ONLINE APPENDIX – NOT FOR PUBLICATION

Labor Market Institutions and the Distribution of Wages:

The Role of Spillover Effects*

By Nicole M. Fortin, Thomas Lemieux, and Neil Lloyd

Appendix A: Issues in the estimation of the direct and spillover effects of the minimum wage for distribution regressions

Figure A1 illustrates the potential importance of heaping issues by presenting histograms of the lower part of the distribution of wages. The figures differentiate the mass (in gray) at actual minimum wages from mass that arises due to heaping issues. As can be gleaned from the figure, mass at minimum wages is substantially more important in the period 1979-88 than in the other periods.

Figure A2 illustrates the estimated wage distributions for 1979-88 with and without the minimum wage. The two distributions are reported in Figure A2, Panel A, correspond to the values of the unconditional probability $\widehat{\mathrm{Q}}_t^k$, where $\hat{Q}_t^k = \frac{1}{N_t}\sum_i \hat{Q}_{ist}^k$. , and the predicted probabilities $\widehat{\mathrm{Q}}_t^{k,c}$ recentered around the median value of the minimum wage. As explained in the text, $\hat{Q}_{ist}^{k,c} = \widehat{\mathrm{P}}_{ist}^{k,c} - \widehat{\mathrm{P}}_{ist}^{k+1,c}$ and $\hat{Q}_t^{k,c}$ is defined as $\hat{Q}_t^{k,c} = \frac{1}{N_t}\sum_i \hat{Q}_{ist}^{k,c}$. Panels B and C of Figure A2 illustrates the impact of the measurement error correction on the wage distribution for the 1979-88 period.

Figure A3 compare our estimated spillover effects where we do control for rounding off with the set of heaping dummies L_{kst}^p to what we would find without controlling for these dummies. The results of this comparison for 1979-88 are reported in Figure A3.

