	0.586	0.593	0.594	0.597
	(330.35)	(311.82)	(291.81)	(267.91)	(247.55)
Year of Experience (Exp)	0.0248	0.0243	0.0220	0.0188	0.0181
	(40.85)	(36.61)	(30.49)	(23.61)	(20.82)
$\text{Exp}^2/100$	-0.0400	-0.0354	-0.0229	-0.00666	-0.00221
	(-14.39)	(-11.59)	(-6.89)	(-1.81)	(-0.55)
$\text{Exp}^3/100000$	0.235	0.174	0.00153	-0.216	-0.283

Continued on next page

Table A5 – continued from previous page

VARIABLES β	(1) $\tilde{w}$ Full Obs.	(2) $\tilde{w}$ excl. 10%	(3) $\tilde{w}$ excl. 20%	(4) $\tilde{w}$ excl. 30%	(5) $\tilde{w}$ excl. 40%
	(6.44)	(4.31)	(0.03)	(-4.43)	(-5.31)
Exp $\times$ Edu	0.00165	0.00169	0.00181	0.00197	0.00200
	(37.30)	(35.07)	(34.65)	(34.30)	(32.10)
$\text{Exp}^2/100 \times$ Edu	-0.00888	-0.00919	-0.00983	-0.0106	-0.0108
	(-43.79)	(-41.37)	(-40.80)	(-39.96)	(-37.47)
$\text{Exp}^3/100000 \times$ Edu	0.107	0.111	0.120	0.130	0.134
	(39.56)	(37.39)	(37.22)	(36.63)	(34.48)
Male $\times$ Exp	0.00581	0.00568	0.00605	0.00657	0.00683
	(22.72)	(20.51)	(20.26)	(20.16)	(19.27)
$\text{Exp}^2/100 \times$ Male	0.00459	0.00356	0.000249	-0.00338	-0.00564
	(3.52)	(2.51)	(0.16)	(-2.02)	(-3.10)
$\text{Exp}^3/100000 \times$ Male	-0.233	-0.216	-0.168	-0.119	-0.0838
	(-12.13)	(-10.31)	(-7.41)	(-4.81)	(-3.11)
Observations	5,803,821	5,146,824	4,541,693	3,916,580	3,351,354
R-squared	0.584	0.547	0.515	0.489	0.474
R-MSE	0.449	0.453	0.456	0.458	0.460
Time x Region Dummy	Y	Y	Y	Y	Y

t-statistics in parentheses

Table A2: Summary Statistics

Year	s.d. g_i	mean $q_s(=)$	mean $q_s(<)$	#Region $q_s > 1.5\%$	#≤FedMW
1979	0.346	0.045	0.060	79	80
1980	0.346	0.054	0.061	81	80
1981	0.345	0.050	0.066	81	80
1982	0.345	0.047	0.052	78	78
1983	0.344	0.048	0.045	79	80
1984	0.344	0.045	0.040	80	80
1985	0.345	0.038	0.035	76	79
1986	0.345	0.036	0.035	69	75
1987	0.347	0.032	0.032	69	74
1988	0.346	0.026	0.029	55	72
1989	0.359	0.018	0.025	41	56
1990	0.360	0.012	0.038	26	0
1991	0.360	0.006	0.024	7	63
1992	0.353	0.023	0.026	50	76
1993	0.354	0.020	0.024	50	75
1994	0.361	0.012	0.036	22	73
1995	0.366	0.011	0.030	18	72
1996	0.367	0.008	0.027	7	69
1997	0.368	0.008	0.033	10	70
1998	0.369	0.013	0.041	24	0
1999	0.371	0.011	0.034	19	0
2000	0.371	0.010	0.030	14	0
2001	0.370	0.007	0.030	10	0
2002	0.371	0.006	0.029	7	0
2003	0.372	0.007	0.026	7	0
2004	0.372	0.006	0.025	6	0
2005	0.371	0.007	0.025	9	0
2006	0.371	0.006	0.026	8	0
2007	0.371	0.008	0.032	12	0
2008	0.370	0.009	0.034	16	25
2009	0.368	0.011	0.038	19	34
2010	0.367	0.018	0.040	41	55
2011	0.370	0.016	0.039	38	51
2012	0.377	0.016	0.040	39	48
2013	0.368	0.015	0.038	33	46
2014	0.369	0.014	0.052	29	32
2015	0.369	0.014	0.043	28	32
2016	0.369	0.015	0.043	22	32
2017	0.368	0.013	0.040	21	32
2018	0.368	0.011	0.045	18	32
2019	0.368	0.010	0.055	16	32

