Estimating Effects of Wages on Smoking Prevalence Using Labor Unions as Instrumental Variables

J. Paul Leigh, PhD and Bozhidar T. Chakalov, MA

Objectives: To test for the effects of wages on smoking using labor unions as instrumental variables. **Methods:** We analyzed four waves of the Panel Study of Income Dynamics (2013 to 2019 alternate years). The overall sample included workers aged 18 to 70 years in 2013 and subsamples within blue + clerical/white-collar and private/public sector jobs (N = 37,117 to 8446 person-years). We used two instrumental variables: worker's union membership and states' right-to-work laws. **Results:** $1 (2019 US dollars) increases in wages-per-hour resulted in 1.3 ($P < 0.001$) percentage point decreases in smoking prevalence (8.2% decreases at the smoking mean). Larger effect sizes and strong statistical significance were found for blue-collar + clerical and private-sector subsamples; smaller sizes and insignificance were found for public-sector and white-collar subsamples. **Conclusions:** Unions increase wages, and higher wages, in turn, reduce smoking. Wages and labor unions are underappreciated social determinants of health.

Keywords: causal model, income

LEARNING OUTCOMES

1. Demonstrate the usefulness of instrumental variable analysis when testing for causation.
2. Analyze the effects of wages on smoking prevalence using the instrumental variable analysis.
3. Appraise the varying effects across blue-collar + clerical, white-collar, private, and public sector subsamples.

Research into the effects of minimum wage laws on public health has been rapidly expanding; more than 60 articles have been published in recent years.[1] Minimum-wage-and-health research is popular, in part, because minimum wage laws are viewed as exerting exogenous, causal forces on wages. But laws are not the only way to raise wages; unions can too, as extensive economic research demonstrates.[2] Yet, we are aware of only two studies that even partially investigate the effects of union wage premiums on measures of health,[3,4] but one of these relies on ecological data, and neither uses causal models.

A significant problem confronts investigations into the effects of wages (or income) on health or behavior. Whereas inverse correlations between wages (or income) and health and behavior are well known, the causes of these correlations are disputed. Some argue that higher wages result in better health or behavior, whereas others argue that bad health or behaviors harm wages.[5] Still others allege that some unobserved "third variable" such as time preference (ability to delay gratification) explains both high wages and good health or behavior.

In this article, we advance research into the effects of wages by (1) using the instrumental variable (IVs) technique and (2) drawing on economic research on the effects of unions. The IV technique directly addresses issues surrounding mutual causality and "third variables." Labor unions, although well-researched within economics, receive scant public health research attention. Health can be measured in numerous ways. We select smoking for several reasons. First, even during the COVID-19 pandemic, smoking killed more people in the United States than COVID-19. Moreover, smoking increases the risks of hospitalizations and deaths among people with COVID-19. Second, there are well-documented negative correlations between both income and wages on the one hand and smoking on the other.[6,7] Third, smoking is a popular health measure in the IV literature on wages and health. However, most IV studies test for effects of smoking on wages. We are aware of only one study[8] using IVs to consider the effects of wages on smoking, but this study is restricted to people who are or once were smokers and full-time workers and does not focus on labor unions as IVs.

We use two measures of labor unions as IVs: individual worker's union membership and indicators of statewide laws on right-to-work (RTW). Labor unions are likely powerful IVs. Virtually all economists agree that unions raise wages. A leading labor economics text reports that unions raise wages for private-sector workers by 20% and for public-sector workers by 10% with blue-collar workers in either sector enjoying the largest wage premiums.[2]

There are additional reasons for considering unions and wages. Unions have strongly advocated for safe working conditions and hazard pay for decades and, most recently, for essential workers during COVID-19. Inflation-adjusted wages have been stagnant or falling for large shares of American workers for over 40 years, particularly those without college educations.[9] The percent of American workers in unions has declined by more than half during these same years.[10] In 2020, private sector membership was 6%; public sector, 35%; and combined was 11% of the workforce.[11] Both stagnant wages and declines in union membership have been linked to rising income inequality,[10,12] and inequality, in turn, likely harms population health.[13] Both have also been linked to rising "deaths of despair"—liver cirrhosis, suicide, and drug overdoses—in the United States and United Kingdom.[14] "Deaths of despair" and falling wages predict votes for Trump and Brexit.[15,16] Finally, epidemiologic interest in unions and wages is increasing.[3,4,17]

METHODS

Data

The Panel Study of Income Dynamics (PSID) is a longitudinal sample, representative of the US population. It contains extensive data including information on subjects' smoking status, wages, union status, and state of residence. Our data were for "household heads" and "spouses," if any, for four recent waves: 2013, 2015, 2017, and 2019 (PSID does not collect data in even years). Because one critical variable

From the Department of Public Health Sciences, University of California Davis School of Medicine, Davis, California (J.P.L., B.T.C.); Center for Healthcare Policy and Research, University of California Davis School of Medicine, Sacramento, California (J.P.L.); Center for Poverty & Inequality Research, University of California, Davis, California (J.P.L.).

