NURSE UNIONS AND PATIENT OUTCOMES

ARINDRAJIT DUBE, ETHAN KAPLAN, AND OWEN THOMPSON*

The authors estimate the impact of nurse unions on health care quality using patient-discharge data and the universe of hospital unionization in California between 1996 and 2005. They find that hospitals with a successful union election outperform hospitals with a failed election in 12 of 13 potentially nurse-sensitive patient outcomes. Hospitals were more likely to have a unionization attempt if they were of declining quality, as measured by patient outcomes. When such differential trends are accounted for, unionized hospitals also outperform hospitals without any union election in the same 12 of 13 outcome measures. Consistent with a causal impact, the largest changes occur precisely in the year of unionization. The biggest improvements are found in the incidence of metabolic derangement, pulmonary failure, and central nervous system disorders such as depression and delusion, in which the estimated changes are between 15% and 60% of the mean incidence for those measures.

Economists have long recognized the possibly contradictory effects of trade unions on worker productivity and product quality. On the one hand, unionization can improve worker productivity or product quality through a variety of mechanisms induced through greater levels of worker voice (Freeman and Medoff 1984). On the other hand, product quality or productivity may suffer from restrictive union-imposed workplace rules, reduced skill investment, and weaker employer incentives to screen for better workers, among other considerations (Lee 1978; Wessels 1994; Card 1996). Empirically, the evidence on this topic has been mixed, and the causal relationship between labor relations and productivity-related outcomes has been difficult to discern because of the dual challenge of reliably

*Arindrajit Dube is an Associate Professor at the University of Massachusetts Amherst. Ethan Kaplan is an Associate Professor at the University of Maryland, College Park. Owen Thompson is an Assistant Professor at the University of Wisconsin-Milwaukee. We wish to thank Jonathan Gruber, Alex Mas, and participants at the Employment Relations in Healthcare Conference at Rutgers and the 2015 Allied Social Sciences Association (ASSA) session on Employment Relations in the Healthcare Setting for helpful comments. We also thank Biswo Poudel and Simon Sturn for excellent research assistance and the staff at OSHPD for their help with use of the data. Finally, we acknowledge support from the University of California Labor and Employment Research Fund. Computer programs used to generate the results presented in the article are available from the authors at adube@econs.umass.edu.

Keywords: unions, health care quality, nursing, productivity

DOI: 10.1177/0019793916644251. © The Author(s) 2016
Journal website: ilr.sagepub.com
Reprints and permissions: sagepub.com/journalsPermissions.nav

measuring productivity and isolating plausibly exogenous variation in unionization.

In this article, we estimate the impact of the presence of nurses' unions on a host of patient outcomes considered in the medical nursing literature to be potentially sensitive to the performance of nurses. Understanding the impact of nurses' unions is important for a number of reasons. First, nurses play a critical role in health care delivery, and health care is not only important in and of itself but also accounted for approximately 17% of GDP in 2013. Second, union coverage for registered nurses (RNs) has not fallen in the same way as it has for the workforce overall. For the time period we study (1988 to 2005), the nationwide union coverage rate dropped from 19.0 to 13.7% for the workforce overall. In contrast, the coverage rate for RNs was much stabler, falling from 19.8 to 18.7%. In part, nurses' union coverage rates have been stabler because of successful organizing drives; this makes studying the impact of unionization for this group more feasible than for most other occupations.

Third, the availability of patient-outcome data makes nurses' unions a natural place to look to better understand the impact of unionization on product or service quality. We use nurses' union election data matched to a panel of administrative patient-outcome data from the universe of hospitals that report to the state of California. Different from most workers, nurses work for employers who are required to report detailed outcome measures. Thus, reliable data are available from which to measure the impact of nurse unionization on quality, and we specifically use the 13 outcomes identified as being potentially nurse-sensitive in a seminal article by Needleman et al. (2002) in the *New England Journal of Medicine*.

We use a difference-in-differences research design, and our preferred empirical strategy compares changes in patient outcomes at hospitals undergoing successful unionization with changes where the effort failed. As an alternative strategy, we also compare hospitals with union victories to the (much larger) set of hospitals that did not hold a unionization election during our study period. To help account for different trends in hospitals' pre-unionization patient outcomes, we present results that include hospital-specific time trends in addition to standard difference-in-differences specifications, and we also conduct an extensive battery of robustness checks to account for various threats to our approach. Finally, we estimate dynamic models that allow us to observe the timing of changes in patient outcomes relative to unionization.

Existing Literature

This article contributes to the broader literature on labor relations as a potential determinant of productivity and product quality. For the most part, existing literature has either documented cross-sectional differences between unionized and non-union firms, or considered changes in the output or service quality surrounding strike activity. Clark (1980a) found unionization has a positive 6 to 8% impact on productivity using cross-sectional regressions of output on capital and labor use. Krueger and Mas (2004) found that, when unionized workers returning from a strike worked side by side with the replacement workers who had been used during the strike, the error rates in tire production increased significantly. Mas (2008) similarly documented that the resale price of Caterpillar equipment was substantially lower when it was produced during contract negotiations. Finally, Mas (2006) showed that police unions in New Jersey that won final offer arbitrations experienced increases in apprehension relative to those police unions that lost. In contrast to these studies, we use panel methods to look at the medium-run impact of unionization itself.

A relatively small number of recent studies have also specifically estimated the impact of nurse's unions on various measures of patient-care quality. Gruber and Kleiner (2010) used an event study methodology to estimate the impact of nurses' strikes on mortality and found that the average strike increases mortality by more than 18%. Strike days are a very small portion of the total days of employment for a unionized workforce, however; therefore, looking at the impact of strikes is not sufficient for understanding the quality impact of nurses' unionization. Similar to this article, Ash and Seago (2004) estimated the impact of nurses' unions on patient health, specifically considering mortality attributable to cardiac arrest. They found that patients in hospitals where nurses are unionized are 5.5% less likely to die from a myocardial infarction. Although they use a variety of methods to account for the selectivity of unionization, their identification is cross sectional. Given the nonrandom nature of unionization documented here, a concern arises about whether these estimates necessarily reflect a causal impact of unionization.

The previous work most closely related to ours is by Sojourner, Town, Grabowski, and Chen (2015), who used a regression discontinuity design to study patient-care quality measures in a national sample of nursing homes following unionization events. The authors found no impact of unionization on care quality, although they did find a decline in employment, which they argued indicates increased productivity. Given that the patient profiles and basic nature of nursing tasks in the convalescent settings studied by Sojourner et al. (ibid.) are substantively different from those in the general hospital settings that we study here, as well as differences in sample frame, outcome measures, and methodological approach, we view our work as complementary to that of Sojourner et al.