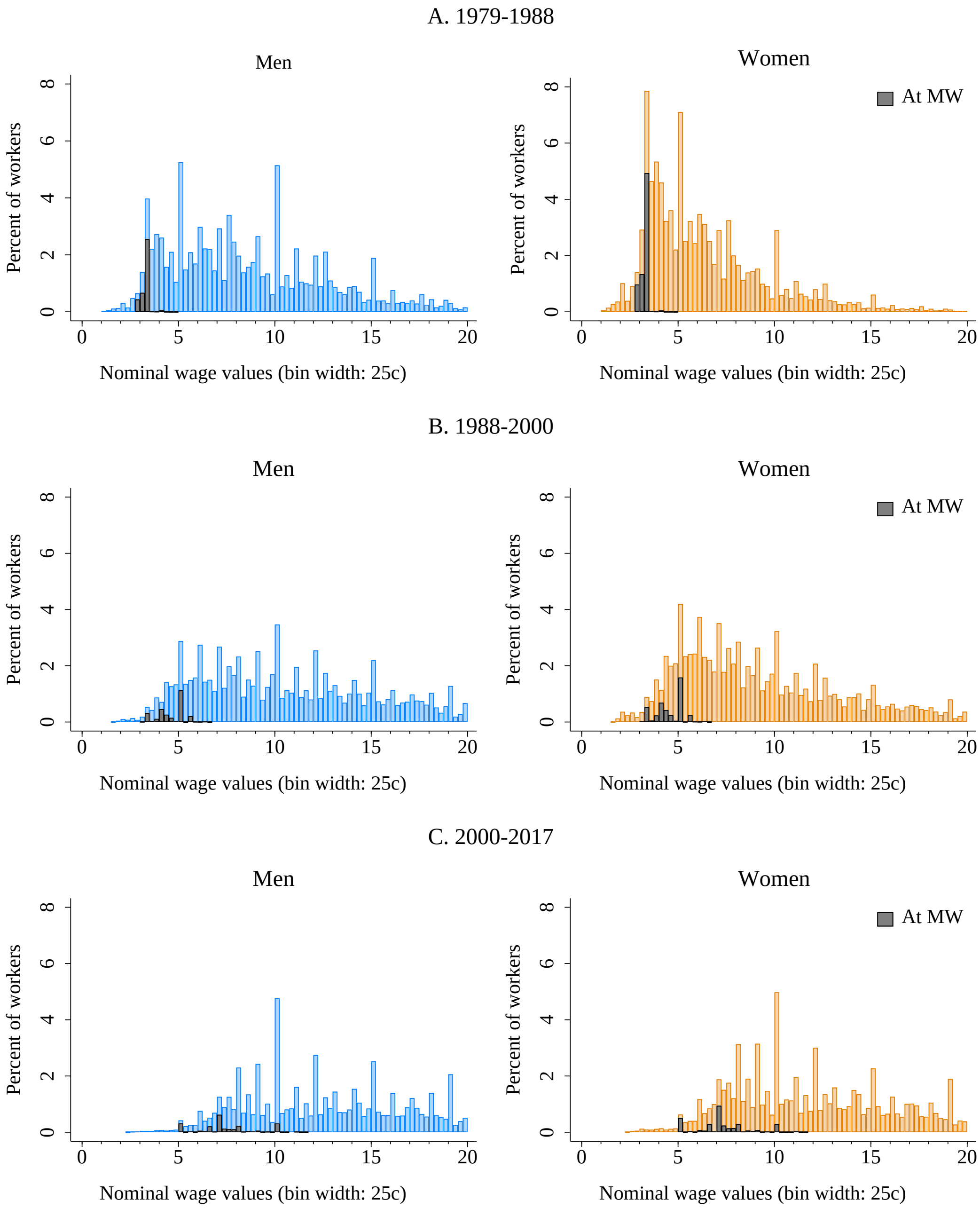

Figure A1: Histogram of Nominal Wage Bins

NOTE: The histogram uses 25 cents wage bins (starting from 0) and displays the distribution only up to $20/ hrs, but remains proportional to total distribution. The darker wage bins show the share of workers earning their state's minimum wage. This fraction is calculated by identifying workers who report an hourly wage within 10 cents of their state's minimum wage.

