

MIT Graduate Labor Economics 14.662 Spring 2009
Lecture Note 8: The Peculiar Case of the US Minimum
Wage

David H. Autor
MIT and NBER

May 14, 2009

1 THE PECULIAR CASE OF THE U.S. MINIMUM WAGE

These notes draw on Lee (1999) and Autor, Manning and Smith (2009).

As noted in class, an important limitation of the DiNardo, Fortin and Lemieux paper (DFL) is identification—there is no experiment, no contrast between a ‘treated’ and ‘untreated group,’ and no statistical inference (i.e., no standard errors). It’s simply a reweighting of the data, albeit an ingenious one.

David Lee’s paper improves on DFL in all these dimensions. The U.S. minimum wage is not indexed, meaning that it is set in nominal terms. As earlier figures have shown, the real value of the minimum wage eroded rapidly during the 1980s due to a combination of high inflation and a conservative administration that was happy to see it fall. It is clear that a time series plot would show a strong correlation between the U.S. minimum wage and U.S. wage inequality. But most economists would not find this evidence convincing. Is it possible to improve on this macro approach?

David Lee’s idea was to use cross-state differences in the ‘effective minimum wage’ to study the impact of the erosion of the minimum wage on earnings inequality. Although the U.S. Federal minimum wage is set at a uniform national level (with certain states setting slightly higher minima) prevailing wage levels vary considerably across regions of the countries. In high wage states like Massachusetts and California, few workers are paid at the minimum wage. In low wage states like Mississippi and Alabama, a kernel density plot strongly suggests that in 1979, the minimum wage ‘held up’ the lower tail of the distribution.

Motivated by this observation, Lee defines the ‘effective minimum wage’ as the log difference between the nominal value of the minimum wage and a state’s median wage. He then examines the cross-state relationship between the decline in the minimum wage over 1979 to 1989 and the growth in lower-tail inequality. The key identifying assumption for this approach is that cross-state variation in the ‘effective minimum’ is not systematically related to underlying variation in states’ wage distributions (i.e., the distribution that would prevail in the absence of a minimum wage).

1.1 FIRST ORDER EVIDENCE

- Some initial evidence on the relevance of this hypothesis is provided by the initial figures of Lee. Figures Ia - Ic are similar to DFL.
- Of some concern is the fact that Figure II seems to offer nearly as much evidence that the minimum wage raised the 90-50 differential as lowered the 10-50 differential.
- But Figure III which plots the 10-50 and 75-50 differentials for three high wage and three low wage states makes a compelling case that the ‘effective minimum wage’ construct has empirical traction. The drop in the lower tail of the distribution is *much more pronounced* in low wage states—but this is not so for the 75-50 differential. Figure IV adds to this case.

These observations suggest a simple empirical approach: a regression of the change in state level wage inequality on the change in the state’s effective minimum wage.

Figure VIa plots the effective minimum wage against the 10-50 log wage differential in 1979 and 1989. Consistent with Lee’s story, the minimum wage appears *very binding* in 1979—but not fully binding for high wage states. By 1989, it looks quite non-binding. Interestingly, the estimated locus of the 10-50 also shifts up slightly by 1989, suggesting in Lee’s interpretation a modest decline in ‘latent’ wage inequality.

Figure VIb shows a strong relationship between the effective median and the 20-50 differential. This is what’s predicted by the hypothesis. (It’s not clear to me that an almost equally significant relationship is *not* visible for the 75-50. If so, that’s problematic.)

1.2 ECONOMETRIC FRAMEWORK (FROM AUTOR, MANNING AND SMITH 2009)

To make inferences about the impact of the minimum wage at all percentiles of the wage distribution, we want to estimate models where the impact of the minimum wage is a function of not only the real value of the minimum wage, but is also a function of the overall shape or location of the wage distribution. One way to do this is to scale the minimum wage by some measure of the general level of wages. Lee (1999) estimated minimum wage effects in this

spirit, and used the log of the minimum relative to the median as his measure of the ‘bite’ of the minimum wage, which he called the ‘effective minimum.’