Author Contributions: J.P.L. conceived idea and wrote most of article; B.T.C. carried out computer work, analyzed data, and wrote some of the article.
Ethical considerations: Secondary data are publicly available and de-identified.
Funding sources: None to disclose.
Conflicts of interest: None declared.
Supplemental digital contents are available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.joem.org).

Address correspondence to: J. Paul Leigh, PhD, Department of Public Health Sciences, MS1C, UC Davis, Davis, CA 95616 (pleigh@ucdavis.edu).

DOI: 10.1097/JOM.0000000000002792

TABLE 1. Descriptive Statistics; Means (SDs) on Key Variables

Variables	Overall Sample	Blue Collar + Clerical	White Collar	Private Sector	Public Sector
Sample size	37,117 (person-years)	20,958 (person-years)	16,159 (person-years)	28,671 (person-years)	8,446 (person-years
Dependent variables					
Wage-per-hour ($2019)	$28.9 (28.6)	$19.4 (14.7)	$34.3 (38.3)	$25.7 (30.28)	$26.4 (21.7)
Smoker	15.9% (36.5)	21.0% (40.8)	9.19% (28.9)	17.7% (38.2)	9.7% (29.5)
Instruments					
Union member	11.2% (31.5)	12.8% (33.5)	9.02% (28.6)	6.7% (25.0)	26.4% (44.1)
Right to work	56.1% (49.6)	58.3% (49.3)	53.3% (49.9)	56.4% (49.6)	55.3% (49.7)
Covariates					
Age	41.1 (12.8)	40.9 (12.9)	41.4 (12.5)	40.6 (12.7)	42.9 (12.9)
Female	52.3% (49.9)	49.2% (49.9)	56.2 (49.6)	50.3% (49.9)	58.9% (49.2)
Black, non-Hispanic	32.7% (46.9)	39.9% (48.9)	23.3% (42.3)	31.6% (46.5)	36.3% (48.1)
Hispanic	10.3% (30.4)	12.7% (33.3)	7.17% (25.8)	10.9% (31.2)	8.21% (27.4)
Other, non-Hispanic	3.4% (18.2)	3.0% (17.1)	3.99% (19.8)	3.5% (18.4)	3.17% (17.5)
White, non-Hispanic	53.6% (49.9)	44.3% (49.7)	65.5% (47.5)	53.9% (49.8)	52.3% (49.9)
<High school	10.2% (30.3)	15.9% (36.6)	2.75% (16.4)	11.7% (32.1)	5.13% (22.1)
Exactly high school (12 y of schooling)	25.6% (43.6)	36.3% (48.1)	11.6% (32.0)	27.1% (44.4)	20.4% (40.3)
Some college	28.9% (45.3)	33.3% (47.1)	23.2% (42.2)	30.0% (45.8)	24.9% (43.3)
College graduate or more	35.4% (47.8)	14.5% (35.2)	62.5% (48.4)	31.2% (46.3)	49.5% (49.9)
Married	60.7% (48.9)	54.3% (49.8)	68.9% (46.3)	59.5% (49.1)	64.6% (47.9)

Appendix Table 1, http://links.lww.com/JOM/B259, contains all variables.

was wages, we selected only people with jobs indicated by subjects' reporting annual work hours greater than 100. We considered employees aged 18 to 70 years in 2013 (<1% of our 2015 to 2019 samples were >70). Because another critical variable was unionization, we selected only employees, thereby excluding the self-employed. The smoker variable reflected answers to this question: "Do you smoke cigarettes?" Our wage variable measured annual earnings from work including wages, salary, bonuses, overtime, tips, and commissions divided by annual work hours. Wages were adjusted for inflation with the Consumer Price Index using 2019 as the base year.

We used two measures for unions. The first indicated whether the subject belonged to a union. The second measured whether the subject resided in a state with an RTW law. Right-to-work refers to laws that states may pass requiring "open shops" whereby workers can take jobs covered by union contracts but are not required to join unions; that is, they can "free ride." One intent of RTW laws is to discourage unionization, and the laws are successful in doing that.[18] Subjects with missing data on any variable were excluded.