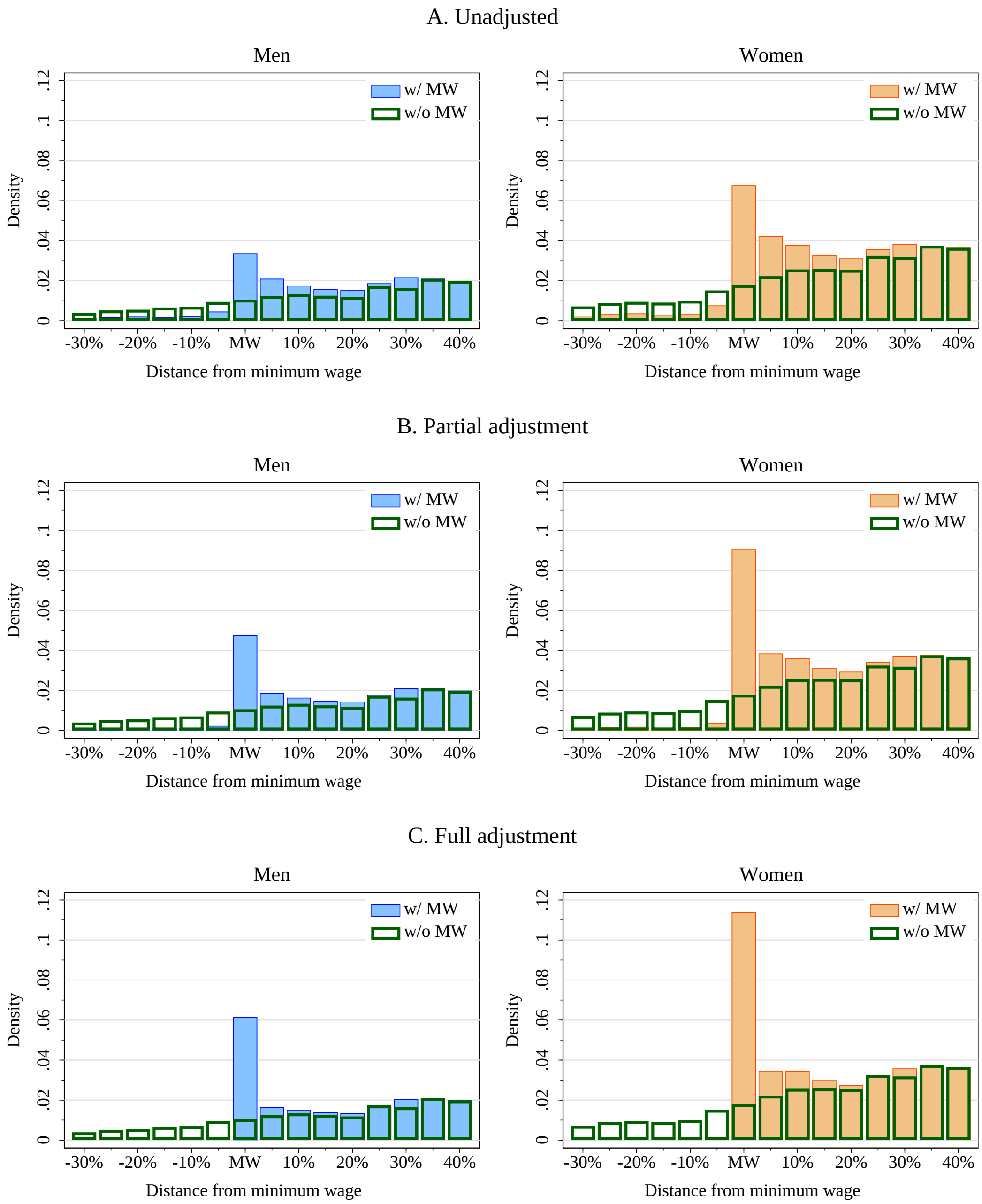

Figure A2. Estimates of Minimum Wages Impact - 1979-88

NOTE: The bars depict the predicted density of each wage bin around the minimum wage, given the estimates of the model. In each panel, the counterfactual probabilities without a minimum wage ("w/o MW") are obtained by setting the minimum wage coefficients to zero (see discussion in Section 4.2). Panels B and C reflect the partial and full adjustments made to account for measurement error discussed in the text (Section 4.3).

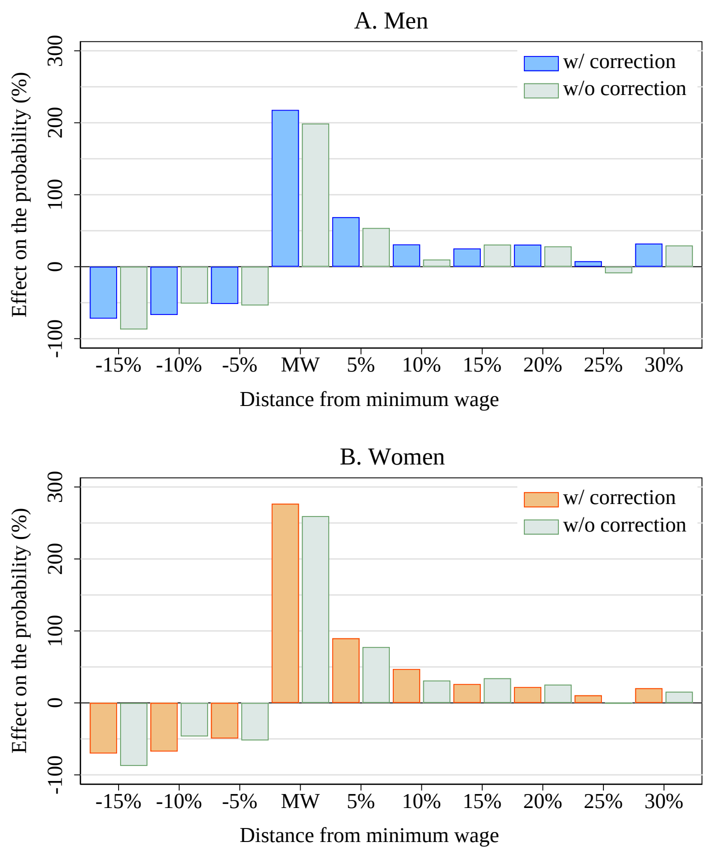

Figure A3: Marginal effects with and without the heaping correction, 1979-88

NOTE: As in Figure 3, each bar represents the marginal effect of the minimum wage, shown here for the 1979-1988 period. Each panel includes two sets of estimates: with and without the correction for heaping. The model “with correction” includes the controls for bunching at integer, nominal wage-values (as in Figure 3). In contrast, the set of marginal effects “without correction” is constructed from a separate model estimated without these additional controls. See discussion in Section 4.3.

