Does regional variation in wage levels identify the effects of a national minimum wage?

Daniel Haanwinckel *

September 8, 2023

Abstract

This paper examines the identification assumptions underlying two types of estimators of the causal effects of minimum wages based on regional variation in wage levels: the "effective minimum wage" and the "fraction affected/gap" designs. For the effective minimum wage design, I show that the identification assumptions emphasized by Lee (1999) are crucial for unbiased estimation but difficult to satisfy in empirical applications. Even in best case scenarios, this estimator may suffer from significant biases arising from measurement error. The fraction affected/gap designs also exhibit biases from measurement error that lead to inflated rejection rates of true causal effect. Biases are exacerbated by trends in the dispersion of latent wages or equilibrium responses to the minimum wage.

*Department of Economics, UCLA and NBER. Contact information: haanwinckel@econ.ucla.edu
This paper benefited from conversations with Maurizio Mazzocco, Yotam Shem-tov, and many other faculty at UCLA.

1 Introduction

One approach to measuring the effects of minimum wage regulations involves leveraging variation in wage levels across different regions within a country. For instance, the influence of the US federal minimum wage is more pronounced in states like Mississippi or Arkansas than in Texas or Georgia due to the higher median wages in the latter states. Thus, one can use wage distributions to construct treatment intensity measures for a "differences-in-differences" analysis. Classic applications using cross-state variation from the US are Card (1992), who introduced the "fraction affected" design, and Lee (1999), who introduced the "effective minimum wage" design. This strategy is convenient in countries with no variation in minimum wage levels across space. Examples of such applications published by leading economic journals include Mexico (Bosch and Manacorda, 2010), South Africa (Dinkelman and Ranchhod, 2012), Germany (Dustmann et al., 2021), the US in the 1960s and 1970s (Bailey, DiNardo and Stuart, 2021), and Brazil (Engbom and Moser, 2022).

This paper examines the identification assumptions underlying those econometric approaches. Specifically, it uses a combination of economic theory and simulation exercises to investigate whether fraction affected and effective minimum wage designs can accurately capture the causal effects of the minimum wage on employment and wages, as predicted by specific economic models.

I start by examining one estimator at a time. For the effective minimum wage design, I show that two identification assumptions emphasized by Lee (1999) are crucial for unbiased estimation of causal effects but challenging to satisfy in typical applications. The first assumption is that the effective minimum wage should be constructed as the gap between the actual minimum wage and a measure of the centrality of the latent log wage distribution—that is, the distribution of wages that would prevail with no minimum wage in place. Practitioners typically construct the effective minimum wage using the median wage as the centrality measure. I show that this choice normally introduces correlated measurement error, leading to economically relevant biases. The second assumption is that overall wage levels should be uncorrelated with the dispersion of latent log wages across regions. This assumption may be violated if, for example, regions differ in the share of skilled workers, and there is more latent wage dispersion for skilled workers (Lemieux, 2006). I show that even small correlations introduce economically significant biases.

I discuss a series of potential diagnostics and solutions to those problems. Those include

testing for spillovers in the upper tail, using higher quantiles of the wage distribution to construct the effective minimum wage, or using the instrumental variables approach developed by Autor, Manning and Smith (2016). I conclude that the application is most likely to be successful when the econometrician can precisely point out the source of arguably exogenous variation in the effective minimum wage, taking into account institutional knowledge and the set of fixed effects used in the design.

Next, I discuss the fraction affected and gap designs at the region-time level.[1] I find that they exhibit small biases even in cases where optimal performance might be expected, such as when the latent log wage distribution solely differs between regions due to a "location" parameter and the minimum wage is the only time-varying factor. These small biases result in confidence intervals rejecting the true causal effect more frequently than the desired confidence level. Biases can be more prominent in two scenarios. First, when there are trends in the dispersion of log wages (such as those coming from technical change or trade shocks), even if they affect all regions uniformly. Second, when the minimum wage induces substantial changes in returns to skill.

Finally, I evaluate the performance of the estimators in the context of a specific application: the increase in the federal minimum wage in Brazil between 1998 and 2012. I conduct a simulation exercise based on the model developed and estimated in Haanwinckel (2023). Specifically, I assess to which extent the estimators can recover the true causal effects implied by the model using simulated data. That exercise is more realistic than the previous ones in two ways. First, the model features multiple causal pathways for minimum wage effects, including truncation, censoring, worker reallocation from low- to high-wage firms, changes in the assignment of workers to tasks within firms, and general equilibrium effects. Second, the simulated data closely tracks the original Brazilian data, capturing regional and time variation in the shape of the wage distribution (including the minimum wage spike, dispersion, and skew) and the contributions of firm wage premiums and firm-worker sorting patterns to wage inequality.

None of the estimators comes close to recovering the true employment effects. However, the effective minimum wage design *without* region fixed effects provides a surprisingly good approximation for the effects on quantile gaps. This result resonates with Lee (1999), who argues that the decision to include region-fixed effects in the effective minimum wage design

[1] This paper does not address within-region designs that compare firms based on the share of affected workers, such as Card and Krueger (1994) or Harasztosi and Lindner (2019).

should be considered carefully.

This paper contributes to the growing literature that points out econometric challenges associated with difference-in-differences designs in general and specifically regarding minimum wages. Sorkin (2015) and Vogel (2023) discuss the difficulty of detecting long-run minimum wage effects in reduced-form designs. Autor, Manning and Smith (2016) highlight issues with the effective minimum wage design in the presence of measurement error. de Chaisemartin and D'Haultfœuille (2020) raise concerns regarding difference-in-differences designs with heterogeneous treatment effects and non-uniform timing of treatment that are relevant to the effective minimum wage design. Callaway, Goodman-Bacon and Sant'Anna (2021) explore identification challenges when the treatment variable in a difference-in-differences design is continuous, as is the case with all estimators examined in this paper (see also Roth et al., 2023, for a comprehensive review of recent advances in the econometrics of difference-in-differences designs). The issues discussed in this paper complement and differ from those addressed in the studies above.

The paper is structured as follows. Section 2 introduces the setup and defines the causal effects of interest. Sections 3 and 4 analyze the effective minimum wage and the fraction affected estimators, respectively. Section 5 presents the simulation exercise based on Brazilian data. The final section concludes with recommendations for researchers interested in measuring the effects of national minimum wage changes in contexts without state variation in minimum wage laws.

2 Setup

I consider two-period data generating processes of the following form:

$$\boldsymbol{y}_{r,t} = f\left(mw_t, \boldsymbol{\theta}_{r,t}\right) \tag{1}$$

where $r \in \{1, \ldots, R\}$ denotes region, $t \in \{0, 1\}$ denotes time, $\boldsymbol{y}_{r,t}$ is a vector of equilibrium outcomes (such as employment to population ratio or quantiles of the log wage distribution), mw_t denotes the logarithm of the national minimum wage, $\boldsymbol{\theta}_{r,t}$ is a vector of region-time-specific parameters, and f is a function that depends on the particular economic model (to be described later in the paper). Minimum wage levels are different between periods. Without loss of generality, I assume henceforth that $mw_1 > mw_0$.

Equation (1) is fairly general, but it imposes important constraints that limit the scope of my analysis. First, because equilibrium outcomes in one region cannot affect outcomes in another, this model cannot be used to study econometric issues such as minimum-wage-induced migration or capital mobility. Second, the restriction to two periods means that I do not address econometric problems arising from the combination of treatment effect heterogeneity and differential timing of treatment assignment (see de Chaisemartin and D'Haultfœuille, 2020, for a discussion of this issue). Third, I also do not discuss the econometric difficulties associated with distinguishing between short- and long-run minimum wage effects (Sorkin, 2015; Hurst et al., 2022; Vogel, 2023). Thus, this paper does not aim at providing a comprehensive account of econometric issues affecting minimum wage designs; instead, the contribution is pointing out some previously undocumented concerns.

The motivation for my study are papers that estimate the aggregate effects of a national minimum wage, like the ones cited in the beginning of the introduction. Given this purpose and the data-generating process described above, there are two natural ways to define the ceteris paribus causal effects to be identified:

$$\begin{aligned}
\boldsymbol{ATE}_0 &= \frac{1}{R}\sum_r f\left(mw_1, \theta_{r,0}\right) - f\left(mw_0, \theta_{r,0}\right) \\
&= \frac{1}{R}\sum_r f\left(mw_1, \theta_{r,0}\right) - \boldsymbol{y}_{r,0} \\
\boldsymbol{ATE}_1 &= \frac{1}{R}\sum_r f\left(mw_1, \theta_{r,1}\right) - f\left(mw_0, \theta_{r,1}\right) \\
&= \frac{1}{R}\sum_r \boldsymbol{y}_{r,1} - f\left(mw_0, \theta_{r,1}\right)
\end{aligned}$$

The first considers a counterfactual where the minimum wage is increased to the level valid in period 1, keeping other characteristics constant at their $t = 0$ levels. The other compares the outcomes as of $t = 1$ to a counterfactual scenario where the minimum wage remained at the $t = 0$ level. The two definitions are identical if the minimum wage is the only time-varying component of the model, and are equal in expectation if the parameters $\boldsymbol{\theta}_{r,t}$ have the same distribution across regions in both periods. This is true for some of the models I study, but not all. To keep the analysis as simple as possible, I use the average of these two definitions as the "true" causal effects to be recovered by the econometric designs:

$$\boldsymbol{ATE} = \frac{\boldsymbol{ATE}_0 + \boldsymbol{ATE}_1}{2}$$

which gives equal weight to all observations in the data, and is thus closely related to the usual definition of average treatment effects in potential outcomes models.