Existing research also identified and studied a variety of possible mechanisms through which aspects of labor relations could affect worker productivity, both in general work settings and specifically for nurses. In the general case, union presence may increase productivity or product quality through reduced turnover, increased worker effort, and improved worker morale, especially when unionization raises the relative wages of the unionized workers (Freeman and Medoff 1984; Mas 2006, 2008). Related mechanisms connecting union activity and productivity that have been studied in the industrial relations literature include skill upgrading (Lewis 1963), greater capital intensity of production, and changes in managerial practices in response to unionization (Slichter, Healy, and Livernash 1960; Clark 1980b; Kochan, Eaton, McKersie, and Adler 2009; Litwin 2011).

With respect to the work settings of nurses specifically, health policy researchers have found that several of their working conditions—as measured by indices of the nursing work environment, as in Lake (2002)—have detectable effects on nurse productivity and patient outcomes (Aiken et al. 2002; Kuokkanen, Leino-Kilpi, and Katajisto 2003; Friese et al. 2008). Unionization may directly affect important aspects of the work environment, such as nurses' involvement in hospital governance, ongoing professional development opportunities, collegial nurse–physician relations, and fair procedures for resolving nurse–supervisor disputes; unionization may also affect aspects of nursing that have been shown to interact with the nurses' work environment, such as length of nursing shifts, staffing levels, and retention rates (Aiken et al. 2011). (We return to these potential mechanisms later in the article.)

Data

Our data on patient outcomes come from the Patient Discharge Database (PDD) maintained by the Office of Statewide Health Planning and Development (OSHPD) in California. The PDD is a confidential data set covering all individuals discharged from regulated California hospitals between 1988 and 2005.

The PDD contains information on a large number of patient diagnoses. To ensure that our selection of patient-outcome measures was not inappropriately influenced by the corresponding end results, we pre-committed to extracting only 13 diagnosis measures that were previously studied by Needleman et al. (2002), a seminal article in the nursing-quality literature. We used all the measures reviewed by Needleman et al. with the exception of Length of stay, which is not a patient-welfare measure. We then extracted only these 13 outcomes from the PDD microdata.

The specific patient-outcome measures and corresponding acronyms are urinary tract infection (UTI), pressure ulcer (PRU), hospital-acquired pneumonia (HAP), hospital-acquired sepsis (HAS), shock or cardiac arrest (SCA), upper gastrointestinal bleeding (UGB), metabolic derangement (MDB), deep vein thrombosis (DVT), central nervous system disorder (CNS), wound infection (WIN), failure to rescue (FTR), and in-hospital death (IHD). Each of these measures is constructed using the *International Classification of Diseases*, Ninth Revision (ICD9) diagnostic codes described in Needleman et al. (2002). We received permission from OSHPD to extract the counts of these outcomes at the hospital level, disaggregated by patient demographic categories (race, age, and gender), the major diagnostic code for admission, and the month of admission.

Note that the original research by Needleman et al. (2002) focused on the effect of nurses' staffing levels on patient outcomes and found consistently significant associations only between staffing levels and 5 of the 13 patient-outcome measures. Nevertheless, we report the results for all 13 measures because Needleman et al.(as well as other leading experts in nursing quality) consider all 13 measures to be potentially nurse-sensitive and specifically use the phrase "outcomes potentially sensitive to nursing" to describe the measures. The inclusion of all 13 measures in the present article, rather than just the outcomes for which Needleman et al. found associations, seems appropriate given that we are studying the effects of nurse unionization, which could affect factors such as morale, retention, and capital investments by hospitals, in addition to shift lengths and staffing ratios. We also note that numerous recent studies (e.g., Pappas 2008; Twigg et al. 2011; Blegen et al. 2013) used overlapping sets of measures to assess how various aspects of nursing quality affect patient outcomes (for a detailed discussion of classifying the nurse-sensitivity of different patient outcomes, see Laschinger and Almost 2003). Moreover, by considering all 13 outcomes we originally obtained, we avoid the pretest bias that might occur from selective omission of some of the outcomes.

A distinct strength of the PDD data is that, from 1996 forward, they contain counts for each outcome that was present on admission. In most of our results, we report the incidence of these outcomes when they were probably obtained while the patient was in the hospital (i.e., they were not present on admission [NPOA]). We collapse our data to the hospital-year level so that our working data set is an annual panel of hospitals from 1996 to 2005 with incidence of the 13 conditions that were not present on admission, along with the share of patients by gender, 4 race and ethnicity categories; 8 age categories; 25 major diagnostic codes; and present-on-admission levels for each of the outcomes.

Using the total number of admitted patients, we express each of our 13 hospital-year outcome measures in incidence rate per 1,000 patients. In addition, we construct an aggregate measure of disease incidence across outcomes for each hospital-year. To do so, we first standardize each specific measure by subtracting its mean and dividing by the standard deviation taken over the entire panel of hospitals. We then take the simple average of these standardized outcomes across all 13 measures. That is, letting h index hospitals, t index time, and j index outcomes, so that y^j_{ht} corresponds to outcome j at hospital h at time t, our combined measure is defined as

$$\text{All}_{ht} = \frac{1}{13}\sum_{j=1}^{13} z^j_{ht} \qquad (1)$$

where z_{ht}^{j} is the number of standard deviations from the mean observed for outcome j at hospital h in time t (i.e., the outcome measured as a z-score).

With our patient-outcome data, we merge information on the universe of National Labor Relations Board (NLRB)-conducted union-representation elections for bargaining units that included RNs in California occurring during the sample period. The majority of the information in our election data originated in monthly NLRB election reports; however, these reports contain only broad industry and bargaining-unit classifications and differentiating RN unionization elections from elections involving, for instance, nursing assistants or hospital clerical, food service, or janitorial workers is often impossible. To gain more precise information on the categories of workers in the hospitals that were included in each NLRB election, we purchased supplemental data from a private analytics firm, the Bureau of National Affairs (BNA) Employment and Labor Division. The BNA data were compiled through systematic searches of periodicals and court documents and specifically identify which certification elections in hospitals included RNs. In all but three instances, the bargaining units containing RNs consisted exclusively of RNs. In the remaining three instances, RNs were grouped with other skilled hospital staff, such as pharmacists, dieticians, and lab technicians. The results we present here are not substantively changed when the three cases are excluded.