Appendix B: An illustrative example of the identification of minimum wage effects

This appendix is offered as a complement to section 2.3. Here, we illustrate how a new minimum wage affects wages below (loss of mass), at (spike) and above (spillover) the new minimum. Importantly, we show how these changes to the wage distribution map into the parameters φ_m of the model. Consider a latent normal wage distribution in Figure B1a (blue line). We now add a minimum wage (red line) that creates a large spike at the minimum, adds some mass slightly above the minimum wage (spillover effects), and dramatically reduces the probability of being at values below the minimum wage.

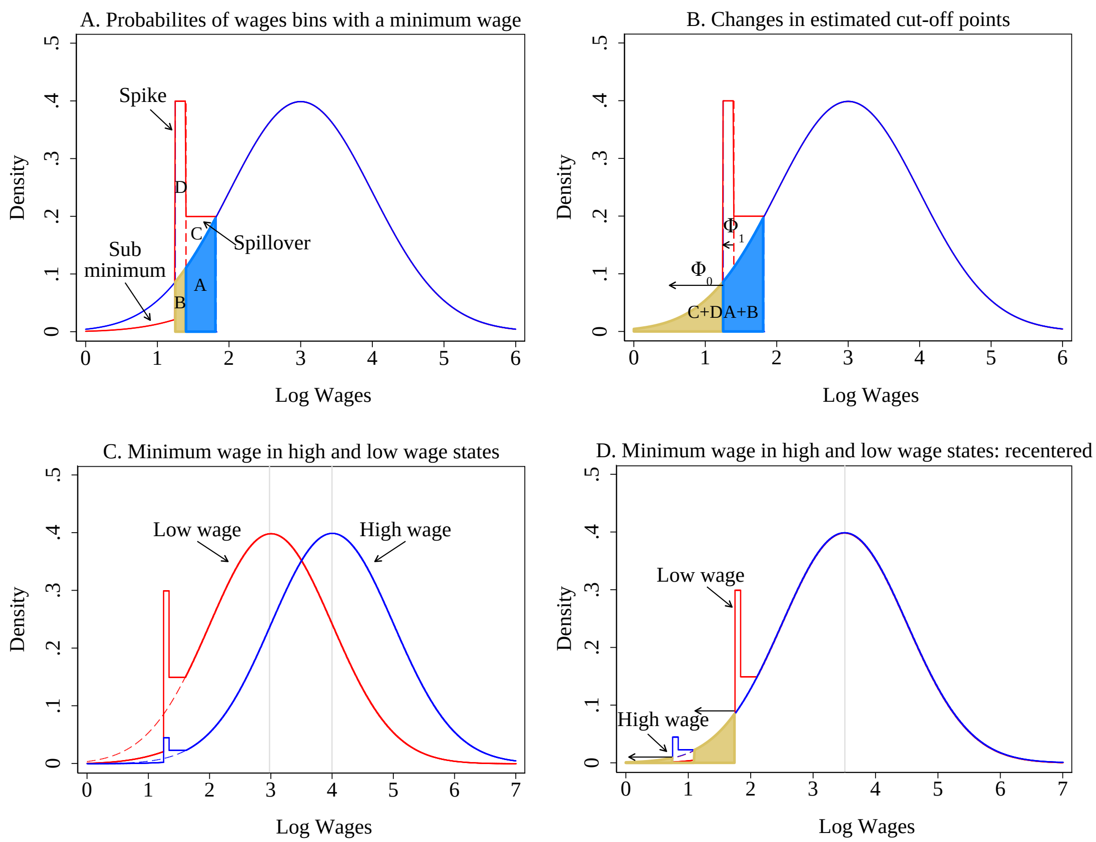

Figure B1. Illustrative Example of the Identification of Minimum Wages Effects

Figure B1a shows that the probability of being in the “spillover zone” just above the minimum wage increases from A to A+C, while the probability of being at the spike increases from B to B+D. In this simple example, the parameters φ_1 and φ_0 are the horizontal values

(illustrated by arrows in Figure B1b) by which the cutoff points have to be moved to increase the two probabilities by an amount of C and D, respectively.