Use of the effective minimum can be justified in the following way, which fleshes out the arguments used by Lee (1999). Denote by $w_{st}^*(p)$ the log wage in state s at time t for percentile p in the absence of the minimum wage—call this the latent wage distribution. With a minimum wage, denoted in log form by w_{st}^m , the actual log wage at percentile p , which we will denote by $w_{st}(p)$ will deviate from the latent distribution for at least some percentiles. If, for example, the minimum wage had no effect on employment rates, and no spillovers then we would have the relationship:

$$w_{st}(p) = \max[w_{st}^m, w_{st}^*(p)].$$

However, if there are spillovers or some employment effects, then the minimum wage will have an effect on percentiles above where it binds (see Lee, 1999, for more discussion of these arguments, or Teulings, 2000, for an explicit supply and demand model with this feature). So, generalize the expression above to the form:

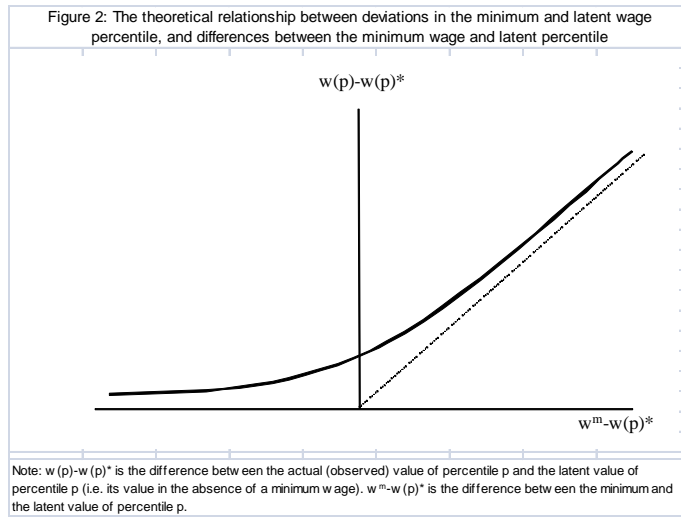
$$w_{st}(p) = \phi[w_{st}^m, w_{st}^*(p)].$$

What are plausible restrictions on the function $\phi(\cdot)$? We would expect it to be increasing in both its arguments and that it also satisfies a homogeneity property, specifically: if the latent percentile and the minimum wage both rise in the same proportion, the actual percentile also rises in that proportion. As the model is expressed in logs this restriction can be written as:

$$\phi[w_{st}^m + a, w_{st}^*(p) + a] = a + \phi[w_{st}^m, w_{st}^*(p)].$$

Now set $a = -w_{st}^*(p)$ and substitute into the above:

$$\begin{aligned} \phi[w_{st}^m - w_{st}^*(p), 0] &= -w_{st}^*(p) + \phi[w_{st}^m, w_{st}^*(p)] \\ \phi[w_{st}^m - w_{st}^*(p), 0] &= -w_{st}^*(p) + w_{st} \\ w_{st} &= w_{st}^*(p) + \psi[w_{st}^m - w_{st}^*(p)] \end{aligned}$$



In words, the deviation of the actual percentile from the latent percentile depends (via some unknown function, $\psi(\cdot)$) on the gap between the minimum and the latent percentile.

What are plausible restrictions on $\psi(\cdot)$? We would expect it to be positive everywhere (otherwise the minimum wage would reduce wages at some percentiles) and to have a positive first derivative. In addition, if the minimum wage is very low (or non-existent) we would expect the actual percentile to be very close to the latent percentile so that we have $\lim_{x \rightarrow -\infty} \psi(x) = 0$.

On the other hand, if the minimum wage gets very high we would expect the actual percentile to be very close to the minimum wage so that we have $\lim_{x \rightarrow \infty} \psi(x) = x$. Graphically, we might expect that the relationship between deviations of the actual from the latent percentile and the difference between the minimum wage and the latent percentile to look something like that presented in the figure below. In this figure, the x-axis plots the difference between the minimum and the latent value of percentile p . The y-axis plots the difference between the observed and latent values of percentile p . For low values of the minimum wage relative to the latent percentile, the minimum wage has no effect on the wage distribution so the observed value of the percentile is the latent value. For percentiles for which the minimum wage exceeds the latent percentile, the observed percentile will be equal to the minimum wage.

This figure allows for the possibility of ‘spillovers’ where the minimum either raises wages that are latently below the minimum to a value exceeding the minimum, or raises wages that are

latently above the minimum to a value exceeding their latent level. If present, such spillovers would be largest at the location where the minimum wage exactly equals the latent wage value (in the figure, this is the intersection of the x and y axes). Spillovers would be expected to attenuate in either direction from this point: further down the wage distribution, the minimum becomes extremely binding and so the mechanical effect dominates; further upward, the minimum wage becomes increasingly irrelevant.