3 The effective minimum wage design

3.1 Definition

Let $w_{q,r,t}$ denote quantile q of the log wage distribution in region r at time t. The log effective minimum wage is defined as $mw_t - w_{p,r,t}$, where p is a benchmark quantile. For example, if $p = 0.5$, then the effective minimum wage is the minimum wage relative to the median wage in region r at time t. The effect of the minimum wage on the shape of the wage distribution is recovered by running separate regressions for different quantiles q of the wage distribution:

$$w_{q,r,t} - w_{p,r,t} = \alpha_{q,r}^{Eff} + \delta_{q,t}^{Eff} + \beta_q^{Eff}\left[mw_t - w_{p,r,t}\right] + \gamma_q^{Eff}\left[mw_t - w_{p,r,t}\right]^2 + \varepsilon_{q,r,t}^{Eff} \quad (2)$$

The average treatment effect on quantile q of the log wage distribution can be calculated by multiplying the average change in the "effective minimum wage" $mw_t - w_{p,r,t}$ by β_q^{Eff}, then multiplying the average change in the squared effective minimum wage by γ_q^{Eff}, and finally adding up both products. Note that each q corresponds to a separate regression.

The effective minimum wage design has been also used to estimate employment effects; Engbom and Moser (2022) is one example. In this case, the model is identical to Equation (2), except that the left-hand side contains the employment to population ratio. These regressions follow in the tradition of earlier papers that used variation in wage levels to measure employment effects of minimum wages. A well-known example is Neumark and Wascher (1992), who use as the treatment variable the nominal minimum wage in a state-year multiplied by the state-specific minimum wage coverage and divided by the state-specific average wage. In my analysis, I focus on the quantile-based effective minimum wage because it is more common in recent work.

To state the identification assumptions emphasized by Lee (1999), it is useful to introduce some of his notation. Suppose that each region has a latent distribution of log wage in each period—that is, the distribution of log wages that would prevail with no minimum wage regulation in place. Assume further that the cumulative distribution function for those latent log wages have the form:

$$F_t\left(\frac{w - \mu_{r,t}}{\sigma_{r,t}}\right)$$

where $\mu_{r,t}$ and $\sigma_{r,t}$ are the *centrality* and *dispersion* parameters, respectively.

Using this notation, Lee (1999) emphasizes two identification assumptions. First, the statistic used to calculate the effective minimum wage—that is, $w_{p,r,t}$ above—should provide an approximation for the centrality parameter $\mu_{r,t}$. That is why he chooses the median, $p = 0.5$ (and so do **?**, in their reassessment of Lee's estimates) Second, the location and dispersion parameters should be uncorrelated across regions conditional on r. The central message of my analysis below is that, for economic reasons, it may be hard to find applications where both of these assumptions are satisfied.

In the following subsections, I begin with the estimator based on the median. Later on I discuss the use of other quantiles to construct the effective minimum wage.

3.2 Issue #1: Correlated measurement error in the centrality measure

Whenever the minimum wage causes employment effects or wage spillovers extending to the middle of the wage distribution, the median wage will not be a perfect measure of the centrality of the latent wage distribution. Below, I argue that this measurement error is not classic and may lead to point estimates that have the opposite sign compared to the true causal effects.

This issue can be illustrated using a simple model with two regions. Suppose that latent log wages in region r are uniformly distributed between $-L_r$ and L_r, with $L_A > L_B > 0$, in both periods. The minimum wage is at $-L_A$ in period $t = 0$ and moves to $-L_B$ in the next period. It only causes truncation. That means the observed log wage distribution in r,t is $U(\max(-L_r, mw_t), L_r)$.

What would one recover using a linear version of the effective minimum wage estimator in this environment? Due to the inclusion of region and time fixed effects, the coefficient of interest is given by the double difference in outcomes divided by the double difference in the effective minimum wage:

$$\begin{aligned}
\beta &= \frac{(m_{A,1} - m_{A,0}) - (m_{B,1} - m_{B,0})}{\left[(mw_1 - w_{0.5,A,1}) - (mw_0 - w_{0.5,A,0})\right] - \left[(mw_1 - w_{0.5,B,1}) - (mw_0 - w_{0.5,B,0})\right]} \\
&= \frac{(m_{A,1} - m_{A,0}) - (m_{B,1} - m_{B,0})}{-\left[(w_{0.5,A,1} - w_{0.5,A,0}) - (w_{0.5,B,1} - w_{0.5,B,0})\right]} \\
&= -\frac{m_{A,1} - m_{A,0}}{w_{0.5,A,1} - w_{0.5,A,0}}
\end{aligned}$$

where $m_{r,t}$ denotes the left-hand side variable in the regression. The second line follows from eliminating the common minimum wage terms. The third line follows from noting that the minimum wage never has any effect on region B in this model.

Now suppose the left-hand side variable is quantile 25 of the observed wage distribution relative to the mean, that is, $m =_{r,t} = w_{0.25,r,t} - w_{0.5,r,t}$:

$$\begin{aligned}\beta &= -\frac{(w_{0.25,A,1} - w_{0.5,A,1}) - (w_{0.25,A,0} - w_{0.5,A,0})}{w_{0.5,A,1} - w_{0.5,A,0}} \\ &= -\frac{(w_{0.25,A,1} - w_{0.25,A,0}) - (w_{0.5,A,1} - w_{0.5,A,0})}{w_{0.5,A,1} - w_{0.5,A,0}}\end{aligned}$$

The denominator is clearly positive, since truncation of the left tail of the wage distribution moves all quantiles up. But lower quantiles move more than higher ones. So the numerator is also positive, and thus $\beta < 0$. Thus, the estimate would imply that the minimum wage causes a *decline* in $w_{0,25,r,t}$ relative to the median, which is the opposite of the causal effect in the model.

If the outcome is the employment to population ratio, one would obtain a positive β, suggesting that the minimum wage increases employment. Again, that is the opposite of what happens in the underlying causal model.

Before analyzing whether this issue is empirically relevant, I make two remarks. First, this is not an artifact of the uniform distribution. It is possible to derive similar analytical results using a model where latent log wages are Normally distributed, and where the two regions differ in the standard deviation parameter.[2] Second, disemployment is not crucial for this result. With monopsony power, the minimum wage can cause net inflow of workers earning low wages. That would move the median wage downwards in more treated areas, leading to bias in the opposite direction. Models with strong spillovers reaching into the median in some regions also suffer from correlated measurement error.

Are those biases relevant in empirical applications? The example above is unrealistic in that it does not include any "good" variation for the effective minimum wage to use. Specifically, the ideal variation for this estimator comes from differential *changes* in the location parameter that are uncorrelated with other structural components of the model. In this case, it is clear that the regions whose latent distributions moved "to the right" provide good coun-

[2] In that model, there is always a bias, which is stronger when the initial minimum wage is low (keeping the size of the minimum wage increase constant).

terfactuals for those whose latent distributions moved "to the left," if the national minimum wage were lower. In more realistic scenarios, we should expect the data to contain a mix of "good" and "bad" variation, with biases growing in the share of bad variation.

I end this section with a simulation exercise aimed at assessing the magnitude of those biases in a more realistic scenario. I assume that the underlying distribution of latent log wages is Normal. Each region is described by a vector $[\mu_{r,0}, \sigma_{r,0}, \Delta\mu_r, \Delta\sigma_r]$, which is itself drawn from a multivariate Normal distribution. Each particular simulation exercise makes different assumptions about that meta-distribution of parameters across regions, which are shown at the bottom of the corresponding tables. Appendix A.2 provides further details.

How does the minimum wage affects the economies in this model? I assume that there is a "markdown" parameter $m \in [0,1]$ such that the latent distribution is truncated at $mw_t + \log m$, and censored at mw_t. That is: workers who would earn less than the minimum wage times the markdown are disemployed, and those with latent wages earn above that cutoff but below the minimum earn exactly the minimum wage. That model allows for spillovers arising from mechanical increases in wages even when disemployment effects are relatively small.

The first panel in Table 1 shows a case where the only difference between regions is that they receive different shocks to the location parameter, that is, they have different $\Delta\mu_r$. This model corresponds to the ideal scenario with only "good" variation. Correspondingly, the estimator performs with essentially no bias and very precise estimates.

The other panels show richer scenarios where regions differ in both location and dispersion parameters, and both in initial levels and shocks. All of them satisfy the structural assumptions emphasized by Lee (1999): distributions only differ in location and dispersion parameters, not shape, and the location and dispersion parameters are uncorrelated. Still, the estimators display biases toward stronger spillovers and weaker employment effects.