Next, Figure B1c and B1d illustrate a case with two states that differ in terms of mean wages. If we use the dummy variable *X* in equation (2) to indicate if an observation comes from the high-wage state, the parameter β will capture the mean wage differences between the two states. The three key parameters to be estimated in this example are β (the difference in means) and the minimum wage parameters φ_1 and φ_0. As discussed at the beginning of this section, these parameters are jointly estimated in our estimation approach, while corresponding parameters are estimated in two separate steps in Lee (1999) and Autor, Manning, and Smith (2016).[1]

To better understand how φ_1 and φ_0 are estimated in the two states example, Appendix Figure B1d shows the recentered densities obtained using the parameter – or adjustment factor – β. The recentering clearly shows how the same federal minimum wage bites at different points of the distribution in the two states. A precisely similar graph would be obtained if the two states had the same latent wage distribution but different state wage minimum wages. Thus, from an identification perspective, it does not matter whether the variation in the relative minimum wage is driven by differences in mean wages across states (as in Lee, 1999), or difference in state minimum wages (as in Autor, Manning and Smith, 2016)[2]. The parameters φ_1 and φ_0 correspond again to horizontal moves in cutoff values (arrows in Appendix Figure B1d) required to fit the change in probabilities induced by the minimum wage. Interestingly, the same horizontal shift has a larger impact on probabilities when the minimum wage is relatively higher up in the distribution (low-wage state case in Figure B1d). This convenient property is linked to the well-known fact that marginal effects in a probit model are directly proportional to the density at the point where the marginal effects are computed. As in Lee (1999), the relative bite of the minimum wage —its distance relative to the median— also plays a central role in the estimation in the distribution regression model.

---

[1] The model parameters are quite different in the two approaches since we are modeling the probability distribution, while Lee (1999) and Autor, Manning, and Smith (2016) are modeling quantiles of the wage distribution. The β parameters in equation (2) are, nonetheless, closely connected to the "first-step" median used in these two papers to compute the relative value of the minimum wage. The measurement error linked to plugging in estimates of the medians does not apply given that we are jointly estimating similar centrality parameters and minimum wage effects.

[2] As long as the state-fixed effect is constant, or changes smoothly with the wage-distribution.

Appendix C: RIF-regression estimates of the impact of unionization

As discussed in Section 3, the estimates of union threat effects used in the decompositions are based on estimates of the distribution regressions where the unionization rate at the state-industry-year level is included as an additional regressor. Here we present more straightforward estimates based on OLS and RIF-regression models where it is easier to estimate the effects of the unionization rate at different points of the distribution. The results from these simple regressions are reported in Figure C1.

To compare our results with earlier studies, we report in the first panel estimates of the effect of the union status on wages. The OLS estimates yield the typical union wage premium, while the RIF-regression coefficients indicate how the union effect varies at different point of the distribution.[3] We use the same set of covariates but estimate the model over the entire 1979-2017 period. These results are comparable to those of the Figure 4 which provides marginal effects of a 1 %-point increase in union coverage ("shift share") and a 1 %-point increase the industry-state unionization rate ("union threat").

Consistent with the existing literature, Panel A of Figure C1 shows that the union wage premium (horizontal red line) is about 20% for men, and a bit smaller for women.[4] As in Firpo, Fortin, and Lemieux (2009), the union effect estimates obtained using RIF-regressions are hump-shaped. For both men and women, they peak around the middle of the distribution, and steadily decline in the upper part of the distribution.

Intuitively, the pattern of union wage effects – positive on average but declining in the upper part of the distribution – is consistent with other evidence on the effect of unions on the wage structure. For instance, Card (1996) shows that the union wage premium is positive on average, but declines over the skill distribution.

[3] As discussed in Firpo, Fortin, and Lemieux (2009), RIF-regression estimates can be interpreted as the impact of a small change in the probability of unionization on the unconditional quantiles of the wage distribution. As such, RIF-regressions are one among several possible ways of computing the counterfactual distribution obtained by changing the probability of unionization. The alternative approach used in Section 4.3 consists of reweighting the data to slightly increase the fraction of union workers (as in DFL), and see how it affects the various wage quantiles.