This discussion should make it clear that non-linearity is likely to be an important feature of $\psi [w_{st}^m - w_{st}^*(p)]$, so that some thought needs to be given to the functional form of the estimating equation. The main specification in Lee (1999) and AMS (2009) use a quadratic approximation:

$$w_{st} = w_{st}^*(p) + \alpha_0 + \alpha_1 (w_{st}^*(p) - w_{st}^m) + \alpha_2 (w_{st}^*(p) - w_{st}^m)^2$$

It should be noted, however, that a quadratic cannot have a shape similar to that drawn in the above figure over the whole of its range so that we have to exercise caution in estimating minimum wage effects outside the observed sample. In particular, this specification cannot be used for an assessment of what the distribution of wages would be like if there was no minimum wage. (Among other issues, this would require taking the log of zero.)

To make the above equation estimable, we need to put some additional structure on the form taken by the latent wage distribution. Lee (1999) proposes that the latent wage distribution can be summarized by 2 parameters – the median and the variance – so that we can write:

$$w_{st}^*(p) = \mu_{st} + \sigma_{st} F^{-1}(p),$$

where μ is the centrality parameter, σ is the scale parameter, and $F^{-1}(\cdot)$ is the inverse of the CDF of the latent wage distribution. Assuming that the mean and median are identical, it will be the case that $F^{-1}(\cdot) = 0$, so that μ_{st} is the median log wage in state s at time t . Substituting:

$$\begin{aligned} w_{st}(p) - \mu_{st} &= \alpha_0 + \sigma_{st} (1 - \alpha_1) F^{-1}(p) + \alpha_2 [\sigma_{st} (1 - \alpha_1) F^{-1}(p)]^2 \\ &\quad + [\alpha_1 - 2\alpha_2 \sigma_{st} F^{-1}(p)] (w_{st}^m - \mu_{st}) + \alpha_3 (w_{st}^m - \mu_{st})^2, \end{aligned}$$

The first two terms are related to the overall evolution of wage inequality and last two terms to the effect of the effective minimum. Note that the coefficients in this equation will vary with the percentile, not just because p appears in the linear term of the effective minimum but also because, as pointed out by White (1980), the coefficients α will vary with the data. Intuitively we would expect that a rise in latent wage inequality leads to a larger impact on lower percentiles for a given effective minimum.

For (7) to be estimable one also needs models for the median and variance. There are a number of potential options, and we start our discussion with the choices made by Lee. Lee replaces μ_{st} by the observed median, σ_{st} by a set of time dummies and assumes that any cross-state variation in latent wage inequality is uncorrelated with the median and can therefore be subsumed into the error without causing bias in the estimated impact of the minimum wage. In other words:

$$\sigma_{jt}|t \perp \mu_{jt}$$

Under these assumptions, the estimating equation becomes:

$$w_{st} = \alpha_0 + \alpha_t + \alpha_1 (w_{st}^m - w_{st}(50)) + \alpha_2 (w_{st}^m - w_{st}(50))^2 + \varepsilon_{st}$$

1.3 TWO CONCERNS: VIOLATION OF ID AND MEASUREMENT ERROR

But a practical problem immediately arises in estimating the following regression:

$$(w_{jt}^{10} - w_{jt}^{50}) = \alpha_t + \beta (mw_t - w_{jt}^{50}) + \varepsilon_{jt}, \quad (1)$$

which is that there is a potential mechanical relationship between the dependent and independent variables induced by the presence of w_{jt}^{50} in both expressions. This relationship may induce an upward bias in estimates of β , giving rise to the spurious conclusion that increases in minimum wages reduce inequality (note: the dependent variable is by construction negative and factors that raise it towards zero reduce inequality).

More formally, the covariance between the independent and dependent variables is:

$$cov[w_{jt}^{p*} - w_{jt}^{50}, (mw_t - w_{jt}^{50})|t] = cov[\mu_{jt} - w_{jt}^{50} + \sigma_{jt}F_t^{-1}(p), (mw_t - w_{jt}^{50})|t] \quad (2)$$

Under the identifying assumption that the shape of the wage distribution is identical in all states, we have:

1. First, $\mu_{jt} = w_{jt}^{50}$, so that $F_t^{-1}(50) = 0$.
2. Second, $F_t^{-1}(p)$ does not vary across states—that is, the latent wage percentile is identical (again up to the scale parameter σ_{jt}).

If these assumptions are correct, (2) reduces to:

$$\text{cov}[\mu_{jt} - w_{jt}^{50} + \sigma_{jt}F_t^{-1}(p), (mw_t - w_{jt}^{50}) | t] = \text{cov}[\sigma_{jt} \cdot F_t^{-1}(p), mw_t - w_{jt}^{50} | t] = 0. \quad (3)$$

Note that this expression equals zero because of the assumption that *latent* wage dispersion ($F_t^{-1}(p)$) is uncorrelated with the minimum wage, not because *actual* dispersion is uncorrelated with the minimum wage. (It’s the latter thing that we want to test).