We then hand-matched hospitals that had RN elections to the corresponding patient-discharge data using the municipality where the hospital was located and the hospital's commercial name and/or parent company. This procedure yielded matches for 50 RN union-certification elections, 39 of which resulted in the certification of a union and 11 of which did not. In two instances, a hospital had a failed RN union-certification election followed by a successful election at a later date. We retain these cases and categorize them as union wins after the date of the successful election, but our results are not substantively changed if we, instead, classify them as union losses in the interim period or exclude them.

Although a successful unionization election is the most common source of a change in union status in our study's context, changes in union status can also occur because of decertification elections or voluntary union recognition by an employer, commonly referred to as card-check. Our data indicate that five decertification elections were held in RN bargaining units during our study period and that all were unsuccessful (i.e., no change in union status occurred). Likewise, one instance of a card-check agreement is listed in our BNA data, but whether a union was actually formed is unknown to us because card-check procedures occur outside of the formal NLRB process. Because the decertifications and card-check did not result in any confirmed changes in union status in our data, we focus exclusively on the unionization elections. Moreover, our main results are very similar if the hospitals that had decertification elections are simply excluded. A final important aspect of the unionization process is the negotiation of an initial contract. Unfortunately our data do not contain information on the existence or characteristics of first contracts. (We consider the implications of lags between union recognition and first contracts later in the article.)

Figure A.1 (in the Appendix) displays a histogram of the years in which our sample of unionization elections took place. As the figure indicates, elections took place at relatively uniform time intervals over the sample period, allowing us to reliably discern time effects from the impact of unionization. Figure A.2 displays the histogram of the vote shares in favor of unionization. The limited number of union losses makes our sample unsuitable for a regression discontinuity research design, but Figure A.2 does indicate that a majority of the successful elections in our sample were at least reasonably close, and only in relatively few cases did the union vote share exceed 65%. Although not a substitute for a full analysis based on the discontinuity in assignment of union status occurring at 50% of the vote share, the fact that most union victories are relatively close in vote share to union losses increases the likelihood that our difference-in-differences estimates reflect the causal impact of unionization as opposed to trends in hospitals associated with union election victories.

Empirical Specification

We employ a difference-in-differences research design using two sets of control hospitals. Our preferred control group consists of hospitals that experience a failed unionization attempt. We believe that this control group better accounts for unobserved confounders that may be correlated with unionization status. Our second control group consists of all hospitals in the OSHPD data.

For each of these samples, we use two main specifications. In the first specification, we regress the outcome measure j at time t in hospital h, denoted by y^j_{ht}, onto the union status of the hospital's RNs, hospital fixed effects, year effects, and a vector of control variables:

$$y^j_{ht} = \propto + \beta \text{Union}_{ht} + X_{ht}\Gamma + D_h + I_t + \varepsilon_{ht} \tag{2}$$

where Union_{ht} is a dummy variable that is equal to 1 when an RN union is present, D_h is a set of hospital dummies, I_t is a set of year dummies, and X_{ht} is a vector of controls that contains the percentage of patients discharged in each year by gender, 4 race and ethnicity categories, 8 age categories, 25 major diagnostic codes, and present-on-admission levels for each of the outcomes.

Our second specification is similar to Equation (2), but also allows for a hospital-specific time trend, $\theta_h t$:

$$y^j_{ht} = \propto + \beta \text{Union}_{ht} + X_{ht}\Gamma + D_h + I_t + \theta_h t + \varepsilon_{ht} \tag{3}$$

Finally, to examine the timing of the impact, we also estimate dynamic specifications in which we include leads and lags in union presence as regressors:

$$(4) \qquad y^{j}_{ht} = \propto + \sum_{k=-F}^{L} \beta_k \mathrm{Union}_{ht+k} + X_{ht}\Gamma + D_h + I_t + \varepsilon_{ht}$$

All other regressions reported are modifications of these basic specifications.

Results

In this section, we present the descriptive statistics by unionization status, our baseline difference-in-differences estimates, as well as dynamic evidence and a variety of robustness checks.

Descriptive Statistics on Hospitals

We consider three types of hospitals in our analysis, based on the incidence of, and results from, union elections during our sample period: hospitals that did not have any nurses' union elections during our sample, hospitals that had an election in which the union won, and hospitals that had an election in which the union lost. This leaves us with a total of 616 hospitals in our sample with no elections, 39 with union wins, and 11 with union losses. We present the descriptive statistics for the full sample of hospitals as well as these three subsamples in Table 1.

Table 1 reveals substantial differences in the racial composition of patients across hospitals with no union election, those with a successful election, and those with a failed election. In particular, the proportion of white patients in hospitals with a failed union election is more than 10 percentage points greater than in hospitals with no union election (0.587 compared to 0.483), while the proportion of white patients in hospitals with a successful union election is more than 5 percentage points greater than in hospitals with no election (0.538 compared to 0.483). Hospitals that held elections have correspondingly lower proportions of black and Latino patients. These differences are likely to reflect geographical differences in union activity, with RN unionization campaigns being more common in wealthier urban areas over our study period.

Turning to other patient demographics, small differences in the gender and age composition of patients are apparent between hospitals that held an election and those that did not: hospitals that held an election had a somewhat older and more heavily female patient mix. Comparing hospitals with successful unionization elections to those with failed elections, we note that hospitals that held a losing election had a substantially higher proportion of patients over age 65 than did hospitals that held a winning election (0.460 compared to 0.411). Similarly, hospitals in which the union lost the