Note that in Panels B and C, biases are economically meaningful even though the disemployment effects in the model are not dramatic. In all simulations, initial employment levels are close to 100%. Thus, given that the increase in the log minimum wage is of 0.2, the elasticity of employment with respect to the minimum wage is about -0.06 in Panel C. Yet, in that scenario, the measured employment effects are less than half of the average causal effects, and spillovers are over-estimated by more than 50%.

The exercise in Panel D is somewhat stylized, but serves to illustrate that the issues are not dependent on disemployment effects at all. Disemployment effects are ruled out by

Table 1: Effective minimum wage design: good vs. bad variation

	Outcome			
	Emp.	p10 - p50	p25 - p50	p90 - p50
Panel A: Only shocks to location				
Mean causal effect	-0.007	0.014	0.004	-0.002
Effective min. wage	-0.007	0.012	0.004	-0.002
	(0.001)	(0.003)	(0.000)	(0.000)
Panel B: Baseline scenario				
Mean causal effect	-0.010	0.024	0.006	-0.004
Effective min. wage	-0.009	0.029	0.009	-0.010
	(0.003)	(0.016)	(0.009)	(0.017)
Panel C: Larger differences in dispersion				
Mean causal effect	-0.012	0.029	0.008	-0.006
Effective min. wage	-0.005	0.043	0.019	-0.034
	(0.005)	(0.027)	(0.015)	(0.031)
Panel D: Censoring only, decline in average dispersion				
Mean causal effect	0.000	0.033	0.002	-0.000
Effective min. wage	0.000	0.021	0.002	-0.000
	(0.000)	(0.018)	(0.010)	(0.019)

Notes: All panels illustrate scenarios where the only time-varying factor is an increase in the national minimum wage of 20 log points. In Panel A, regions differ only in the shock to the location parameter, $\Delta\mu_r \sim \mathcal{N}(0, 0.2^2)$. In Panel B (Baseline), $\mu_{r,0} \sim \mathcal{N}(0, 0.2^2)$, $\sigma_{r,0} \sim \mathcal{N}(0.5, 0.1^2)$ (truncated below at 0.1), $\Delta\mu_r \sim \mathcal{N}(0, 0.2^2)$, and $\Delta\sigma_r \sim \mathcal{N}(0, 0.1^2)$ (truncated below at $0.1 - \sigma_{r,0}$). Panel C is similar to Panel B, except that the standard deviations of initial dispersion and dispersion shocks are 0.2 instead of 0.1. Panel D is similar to Panel B, except that $\sigma_{r,0} \sim \mathcal{N}(0.6, 0.1^2)$ and $\Delta\sigma_r \sim \mathcal{N}(-0.2, 0.1^2)$, and the model has markdown $m = 0$ instead of $m = 0.7$. Each panel displays average results for 5,000 simulations, each with 50 regions and two periods. For each outcome, the numbers correspond to the mean true ATE across simulations, the mean estimates of causal effects based on the regressions listed on the left, and the average standard error associated with the estimates (in parentheses, clustered at the region level).

imposing a "markdown" of 0, such that there is no truncation, only censoring. In addition, it features increasing dispersion of latent productivity distributions over time, which could be interpreted for example as resulting from skill-biased technical change. That combination is sufficient to introduce economically significant biases in the estimation of spillovers.

3.3 Issue #2: Correlation between location and dispersion parameters

The second assumption emphasized by Lee (1999) is independence between the location and dispersion parameters, $\mu_{r,t}$ and $\sigma_{r,t}$, conditional on t. I start by showing through simulations that this assumption is indeed crucial for unbiased estimation. In the remainder of this subsection, I argue that there are plausible economic reasons why we should expect such

Table 2: Correlation between location and dispersion parameters

	Outcome			
	Emp.	p10 - p50	p25 - p50	p90 - p50
Panel A: Baseline scenario				
Mean causal effect	-0.010	0.024	0.006	-0.004
Effective min. wage	-0.009	0.029	0.009	-0.010
	(0.003)	(0.016)	(0.009)	(0.017)
Panel B: Correlation between location and dispersion = 0.2				
Mean causal effect	-0.009	0.023	0.006	-0.004
Effective min. wage	-0.007	0.048	0.020	-0.034
	(0.003)	(0.015)	(0.008)	(0.017)
Panel C: Correlation between location and dispersion = 0.5				
Mean causal effect	-0.008	0.019	0.005	-0.003
Effective min. wage	-0.004	0.075	0.036	-0.067
	(0.002)	(0.013)	(0.007)	(0.015)

Notes: All panels illustrate scenarios where the only time-varying factor is an increase in the national minimum wage of 20 log points. In Panel A (Baseline), $\mu_{r,0} \sim \mathcal{N}(0, 0.2^2)$, $\sigma_{r,0} \sim \mathcal{N}(0.5, 0.1^2)$ (truncated below at 0.1), $\Delta\mu_r \sim \mathcal{N}(0, 0.2^2)$, and $\Delta\sigma_r \sim \mathcal{N}(0, 0.1^2)$ (truncated below at $0.1 - \sigma_{r,0}$). Panel B introduces a correlation of 0.2 between $\mu_{r,0}$ and $\sigma_{r,0}$, and also between $\Delta\mu_r$ and $\Delta\sigma_r$. Panel C increases that correlation to 0.5. Each panel displays average results for 5,000 simulations, each with 50 regions and two periods. For each outcome, the numbers correspond to the mean true ATE across simulations, the mean estimates of causal effects based on the regressions listed on the left, and the average standard error associated with the estimates (in parentheses, clustered at the region level).

correlations to occur.

Table 2 shows that even moderate correlations can introduce economically significant biases. When the correlation is 0.5, the biases are large. With a positive correlation, disemployment effects are underestimated, while spillovers are overestimated.

There are economic reasons why we should expect such correlations in some contexts. One is the observation that, if workers are grouped by observable skill measures such as formal education, higher skill is associated with more within-group inequality. This fact is discussed in detail by Lemieux (2006), who argues that much of the increase in inequality observed in the US from 1973 to 2003 is a compositional effect deriving from increased educational achievement. The same result has been found in other contexts, such as Brazil (Ferreira, Firpo and Messina, 2017). Those differences in skill dispersion can in turn be justified by a range of economic fundamentals, such as larger cross-firm wage premiums for skilled workers, skilled workers placing more value on non-wage amenities, or log-complementarities between intrinsic ability and human capital investments.

This fact is relevant for the effective minimum wage design because there might be regional differences in educational achievement. Regions with more skilled workforces may thus have both higher location and higher dispersion of latent log wages. Furthermore, changes in workforce composition may also be heterogeneous, leading to correlations in *changes*, in addition to in levels.

Regional differences in endowments, leading to heterogeneity in industrial composition, may also generate correlation between location and dispersion parameters. That's because industries—or clusters of connected industries that tend to co-locate—may differ not only on average wages but also on the breadth of occupations and skill levels used in production.

3.4 Issue #3: Confounders

Lee (1999) devotes significant attention to a version of the model without region fixed effects and argues that the decision of whether to include those fixed effects is not an obvious one. It is easy to construct examples where the estimator *without* region fixed effects does perform better because it uses more "good" variation, using the language introduced in Subsection 3.2. However, it is equally easy to contemplate omitted variable bias that can be corrected for using region fixed effects. One example would be that unregistered employment, not visible in the data, is more relevant in low-wage areas, generating a spurious negative correlation between measured employment to population and the effective minimum wage. This is why I focused on the version with region fixed effects.

Indeed, the specifications in recent papers such as Bosch and Manacorda (2010), Autor, Manning and Smith (2016), and Engbom and Moser (2022) go beyond region fixed effects, including region-specific trends. Why would one want to include those trends? The reason is that regions might be differentially affected by supply and demand shocks that affect both the outcome of interest and the median wage in the region, thus introducing omitted variable bias in the effective minimum wage design even if it controls for time-invariant region fixed effects. One example are changes in educational composition. As mentioned in the previous subsection, an increase in educational achievement may boost the median wage while also changing the overall shape and dispersion of the latent wage distribution. That may introduce nontrivial changes in quantiles of the wage distribution which are fundamentally unrelated from the effect of the minimum wage. Because employment and participation rates also vary systematically across educational groups, the measurement of employment effects is also affected. Similar arguments can be made for demand-side shocks such as "the China syn-

drome," which may cause disemployment in some industries and thus simultanously change employment rates, the shape of the local wage distribution, and the median wage (Autor, Dorn and Hanson, 2013).

It is not clear, though, that region-specific trends help. The reason is analogous to Lee's discussion of whether to include region fixed effects: by including the trends, the econometrician may be throwing the "good variation" out with the bathwater.[3] Region-specific trends are helpful if the effects of the confounders on outcomes are well approximated by a linear trends over time (after netting out the national-level time fixed effects), while the minimum wage variation is less linear. If the opposite is true, then the region-specific trends may change the balance of "good" to "bad" variation towards the latter, thus increasing the amount of bias in the estimator.