[4] For instance, Card, Lemieux, and Riddell (2018) find a union wage premium of 0.16 for men and 0.09 for women in 2015.

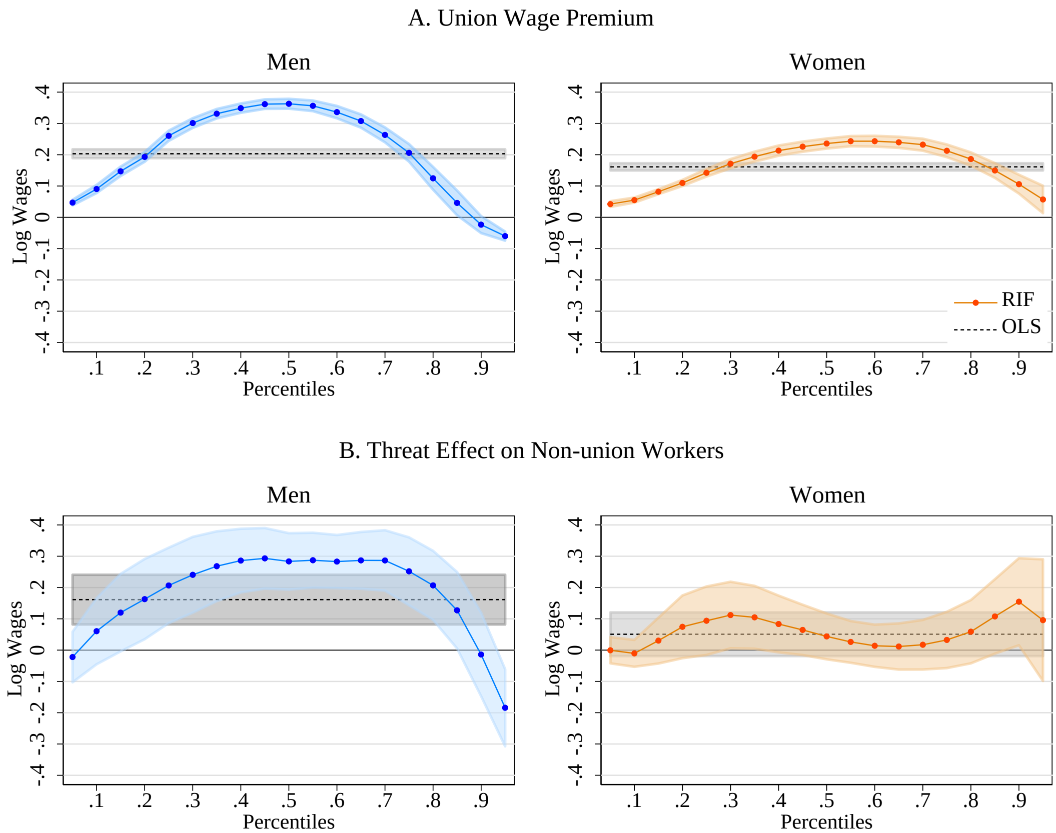

Figure C1: OLS and RIF regression model: Spillover Effects Based on Unionization Rates

NOTE: Each figure plots the coefficient and 95% confidence interval for the OLS and RIF regression model. Panel A corresponds to the direct union-wage gap using the sample of non-allocated wages of covered and uncovered salary workers. Panel B corresponds to the coefficient on the unionization rate at the industry-state level (3 year moving average) using the sample of uncovered salary workers. All models pool data from 1979-2017, for which union coverage and non-allocated wages are observed; this excludes data from 1980-1982, 1994, and part of 1995. The density and RIF are estimated separately by year, while the RIF regression is estimated once, pooling all available years of data. Each model includes state and industry FE, as well as state- and industry-specific trends, as in the larger distribution model. Additional covariates include our standard set of controls: year FE, marital status, part-time status, experience (quartic), education, education-experience bins, public sector, part-time, occupation, and CMSA location.