Table 1. Means and Standard Deviations by Union Event Status

Variable	*Full sample*	*No election*	*Won election*	*Lost election*
Female	0.592	0.590	0.600	0.599
	(0.0581)	(0.0602)	(0.0464)	(0.0203)
White	0.492	0.483	0.538	0.587
	(0.269)	(0.269)	(0.272)	(0.245)
Black	0.0752	0.0802	0.0482	0.0391
	(0.0943)	(0.0995)	(0.0470)	(0.0434)
Latino	0.270	0.274	0.249	0.218
	(0.196)	(0.203)	(0.157)	(0.0959)
Asian	0.0633	0.0633	0.0680	0.0388
	(0.0776)	(0.0742)	(0.102)	(0.0385)
Under 18	0.220	0.219	0.232	0.179
	(0.138)	(0.133)	(0.174)	(0.0569)
Over 65	0.391	0.386	0.411	0.460
	(0.155)	(0.158)	(0.147)	(0.0794)
Urinary tract infection (UTI)	5.936	5.684	7.719	5.760
	(7.101)	(7.336)	(5.671)	(2.920)
Pressure ulcer (PRU)	0.871	0.863	0.959	0.744
	(1.584)	(1.675)	(0.947)	(0.665)
Hospital-acquired pneumonia (HAP)	8.489	8.242	10.16	8.693
	(5.950)	(6.061)	(5.222)	(3.861)
Hospital-acquired sepsis (HAS)	0.527	0.508	0.690	0.375
	(1.044)	(1.065)	(0.977)	(0.372)
Shock or cardiac arrest (SCA)	3.270	3.224	3.588	3.261
	(2.823)	(2.957)	(1.848)	(1.901)
Upper gastrointestinal bleeding (UGB)	1.003	0.960	1.302	0.995
	(1.529)	(1.600)	(1.051)	(0.636)
Pulmonary failure (PNF)	3.410	3.241	4.540	3.589
	(3.082)	(3.008)	(3.387)	(2.854)
Metabolic derangement (MDB)	0.650	0.632	0.728	0.869
	(0.639)	(0.645)	(0.528)	(0.822)
Deep vein thrombosis (DVT)	1.158	1.108	1.492	1.222
	(1.259)	(1.252)	(1.350)	(0.666)
Central nervous complication (CNS)	1.041	0.988	1.379	1.184
	(1.698)	(1.696)	(1.755)	(1.262)
Wound infection (WIN)	1.521	1.483	1.761	1.619
	(1.288)	(1.277)	(1.399)	(0.906)
Failure to rescue (FTR)	6.766	6.727	6.974	7.077
	(4.583)	(4.870)	(2.342)	(2.232)
In-hospital death (IHD)	22.51	22.42	23.02	23.01
	(18.50)	(19.85)	(6.948)	(4.975)
Number of hospitals	666	616	39	11
Hospital-years	4,987	4,522	385	80

Notes: Standard deviations are in parentheses. Patient outcomes are measured as incidence per 1,000 patients. Statistics are weighted by total number of patients.

election had a substantially lower proportion of patients under age 18 (0.179 compared to 0.232).

The remaining rows of Table 1 compare the disease prevalence of our patient-outcome measures across the three subsamples. Recall that these are prevalence rates for poor health conditions—so higher rates indicate worse outcomes. Although the exact differences across the three groups depend on the outcome measure, disease incidence tends to be lowest in hospitals with no election, followed by hospitals with a losing election, and finally hospitals with a winning election, which tend to have the highest incidence of poor health conditions. Specifically, incidence rates at hospitals with no election are the lowest of the three groups for all but two of our 13 measures (the exceptions are pressure ulcers and hospital-acquired sepsis), and incidence rates at hospitals with a winning election are the highest of the three groups for all but two of our 13 measures (the exceptions are metabolic derangement and failure to rescue).

The reported differences in both demographics and mean patient outcome levels across the three groups of hospitals suggest some caution is warranted in our subsequent analysis because differences in observable patient characteristics may be indicative of unobserved differences as well, which could bias comparisons. As we show below, however, no differences seem to exist in the pre-existing patient-outcome trends between our sample of union wins and our sample of union losses, or in the trajectory of patient outcomes between our sample of union wins and our full sample after hospital-specific time trends are accounted for.

Baseline Results

Our baseline estimates of the impact of nurse unionization on potentially nurse-sensitive patient outcomes show that hospitals with union victories in California during the 1990s and early 2000s were hospitals of poor and declining quality, as measured by patient outcomes. But hospitals with successful unionization elections performed better subsequently relative to those in which the unionization drive failed. In addition, we find that hospitals with successful unionization elections outperformed hospitals in which no unionization attempt was made once we account for hospital-specific time trends.

In Table 2, we show estimates for both the full sample and the sample of hospitals with union elections, using models with and without hospital-specific time trends. The coefficients in Table 2 represent the impact of unionization on disease per 1,000 patients. To aid in the visual interpretation of the results, in Figure 1 the coefficients from Table 2, columns (2) and (4), are converted to percentage changes using the sample-wide mean of each outcome; these are plotted along with the associated 90% confidence intervals (CIs). The full sample is almost 10 times the size of the election sample, with nearly 5,000 observations for each outcome measure, compared to 465 in the election sample. Recall that in the election sample, the key coefficient measures the change in outcome in hospitals following a successful union election relative to hospitals that had a failed union election. In contrast, in the full sample the coefficient represents the change in outcome in hospitals following successful union election relative to all other hospitals—including those that had no elections during our sample period.

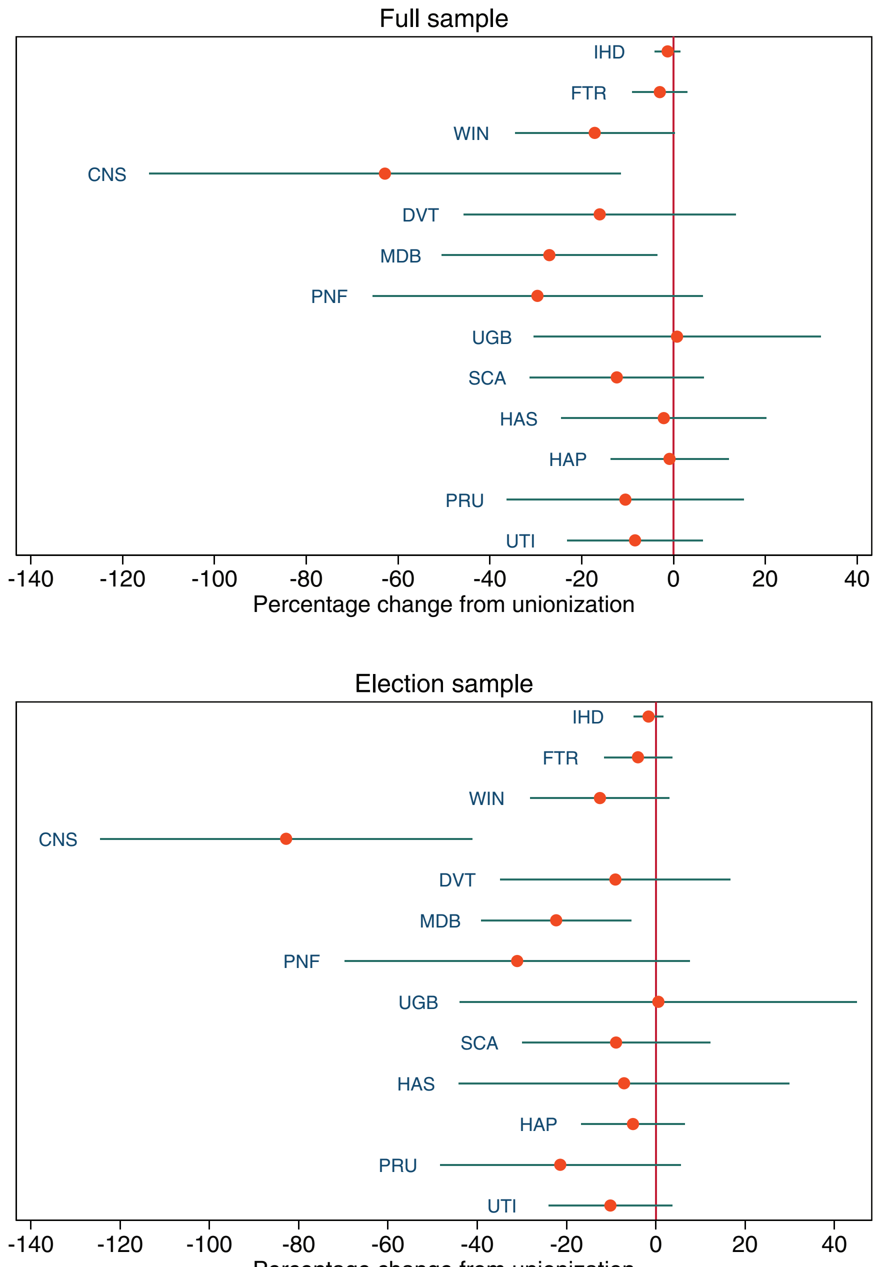
Table 2. Impact of Unionization on Nurse-Sensitive Patient Outcomes