To evaluate which is more likely to be the case in a given application, the econometrician must have a clear sense of where the "good" variation comes from. Consider the case of the US from the 1980's to recent years, where there are state-specific changes in the minimum wage that do not occur at the same time and that are not well approximated by a linear trend. This is a good source of variation, provided policymakers are not setting minimum wages based on concurrent or expected changes in employment or the shape of the wage distribution. Now consider scenarios where the minimum wage is set at the federal level only. Given that the time fixed effects absorb much of the variation coming from the national minimum, the good variation exploited by the effective minimum wage can only come from "pure" region-specific shocks the location parameters $\mu_{r,t}$. In such cases, a region-specific trend is only helpful if there is reason to believe that those shocks are significantly non-linear over time, while the influence of counfounding shocks (e.g., demand shocks that simultaneously change median wages and the shape of the distribution, or simultaneously change median wages and employment) are approximately linear.[4]

Where the minimum wage is set at the national level only, one may be tempted to abandon the time fixed effects so that the estimator exploits more "good" variation coming from the national changes. Lee (1999) explains why this choice may be unwise in the presence of secular trends affecting the shape of the latent distribution of wages, such as those coming from technical change or international trade. Table 3 makes the problem concrete. As estimator

[3] Lee writes: "*... the reduced identifying variation resulting from eliminating the "permanent" state effects may magnify biases due to misspecification, in the same way biases stemming from measurement error in the independent variable are magnified when true variation in the independent variable is reduced.*"

[4] I discuss this issue further, in the context of Brazil, in Haanwinckel (2023). See Appendix D.6.4.

Table 3: Should one drop the time fixed effects?

	Outcome			
	Emp.	p10 - p50	p25 - p50	p90 - p50
Panel A: Baseline scenario				
Mean causal effect	-0.010	0.024	0.006	-0.004
Effective min. wage	-0.009	0.029	0.009	-0.010
	(0.003)	(0.016)	(0.009)	(0.017)
Effective min. wage, no time FE	-0.011	0.029	0.008	-0.007
	(0.002)	(0.011)	(0.006)	(0.012)
Panel B: Dispersion declines over time				
Mean causal effect	-0.011	0.026	0.007	-0.004
Effective min. wage	-0.007	0.022	0.006	-0.006
	(0.003)	(0.018)	(0.010)	(0.019)
Effective min. wage, no time FE	-0.001	0.139	0.072	-0.137
	(0.002)	(0.019)	(0.010)	(0.021)

Notes: All panels illustrate scenarios where the only time-varying factor is an increase in the national minimum wage of 20 log points. Each panel displays average results for 5,000 simulations, each with 50 regions and two periods. For each outcome, the numbers correspond to the mean true ATE across simulations, the mean estimates of causal effects based on the regressions listed on the left, and the average standard error associated with the estimates (in parentheses, clustered at the region level). The baseline scenario is described in the notes to Tables 1. The scenario in Panel B has average dispersion falling from 0.6 to 0.4.

with no time fixed effects may perform better when the environment is stable. However, if there is a secular trend in average dispersion, the estimator with no time fixed effects displays much more significant biases than the one that includes them.

3.5 Is the standard diagnostic test in the literature an effective tool?

Lee (1999) proposes estimating the effects on "high" quantiles $q > p$ to validate the model. The justification for that approach is that, in many applications (such as in the US), the econometrician may have a strong prior that the minimum wage should have minimal effects on the upper tail of the wage distribution. Autor, Manning and Smith (2016) use the same specification test to validate their instrumental variables implementation of the effective minimum wage design.

This test is not perfect because it is subject to both false positives and false negatives. False positives—detecting a problem where none exists—may arise because there are many plausible mechanisms that could lead to minimum wage spillovers that extend beyond the median wage. Engbom and Moser (2022), for example, develop and estimate an on-the-job search model where minimum wages cause spillovers that extend far into the upper tail of the wage

distribution, primarily due to worker reallocation from low- to high-wage firms. In Haanwinckel (2023), I argue that those spillovers can be compounded by endogenous changes in within-firm returns to skill in response to reallocation flows. Notably, those channels may be quantitatively important even when net disemployment effects are small. Thus, an econometrician with a strict rejection rule based on effects in the upper tail may reject a model that is valid.

There are two concerns regarding false negatives. One is that the estimator may be biased in the lower tail but not in the upper tail. One example is provided in Table 1, Panel D. Still, in the simulations shown above, an estimator that is biased for employment and lower-tail spillovers also detects spillovers in the upper tail. That leads to the second concern. Even if such upper tail spillovers are detected, the econometrician may still interpret those results as a negative, if they believe that such spillovers are economically plausible.[5] That scenario is the opposite from the one mentioned above. The fundamental issue is that, because theoretical predictions are ambiguous, there is typically not a consensus on whether we should expect upper-tail spillovers or not in any given context. That precludes a consensus on the correct heuristics for interpreting those tests.

3.6 Does using a higher quantiles as benchmark wages help?

In some applications, the econometrician may have a strong prior that the minimum wage may have significant effects on the median wage, making it a poor measure of centrality. In those cases, the researcher may consider using a higher quantile of the wage distribution to construct the effective minimum wage. For example, Bosch and Manacorda (2010) use $p = 0.7$ in a study of Mexico, and Engbom and Moser (2022) use $p = 0.9$ when studying Brazil. The potential problem with that strategy is that higher quantiles may be more sensitive to differences in the dispersion of latent productivities.

Table 4 shows simulation results for the previous models using the alternative effective minimum wage estimator based on percentile 90 of the wage distribution. As before, Panel A—containing the "ideal" scenario for the estimator—yields unbiased results. For all others, there are economically significant biases. Indeed, in all cases, the biases are larger than when using the estimator based on the median.

[5] Engbom and Moser (2022) provide one example.

Table 4: Effective minimum wage based on percentile 90

	Outcome			
	Emp.	p10 - p50	p25 - p50	p90 - p50
Panel A: Only shocks to location				
Mean causal effect	-0.007	0.014	0.004	-0.002
Effective min. wage, p90	-0.007	0.012	0.004	-0.002
	(0.001)	(0.003)	(0.001)	(0.000)
Panel B: Baseline scenario				
Mean causal effect	-0.010	0.024	0.006	-0.004
Effective min. wage, p90	-0.003	0.066	0.032	-0.060
	(0.003)	(0.021)	(0.017)	(0.000)
Panel C: Larger differences in dispersion				
Mean causal effect	-0.012	0.029	0.008	-0.006
Effective min. wage, p90	0.006	0.089	0.050	-0.103
	(0.003)	(0.018)	(0.015)	(0.000)
Panel D: Censoring only, decline in average dispersion				
Mean causal effect	0.000	0.033	0.002	-0.000
Effective min. wage, p90	0.000	0.154	0.071	-0.131
	(0.000)	(0.058)	(0.048)	(0.000)
Panel E: Correlation between location and dispersion = 0.2				
Mean causal effect	-0.009	0.023	0.006	-0.004
Effective min. wage, p90	-0.002	0.070	0.034	-0.065
	(0.003)	(0.017)	(0.014)	(0.000)
Panel F: Correlation between location and dispersion = 0.5				
Mean causal effect	-0.008	0.019	0.005	-0.003
Effective min. wage, p90	-0.001	0.073	0.037	-0.071
	(0.002)	(0.012)	(0.010)	(0.000)

Notes: All panels illustrate scenarios where the only time-varying factor is an increase in the national minimum wage of 20 log points. Each panel displays average results for 5,000 simulations, each with 50 regions and two periods. For each outcome, the numbers correspond to the mean true ATE across simulations, the mean estimates of causal effects based on the regressions listed on the left, and the average standard error associated with the estimates (in parentheses, clustered at the region level). See the notes for Tables 1 and 2 for a description of each panel.

3.7 Does the instrumental variables approach of Autor, Manning and Smith (2016) help?

Autor, Manning and Smith (2016) discuss how the effective minimum wage design is subject to bias whenever there is imperfect measurement of the median wage. That mechanical issue arises from the fact that the median wage is used to construct both the left- and right-hand side variables in the regressions. As a solution, they propose an instrumental variables estimator that leverages variation in statutory minimum wages as instruments for the effective minimum wage. That instrumental variables estimator can only include time fixed effects if statutory minimum wages are set at the region level and change over time.

In unreported simulations (to be included in a future version of this paper), I find that the instrumental variables estimator can reduce the biases discussed above. The reason is that it discards much of the "bad" variation introduced by correlated measurement error, as discussed above. However, I have also discussed how dropping time fixed effects is likely to introduce significant bias in the estimator. Thus, effective use of the instrumental variables approach is only warranted in contexts with a significant amount of region-level variation in minimum wage laws.

3.8 Taking stock

The core message of this discussion is that, to evaluate whether the effective minimum wage strategy is likely to be successful, the econometrician should have a clear sense of where the identifying variation is coming from. Ideally, one needs an underlying shifter of the effective minimum wage that is arguably exogenous. For the measurement of spillovers, that means the shifter must change the effective minimum wage but have no effect on the shape or dispersion of the latent wage distribution that is not driven by the minimum wage itself. For the measurement of employment effects, the shifter should have no independent effect on employment rates.