It is not as intuitive, however, to see why the RIF-regression estimates first grow before reaching a peak around the middle of the distribution. Part of the story is that changes in the rate of unionization have little impact at the bottom of the distribution where wages mostly depend on the minimum wage. Another part of the story is that very few workers are unionized at the bottom of the distribution. The issue is discussed in more detail using an example with uniform

distributions in Appendix D. Note that the hump-shaped pattern of RIF-regression coefficients has important implications on how de-unionization affects the shape of the wage distribution. Panel A of Figure B1 indeed indicates that unionization substantially reduces the 90-50 gap, but also increases the 50-10 gap. Interestingly, DFL reach a similar conclusion using a reweighting approach, as we do with the distribution regression method (see Section 4.3).

Panel B shows corresponding estimates of the effect of the state-industry-year unionization rate on the wages of non-union workers. Interestingly, in the case of men the shape and magnitude of the estimated effects are qualitatively similar to those for the union status reported in Panel A. In the case of women, the OLS estimate is substantially smaller, and the RIF-regression estimates are a bit unstable across the various percentiles of the distribution. This is consistent with the main model, where we find that the threat effect is much smaller. Moreover, here we pool the three periods. Given the rapidly changing composition of female labour supply over this period, it may be preferable to estimate separate models.

Taken together, the results reported in Figure B1 support the view that the threat of unionization has a positive effect on the wages of non-union workers. Although the shape of the RIF-regression coefficients varies across the specifications reported in Panels B, the estimates tend to be small and often negative at the top of the distribution. As discussed in Section 3, this supports the view that declining unionization rates (or success rates of union elections) capture declining threat effects instead of spurious state-industry shocks that both reduce wages and unionization rates.

Appendix D: Understanding the hump shaped effect of unionization on the wage distribution

In this appendix, we discuss a simple example to illustrate why a change in the rate of unionization is likely to have a “hump shape” or “inverse U-shape” impact on wage quantiles. The hump shape effect has been documented empirically using RIF-regressions (e.g. Appendix Figure C1 or Firpo, Fortin, and Lemieux, 2009) and distribution regressions (Figure 4).

For the sake of simplicity, consider a case where non-union wages follow a uniform distribution between zero and one ($Y \sim U(0,1)$). Union wages follow a $U(0.6,0.8)$ distribution, which has a higher mean but lower variance than the non-union distribution. The two distributions are illustrated in Figure D1a.