	Full sample		*Election sample*	
Variable	*(1)*	*(2)*	*(3)*	*(4)*
UTI	0.339	−0.712	−1.123*	−0.870
	(0.722)	(0.766)	(0.628)	(0.711)
PRU	0.079	−0.134	−0.190	−0.279
	(0.140)	(0.201)	(0.162)	(0.208)
HAP	0.393	−0.078	−1.564**	−0.454
	(0.698)	(0.705)	(0.710)	(0.632)
HAS	0.050	−0.012	−0.012	−0.042
	(0.051)	(0.076)	(0.091)	(0.126)
SCA	−0.063	−0.418	−0.307	−0.276
	(0.267)	(0.390)	(0.298)	(0.428)
UGB	0.022	0.011	−0.118	0.002
	(0.195)	(0.266)	(0.256)	(0.372)
PNF	0.058	−0.735	−1.132**	−0.757
	(0.439)	(0.543)	(0.484)	(0.585)
MDB	0.012	−0.154*	−0.114*	−0.128**
	(0.037)	(0.081)	(0.062)	(0.059)
DVT	−0.011	−0.180	−0.213	−0.104
	(0.099)	(0.202)	(0.153)	(0.174)
CNS	−0.064	−0.604**	−0.702**	−0.781***
	(0.214)	(0.300)	(0.272)	(0.242)
WIN	−0.086	−0.204	−0.272**	−0.146
	(0.078)	(0.126)	(0.104)	(0.114)
FTR	−0.573***	−0.241	−0.370	−0.301
	(0.213)	(0.298)	(0.357)	(0.379)
IHD	−1.196***	−0.395	−1.231**	−0.478
	(0.410)	(0.530)	(0.588)	(0.640)
All	0.004	−0.056**	−0.091***	−0.060**
	(0.025)	(0.025)	(0.023)	(0.024)
N	4,987	4,987	465	465
Hospital-specific time trends	No	Yes	No	Yes

Notes: Dependent variables are measured as the incidence rate per 1,000 patients for the specified condition, except for the combined measure, All, which is measured in standard deviation units. Full sample results are estimated using all hospitals, while election sample results are estimated using only hospitals that had either a winning or a losing unionization attempt. All specifications contain hospital and year fixed effects; in addition, all specifications control for the proportion of patients who are female, white, black, Hispanic, and Asian; 8 age categories; 25 major diagnostic categories; the proportion of patients who suffered from the specified condition when admitted to the hospital (where applicable);and interactions among the major diagnostic categories, age, and gender. All models are weighted by the total number of patients. Robust standard errors, clustered by hospital, are in parentheses.
$^{*}p < 0.10$; $^{**}p < 0.5$; $^{***}p < 0.01$.

Table 2, column (1), shows results from the full sample that do not control for hospital-specific time trends. These results indicate that hospitals with successful union elections experienced a reduction in 6 of the 13 measures of poor health outcomes but an increase for the other 7 measures. Only two of the measures are statistically significant at conventional levels: in-hospital death and failure to rescue. These two measures both significantly decreased in hospitals following a successful union election. When

Figure 1. Percentage Changes in Incidence of Nurse-Sensitive Outcomes from Unionization



Notes: Figures show the point estimates of unionization effects from Table 2 (columns (2) and (4)) converted to percentages using the full sample means of each outcome; the bars display the corresponding 90% confidence intervals. All specifications contain hospital and year fixed effects, hospital-specific time trends, and demographic controls and are weighted by the total number of patients.

we pool across the outcomes using standardized measures, the aggregate index of disease prevalence shows a very small increase.

We may worry, however, that the assumption of parallel trends in the first specification does not hold across hospitals with and without union elections. That is, hospitals that held an election may have been on a different patient outcome trajectory than those that did not. Indeed, unobserved hospital characteristics associated with poor patient outcomes may have been one of the reasons RNs sought to form a union in the first place—such as ineffectual management or a poor working environment for nurses. The potential for differential trajectories can be partially accounted for by adding a hospital-specific time trend to our base specification, as in Equation (3). The inclusion of a hospital-specific time trend allows "treated" hospitals (those with a successful unionization election) to follow trends that are different from the control hospitals, which had no election, although these differential trends are constrained to be linear.

Table 2, column (2), shows the results using the full sample with hospital-specific time trends. Consistent with a bias stemming from differential trends, we observe qualitatively different results. Although we have only one reduction in disease prevalence that is significant at the 5% level (central nervous system disorders) and one significant at the 10% level (metabolic derangement), we now find that estimates for 12 of the 13 measures are negative. If the sign of the impact on these different outcome measures were independently distributed with a 50% probability, this would happen by random chance with a probability of less than 0.2%. Notably, when we pool across the standardized outcomes, our aggregate index of disease prevalence shows a drop of 5.6% of a standard deviation and is significant at the 5% level of significance.