1.3.1 CONCERN 1: VIOLATION OF ID

But if $F_t^{-1}(50) \neq 0$, we will have a problem. Imagine, for example, that there is no movement in the actual 10th and 90th wage percentiles but there is substantial time variability in state median wages.

We would then estimate that β is positive in both of the following regressions due to the mechanical covariance between w_{jt}^{50} on both sides of the equation:

$$\begin{aligned} (w_{jt}^{10} - w_{jt}^{50}) &= \alpha_t + \beta_{10} (mw_t - w_{jt}^{50}) + \varepsilon_{jt}, \\ (w_{jt}^{90} - w_{jt}^{50}) &= \alpha_t + \beta_{90} (mw_t - w_{jt}^{50}) + \varepsilon_{jt}. \end{aligned}$$

This would yield the spurious inference that a decline in the effective minimum wage raises lower tail inequality and reduces upper tail inequality.

This story is actually a simple violation of the identifying assumption that state wage levels are uncorrelated with latent wage dispersion. The plausibility of this story suggests that the concern is significant.

The concern also suggests a specification test: a regression of the 90-50 and 75-50 differential on the effective minimum wage. Under the maintained hypothesis that the ‘shape’ of the latent

wage distribution in year t is common across state, the minimum wage should not impact wages above the median of the distribution. If we find evidence that the minimum wage does affect inequality above the median, this will cast doubt on the interpretation of the results for wage compression below the median. Hence, the test provides a check against spurious causation.

1.3.2 CONCERN 2: MEASUREMENT ERROR

Even if the ID assumption is valid, measurement error in the centrality measure, w_{jt}^{50} , may still give rise to a spurious correlation between the dependent and independent variables. Hence, we want a ‘strong’ measure of central tendency to increase the likelihood that $F_t^{-1}(50) = 0$.

Lee is cognizant of this problem and quantifies it. The sampling error for the typical state’s median wage is about 0.01, less than 1 percent of variation, seemingly not a concern. However, what happens when state dummy variables are added to (1)?

The observed variance in a given variable can be written as

$$\sigma_v^2 = \sigma_s^2 + \sigma_e^2, \tag{4}$$

where s and e refer to ‘signal’ and ‘error’ variance respectively. The first is true variance in the outcome variable. The latter is measurement error. The ratio of signal to total variance is

$$\gamma = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2}. \tag{5}$$

When explanatory variables are added to a regression, these will *only* explain the signal component of the observed variance; the noise is by definition unsystematic. So, as explanatory variables are added, the noise share of residual variance $(1 - \gamma)$ tends towards 100 percent. This will turn out to be problem.

1.4 RESULTS: 1979 - 1988

Table IA.

- Striking estimates implying that the ‘effective minimum’ substantially reduces the 10–50 differential. Put differently, the effective minimum is becoming more negative over time.

So, the positive point estimate for β means that this decline in the effective minimum causes the $w^{10} - w^{50}$ gap to grow (become more negative).

- Notice the pattern of the time dummies. These are trending strongly downward in the raw data—the 10-50 gap is growing more negative. But conditional on the ‘effective minimum,’ there is no negative—and perhaps even a weakly positive—trend.

Table 1B shows that these impacts are much more pronounced for the young than the old and for women than men, which is logical.

An obvious next step is to add state fixed effects to the model. But here the aforementioned problem arises.

- A regression of the effective minimum wage \widetilde{mw}_{jt} on year dummies alone gives a root mean squared error (residual standard deviation) of 0.125.
- Adding state dummies to the model reduces this to 0.031.
- Hence, the fraction of variance remaining is $(0.031/0.125)^2 = 0.062$, i.e., only 6 percent!
- Since *only* the signal component of variance is absorbed by the state dummies, is likely that some large part of this remaining variance is measurement error. Due to (3), this suggests a high probability of spurious results.

In Table II, these problems are evident:

- For males and, to a lesser degree for the combined-gender sample, the estimated impacts of \widetilde{mw} on the 60-50, 70-50, 80-50 and 90-50 differentials are all positive and significant—suggesting, implausibly that a higher effective minimum wage *raises* inequality in the upper half of the distribution. (*Given that the effective minimum is falling, it may be more accurate to say that the estimates imply that a falling effective minimum causes compression in the upper half of the distribution*). For women alone, these problems are not evident.

- When state dummies are included, all models fail the specification test. These results are therefore dubious.

What is going on with the results for males? Lee does not speculate. One interpretation is that these estimates are driven by movements in state medians rather than the minimum wage itself. A decline in the effective minimum wage can also be read as a rise in the median—which both raises inequality at the bottom and compresses it at the top. (More on this below.)