Table A3: Frequency table by year with the spike in the bottom 70%

Year	Regions	Year	Regions	Year	Regions
1979	7	1993	54	2007	75
1980	1	1994	71	2008	72
1981	4	1995	75	2009	73
1982	8	1996	81	2010	62
1983	4	1997	77	2011	67
1984	7	1998	70	2012	69
1985	14	1999	73	2013	71
1986	17	2000	74	2014	73
1987	24	2001	77	2015	69
1988	42	2002	77	2016	68
1989	57	2003	77	2017	68
1990	72	2004	77	2018	70
1991	76	2005	77	2019	71
1992	45	2006	78	Total	2324

Table A4: Frequency table by region with the spike in the bottom 70%

Region ID	Freq.	Region ID	Freq.	Region ID	Freq.
Atlanta, GA	35	St Louis, MO	31	Delaware	31
Baltimore, MD	33	San Diego, CA	18	Maryland	29
Boston, MA	35	San Francisco, CA	35	Virginia	31
Buffalo, NY	28	San Jose, CA	33	West Virginia	20
Chicago, IL	36	Seattle, WA	39	North Carolina	31
Cincinnati, OH	32	Tampa, FL	31	South Carolina	28
Cleveland, OH	31	Virginia Beach, VA	31	Georgia	24
Columbus, OH	33	Maine	30	Florida	31
Dallas, TX	37	New Hampshire	35	Kentucky	26
Denver, CO	35	Vermont	32	Tennessee	27
Detroit, MI	32	Massachusetts	29	Alabama	25
Greensboro, NC	33	Rhode Island	32	Mississippi	24
Houston, TX	33	Connecticut	34	Arkansas	24
Indianapolis, IN	32	New York	31	Louisiana	22
Kansas City, MO	33	Pennsylvania	30	Oklahoma	26
Los Angeles, CA	5	Ohio	26	Texas	21
Miami, FL	32	Indiana	30	Montana	28
Milwaukee, WI	33	Illinois	23	Idaho	28
Minneapolis, MN	39	Michigan	28	Wyoming	29
New Orleans, LA	30	Wisconsin	31	Colorado	29
New York, NY	36	Minnesota	28	New Mexico	21
New Jersey, NJ	33	Iowa	28	Arizona	28
Philadelphia, PA	34	Missouri	26	Utah	29
Pittsburgh, PA	29	North Dakota	29	Nevada	31
Portland, OR	28	South Dakota	28	Washington	18
Riverside, CA	10	Nebraska	30	Oregon	13
Rochester, NY	29	Kansas	27	California	9

Table A6: Employment Regression with $\ln \Delta g_{sp} = \ln(0.01) - \ln(g_{s,p} - g_{s,p-0.01})$