Turning to the sample of hospitals with union-recognition elections, our specification without trends shows a decline in disease prevalence of all our outcome measures after successful unionization elections compared to hospitals with failed unionization attempts. Moreover, five of these measures show significant declines at the 5% level. If the sign on the impact of these measures were independently distributed, this would happen by random chance with less than a 0.03% probability. In addition, 7 of the 13 measures show significant declines at the 10% level of significance, which would happen by random chance with a probability of less than 0.01%. When we pool across measures, our aggregate index of disease prevalence falls by 9.1% of a standard deviation and is significantly different from 0 at the 1% level of significance.

Adding hospital-specific time trends to this specification produces qualitatively similar findings, although the magnitudes of the estimates are typically smaller. The overall standardized aggregate shows a 6.1% of a standard deviation decline following unionization, and this impact is statistically significant at the 5% level. Of the 13 outcomes, 12 show a reduction in disease prevalence; the only measure with a positive coefficient, upper gastrointestinal bleeding, has an estimated effect that is nearly identical to 0 (0.008) and with standard errors more than 20 times the coefficient size. Similar to the full sample with hospital-specific trends, only two measures are statistically significant: central nervous system disorders and metabolic derangement.

The fact that adding hospital-specific time trends substantively affect our results in the full sample but not in the election sample warrants a discussion. As shown in Table 1, hospitals that held any unionization election—successful or not—tended to be the hospitals with worse patient outcomes, suggesting that significant negative selection into the election group existed and raising the possibility that hospitals that held an election were on a different trajectory than those that did not. If this was the case, then including hospital-specific trends would substantively affect the findings, which is indeed what we observe in Table 2, columns (1) and (2). By contrast, for the models using only the election sample, any differential trends would most likely be attenuated or eliminated because the winning hospitals and losing hospitals share more in common than the winning hospitals share with those that never held an election. In such a case, the inclusion of hospital-specific trends would not strongly affect the results, which is what we observe in Table 2, columns (3) and (4).

As noted, we find particularly large and precise effects for metabolic derangement and central nervous system disorders. Although the percentage reduction in pulmonary failure was also quite large in magnitude, it was less precise and in many cases statistically insignificant, so we focus our discussion here on metabolic derangement and central nervous system disorders.

Metabolic derangement includes ICD9 codes 250.10, 250.11, and 998.0. Included in these diagnostic categories are sugar shock from diabetes and post-operative metabolic shock. Using our baseline estimates with trends, the number of patients with metabolic derangement drops by 13 per 1,000 patients in the election sample and by 15 per 1,000 patients in the full sample. These represent drops of 17% and 21% of the mean number of patients with metabolic derangement in the sample of hospitals with union victories.

Central nervous system disorders include ICD9 codes 780.0, 293.0, 298.2, and 309.1 to 309.9. This category includes delusion, disorientation, and depression. The drop in incidence for central nervous system disorders are even more substantial than those for metabolic derangement, exceeding 50% in both the full sample and the election sample.

In most specifications, we do not find effects on the most serious conditions, such as in-hospital death and failure to rescue. We note, however, that the statistical power to detect an effect for these measures is more limited because they are less prevalent (see Table 1). Furthermore, although we use a prespecified set of outcomes to minimize multiple-testing bias, we do not consider all of them to be equal in terms of sensitivity to nursing quality. Therefore, it is reassuring that we see our strongest effects on the measures in which we think the role of nurses is more critical.

Overall, these findings suggest that hospitals with successful union elections in California during the 1990s and early 2000s had been experiencing declines in patient health outcomes relative to the average hospital prior to the election. But following the election, hospitals with union victories performed better relative to those in which the union lost, and relative to the full sample of control hospitals after we accounted for hospital-specific time trends.

Dynamic Evidence

We also present dynamic evidence on the timing of improvements in health care outcomes following unionization. For simplicity, we restrict our analysis to the aggregate outcome measure. Thus, the coefficient should be interpreted as the effect of unionization on the average number of standard deviations in disease prevalence, pooled across all 13 outcomes. Recall that a negative estimate indicates an improvement in overall patient health. As shown in Equation (4), we regress this patient outcome measure on two leads and four lags of the union status indicator variable. The coefficients on these indicators therefore estimate the effect of unionization from two years prior to the election to four years after the election relative to the omitted category of three or more years before the election (note that the fourth lag captures the effect four *or more* years after the election). We again show four specifications: full and election samples, with and without hospital-specific time trends.

The numerical coefficients on leads and lags of the unionization dummy are reported in Table 3, and in Figure 2 we plot the running sum of the coefficients beginning with the two-year lead, which represents the cumulative change in the mean outcome level compared to the baseline period of three or more years prior to the election. Figure 2 shows two sets of 95 percent CIs associated with two different baselines. The first, lighter-shaded CI is for the response at year t in event time relative to the baseline of year -3 or earlier. The change between the baseline and year t is statistically indistinguishable from 0 if the lighter-shaded 95% CI for year t does not contain 0. The specification with the full sample without hospital-specific time trends (Table 3, column (1)) shows changes between the baseline and years -2 and -1 that are positive and statistically significant. This confirms that unionization tended to occur more often in hospitals undergoing a decline in patient health quality. In contrast, the other three specifications do not show statistically significant or medically sizable changes between the baseline and years -2 and -1. Better comparison groups or parametric trend controls account for these pre-existing trends, and these three specifications show stable relative outcomes prior to the election. The same conclusion can be reached by looking at the individual leading coefficients in Table 3. Overall, the leading-effects falsification test suggests that the specifications in Table 3, columns (2) to (4), are preferred based on how well they match pre-existing trends in the treatment and control groups.

