Appendix A: Measurement of Occupation, Industry, and Regional-Level Industry Coverage

In our regression models, we control for the occupation and industry of respondents' jobs. For occupations, we use the four-digit OCC2010 variable provided by IPUMS, and organize these more detailed occupational codes into 26 broader groups created by IPUMS. These groups are: management in business, science, and arts (10-430); business operations specialists (500-730); financial specialists (800-950); computer and mathematical (1000-1240); architecture and engineering (1300-1540); technicians (1550-1560); life, physical, and social science (1600-1980); community and social services (2000-2060); legal (2100-2150); education, training, and library (2200-2550); arts, design, entertainment, sports, and media (2600-2920); healthcare practitioners and technicians (3000-3540); healthcare support (3600-3650); protective service (3700-3950); food preparation and serving (4000-4150); building and grounds cleaning and maintenance (4200-4250); personal care and service (4300-4650); sales and related (4700-4965); office and administrative support (5000-5940); farming, fisheries, and forestry (6005-6130); construction (6200-6765); extraction (6800-6940); installation, maintenance, and repair (7000-7630); production (7700-8965); transportation and material moving (9000-9750); and military (9800-9830).

For industries, we use the three-digit IND1990 variable provided by IPUMS, and organize these more detailed industrial codes into 14 larger groups provided by IPUMS: These groups are: agriculture, forestry, and fisheries (10-32); mining (40-50); construction (60); manufacturing (100-392); transportation, communications, and other public utilities (400-472); wholesale trade (500-571); retail trade (580-691); finance, insurance, and real estate (700-712); business and repair services (721-760); personal services (761-791); entertainment and recreation services (800-810); professional and related services (812-893); public administration (900-932); and active duty military (940-960).

We create industry-specific regional union coverage using these industry codes, and the regional divisions collected by the CPS: New England Division, Middle Atlantic Division, East North Central Division, West North Central Division, South Atlantic Division, East South Central Division, West South Central Division, Mountain Division, and Pacific Division. We average industry-specific regional union coverage over the twelve Outgoing Rotation Groups (months 4 and 8 in the CPS survey rotation, in which respondents are asked about union coverage and membership) in a calendar year. Our method is thus similar to that of Western and Rosenfeld (2011), though we use slightly fewer industrial categories and more detailed regional categories.

References

Western, Bruce and Jake Rosenfeld. 2011. "Unions, Norms, and the Rise in US Wage Inequality." *American Sociological Review*. 76(4):513-537.

Appendix B: Creating Analysis Weights

We base our analysis weights on the IPUMS-provided weight EARNWT, which is a person-level weight calculated for CPS respondents who fill out the earner study (those who are 15 or older, not in the armed forces, and who are in the fourth or eighth survey month in the CPS sample). We use EARNWT rather than the final basic weight (labeled WTFINL on the IPUMS website) because our analysis uses several variables from the earner study, including workers' hourly wages, weekly earnings, whether they are paid hourly or not, and/or whether they are covered by a union at work. We then augment this weight in three ways.

First, we account for respondent attrition between consecutive monthly surveys: namely, the differential likelihood that a CPS respondent in the first or fifth survey month will respond to the next three monthly surveys and thus enter our sample. The IPUMS-created survey weights that account for this attrition, LNKFWMIS14WT and LNKFWMIS58WT, are both based on WTFINL. We first create a variable, $w^0_{i,t}$, which is equal to LNKFWMIS14WT for respondents in their first set of four consecutive monthly surveys, and equal to LNKFWMIS58WT for respondents in their second set of four consecutive monthly surveys. Next, in order to base this weight on EARNWT, then, we divide this weight by WTFINL and multiply it by EARNWT. Formally, we define our new attrition-based weight $w^1_{i,t}$ for respondent i at time t as follows:

$$w^1_{i,t} = \frac{w^0_{i,t} * EARNWT}{WTFINL}. \tag{4}$$

Second, we sequentially account for the differential likelihood that CPS respondents in the labor force will a) be employed for four consecutive months, b) be employed at the same job for four consecutive months, and c) additionally not report that they missed work or worked part-time during any of the reference weeks pertaining for the four consecutive survey months for what the CPS terms 'non-economic reasons': performing kin care or other family obligations, having an illness or medical limitation, taking maternity / paternity leave, attending school or training, performing civic or military duty, or taking time off for holiday or vacation. This weighting accounts for any correlation that could exist between a worker's propensity to experience work hour volatility and their propensity to become unemployed, switch between jobs, and/or miss some work for 'non-economic reasons' during a survey reference week.

We account for the propensity of workers to experience work hour volatility using categorical model controls: wage quartile, educational attainment, sex, age, race, marital status, the number of household children, union coverage, and the broad occupational and industry categories listed in Appendix A. Let $x_{i,t}$ be a vector of these characteristics for respondent i. Then, we create $w^2_{i,t}$, the weight for CPS respondents who remain in the sample after the above restrictions (which we term "sample restrictions"), as follows:

$$w^2_{i,t} = w^1_{i,t} \prod_{n=1}^{N} \frac{\Pr(x^n_t = x^n_{i,t} | \text{In labor force})}{\Pr(x^n_t = x^n_{i,t} | \text{Under sample restrictions})}, \tag{5}$$

where the probabilities in the numerator and denominator of the above expression are both weighted based on $w^0_{i,t}$. This procedure gives more weight to respondents whose attributes for the model variables are less common under this sample restriction than among all those in the labor force.

Third, we augment $w^2_{i,t}$ for the likelihood that a CPS respondent will self-report their labor force information in all four consecutive monthly surveys. Let SR_i be an indicator variable equal to 1 if respondent i self-reports their labor force information in four consecutive months and 0 if not. We estimate this final weight that accounts for respondent self-response, $w_{i,t}^3$, in two steps. First, we estimate a probit regression (using $w_{i,t}^2$ as regression weights) to predict SR_i as a function of the categorical model controls above:

$$\Pr(SR_i = 1) = \Phi(x_i^T \beta). \tag{6}$$

Then, we create our final weight $w_{i,t}^3$ by dividing $w_{i,t}^2$ by the predicted probability of self-reporting labor force information in all four monthly surveys:

$$w_{i,t}^3 = \frac{w_{i,t}^2}{\Pr(SR_i = 1)}. \tag{7}$$

Appendix C: Creating Figures 2 and 3

Figures 2 and 3 show trends (by wage quartile and educational attainment, respectively) in intra-year work hour volatility from 1995 to 2017, scaled so the mean work hour volatility in 1996 is equal to 100. Because reports of how many hours a respondent worked in the month's reference week systematically vary by month, we seasonally adjust these time series as follows. First, we apply a log transformation to the time series and calculate the difference between each month's logged time series value and the 12-month moving average of the logged time series. Next, we calculate the average difference across years for each month, which gives a seasonal factor associated with each month. Finally, we adjust each month's logged time series value by that month's seasonal factor and then exponentiate the time series to return the seasonally adjusted time series to its original units. We then smooth these seasonally adjusted trend lines with an Epanechnikov kernel with bandwidth equal to six months.