VARIABLES	(1) $\ln \Delta g_{sp}$	(2) $\ln \Delta g_{sp}$	(3) $\ln \Delta g_{sp}$	(4) $\ln \Delta g_{sp}$
$\widehat{q}_s$	4.380	-0.874	2.388	-2.053
	(11.78)	(-1.53)	(4.82)	(-4.19)
$\widehat{q}_s \times (p-0.5)$	-27.71	-25.84	-23.34	-20.68
	(-15.11)	(-15.06)	(-13.22)	(-11.89)
$\widehat{q}_s \times (p-0.5)^2$	-23.85	-16.33	-8.053	9.152
	(-3.17)	(-1.59)	(-0.83)	(0.94)
$\widehat{q}_s \times (p-0.5)^3$	65.86	-3.410	36.83	-18.20
	(5.97)	(-0.32)	(3.27)	(-1.62)
$\widehat{q}_s \times (p-0.5)^4$	-76.48	165.8	-72.08	65.44
	(-2.23)	(3.88)	(-1.58)	(1.44)
$\widehat{q}_s^2$	-49.99	5.322	-53.15	-1.915
	(-11.33)	(0.80)	(-9.16)	(-0.34)
$\widehat{q}_s^2 \times (p-0.5)$	260.0	297.7	160.3	191.8
	(11.37)	(13.37)	(6.61)	(8.13)
$\widehat{q}_s^2 \times (p-0.5)^2$	602.3	248.8	796.8	376.1
	(6.32)	(1.90)	(5.94)	(2.85)
$\widehat{q}_s^2 \times (p-0.5)^3$	-881.6	-668.4	-435.1	-294.3
	(-6.28)	(-4.91)	(-2.81)	(-1.95)
$\widehat{q}_s^2 \times (p-0.5)^4$	-971.6	-1,475	-2,097	-1,912
	(-2.23)	(-2.76)	(-3.35)	(-3.12)
Bartik IV	-2.679	-4.240	-2.934	-4.399
	(-11.56)	(-11.73)	(-8.46)	(-13.05)
Bartik IV $\times(p-0.5)$	-3.100	-0.241	-1.464	2.048
	(-4.68)	(-0.37)	(-2.28)	(3.20)
Bartik IV $\times(p-0.5)^2$	51.18	107.2	45.83	96.46
	(16.14)	(24.60)	(12.93)	(27.38)
Bartik IV $\times(p-0.5)^3$	41.46	49.37	33.20	32.26
	(10.35)	(12.06)	(8.12)	(7.89)
Bartik IV $\times(p-0.5)^4$	-126.1	-408.8	-95.02	-338.3
	(-9.19)	(-22.21)	(-5.75)	(-20.54)
Time	Y	Y	Y	Y
Region	Y	Y	Y	Y
Dummy p	Y	Y	Y	Y
Dummy p_g	N	Y	N	Y
Regression	OLS	OLS	Robust	Robust
Observation	325,346	325,346	325,346	325,346
R-squared	0.423	0.461	0.442	0.487
RMSE	0.691	0.667	0.646	0.625

Table A7: Instrumental Variable First Stage Regression without Bartik IVs

	(1)	(2)	(3)	(4)
VARIABLES	h_{qs}	h_{qs}	h_{qs}	h_{qs}
m_s	1.224	1.340	1.398	1.426
	(21.81)	(23.92)	(34.23)	(40.34)
Observations	1,414	1,414	3,321	3,321
R-squared	0.882	0.884	0.837	0.868
RMSE	0.0737	0.0736	0.147	0.127
Time Dummy	Y	Y	Y	Y
Region Dummy	Y	Y	Y	Y
Regression	OLS	Robust	OLS	Robust

t-statistics in parentheses

Table A8: Regression with Bias$_s$ corrected $\widehat{\omega}_{0s}$ $\widehat{\omega}_{1s}$ and $\widehat{\omega}_{2s}$ without Bartik IVs

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	$\widehat{\omega}_{0s}$	$\widehat{\omega}_{1s}$	$\widehat{\omega}_{2s}$	$\widehat{\omega}_{0s}$	$\widehat{\omega}_{1s}$	$\widehat{\omega}_{2s}$
Panel A: with 2SLS Regression						
$\widehat{q}_s$	2.119	-3.192	4.321	2.511	-5.184	6.002
	(3.18)	(-4.51)	(3.90)	(8.36)	(-14.10)	(11.71)
$\widehat{q}_s^2$	-1.986	36.36	-58.91	-4.075	50.40	-69.48
	(-0.33)	(5.68)	(-5.87)	(-1.32)	(13.32)	(-13.17)
R-squared	0.986	0.686	0.574	0.984	0.571	0.468
RMSE	0.0496	0.0527	0.0825	0.0486	0.0595	0.0829
Panel B: with Robust OLS Regression						
$\widehat{q}_s$	3.045	-1.059	1.288	3.273	-5.020	5.639
	(6.03)	(-1.72)	(1.35)	(12.52)	(-14.36)	(11.44)
$\widehat{q}_s^2$	-9.997	19.30	-36.12	-15.65	56.56	-76.83
	(-2.30)	(3.65)	(-4.41)	(-5.85)	(15.80)	(-15.23)
R-squared	0.991	0.730	0.637	0.990	0.678	0.568
RMSE	0.0420	0.0512	0.0793	0.0391	0.0523	0.0737
Panel C: with Robust OLS Regression						
$\widehat{q}_s$	-0.392	3.982	-7.328	-1.090	7.981	-10.52
	(-0.66)	(5.49)	(-6.52)	(-3.39)	(18.85)	(-17.31)
$\widehat{q}_s \times \ln\widehat{q}_s$	-1.299	1.642	-2.631	-1.425	3.933	-4.617
	(-4.17)	(4.31)	(-4.46)	(-9.73)	(20.39)	(-16.68)
R-squared	0.991	0.731	0.639	0.990	0.692	0.574
RMSE	0.0418	0.0510	0.0791	0.0387	0.0509	0.0731
Time Dummy	Y	Y	Y	Y	Y	Y
Region Dummy	Y	Y	Y	Y	Y	Y
Observations	1,414	1,414	1,414	3,321	3,321	3,321