We conducted main analyses reported in Tables 1-3 and sensitivity analyses (Tables 5 and 6 in the Supplemental Digital Content [SDC]/Appendix, http://links.lww.com/JOM/B259). Our selection of the overall sample and subsamples was informed by labor economics studies. For example, union wage effects are known to be strongest in blue-collar and private-sector jobs but weaker in public-sector and white-collar jobs.[2] In the economics literature, clerical workers are frequently regarded as blue-collar, even though many others regard them as white-collar. We therefore created the category blue-collar + clerical. For all analyses, our overall sample comprised all employees; subsamples included blue-collar + clerical only, white-collar only (without clerical workers), private-sector only, and public-sector only. We defined blue-collar as occupation codes 360 to 465 and 600 to 975 and clerical as 500 to 594 using the 2000 US Census classification codes. Blue-collar occupations included broad categories of "health care support," construction trades, "extraction workers," repair workers, production, and transportation workers. Clerical occupations included the broad category of "office and administration support."

TABLE 2. First-Stage Regressions Dependent Variable was Wages

	Overall Sample	Blue Collar + Clerical	White Collar	Private Sector	Public Sector
Independent Variables	**Coefficient (*P*)**	**Coefficient (*P*)**	**Coefficient (*P*)**	**Coefficient (*P*)**	**Coefficient (*P*)**
Model with both instruments					
Union member	−1.220* (0.001)	4.407* (<0.001)	−4.406* (<0.001)	2.151* (<0.001)	0.192 (0.807)
RTW	−2.075* (<0.001)	−1.447* (<0.001)	−2.151† (0.027)	−1.353† (0.025)	−4.554* (<0.001)
Model with only union member					
Union member	0.939† (0.023)	4.482* (<0.001)	−3.994* (<0.001)	2.159* (<0.001)	0.491 (0.582)
Model with only RTW					
RTW	−2.32* (<0.001)	−2.43* (<0.001)	−2.486* (<0.001)	−1.42† (0.014)	−4.60* (<0.001)
Sample size for all regressions in column	37,117	20,958	16,159	28,671	8,446

Results only for instruments. Results for covariates for union member and RTW (combined) regressions available in Appendix Table 2, http://links.lww.com/JOM/B259. All regressions included age, age-squared, female, non-Hispanic Black, non-Hispanic other, Hispanic (non-Hispanic White omitted), high school graduate, some college, college or more (less than high school omitted), married, indicator (dummy) variables for three of four regions (Northeast, South, Midwest, West), three of four wave-years; random effects for individuals.

*Statistical significance at the 0.01 level, two-tailed test.

†Statistical significance at the 0.05 level, two-tailed test.

RTW, right to work.

TABLE 3. Second-Stage Regressions Dependent Variable was Smoking (0/1)

	Overall Sample	Blue Collar + Clerical	White Collar	Private Sector	Public Sector
Independent Variables	**Coefficient (*P*)**	**Coefficient (*P*)**	**Coefficient (*P*)**	**Coefficient (*P*)**	**Coefficient (*P*)**
Model with both union member and RTW					
Predicted wages	−0.013* (<0.001)	−0.008* (<0.001)	−0.000 (0.785)	−0.011† (0.017)	1.49 (0.225)
F test for strength	12.01* (<0.001)	72.35* (<0.001)	13.75* (<0.001)	8.32* (<0.001)	25.79* (<0.001)
Overidentification test for validity	0.911 (0.339)	1.359 (0.244)	9.93* (0.002)	1.138 (0.2860)	0.292 (0.589)
Model with only union member					
Predicted wages	−0.018* (<0.001)	−0.008* (<0.001)	0.003 (0.194)	−0.007 (0.177)	−0.014 (0.645)
F test for strength	3.97* (<0.001)	105.51* (<0.001)	15.88† (0.011)	10.85* (0.004)	0.30 (0.582)
Model with only RTW					
Predicted wages	−0.0112* (0.006)	−0.012* (0.008)	−0.008 (0.168)	−0.017‡ (0.052)	−0.004 (0.121)
F test for strength	14.82* (<0.001)	49.89* (<0.001)	2.97‡ (0.085)	5.81† (0.016)	51.34* (<0.001)
Sample size for all regressions in column	37,117	20,958	16,159	28,671	8,446

Results only for instruments. Results for covariates for union member and RTW (combined) regressions available in Appendix Table 3, http://links.lww.com/JOM/B259. All regressions included age, age-squared, female, non-Hispanic Black, non-Hispanic other, Hispanic (non-Hispanic White omitted), high school graduate, some college, college or more (less than high school omitted), married, indicator (dummy) variables for three of four regions (Northeast, South, Midwest, West), three of four wave-years; random effects for individuals.