It is clear how exogenous changes in state-level minimum wages satisfy those requirements. Even in those cases, the estimator might have some bias because it uses both "good" and "bad" variation to identify the causal effects. This is why the instrumental variables approach of Autor, Manning and Smith (2016) may be helpful.

If there is not a significant amount of state-level variation in minimum wage laws, then the underlying shifters for the minimum wage need to act through the median wage. It needs to

be a local labor market shock that shifts the location of the latent productivity distribution but has no independent effects on shape, dispersion, or employment. If the econometrician cannot think of what constitute such a shifter, the effective minimum wage design may not be warranted. And if the econometrician does have a clear sense of what is that shifter, then an instrumental variables approach or differences-in-differences estimator directly exploiting that shifter may be a more adequate tool.

4 Fraction affected and Gap estimators

4.1 Definition

Now I study a difference-in-differences model with a time-invariant, continuous measure of treatment intensity based on the initial distribution of wages:

$$y_{o,r,t} = \alpha^{FA}_{o,r} + \delta^{FA}_{o,t} + \beta^{FA}_{o} FA_r \cdot \mathbf{1}\{t = 1\} + \varepsilon^{FA}_{o,r,t} \tag{3}$$

where the subscript o indexes the specific equilibrium outcome, such that different o correspond to separate regressions. The treatment intensity variable FA_r is the "fraction affected," that is, the share of workers in the initial period earning less than mw_1. The α parameters are region fixed effects intended to capture unobserved heterogeneity across regions that may have independent effects on the outcomes. The time fixed effects δ capture shocks concurrent with the minimum wage hike that affect all regions in the same way.

Given the linearity of this model, the estimated average treatment effect for outcome o is given by the product of the regional average of the treatment variable and the β^{FA}_{o} parameter.

The fraction affected design was first introduced by Card (1992), in an analysis of the 1990 increase in the federal minimum wage in the US. Card emphasizes that much of the identifying variation in his application comes from significant heterogeneity in the bindingness of state-level minimum wages in the preceding years. Since that original application, that estimator has been applied in other contexts with no regional variation in nominal minimum wages, such as the introduction of a federal minimum wage in Germany in 2015 (Ahlfeldt, Roth and Seidel, 2018; Fedorets and Shupe, 2021).

This design is typically employed in scenarios where there is a marked increase in the minimum wage after at least a few years without adjustments. In those cases, the econometrician can use pre-treatment data to check for differential trends, which may provide support for

the parallel trends identification assumption.

I also study a closely related design based on the "Gap measure:"

$$y_{o,r,t} = \alpha_{o,r}^{Gap} + \delta_{o,t}^{Gap} + \beta_{o}^{Gap} Gap_r \cdot \mathbf{1}\{t = 1\} + \varepsilon_{o,r,t}^{Gap}$$
$$Gap_r = \frac{\sum_{i=1}^{I_r} \max\{\exp(mw_1) - \exp(w_{i,0}), 0\}}{\sum_{i=1}^{I_r} \exp(w_{i,0})}$$

where $i \in \{1, \ldots, I_r\}$ indexes workers in the initial period and $w_{i,0}$ is their log wage in that period.

That definition of treatment intensity was introduced in Card and Krueger (1994) in a firm-level econometric design. It has later been extended to region-level designs like the ones studied in this paper. Dustmann et al. (2021) provides an example, again in the context of Germany. It corresponds to the resulting relative increase in average wage in each region if all low-wage workers were to receive raises to comply with the new minimum.

It is worth mentioning that this design is subject to concerns described by Callaway, Goodman-Bacon and Sant'Anna (2021). In that paper, the authors study difference-in-difference estimators with a continuous treatment variable. They describe how the parallel trends assumption needs to be stronger in those models. They also point out that, even when their "strong parallel trends" assumption holds, the OLS estimator with region and time fixed effects will not yield the average treatment effect, but instead a weighted average of treatment effects with more weights for intermediate values of the treatment variable. They discuss alternative estimators that may be useful for correcting that problem when the "strong parallel trends" assumption holds.

4.2 Issue #1: Correlated measurement error

The fraction affected and the gap designs do not neatly correspond to reduced forms of typical economic models of minimum wages. The deviations between the underlying model and the regressions may not necessarily be captured solely in the residual term. They may lead to measurement error in the treatment variable, causing bias whose direction may be difficult to predict.

For a simple example, consider a regression of the employment to population ratio on the fraction affected, in a scenario where population is constant at the regional level. The regres-

sion can be written in differences as:

$$\frac{\Delta Emp_r}{Pop_r} = \delta + \beta \frac{NA_{r,0}}{Emp_{r,0}} + \varepsilon_r$$

where $NA_{r,t}$ is the count (number) affected. If one multiplies both sides by population, the model is:

$$\Delta Emp_r = \delta Pop_r + \beta NA_{r,0} \left(\frac{Emp_{r,0}}{Pop_r} \right)^{-1} + \varepsilon_r Pop_r$$

If one ignores the term in parenthesis, this looks like a very reasonable model. The change in employment count is given by a common trend in the share of the population who works (the α term); the minimum wage causes an effect that is on average proportional to the number of affected workers (for example, 1 out of 20 affected workers lose their jobs if $\beta = -0.05$); and there is a random inflow or outflow of workers at each region, proportional to the population.

However, the presence of the term in parenthesis implies that in the fraction affected model, employment effects of the minimum wage are proportional to number of affected workers divided by the initial employment to population ratio. If the true causal model is more similar to the one described above—that is, without the term in parenthesis—then the fraction affected regression uses a regressor with non-classic, non-mean-zero measurement error. The correct regressor should be $NA_{r,0}/Pop_r$, not $NA_{r,0}/Emp_{r,0}$. Using the incorrect regressor will lead to bias.

The solution is obvious in this example. But in more general settings, which regressor to use may depend on the specific outcome variable and assumptions on the underlying economic model. However, one of the most salient advantages of not using a reduced-form approach such as the fraction affected design is precisely not having to take a stance on the economic model. One way to compatibilize this goal with the measurement error issues is to acknowledge that estimates coming from those designs may be more imprecise than what the confidence intervals coming from the regression suggest.

Table 5 illustrates this issue with simulations of models that should be ideal applications for the fraction affected and gap designs. That's because in those simulations, regions only differ in the location parameter of the latent log wage distribution, assumed to be constant over time. The only time-varying factor in the model is the minimum wage. The effects of the minimum wage are a combination of truncation and censoring, exactly as in the simulations run in the previous section of the paper (see Appendix A.2 for details).

Table 5: Correlated measurement error in the fraction affected and gap designs

	Outcome				
	Emp.	p10	p25	p50	p90
Panel A: Initial min. wage is small, $m = 0.7$					
Mean causal effect	-0.006	0.017	0.007	0.004	0.002
Fraction affected	-0.008	0.026	0.009	0.005	0.002
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)
Gap measure	-0.006	0.020	0.007	0.004	0.002
	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)
Panel B: Initial min. wage is large, $m = 0.7$					
Mean causal effect	-0.055	0.166	0.072	0.035	0.017
Fraction affected	-0.069	0.099	0.118	0.046	0.024
	(0.001)	(0.018)	(0.008)	(0.001)	(0.001)
Gap measure	-0.048	0.062	0.084	0.032	0.017
	(0.001)	(0.014)	(0.005)	(0.001)	(0.000)
Panel C: Initial min. wage is small, $m = 0.9$					
Mean causal effect	-0.018	0.039	0.021	0.012	0.005
Fraction affected	-0.019	0.034	0.020	0.012	0.006
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Gap measure	-0.014	0.025	0.015	0.009	0.004
	(0.001)	(0.002)	(0.001)	(0.000)	(0.000)
Panel D: Initial min. wage is large, $m = 0.9$					
Mean causal effect	-0.095	0.139	0.094	0.063	0.033
Fraction affected	-0.080	0.085	0.071	0.059	0.036
	(0.004)	(0.005)	(0.003)	(0.001)	(0.000)
Gap measure	-0.057	0.061	0.051	0.043	0.026
	(0.005)	(0.005)	(0.004)	(0.002)	(0.001)

Notes: In all panels, the national minimum wage increases by 20 log points from the first period to the second. Regions differ only in the time-invariant location parameter $\mu_r \sim \mathcal{N}(0, 0.2^2)$. Each panel displays average results for 5,000 simulations, each with 50 regions. For each outcome, the numbers correspond to the mean true ATE across simulations, the mean estimates of causal effects based on the regressions listed on the left, and the average standard error associated with the estimates (in parentheses, clustered at the region level).