Table 3. Dynamic Effects of Unionization on Mean Standardized Outcome

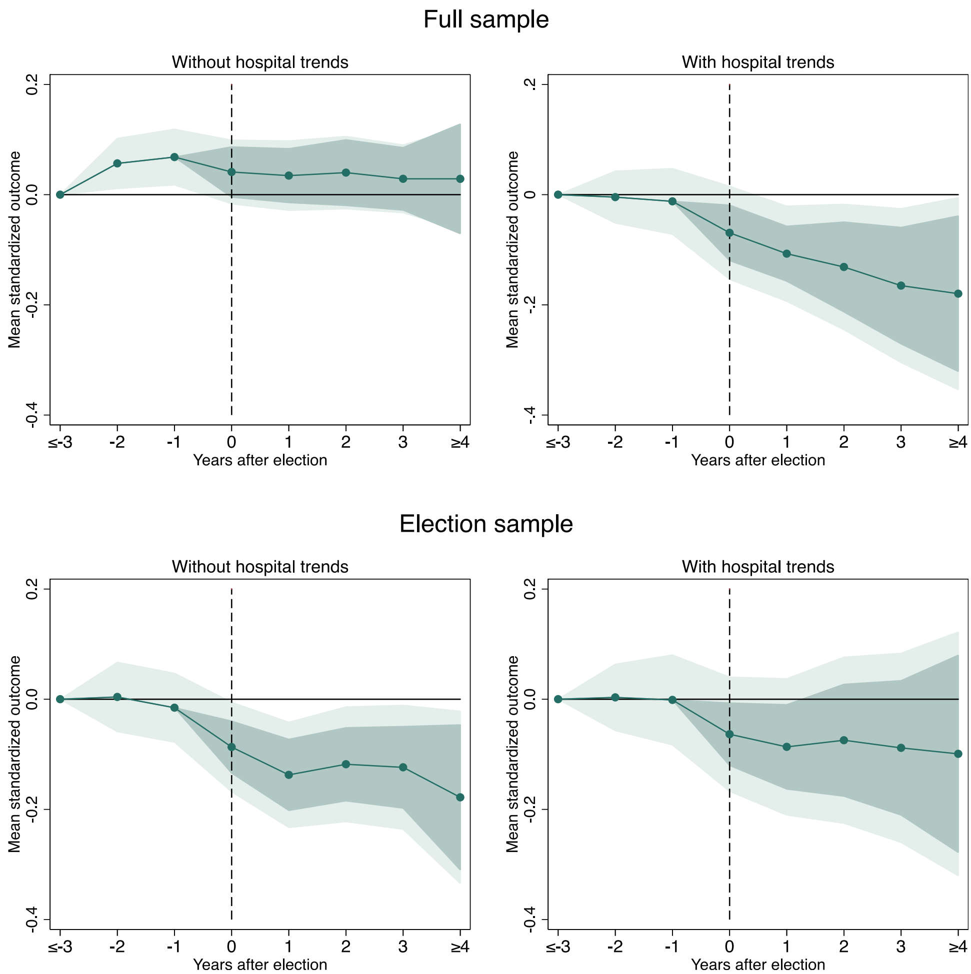
	Full sample		*Election sample*	
Variable	*(1)*	*(2)*	*(3)*	*(4)*
2 years pre-election	0.057**	−0.004	0.004	0.003
	(0.023)	(0.024)	(0.032)	(0.031)
1 year pre-election	0.012	−0.008	−0.020	−0.005
	(0.015)	(0.018)	(0.018)	(0.026)
Year of election	−0.027	−0.057**	−0.071***	−0.062**
	(0.023)	(0.026)	(0.023)	(0.029)
1 year post-election	−0.007	−0.038**	−0.050**	−0.023
	(0.016)	(0.016)	(0.024)	(0.027)
2 years post-election	0.005	−0.024	0.019	0.011
	(0.023)	(0.029)	(0.026)	(0.033)
3 years post-election	−0.011	−0.034	−0.005	−0.014
	(0.013)	(0.022)	(0.022)	(0.028)
$\geq$ 4 years post-election	−0.000	−0.015	−0.055	−0.011
	(0.041)	(0.038)	(0.041)	(0.042)
Short-term effect: 1 year	−0.033	−0.094***	−0.121***	−0.085**
	(0.024)	(0.025)	(0.032)	(0.038)
Long-term effect: $\geq$4 years	−0.039	−0.167**	−0.163**	−0.098
	(0.05)	(0.071)	(0.066)	(0.091)
N	4,564	4,564	452	452
Hospital-specific time trends	No	Yes	No	Yes

Notes: Dependent variable is the mean standardized incidence rate across all conditions. Independent variables are leads and lags of union status, as indicated. The reported short-term effect is the sum of the contemporaneous union status variable and the one-year lag; the reported long-term effect is the sum of the contemporaneous union variable and the full set of lags. Full sample results are estimated using all hospitals, while election sample results are estimated using only hospitals that had either a winning or a losing unionization attempt. All specifications contain hospital and year fixed effects; in addition, all specifications control for the proportion of patients who are female, white, black, Hispanic, and Asian; 8 age categories; 25 major diagnostic categories; the proportion of patients who suffered from the specified condition when admitted to the hospital (where applicable); and interactions among major diagnostic categories, age, and gender. All models are weighted by the total number of patients. Robust standard errors, clustered by hospital, are in parentheses.
*$p < 0.10$; **$p < 0.5$; *** $p < 0.01$.

Figure 2 provides visual evidence that in all three of these preferred specifications (in Table 3, columns (2) to (4)) the aggregate disease prevalence dropped in hospitals with union wins during the year of the election. Table 3 documents that the contemporaneous coefficient is statistically significant in all three specifications; it is also the largest coefficient in magnitude in these three cases. In contrast, the specification using the full sample without trends (column (1)) does not indicate a substantial change following the election.

To statistically test for short- and long-term changes following unionization, we also show in Figure 2 a second, darker-shaded CI for the effect since year −1, the year just prior to the election. The change from year −1 to year *t* is statistically significant if the darker 95% CI in year *t* excludes the point estimate associated with year −1. We can see from the figure that for

Figure 2. Dynamic Response of Mean Standardized Outcome from Unionization



Notes: Dependent variable is the mean standardized incidence rate across all conditions. Figures show the cumulative sum of coefficients (from Table 3) beginning with the two-year lead. All specifications contain hospital and year fixed effects and demographic controls, as well as hospital-specific time trends as indicated, and they are weighted by the total number of patients. The lighter-shaded confidence interval is for year *t* relative to a baseline of three years before unionization, and the darker-shaded confidence interval is for year *t* relative to a baseline immediately prior to the election.

the models associated with columns 2, 3, and 4 in Table 3, the effects are statistically significant at the 5% level during years 0 and 1. For the models associated with columns (2) and (3), the effects remain significant throughout the post-election period, but for the model in column (4) (the election sample with trends), the later lags lose precision.

Table 3 provides the numerical counterpart to the visual evidence: the short-run impact through the year after the election is quite substantial, ranging between −0.085 and −0.121 across the three specifications; the estimates are statistically significant at the 5% level. We do note that we see statistically significant leading effects in hospitals with successful union elections relative to the others in the full sample two years before the election. But when we control for hospital-specific time trends, we eliminate this leading effect statistically and substantively. No such leading effect exists in the hospital election sample with or without trends. Estimates for longer-run impacts (through the fourth year following the election or later) range between −0.097 and −0.167 and are, unsurprisingly, less precise; however, they continue to be statistically significant at the 5% level in two of the three preferred specifications. In all three preferred specifications, we see no indication that the gains in patient health were temporary: the longer-term estimates appear to be larger than the estimates from the contemporary specification.