*Statistical significance at the 0.01 level, 2-tailed test.

†Statistical significance at the 0.05 level, 2-tailed test.

‡Statistical significance at the 0.10 level, 2-tailed test.

RTW, right to work.

Our white-collar workers included managers, professionals, and salespeople. Samples ranged in size from highs of 14,017 persons and 37,117 person-years for all employees to lows of 3833 persons and 8446 person-years for public-sector only in Tables 1-3 and SDC/Appendix Tables 1 to 6, http://links.lww.com/JOM/B259.

Model

Any correlation between smoking and wages could be explained three ways. First, smoking could reduce wages. For example, smokers may be less productive at the job than nonsmokers because smokers might take more breaks (to smoke), and lower productivity might result in lower wages.[7] Second, low wages might increase smoking. For example, low wages might result in low self-esteem and depression that, in turn, might result in smoking.[8] Finally, some unmeasured "third variable" might explain the correlation. Low ability to delay gratification (time preference) has been linked to both low earnings and smoking.[7,8]

The IV method is designed to mimic a randomized trial. In a randomized controlled trial, a fair coin determines who will and will not be treated. In the language of IV, the coin is the instrument. An instrument has two characteristics: strength and validity. Strength measures the ability of the instrument to assign people to treatment. Validity indicates whether the instrument is correlated with the outcome. A fair coin is a perfect instrument because it precisely determines who gets treatment (strong), and by definition, it is not correlated with the outcome, for example, smoking (valid). A good instrument is a good—not perfect—predictor of who gets treatment, and although it does not randomly assign people to treatment, the only correlation it may have with the outcome is through the indirect effect it has on the treatment, for example, wages. Good instruments are best identified by researchers with knowledge of the relevant literature. We believe that measures of labor unions are good instruments because unions are well-known predictors of wages but have never, as far as we know, appeared in any study as correlated with smoking. In addition to this knowledge-of-literature argument, tests can be run to assess strength and validity. An *F* test can be constructed that compares regressions with and without the instruments in which the treatment is the dependent variable (eg, wages); the *F* test can be used to assess the strength of the instrument. *F* statistics >10 indicate strong instruments, *F* statistics 5 to 10 are moderate, and <5, weak ones.[19] But even weak IVs can be statistically significant and, arguably, adequate. When more than one instrument is available, an overidentification test can test for validity.[20] The null hypothesis in the overidentification test is that all instruments as a group are exogenous; statistical insignificance is therefore consistent with but does not "prove" that the IVs are valid; statistical significance strongly suggests that one or more of the IVs are not valid. There is no test for validity when only one IV is available.

The IV method proceeds with three steps involving two regressions. In step 1, the instruments—union membership and RTW—together with other exogenous variables such as subject's age, education, and marital status enter as covariates in the first-stage regression to predict wages. In step 2, the predicted values of wages are obtained. In step 3, the predicted values together with all other exogenous variables enter as covariates in the second-stage regression to predict smoking. We used Stata 16.1 (StataCorp, College Station, TX) throughout.

There were challenges to implementing the IV method when the predicted variable (wages) was continuous, the outcome variable was binary (smoker), and data were longitudinal. We are not aware of joint linear and logistic IV programs that incorporate longitudinal data. Linear regressions for both the instrumented variable (wages) and the outcome will, nevertheless, yield consistent estimates[21]; moreover, Angrist and Pischke[21(pp197–198)] suggest using linear regressions if sample sizes are large, as ours are. We included random effects to account for lack of independence among data points over the years associated with the same subject. There were challenges regarding standard errors, attrition, and addition of new subjects. The PSID has geographic clusters. Adjusting standard errors for geographic clustering within longitudinal data is most often done at the state level. Within Stata, these adjustments required that people remain in the same state over the four waves (8 years) of PSID data. Some attrition undoubtedly occurred, even though one study finds attrition bias is not large in the PSID.[22] In our preferred larger sample, everyone was included—even those who moved to different states or dropped out or entered the samples over 8 years. A smaller sample included only those who resided in the same state every year. The larger sample (Tables 1-3) was not adjusted for state clustering, but the smaller one was (SDC/Appendix Table 5, http://links.lww.com/JOM/B259). We prefer the larger sample because it minimizes moving and attrition bias.

A different challenge involves choice of instruments. Whereas the union member variable conveys the most obvious measure of unionization, it may not be best for IV purposes. Unmeasured personal

characteristics might influence both the decision to join a union and smoking, rendering the IV invalid. When the outcome variable (smoking) is measured on a personal level, other personal variables are frequently eschewed as IVs; geographic variables are frequently preferred.[23] Accordingly, we run some regressions with both IVs and others with only one IV at a time.