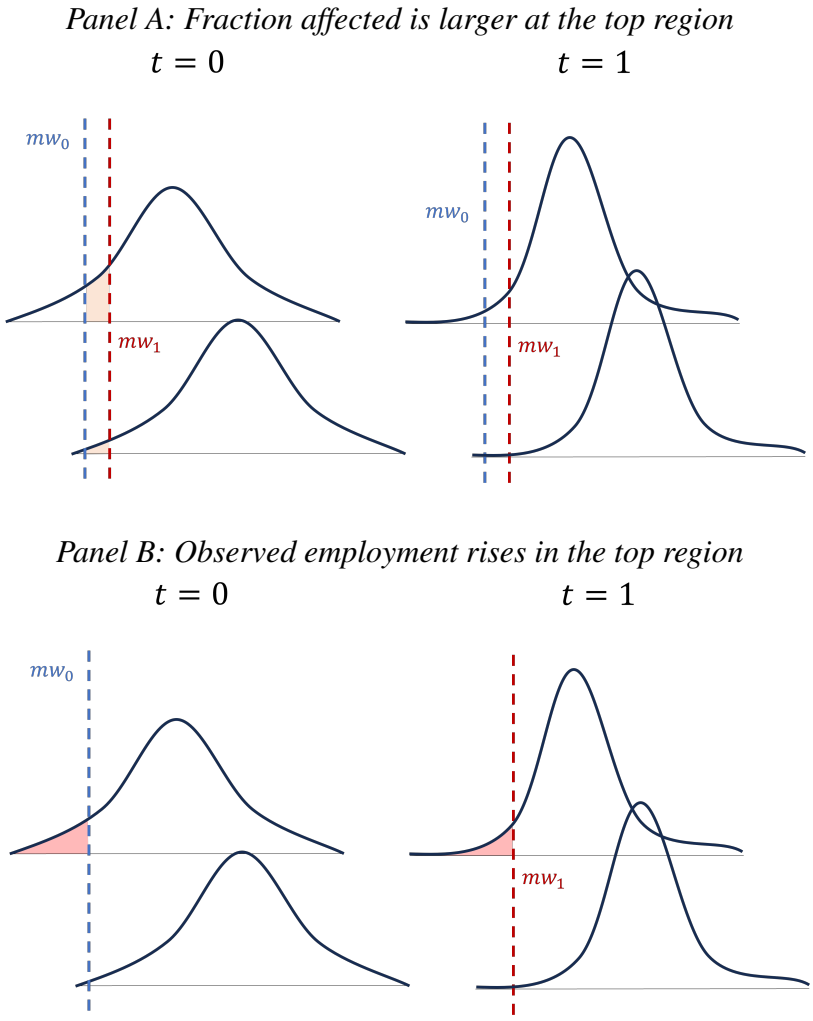


Figure 1: A truncation model with a secular decline in latent log wage dispersion

The simulations show that, in general, the fraction affected and gap designs have biases even when the employment effects of the minimum wage are small. There is no clear pattern on what is the direction of the bias for each outcome or estimator. In fact, the direction of the bias may change depending on the size of the markdown in the structural model (which determines the extent of censoring versus truncation effects). The biases can be significant when the minimum wage generates large disemployment effects.

4.3 Issue #2: Trends in the dispersion of productivity

Consider the model illustrated in Figure 1. Regions differ in a location parameters μ_r that is constant over time. There is a common shape parameter σ_t that decreases over time. Suppose

Table 6: Trend in the dispersion of latent log wages

	Outcome				
	Emp.	p10	p25	p50	p90
Panel A: Baseline scenario					
Mean causal effect	-0.010	0.032	0.013	0.007	0.004
Fraction affected	-0.015	0.053	0.021	0.012	0.006
	(0.006)	(0.025)	(0.028)	(0.030)	(0.038)
Gap measure	-0.011	0.041	0.016	0.009	0.004
	(0.005)	(0.019)	(0.021)	(0.022)	(0.029)
Panel B: Declining dispersion of latent log wages					
Mean causal effect	-0.011	0.035	0.015	0.008	0.004
Fraction affected	0.004	-0.007	-0.010	-0.006	-0.005
	(0.007)	(0.041)	(0.042)	(0.043)	(0.053)
Gap measure	0.002	-0.003	-0.006	-0.004	-0.003
	(0.005)	(0.029)	(0.031)	(0.031)	(0.039)

Notes: All panels illustrate scenarios where the only time-varying factor is an increase in the national minimum wage of 20 log points. Each panel displays average results for 5,000 simulations, each with 50 regions and two periods. For each outcome, the numbers correspond to the mean true ATE across simulations, the mean estimates of causal effects based on the regressions listed on the left, and the average standard error associated with the estimates (in parentheses, clustered at the region level). See the notes for Tables 1 and 2 for a description of Panels A and B.

that the minimum wage change is not that large, such that the new minimum is always to left of the peak of the latent log wage distribution. All else equal, regions with low μ_r a higher fraction affected and a higher causal effect of the minimum wage. But with the fall in σ_t, these regions may actually see a relative *increase* in employment. That is because as latent wages become more concentrated around the median, they may become less truncated. This effect may be negligible for regions with high μ_r. As a result, the estimator may recover a positive relationship between the fraction affected and employment, despite the causal effects are negative for all observations in all periods.

Table 6 shows simulations that explore this issue. In the baseline scenario shown in Panel A, the models perform reasonably well (the reasons for the biases in this Panel are described in the previous Subsection 4.2). In Panel B, the average dispersion is falling. That is sufficient to flip the sign of the employment results, consistent with the logic described above.

Th model just described satisfies the "parallel trends" assumption in the sense that there is no differential trend in structural parameters between more and less treated regions. However, one may argue that there is no clear reason why dispersion parameters should change exactly at the same time of a change in the minimum wage. If the trend in dispersion is a more long-

run phenomenon, having started before the minimum wage change, then it is possible that a standard test for differential pre-trends would detect an issue. That's because, if the initial minimum wage (assumed to be constant over the pre-period) is binding in low-wage regions, then the change in dispersion should be reflected in changing employment in the pre-period as well.

There is, however, one reason why we should expect changes in the shape of the latent wage distribution happening at the same time as the minimum wage. I turn to this issue in the next subsection.

4.4 Issue #3: Endogenous responses in returns to skill

Consider a competitive economy with two types of labor, low- or high-skill. Each worker is characterized not only by their type but also by their amount of efficiency units of labor. Wages are determined by marginal products of labor, which are determined by a representative production function with constant returns and constant elasticity of substitution. This is the basic setup in Katz and Murphy (1992).

Now consider the implications of a binding minimum wage. The representative firm will not employ workers whose productivity is below the minimum wage. Thus, one can think of this model as one where each worker has a latent log wage given by the sum of the log price of the efficiency unit of their type (low- or high-skill) and the log of their amount of efficiency units. The observed wage distribution is the truncated version of the latent one. See Appendix A.1 for details on the model.

The key difference compared to the model presented before is that the latent log wage distribution responds to the minimum wage. As the minimum wage causes more disemployment for low-skill workers, returns to skill are expected to fall. That generates wage spillovers for other low-skill workers and attenuates the negative disemployment effects.

Table 7 shows simulations of this model. In those simulations, regions only differ in the initial share of workers that belong to the high-skill group, and the only time-varying factor is the minimum wage. When the initial minimum wage is small, then the fraction affected estimator works well. However, as disemployment effects become more significant, biases become larger. Results for the Gap estimator are similar. The underlying reason for these biases is that the endogenous response in returns to skill acts as a reduction in the dispersion of latent productivities, discussed in the previous subsection. But the biases can be exacerbated

Table 7: Canonical model of labor demand

	Outcome				
	Emp.	p10	p25	p50	p90
Panel A: Starting minimum wage is low					
Mean causal effect	-0.006	0.009	0.006	0.004	0.005
Fraction affected	-0.005	0.007	0.005	0.004	0.006
	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Panel B: Starting minimum wage is high					
Mean causal effect	-0.100	0.152	0.129	0.128	0.027
Fraction affected	-0.045	0.052	0.061	0.060	0.044
	(0.009)	(0.007)	(0.006)	(0.007)	(0.009)
Panel C: Starting minimum wage is very high					
Mean causal effect	-0.098	0.190	0.185	0.175	0.020
Fraction affected	0.014	0.028	0.054	0.162	0.026
	(0.009)	(0.005)	(0.009)	(0.019)	(0.001)

Notes: All panels illustrate scenarios where the only time-varying factor is an increase in the national minimum wage of 20 log points. Each panel displays average results for 5,000 simulations, each with 50 regions and two periods. For each outcome, the numbers correspond to the mean true ATE across simulations, the mean estimates of causal effects based on the regressions listed on the left, and the average standard error associated with the estimates (in parentheses, clustered at the region level). Simulations are based on the "canonical" model of labor demand with competitive labor markets and an aggregate production function with constant elasticity of substitution between low- and high-skill. Panels differ in the initial level of the minimum wage. See Appendix A.1 for details on the model.

because the reduction is specific to regions where the fraction affected is largest.

One interesting aspect of Table 7 is that there is an inverse-U relationship between the true causal effects and the estimated ones. A strong prior that employment effects are negligible is sufficient to validate the model. However, an estimated effect of zero cannot be used as independent evidence that the true employment effects are zero.

5 Simulation exercise based on Brazilian data

The previous simulation exercises have two important limitations. First, the causal channels from the minimum wage to employment and wages are rather limited, excluding margins that have been shown to be important such as reallocation of workers from low- to high-wage jobs (Dustmann et al., 2021). Second, they relied on strong functional form assumptions on the shape of the underlying worker productivity distribution and how it varies across regions and over time.