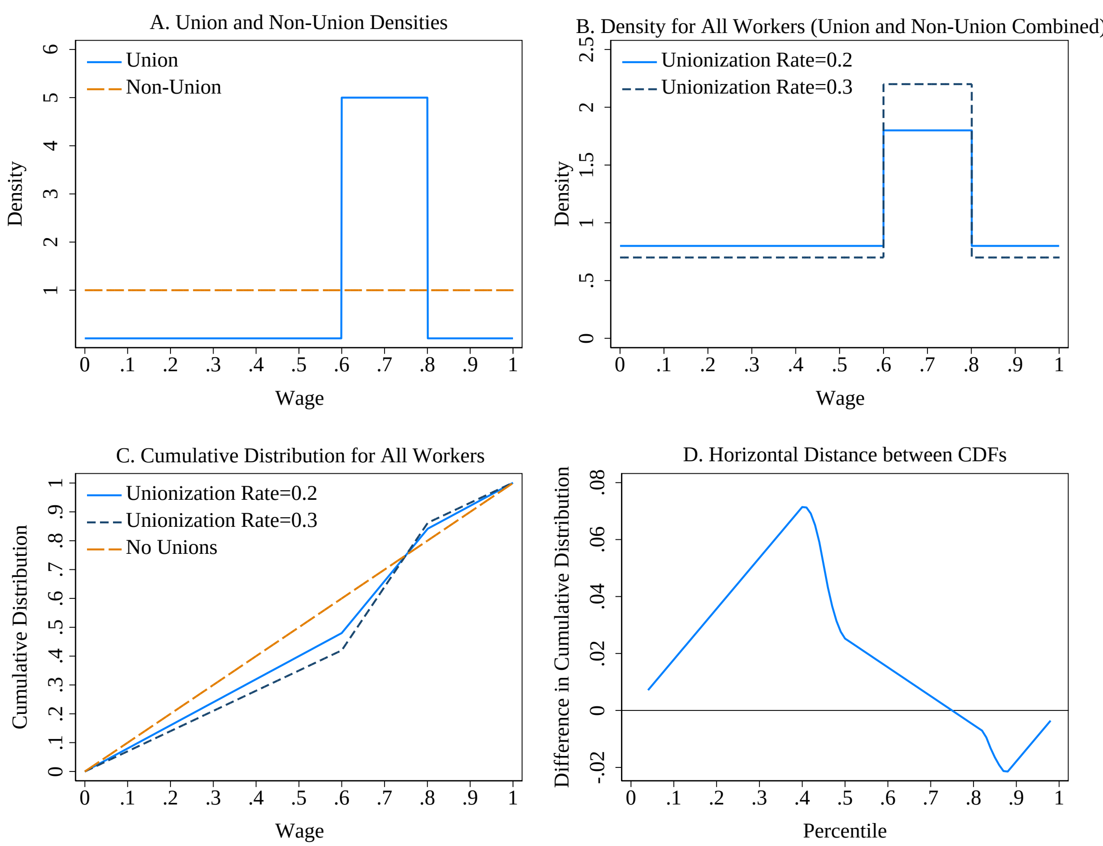

Figure D1. Understanding the Hump-Shaped Effects of Unionization

Now consider a counterfactual experiment where the unionization rate increases from 0.2 to 0.3. Figure D1b shows the wage densities for all workers combined, while Figure D1c shows

the corresponding cumulative distribution functions (CDF). Raising the rate of unionization increases the mass in the upper middle of the distribution and reduces the mass in the two tails of the distribution. While this reduces overall wage dispersion (the variance goes from 0.074 to 0.068), the impact is uneven at different points of the wage distribution.[5] To see this, recall that the effect of increasing the unionization rate on wage quantiles is the horizontal distance between the two CDFs plotted in Figure D1c. The effect on wage quantiles is zero at the very bottom of the distribution, but grows linearly until the 40th percentile. The effect of changing the unionization rate on wage quantiles then starts declining before turning negative around the 80th percentile. This non-monotonic effect of the unionization rate on wage quantiles is illustrated in Figure D1d that plots the (smoothed) change in wage quantiles over the whole distribution, and exhibits the hump-shaped feature discussed above.

The intuition for why unionization increases wage quantiles at the bottom of the distribution, but reduces wage quantiles at the top is straightforward. Increasing the rate of unionization shrinks the wage distribution towards the upper middle (0.6-0.8 range in Figure D1b), which pulls up wage quantiles at the bottom and pulls down wage quantiles at the top. What is not as intuitive is why the effect first grows at the bottom of the distribution before declining later on. In the case of the uniform distribution, the lowest quantiles cannot move much in response to a change in the rate of unionization as they are "pinned down" at the lower bound of the distribution (0 in this example). Likewise, a binding minimum wage that creates a sharp lower bound would generate the same phenomena. For example, if 10 percent of non-union workers are bunched at the minimum wage, the 0th to the 7th (8th) quantiles will be equal to the minimum wage when the unionization rate is 30% (20%). As a result, wage quantiles up to the 7th quantile won't change when the unionization rate increases, while quantiles slightly higher up will increase for the reason discussed above (overall distribution shrinking towards the upper middle).

As it turns out, other distributions like the normal distribution also yield the hump-shaped curve illustrated in the case of the uniform distribution. To see this, note that in Figure C1c, vertical distance between the two CDFs (20% and 30% unionization rates) is a linear function of

[5] The overall variance can be computed using the well-known analysis of variance formula $Var(Y) = \overline{U} \cdot V^U + (1-\overline{U}) \cdot V^{NU} + \overline{U} \cdot (1-\overline{U}) \cdot (\mu^U - \mu^{NU})^2$, where the mean and variance of wages in the union and non-union sectors are, $V^U = \frac{1}{12*25}$, $V^{NU} = \frac{1}{12}$, $\mu^U = .7$, and $\mu^{NU} = .5$, respectively.

the wage. The horizontal distance is equal to the vertical distance divided by the slope of the CDF (the wage density $f(Y)$) evaluated at this point. Thus, the effect is increasing in Y as long as the derivative of $Y/f(Y)$ with respect to Y is positive. This trivially holds in the case of the uniform distribution since $f(Y)$ is a constant, and holds for more general distributions as long as $f(Y)$ is not growing “too fast” as a function of Y at the bottom end of the distribution.