Regressions included the following covariates: age, age-squared, gender, race/ethnicity, marital status, education, and indicator variables for US regions (Northeast, South, Midwest, West) and years (2013, 2015, 2017, 2019). These covariates were selected based on our reading of the labor economics and epidemiology literatures on predictors of wages and smoking. We did not apply PSID population weights. Weights can be useful when attempting to describe samples and comparing with populations, for example, *x*% of women are in public unions, but *y*% are in private unions. We want to estimate causal effects using models, however; for these purposes, weighting can be harmful for precision.[24]

RESULTS

Descriptive statistics (Table 1) support the claim by PSID authorities that their data represent the United States. There were far more workers in the private than public sectors. Wages were higher, and smoking prevalence was lower in the overall sample—which included white-collar workers with college degrees—than in the blue-collar + clerical subsample. Union membership was highest among blue-collar + clerical workers and lowest among private-sector workers. Percent of the states with RTWs was approximately 56%. Percent female was higher in the public sector—which uses disproportionate numbers of teachers and nurses—than the private sector. Blacks, Hispanics, and people without college degrees had higher percentages in the blue-collar + clerical than the overall sample. The high percentage of Blacks reflected PSID's intentional oversampling.

Table 2 presents results for the first-stage regressions in which wage was the dependent variable. The top panel contains regressions with both IVs combined: the middle and bottom panels for union member and RTW separately. Results were consistent with labor economics.[2] Union member predicted higher wages in all but the white-collar subsample; RTW predicted lower wages across all samples; statistical significance was achieved in 18 of 20 cases. Coefficients are effect sizes. In the first panel of regressions, blue-collar + clerical workers who were union members earned $4.41 more per-hour than blue-collar + clerical workers not in unions; the union wage premium was 22.7% evaluated at the mean (4.41/19.4). Results on all covariates for regressions using both IVs are available in SDC/Appendix Table 2, http://links.lww.com/JOM/B259.

Table 3 provides IV results for the effects of wages on smoking prevalence. The top panel contains regressions with both IVs combined: the middle and bottom panels, for union member and RTW separately. For the overall sample in the first panel, a $1 increase in wages reduced smoking prevalence by 1.3 ($P < 0.001$) percentage points or by 8.2% (1.3/15.9) of the mean of smoking (15.9%). Results for the blue-collar + clerical sample were 0.8 ($P < 0.001$) percentage points and 3.8% (0.8/21.0) of the smoking mean (21.0); effects for the private subsample were negative and significant at the 0.05 but insignificant for both the public sector and white-collar subsamples. *F* statistics for the strength of the IVs were significant ($P < 0.001$) in all subsamples and larger than 10.0 in all but the private sector for the first panel. The overidentification tests indicated that we failed to reject the null hypothesis that instruments as a group were exogenous, suggesting validity, in all but the white-collar subsample. Next, consider the middle panel that corresponds to using only union member as the IV. For the overall sample, a $1 increase in wages reduced smoking prevalence by 1.8 ($P < 0.001$) percentage points or by 11.3% (1.8/15.9) of the mean of smoking (15.9%). Results for the blue-collar + clerical subsample were 0.8 ($P < 0.001$) percentage points and 3.8% (0.8/21.0) of the mean (21.0); effects for the white-collar, private, and public subsamples were insignificant. Similar findings emerged in the bottom panel for RTW only. Considering the middle and bottom panels together, *F* tests for strength of the IVs were significant in 9 of 10 analyses and exceeded 10.0 in 6 of 10 analyses. Considering findings from all three panels—union member and RTW combined, member only, and RTW only—effect sizes and *P* values were larger (in absolute value) and smaller, respectively, in the blue-collar + clerical and private-sector subsamples than the white-collar and public-sector subsamples; SDC/Appendix Tables 2 and 3, http://links.lww.com/JOM/B259, corresponding to text Tables 2 and 3, provide results on all covariates. Results were consistent across samples and mirror findings in the literature. SDC/Appendix Table 4, http://links.lww.com/JOM/B259, provides results for regressions in which the IV method was not used; these findings are useful to compare with those in Table 3. We conducted three sensitivity analysis: (1) samples were restricted to subjects who stayed in the same state over 8 years, allowing us to correct standard errors for geographic clustering within states; (2) fixed effects for states were included; (3) logistic regressions were run in the second stage. Discussion and results are shown in the SDC/Appendix, http://links.lww.com/JOM/B259. Findings, especially those in Appendix Table 5, http://links.lww.com/JOM/B259 with geographic clustering corrections, support those in text (Tables 2 and 3).