Table 8: Quality of fit of the task-based, monopsonistic model

	Data		Model		
	1998	2012	1998	2012	R2
Moments	(1)	(2)	(3)	(4)	(5)
Wage inequality measures					
Secondary / less than secondary	0.498	0.168	0.486	0.15	0.77
Tertiary / secondary	0.965	1.038	0.995	0.932	0.131
Within less than secondary	0.41	0.241	0.387	0.225	0.575
Within secondary	0.684	0.355	0.647	0.335	0.831
Within tertiary	0.702	0.624	0.69	0.644	0.051
Total variance of log wages	0.715	0.544	0.722	0.504	0.749
Two-way fixed effects decomposition					
Variance establishment effects	0.116	0.056	0.117	0.057	0.652
Covariance worker, estab. effects	0.049	0.048	0.058	0.048	0.421
Variance worker effects	0.419	0.384	0.417	0.301	0.293
Correlation worker, estab. effects	0.224	0.315	0.256	0.361	0.196
Formal employment rates					
Less than secondary	0.266	0.337	0.266	0.336	0.951
Secondary	0.435	0.508	0.435	0.508	1.0
Tertiary	0.539	0.629	0.539	0.631	0.878
Minimum wage bindingness					
Log min. wage - mean log wage	-1.418	-0.922	-1.418	-0.922	1.0
Share < log min. wage + 0.05	0.031	0.053	0.03	0.074	0.696
Share < log min. wage + 0.30	0.086	0.212	0.099	0.218	0.892

Notes: This table is adapted from Haanwinckel (2023). There are three observable education groups: Less than secondary, Secondary, and Tertiary. Columns (1) through (4) report national averages of the corresponding moments for each year, calculated using region weights based on total formal employment. Column (5) reports the usual R2 metric $r_e^2 = 1 - \left[\sum_{t=1}^{2}\sum_{r=1}^{151} s_r \left(Y_{e,r,t} - \hat{Y}_{e,r,t}\right)^2\right] / \left[\sum_{t=1}^{2}\sum_{r=1}^{151} s_r \left(Y_{e,r,t} - \bar{Y}_e\right)^2\right]$, where e indexes the specific target moment, $\hat{Y}_{e,r,t}$ is the model prediction, and $\bar{Y}_e$ is the sample average using the region weights employed in the estimation of the model. See Haanwinckel (2023) for details.

Table 9: Task-based, monopsonistic model with two-sided heterogeneity

	Outcome			
	Emp.	p10 - p50	p25 - p50	p90 - p50
Mean causal effect	-0.046	0.198	0.077	-0.013
Fraction affected	-0.021	-0.093	-0.047	-0.150
	(0.007)	(0.012)	(0.015)	(0.026)
Gap measure	-0.012	-0.050	-0.030	-0.068
	(0.004)	(0.017)	(0.019)	(0.014)
Effective min. wage	-0.017	0.231	0.131	0.049
	(0.013)	(0.012)	(0.015)	(0.041)
Effective min. wage, p90	0.025	0.130	0.052	-0.261
	(0.012)	(0.015)	(0.021)	(0.000)
Effective min. wage, no region FE	-0.085	0.210	0.093	-0.036
	(0.007)	(0.005)	(0.005)	(0.015)
Effective min. wage, no time FE	0.123	0.225	0.130	-0.181
	(0.005)	(0.003)	(0.004)	(0.013)

Notes: This Table compares the true average causal effects of the minimum wage from the model of Haanwinckel (2023) to the predicted average treatment effects from different estimators. The model has 151 regions (corresponding to microregions in Brazil) and two time periods (corresponding to 1998 and 2012).

To address those limitations, I use the rich model estimated in Haanwinckel (2023) as the data-generating process. In that model, workers belong to one of ten comparative advantage types, and also differ in their amount of efficiency units of labor within groups. The quantity of workers of each type is associated with the educational composition of the region at each period. Firms differ in their demand for skills because they use task-based production technologies, and some firms may have higher demand for high-complexity tasks that are better performed by skilled workers. Firms also differ in the wages they pay to each worker type; that is, the model features firm wage premiums. Minimum wages have a range of effects including disemployment (truncation), mechanical wage increases (censoring), between-firm reallocation, changes in returns to skill within firms, in the distribution of firm types (that is, the share of high-wage, high-skill firms versus low-wage, low skill ones), in participation decisions of workers, and in prices for goods. Regions differ in not only in educational composition but also total factor productivity (TFP), three types of labor demand parameters (which may correlate with TFP and initial educational composition), and some parameters regulating employment rates (which capture unobserved heterogeneity related to the Brazilian informal sector).

The model is estimated using Brazilian matched employer-employee data, targeting a large set of moments at the region-time levels. There are 151 regions (corresponding to Brazil-

ian *microregions*) and two time periods (corresponding to 1998 and 2012). Table 8 shows that the estimated model fits several dimensions of the data pretty well: inequality measures between and within educational groups, the contribution of firm wage premiums to inequality (measured using the methodology developed by Kline, Saggio and Sølvsten, 2018), the minimum wage spike, and the share of workers earning up to 30 log points of the minimum wage.

Table 9 presents the true average causal effects of the increase in the minimum wage in this model, along with the estimated average treatment effects using the estimators discussed. I find large biases for all of them. With the exception of the effective minimum wage model without region fixed effects, all of them under-estimate the substantial disemployment effects of the national minimum wage in Brazil (as implied by the model) by at least 50%. The effective minimum wage design based on quantile 0.9 predicts positive and statistically significant effects of 2.5 percentage points in the employment to population rate, while the true effect is a decline of 4.6 percentage points. The estimators also fail to capture the spillover effects on the wage distribution, though the effective minimum wage estimator without region fixed effects performs surprisingly well in recovering *relative changes* in log wage quantiles.

6 Conclusion

In this paper, I showed that commonly used estimators of the effects of national minimum wages may suffer from previously unknown econometric issues. The fraction affected/gap designs typically have biases from measurement error that lead to incorrectly sized confidence intervals. In some cases, there may be additional issues, some of which are not detectable via tests for differential pre-trends. Researchers may consider simulation exercises calibrated to their particular application as a tool to investigate whether the designs are subject to bias and which specific formulation is the most appropriate.

For the effective minimum wage design, I show that the identification assumptions in Lee (1999) are essential but difficult to satisfy. I show that deviations from either assumption lead to severe biases. I also show that variation in dispersion parameters between regions and over time is often enough to introduce biases in the estimator, even if they are uncorrelated with location parameters.

My results based on Brazilian data suggest that, in some applications, neither the fraction affected nor the effective minimum wage designs may be reliable. In particular, those strate-

gies may not work well when the increase in the national minimum wage is large. That may lead to significant equilibrium effects through reallocation of labor between firms, changes in returns to skill, and the composition of firms operating in the economy. Alternatively, the strategies may also not work well when the minimum wage increase occurs smoothly over many years, such that testing for pre-trends is impossible and it becomes more likely that there are trends in latent productivity dispersion in the study period (caused, e.g., by technical change or international trade shocks).

In those cases, researchers may consider two alternative approaches while recognizing that neither is a perfect replacement for a valid quasi-experiment at the local labor market level. One is to use within-region difference-in-differences designs where the unit of analysis is either firms (Harasztosi and Lindner, 2019) or workers (Dustmann et al., 2021), and treatment and control groups are defined based on initial wages. This approach is not perfect because, as I point out in Haanwinckel (2023), there are theoretical reasons to believe that minimum wage may affect those control units, even if aggregate employment effects or general equilibrium effects on prices are minor. It also requires careful handling of bias generated by reversion to the mean. The other alternative is to use rich structural models to predict the effects of the national increase in the minimum wage (Engbom and Moser, 2022; Haanwinckel, 2023). But even if well-executed, the counterfactual exercises in such papers can only measure the effects of causal pathways pre-specified by the researcher. Each potential application should be assessed individually to determine the best strategy.

References

Aaronson, Daniel, Eric French, Isaac Sorkin, and Ted To. 2018. "Industry Dynamics and the Minimum Wage: a Putty-Clay Approach." *International Economic Review*, 59(1): 51–84.

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

Autor, David H., Alan Manning, and Christopher L. Smith. 2016. "The Contribution of the Minimum Wage to US Wage Inequality over Three Decades: A Reassessment." *American Economic Journal: Applied Economics*, 8(1): 58–99.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic*

Review, 103(6): 2121–68.

Bailey, Martha J., John DiNardo, and Bryan A. Stuart. 2021. "The Economic Impact of a High National Minimum Wage: Evidence from the 1966 Fair Labor Standards Act." *Journal of Labor Economics*, 39(S2): S329–S367.

Bosch, Mariano, and Marco Manacorda. 2010. "Minimum Wages and Earnings Inequality in Urban Mexico." *American Economic Journal: Applied Economics*, 2(4): 128–49.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with a Continuous Treatment."