Table III presents further results indicating the importance of the minimum wage for the lower tail of the earnings distribution. Unfortunately, of the six panels in the table, only the pooled-cross section results for women can be read with great confidence. But these results are of striking magnitude, suggesting no trend towards greater latent inequality in the lower tail of the female wage distribution during the 1980s.

One thing that is puzzling about these results for women is that women’s earnings relative to men rose substantially during the 1980s. *All else equal, these estimates suggest that the falling minimum wage should have reduced female wage levels.* This makes the rise in women’s earnings all the more surprising—and suggests that there is much to understand about changes in the wage structure that is not explained by the minimum wage.

1.5 A SECOND TEST

The results so far are sufficiently problematic that they probably would not stand on their own. To buttress them, Lee exploits the U.S. minimum wage change from \$3.35 to \$4.25 per hour that occurred between April 1990 and April 1991. Combined with the fact that some states had already legislated their own minimum wage increases earlier, this change gives rise to sharp, within-state, over-time jumps in the effective minimum—and a ‘control group’ of states with already higher minima—which are lacking above. This should allow for a compelling test.

The identifying assumption now becomes that the binding relative minimum wage (that is, the maximum of the state and federal minimum by state) is uncorrelated with *changes* in latent wage inequality across states over 1989 - 1991. This assumption seems weaker, and hence more plausible, than the assumptions needed above.

Analogous to the above signal/noise test, Lee provides the following facts.

- A regression of \widetilde{mw}_{jt} on time dummies and state dummies for 1989 - 1991 reduces the root MSE (relative to a model with only time dummies) from 0.137 to 0.040.
- $(0.040/0.137)^2 = 0.085$. This corresponds to a reduction of 92.5 percent in the signal to variance ratio. This does *not* look promising.
- However, a regression of \widetilde{mw}_{jt} on time dummies, state dummies, and the nominal legislated minimum wage rate leaves a root MSE of 0.013.
- Hence, $(0.013/0.040)^2 = 0.106$, which means that 89.4 percent of the within-state variation in state changes in \widetilde{mw} during these two years is driven by legislated changes. This is very good news because it means we have a lot of ‘signal’ variance to work with.

Remarkably, in Table V, the estimated impacts of the minimum wage on the lower tail of the distribution using this source of variation are quite comparable to those over 1979 - 1988. This improves the credibility of the estimates above.

Surprisingly, the specification test again rejects for males and the combined sample. (This is not tabulated but it is reported in the text.) For example:

- A regression of the 80 - 50 on \widetilde{mw} for 1989 - 1991 yields a coefficient of 0.094 (0.035) for men, 0.040 (0.026) for the combined sample, and -0.003 (0.027) for women.

This failure of the specification test is all the more puzzling in this case, but again cautions against placing great confidence in the results for these samples.

1.6 COUNTERFACTUALS AND INFERENCES

Lee develops simple counterfactuals by using the regression estimates to predict the distribution of wages that would have prevailed in 1979 if the effective minimum wage were at its 1989 level (or vice versa). These counterfactuals must of course ignore general equilibrium impacts. There are many interesting patterns in Table VI of the paper and I mention only a few:

- For women, about 70 percent of the 18.6 log point growth in the 10-50 differential is explained by the minimum wage. Almost none of the 10.9 log point growth in the 90-50 differential is explained.
- The minimum wage has equal explanatory power for male earnings inequality in the 10-50 region, but these estimates are dubious, as Lee acknowledges.
- The minimum wage does not seem particularly important to ‘between group’ inequality; regression estimates of educational wage differentials are only slightly compressed by adjusting for the minimum wage (though the impact on the college-HS dropout differential for women is sizable).
- The minimum wage appears *quite important* to the growth of residual earnings inequality, explaining 66 percent of its growth for women.
- Finally, Figure VII suggests that there was no growth in ‘latent’ inequality for the entire sample below the median—but substantial growth above it.

A fair interpretation of Lee’s findings is given in the conclusion:

- Broad trends in educational and experience differentials are essentially *unaffected* by accounting for the falling minimum wage.
- But the minimum is likely to have contributed to rising residual inequality in the same years, particularly for women.
- The results for males remain a puzzle.

1.7 CONCLUSIONS FROM LEE (1999)

There are clearly a number of problems and puzzles in the results above. AMS (2009) take another look at the Lee (1999) analysis and reach somewhat different conclusions. They identify two econometric issues with Lee’s approach—violation of the mean-variance independence assumption and simultaneity bias due to measurement error—and also provide a detailed analysis of the potential spillovers from minimum wages to non-covered wages. See slides from lecture.