DISCUSSION

Our preferred results indicate a $1 ($2019) increase in wages-per-hour resulted in a 1.3 ($P < 0.001$) percentage point decrease in smoking prevalence—equivalent to an 8.2% decrease at the smoking mean. Among the subsamples, the effect sizes were largest (in absolute value), and statistical significance was strongest for the blue-collar + clerical and private-sector subsamples. These findings conform with expectations because economic research indicates blue-collar + clerical and private-sector workers gain higher wages due to unionization than public-sector and white-collar workers.[2] Moreover, white-collar and public-sector workers have the lowest smoking prevalence, 9.2% and 9.7%, versus 17.7% and 21.0% in the private-sector and blue-collar + clerical subsamples. There may be threshold effects whereby rising wages cannot reduce smoking prevalence below, for example, 10%; there may be diminishing returns to the effects of wages on smoking.

The IV method tests for causality running from wages to smoking and is therefore superior to linear or logistic regression that essentially tests for correlations. Our *F* tests and overidentification tests indicate our instruments are significantly (most frequently) strong (>10) and valid. Even though a recent literature review did not find any studies on direct links between union membership and smoking,[25] we acknowledge that union membership may be a problematic IV. Our results are similar when we exclude that variable, however, and rely only on the geographic IV measuring whether the state has an RTW law.

Our results, however, should not be interpreted as indicating that smoking cannot also result in lower wages. Our application of the IV technique simply tests for the effects of wages on smoking, not the reverse. Our view is that there are likely effects going both ways but that the effects of wages on smoking dominate.

There are hypotheses that explain these findings. Low wages can result in poverty-related stress,[26] and the stress, in turn, could lead to less desire or ability to quit smoking. College students who are smokers are less inclined to quit during stressful finals week than at other times. Low wages could also result in low self-esteem that, in turn, could lead to smoking or less likelihood of quitting.[27] Dr Martin Luther King, Jr, put it this way: "We look around every day and we see thousands and millions of people making inadequate wages. Not only do they work in our hospitals, they work in our hotels, they work in our laundries, they work in domestic service, they find themselves underemployed. You see, no labor is really menial unless you're not getting adequate wages. People are always talking about menial labor. But if

you're getting a good (wage) as I know that through some unions they've brought it up...that isn't menial labor. What makes it menial is the income, the wages."[28]

Our results are consistent with economic studies finding smoking is "inferior," an economic term reflecting negative correlations between consumer products and income.[29] They are also consistent with Reeves'[4] finding of associations between higher unionization rates and lower mortality rates across 22 high-income countries that he attributes to the effects of unions on wages as well as O'Brien and colleagues'[30] finding that RTW predicted mortality across 50 states. Our results align with those from an IV study that finds higher wages reduce smoking prevalence among people with lower socioeconomic status.[8] But that study[8] restricts samples to only people who were once smokers and only full-time workers; our study does not impose these restrictions, and our sample sizes were nearly five times larger. Our results showing larger IV coefficients (effect sizes) in Table 3 but smaller non-IV coefficients in SDC/Appendix Table 4, http://links.lww.com/JOM/B259, which used linear regression, not IV, are consistent with other IV studies involving the health effects of education.[31] Grossman[31] suggests, for example, that spillover effects can be captured by IVs. When one worker quits smoking, ripple effects on coworkers can lead to their decisions to also quit. Our findings suggest that downward bias emanates from estimating effects that do not account for endogeneity. Finally, our results are consistent with Lenhart's[32] study showing higher minimum wages lead to lower smoking prevalence.

The results on the union variable for the first-stage regressions in Table 2 coincide with well-known findings in labor economics. Our union wage effects were larger for blue-collar + clerical and private-sector than public-sector workers. The negative correlations between union member and wages in Table 2, as well as the insignificant findings for the white-collar subsample in Table 3, are easily explained. Unions rarely organize high-paying professionals or managers such as doctors, lawyers, or chief executive officers; they organize, instead, lower-paying professionals such as teachers.[2] Among white-collar workers, unions are a marker, not a cause, of low wages. A proper analysis of the effects of unions among white-collar workers would require a narrower sample, for example, compare teachers with and without unions.

Results were also consistent with well-known findings for all other covariates including age, age-squared, gender, race/ethnicity, marital status, and education in predicting wages (SDC/Appendix Table 2, http://links.lww.com/JOM/B259).[2] We find similarly consistent epidemiologic results in SDC/Appendix Table 3, http://links.lww.com/JOM/B259, with respect to smoking for which prevalence is higher for the middle-aged, for men, the nonmarried, the less educated, and those living in the South.[33] The fact that so many of our results on key and control variables mirror those in economics and epidemiology bolster the credibility of our data and method.