Card, David. 1992. "Using Regional Variation in Wages to Measure the Effects of the Federal Minimum Wage." *Industrial and Labor Relations Review*, 46(1): 22–37.

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

Card, David, and Thomas Lemieux. 1996. "Wage dispersion, returns to skill, and black-white wage differentials." *Journal of Econometrics*, 74(2): 319–361.

Corseuil, Carlos Henrique, Miguel Foguel, and Marcos Hecksher. 2015. "EFEITOS DOS PISOS SALARIAIS ESTADUAIS SOBRE O MERCADO DE TRABALHO: UMA NOVA ABORDAGEM EMPÍRICA." *Economia Aplicada*, 19(1): 131–169.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review*, 110(9): 2964–96.

Dinkelman, Taryn, and Vimal Ranchhod. 2012. "Evidence on the impact of minimum wage laws in an informal sector: Domestic workers in South Africa." *Journal of Development Economics*, 99(1): 27–45.

Dustmann, Christian, Attila Lindner, Uta Schönberg, Matthias Umkehrer, and Philipp vom Berge. 2021. "Reallocation Effects of the Minimum Wage*." *The Quarterly Journal of Economics*, 137(1): 267–328.

Engbom, Niklas, and Christian Moser. 2022. "Earnings Inequality and the Minimum Wage: Evidence from Brazil." *American Economic Review*, 112(12): 3803–47.

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

Ferreira, Francisco H. G., Sergio P. Firpo, and Julián Messina. 2017. "Ageing Poorly?: Accounting for the Decline in Earnings Inequality in Brazil, 1995-2012." Inter-American Development Bank IDB Publications (Working Papers) 8220.

Haanwinckel, Daniel. 2023. "Supply, Demand, Institutions, and Firms: A Theory of Labor Market Sorting and the Wage Distribution." NBER Working Paper 31318.

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

Hurst, Erik, Patrick Kehoe, Elena Pastorino, and Thomas Winberry. 2022. "The Distributional Impact of the Minimum Wage in the Short and Long Run." *Unpublished manuscript.*

Katz, Lawrence F, and Kevin M Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *The Quarterly Journal of Economics*, 107(1): 35–78.

Kline, P., R. Saggio, and M. Sølvsten. 2018. "Leave-out estimation of variance components." *ArXiv e-prints*.

Lee, David S. 1999. "Wage Inequality in the United States during the 1980s: Rising Dispersion or Falling Minimum Wage?" *The Quarterly Journal of Economics*, 114(3): 977–1023.

Lemieux, Thomas. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review*, 96(3): 461–498.

Manning, Alan. 2004. "We Can Work It Out: The Impact of Technological Change on the Demand for Low-Skill Workers." *Scottish Journal of Political Economy*, 51(5): 581–608.

Moura, Rodrigo Leandro de, and Marcelo Côrtes Neri. 2008. "Efetividade do "salário mínimo estadual": uma análise via regressões quantílicas para dados longitudinais." *Economia Aplicada*, 12(2): 239–273.

Neumark, David, and William Wascher. 1992. "Employment Effects of Minimum and Subminimum Wages: Panel Data on State Minimum Wage Laws." *Industrial and Labor Relations Review*, 46(1): 55–81.

Rohlin, Shawn M. 2011. "State minimum wages and business location: Evidence from a refined border approach." *Journal of Urban Economics*, 69(1): 103–117.

Roth, Jonathan, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's trending in difference-in-differences? A synthesis of the recent econometrics literature." *Journal of Econometrics*, 235(2): 2218–2244.

Sorkin, Isaac. 2015. "Are there long-run effects of the minimum wage?" *Review of Economic Dynamics*, 18(2): 306–333.

Vogel, Jonathan. 2023. "The Race Between Education, Technology, and the Minimum Wage." NBER Working Paper 31028.

Appendix

A Simulations

This appendix lists the parameters used in all simulations. Every simulation exercise is repeated 5,000 times. The increase in the log minimum wage is always of 0.2.

A.1 The canonical model of labor demand

Consider a competitive economy where the only production factors are skilled ($i = 1$) and skilled ($i = 2$) labor, both of which have inelastic supply. A representative firm produces the numeraire good in the economy using a constant elasticity of substitution (CES) production function:

$$F(L_1, L_2) = \left[\alpha {L_1}^{\frac{E-1}{E}} + (1-\alpha){L_2}^{\frac{E-1}{E}}\right]^{\frac{E}{E-1}}$$

The measure of workers is normalized to 1, and the region-specific share of skilled workers is s_r. Each worker supplies $\exp(e)$ efficiency units of labor, where e has a Normal distribution with mean zero and standard deviation D. Workers whose log marginal product of labor falls below the log minimum wage mw_t are not employed by the representative firm. That is, the minimum wage truncates the worker productivity distribution. The equilibrium log prices per efficiency unit of labor, p_i, are then given by the solution to a system of two equations:

$$\begin{aligned} p_{i,r,t} &= \log F_i\left(s_r E_t(p_{1,r,t}), (1-s_r)E_t(p_{2,r,t})\right) \qquad i \in \{1,2\} \\ \text{where } E_t(p) &= \int_{mw_t - p}^{\infty} \exp(e)\phi\left(\frac{e}{D}\right) de \\ \text{and } F_i(L_1, L_2) &= \frac{dF(L_1, L_2)}{dL_i} \end{aligned}$$

In the expressions above, ϕ denotes the density of a standard Normal distribution.

The resulting employment to population ratio in a given region and period is given by:

$$emp_{r,t} = s_r\left[1 - \Phi\left(\frac{mw_t - p_{1,r,t}}{D}\right)\right] + (1-s_r)\left[1 - \Phi\left(\frac{mw_t - p_{2,r,t}}{D}\right)\right]$$

And the corresponding cumulative distribution function for log wages is:

$$G_{r,t}(w) = s_r \frac{\Phi\left(\frac{w-p_{1,r,t}}{D}\right) - \Phi\left(\frac{mw_t-p_{1,r,t}}{D}\right)}{1-\Phi\left(\frac{mw_t-p_{1,r,t}}{D}\right)} + (1-s_r)\frac{\Phi\left(\frac{w-p_{2,r,t}}{D}\right) - \Phi\left(\frac{mw_t-p_{2,r,t}}{D}\right)}{1-\Phi\left(\frac{mw_t-p_{2,r,t}}{D}\right)} \qquad \text{for } w \geq mw_t$$

where Φ is the cumulative distribution function of a standard Normal.

I choose parameters such that the price of skilled labor is significantly higher than that of unskilled labor. Thus, we should expect the minimum wage to cause more disemployment for unskilled workers. If changes in relative employment are significant, they will generate an endogenous compression of the skill premium $p_{2,r,t}/p_{1,r,t}$. That endogenous response may attenuate the negative employment effects and generate wage spillovers for other unskilled workers who earn more than the minimum wage.

All simulations have $\alpha = 0.8$, $E = 1.4$, $D = 0.3$, and a total of $R = 50$ regions. The shares of skilled workers s_r are Normally distributed across regions, with mean 0.3 and standard deviation 0.1. The initial value of the log minimum wage is -2.5 in Panel A, -1.7 in Panel B, and -1.3 in Panel C.

A.2 Normal-markdown model

Each region r and each time t has a Normal distribution of latent log wages $G^*_{r,t}(w^*) = \Phi\left(\frac{w^*-\mu_{r,t}}{\sigma_{r,t}}\right)$. The employment to population ratio and the distribution of observed wages depends on latent wages, the level of the national minimum wage, and a "mark-down" parameter $m \in (0,1]$ as follows:

$$\begin{aligned} emp_{r,t} &= 1 - \Phi\left(\frac{mw_t - \log m - \mu_{r,t}}{\sigma_{r,t}}\right) \\ G_{r,t}(w) &= \frac{\Phi\left(\frac{w-\mu_{r,t}}{\sigma_{r,t}}\right) - \Phi\left(\frac{mw_t-\log m-\mu_{r,t}}{\sigma_{r,t}}\right)}{1-\Phi\left(\frac{mw_t-\log m-\mu_{r,t}}{\sigma_{r,t}}\right)} \qquad \text{for } w \geq mw_t \end{aligned}$$

This model generates both truncation and censoring of the latent wage distribution. Workers whose latent log wages are below the minimum wage minus the log markdown become disemployed. Those with latent log wages above the log minimum wage, their observed wage is equal to the latent wage. Finally, those who remain employed but have latent log wages below the log minimum wage see a mechanical increase in their wage. The latter

group corresponds to the minimum wage "spike" in the log wage distribution.

The model can be understood as reflecting an economy with inelastic labor supply, exogenous worker productivities, and identical monopsonistic firms paying wages that are below the marginal products of labor, unless mandated to pay higher wages via the minimum wage. When the markdown m is low, disemployment effects are smaller and positive effects on wages are bigger. Unless otherwise noted, I use $m = 0.7$.

For the specific meta-distribution of parameters across regions, consult the notes at the bottom of the corresponding tables.

A.3 Task-based, monopsonistic model

The model is the one estimated in Haanwinckel (2023), with some region- and time-specific parameters determining differences between the 151 regions. See that paper for details.