t-statistics in parentheses

Table A9: Regression with weighted total employment $\ln(N_s)$ without Bartik IVs

	(1)	(2)	(3)	(4)
VARIABLES	$\ln(N_s)$	$\ln(N_s)$	$\ln(N_s)$	$\ln(N_s)$
$\widehat{q}_s$	0.724	4.835	2.092	0.970
	(0.70)	(3.40)	(3.38)	(1.15)
$\widehat{q}_s^2$	19.43		-5.264	
	(1.81)		(-0.82)	
$\widehat{q}_s \times \ln \widehat{q}_s$		0.984		-0.250
		(1.78)		(-0.76)
Time Dummy	Y	Y	Y	Y
Region Dummy	Y	Y	Y	Y
Regression	OLS	OLS	Robust	Robust
Observations	3,321	3,281	3,321	3,281
R-squared	0.947	0.947	0.980	0.980
RMSE	0.175	0.175	0.105	0.104

t-statistics in parentheses

Table A10: Counterfactual Estimation with $\hat{q}$ using 2010 as base year

						Δlog employment[(3)]					log wage differential[(4)]				
year		$\overline{q}_t{}^{(1)}$	$\overline{q}_t^-$	$\Delta\overline{m}_t{}^{(2)}$		1%	5%	10%	20%		50-10	50-15	50-20	90-50	$\Delta\ln\Sigma W_i^{(5)}$
2010	actual	0.018	0.040							actual	0.684	0.582	0.491	0.788	
1991	actual	0.006	0.024		dens	0.063	0.052	0.041	0.023	actual	-0.025	-0.019	-0.021	-0.064	
	c.fact.	0.018	0.030	0.154	distr	0.001	0.003	0.005	0.009	trunc.	-0.004	0.003	0.000	0.010	-0.013
										compr.	-0.029	-0.007	-0.010	-0.017	0.033
1998	actual	0.013	0.041		dens	0.025	0.021	0.017	0.009	actual	-0.019	-0.010	-0.021	-0.054	
	c.fact.	0.018	0.044	0.070	distr	0.000	0.001	0.002	0.004	trunc.	-0.005	-0.004	0.001	-0.003	-0.006
										compr.	-0.007	-0.010	-0.002	0.001	0.015
2004	actual	0.006	0.025		dens	0.065	0.054	0.042	0.023	actual	-0.004	-0.012	-0.018	-0.025	
	c.fact.	0.018	0.032	0.206	distr	0.001	0.003	0.006	0.009	trunc.	-0.002	-0.006	0.000	0.002	-0.014
										compr.	-0.028	-0.006	-0.012	-0.009	0.042
2019	actual	0.010	0.055		dens	0.038	0.031	0.025	0.014	actual	-0.018	-0.028	-0.034	0.026	
	c.fact.	0.018	0.059	0.104	distr	0.000	0.002	0.003	0.005	trunc.	-0.000	-0.003	0.004	0.002	-0.008
										compr.	-0.016	-0.012	-0.006	-0.015	0.022

Note: 1. The unweighed average of the spike q_s across regions for a particular year. 2. The unweighed average of the log minimum wage m_s across regions for a particular year. 3. Δ ln employment: (c.fact.) see Appendix A. 4. The actual and counter-factual log wage differentials among the workers. 5. The difference between the actual and the counter-factual log of the sum of wage for all workers earning more than the minimum wage.