There are additional advantages to our data and method. First, the PSID is highly regarded, large, longitudinal, and representative of the United States. Second, our analysis of panels that include people who may not have data for all four waves likely reduces attrition bias in random-effects models.[34] Attrition bias might be problematic for the question we address.[35] It could be that disproportionate numbers of people in low-wage jobs and/or those who smoke could die or drop out of the sample over these 8 years. Because these people are critical to our analysis, samples that eliminate people who attrite could produce bias.

Our study has limitations. First, smoking is self-reported, but self-reported smoking has high validity compared with biological measures.[36] Second, what is true for wages may not be true for all forms of income. In addition, our sample was limited to only those people with jobs, thereby perhaps inviting some healthy-worker bias. Third, we did not include employer-provided health insurance (EPHI) as a covariate. In part, this was intentional. We regard EPHI as an additional wage, part of "total employee compensation." In addition, the PSID provides only binary—yes/no subject has EPHI—and not monetary values for EPHI, thereby undermining our ability to add it to the wage. Any bias introduced by omitting EPHI suggests we underestimated the effects of wages on smoking because EPHI and wages are strongly and positively correlated.[2] Finally, although smoking is binary, we used two-stage least-squares. We are not aware of any statistical or econometric article that offers a closed-form solution for using logistic regression with the IV approach and longitudinal data, let alone commands in popular programs such as Stata or R. Joshua Angrist, a Noble Prize–winning economist, suggests using two-stage least-squares when the outcome variable is binary because it is a consistent estimator.[22(pp197–198)] Nevertheless, we conducted a makeshift analysis. We first regressed wages on the two instruments and all other covariates with linear, least-squares regression. We obtained the predicted values of wages, which we labeled wagehat. We then ran logistic regressions with smoking regressed on wagehat and other covariates. Results were strongly consistent with those in the two-stage least-squares findings in text (Table 3). The standard errors are not (cannot be) corrected for the IV technique. The results are available in SDC/Appendix Table 6, http://links.lww.com/JOM/B259.

Unions lift wages. We showed that the union wage advantage is negatively associated with smoking prevalence. Because we used the IV method, this association is likely causal. Smoking is only one measure of health, however; future research might consider other measures. Unions are only one exogenous effect on wages; future research might consider other effects such as minimum wages, monopsony, or laws governing occupational licensing and non-compete clauses. Finally, future research might consider whether the effects differ across gender, race, and ethnic categories.

REFERENCES

1. Leigh JP. Treatment design, health outcomes, and demographic categories in the literature on minimum wages and health. *Econ Hum Biol*. 2021;43:Article 101043.
2. McConnell CR, Brue SL, Macpherson DA. *Contemporary Labor Economics*. 11th ed. New York, NY: McGraw-Hill; 2017.
3. Reeves A, Loopstra R, Tarasuk V. Wage-setting policies, employment, and food insecurity: a multilevel analysis of 492 078 people in 139 countries. *Am J Public Health*. 2021;111:718–725.
4. Reeves A. The health effects of wage setting institutions: how collective bargaining improves health but not because it reduces inequality. *Sociol Health Illn*. 2021;43:1012–1031.
5. Glymour MM, Avendano M, Kawachi I. Socioeconomic status and health. In: Berkman LF, Kawachi I, Glymour MM, eds. *Social Epidemiology*. New York, NY: Oxford University Press; 2014.
6. Casetta B, Videla AJ, Bardach A, et al. Association between cigarette smoking prevalence and income level: a systematic review and meta-analysis. *Nicotine Tob Res*. 2017;19:1401–1407.
7. Levine PB, Gustafson TA, Velenchik AD. More bad news for smokers? The effects of cigarette smoking on wages. *Ind Labor Relat Rev*. 1997;50:493–509.
8. Du J, Leigh JP. Effects of wages on smoking decisions of current and past smokers. *Ann Epidemiol*. 2015;25:575–582.
9. Mishel L. Unions, inequality, and faltering middle-class wages. Washington, DC: Economic Policy Institute. Issue Brief #342. 2012. Available at: https://www.epi.org/publication/ib342-unions-inequality-faltering-middle-class/. Accessed January 27, 2020.
10. Mishel L, Gould E, Bivens J. *Wage Stagnation in Nine Charts*. Washington, DC: Economic Policy Institute; 2015. Available at: https://www.epi.org/publication/charting-wage-stagnation/. Accessed November 21, 2019.
11. Bureau of Labor Statistics, 2020. Economic news release. Union Members Summary. 2020. Available at: https://www.bls.gov/news.release/union2.nr0.htm. Accessed January 23, 2020.
12. Kristal T, Cohen Y. The causes of rising wage inequality: the race between institutions and technology. *Socioecon Rev*. 2017;15:187–212.
13. Kawachi I, Kennedy BP. *The Health of Nations: Why Inequality Is Harmful to Your Health*. New York, NY: The New Press; 2002.
14. Case A, Deaton A. *Deaths of Despair and the Future of Capitalism*. Princeton, NJ: Princeton University Press; 2020.
15. Monnat SM. Deaths of despair and support for Trump in the 2016 election. Pennsylvania State University Department of Agricultural Economics Research Brief, 2016. Available at: https://smmonnat.expressions.syr.edu/wp-content/uploads/ElectionBrief_DeathsofDespair.pdf. Accessed June 27 2022.
16. Blanchflower D. *Not Working: Where Have All the Good Jobs Gone?* Princeton, NJ: Princeton University Press; 2019.

17. Eisenberg-Guyot J, Mooney SJ, Barrington WE, Hajat A. Does the union make us strong? Labor-union membership, self-rated health, and mental illness: a parametric G-formula approach. *Am J Epidemiol*. 2021;190:630–641.
18. Zoorob M. Does 'right to work' imperil the right to health? The effect of labour unions on workplace fatalities. *Occup Environ Med*. 2018;75:736–738.
19. Staiger D, Stock J. Instrumental variables regression with weak instruments. *Econometrica*. 1997;65:557–568.
20. Baltagi BH. Simultaneous equations with error components. *J Econom*. 1981; 17:189–200.
21. Angrist JD, Pischke J-S. *Mostly Harmless Econometrics: An Empiricists Companion*. Princeton, NJ: Princeton University Press; 2009.
22. Zabel JE. An analysis of attrition in the panel study of income dynamics and the survey of income and program participation with an application to a model of labor market behavior. *J Hum Resour*. 1998;33:479–506.
23. French MT, Popovici I. That instrument is lousy! In search of agreement when using instrumental variables estimation in substance use research. *Health Econ*. 2011;20:127–146.
24. Solon G, Haider SJ, Wooldridge JW. What are we weighting for? *J Hum Resour*. 2015;540:301–316.
25. Leigh JP, Chakalov B. Labor unions and health: a literature review of pathways and outcomes in the workplace. *Prev Med Rep*. 2021;24:article number 101502.
26. Kezios KL, Zhang A, Kim S, et al. Association of low hourly wages in middle age with faster memory decline in older age: evidence from the health and retirement study. *Am J Epidemiol*. 2022;191:2051–2062.
27. Theodossiou I. The effects of low-pay and unemployment on psychological well-being: a logistic regression approach. *J Health Econ*. 1998;17:85–104.
28. King ML. Speech to SEIU Local 1199; Salute to Freedom. Memphis, Tennessee. 1968. Available at: https://www.afscme.org/about/history/mlk/dr-martin-luther-king-jr-on-labor. Accessed December 6, 2022.
29. Chaloupka FJ, Warner KE. The economics of smoking. In: Culyer AJ, Newhouse JP, eds. *Handbook of Health Economics, Edition 1, Volume 1*. Amsterdam, Netherlands: Chapter 29. Elsevier; 2000:1539–1627.
30. O'Brien R, Bair EF, Venkataramani AS. Death by robots? Automation and working-age mortality in the United States. *Demography*. 2022;59:607–628.
31. Grossman M. The demand for health, 30 years later: a very personal retrospective and prospective reflection. *J Health Econ*. 2004;23:629–636.
32. Lenhart O. Do higher minimum wages benefit health? Evidence from the UK. *J Policy Anal Manage*. 2017;36:828–852.
33. Centers for Disease Control and Prevention. Burden of cigarette use in the U.S. 2021. Available at: https://www.cdc.gov/tobacco/campaign/tips/resources/data/cigarette-smoking-in-united-states.html. Accessed September 22, 2021.
34. Baltagi BH, Song SH. Unbalanced panel data: a survey. *Stat Pap*. 2006;47:493–523.
35. Siddiqui O, Flay BR, Hu FB. Factors affecting attrition in a longitudinal smoking prevention study. *Prev Med*. 1996;25:554–560.
36. Patrick DL, Cheadle A, Thompson DC, et al. The validity of self-reported smoking: a review and meta-analysis. *Am J Public Health*. 1994;84: 1086–1096.