Overall, the evidence strongly points to a clear and immediate improvement in the average patient health outcome following a successful union election, which appears to grow somewhat over time. Moreover, the dynamic evidence also shows that the one specification without a measured union effect is also the only one that fails the falsification test for pre-existing trends.

Without the explicit inclusion of lags in unionization, a delayed effect of union presence can be mistaken for a hospital trend, thereby attenuating the estimate of the union's impact. In our dynamic specifications, the inclusion of the lags and leads in unionization implies that the hospital-specific time trends are largely identified using data from three or more years before or from five or more years after the union election. The estimated hospital trends in such a model are, therefore, more likely to reflect preexisting trends unaffected by the treatment itself. This suggests that we can compare the dynamic and contemporaneous specifications to assess whether the trends are partly absorbing the dynamic treatment effect. Because we have, on average, 3.2 post-election years in our sample (excluding the year of the election), we should expect the estimated effect roughly through year 3 in the dynamic specifications to be comparable to the estimates from the contemporaneous specifications. The three-year-out effects can be calculated from Table 3, and for the preferred models in columns (2), (3) and (4), they are −0.152, −0.107, and −0.086, respectively. These can be compared to the estimates from Table 2, which are −0.056, −0.091, and −0.061. We find that the estimates from models with hospital-specific time trends (columns (2) and (4)) are much larger in magnitude when lags are included. The model that does not include time trends (column (3)) is the least affected by the inclusion of lags. These results suggest that, if anything, the estimates from the contemporaneous specifications may be somewhat understated because of the presence of lagged effects. (We show additional evidence on this question in the Robustness Checks section.)

Patient Characteristics and Case Load

One possible concern about our results is that they reflect changes in the case mix or case load that occurred at the time of unionization. We test for

Table 4. Impact of Unionization on Patient Demographics

	Full sample		Election sample	
Variable	*(1)*	*(2)*	*(3)*	*(4)*
log(Patient total)	−0.021	−0.019	−0.023	−0.017
	(0.030)	(0.028)	(0.035)	(0.028)
Female	0.002	0.005	−0.003	0.003
	(0.004)	(0.007)	(0.005)	(0.007)
Nonwhite	−0.008	0.009	0.002	0.004
	(0.015)	(0.015)	(0.017)	(0.017)
Under 18	−0.015	−0.007	−0.022	−0.007
	(0.019)	(0.007)	(0.017)	(0.007)
Over 65	0.004	−0.001	0.019	0.001
	(0.015)	(0.008)	(0.014)	(0.008)
N	4,987	4,987	465	465
Hospital-specific time trends	No	Yes	No	Yes

Notes: Dependent variables are patient demographics as indicated. Female, Nonwhite, Under 18, and Over 65 are measured as proportions. Full sample results are estimated using all hospitals; election sample results are estimated using only hospitals that had either a winning or a losing unionization attempt. All specifications contain hospital and year fixed effects, and all specifications are weighted by the total number of patients. Robust standard errors, clustered by hospital, are in parentheses.
*$p < 0.10$; **$p < 0.5$; ***$p < 0.01$.

this by regressing demographic measures and number of patients on hospital fixed effects, time fixed effects, other demographic variables, and a union dummy, as we did in Table 2. As before, we present results for our four main specifications: the full sample and the election sample, with and without hospital-specific time trends. The particular measures that we use are the ones we initially collected: the log of the number of patients, the percentage of patients who are female, the percentage of patients who are nonwhite, the percentage of patients under 18, and the percentage of patients over 65.

The results are shown in Table 4. None of the 20 coefficients are different from 0 at even a 10% level of statistical significance. Two of the measures (the logarithm of the total number of patients and percentage of patients under 18) are negative in all four specifications. This leaves open the possibility that the improvement in patient outcomes was achieved, at least in part, through a lower and easier case mix. Nevertheless, the coefficients are rather small for the patients younger than 18 years old. The coefficients on logarithm of the total number of patients, although never significant, do show a roughly 2% decline in number of patients.

Robustness Checks

Tables 5 and 6 show a number of robustness checks of our main results. Table 5 looks at robustness to the window over which the results are estimated. Because hospitals that experienced union elections were hospitals with worsening patient outcomes, we may be concerned that the

Table 5. Restricted Time Windows around Unionizations

	Full sample			*Election sample*		
	2 years	*4 years*	*6 years*	*2 years*	*4 years*	*6 years*
Panel A. Restricted pre-election period						
All	−0.032	−0.050*	−0.059**	−0.063**	−0.049*	−0.056**
	(0.031)	(0.029)	(0.027)	(0.026)	(0.026)	(0.025)
N	4,867	4,918	4,953	345	396	431
Panel B. Restricted post-election period						
All	−0.048*	−0.035	−0.051**	−0.049	−0.053	−0.064**
	(0.027)	(0.026)	(0.025)	(0.04)	(0.033)	(0.028)
N	4,864	4,908	4,942	342	386	420

Notes: Dependent variable is the mean standardized incidence rate across all conditions. In panel A, the sample of unionizing hospitals is restricted to the specified number of years before the winning election took place, while in panel B it is restricted to the specified number of years after the election took place. Full sample results are estimated using all hospitals; election sample results are estimated using only hospitals that had either a winning or a losing unionization attempt. All specifications contain hospital fixed effects, year fixed effects, and hospital-specific time trends. In addition, All specifications control for the proportion of patients who are female, white, black, Hispanic, and Asian; 8 age categories; 25 major diagnostic categories; the proportion of patients who suffered from the specified condition when admitted to the hospital (where applicable); and interactions among major diagnostic categories, age, and gender. All models are weighted by the total number of patients. Robust standard errors, clustered by hospital, are in parentheses.

*$p < 0.10$; **$p < 0.5$; *** $p < 0.01$.

relationship between the control variables and outcomes was changing over time—possibly biasing our estimates. To address this issue, we estimate our effects over shorter windows, in which the coefficients for the controls (including fixed effects) were less likely to be changing. Again, we focus on the overall measure of potentially nurse-sensitive outcome. We show the results in Table 5 for our preferred specification with hospital-specific time trends.

In Table 5, panel A, we include only two, four, or six years of data prior to the election in hospitals with successful unionization, but we continue to use the full panel of data from our control hospitals. These restrictions allow us to check the robustness of our findings to alternative pre-unionization baselines in the levels and trends in the outcomes.

We find that the results are largely similar when we omit the years before the election event. The election sample is more robust when we exclude pre-election periods, probably because the pre-election periods are more similar between the hospitals with failed and successful unionization attempts than they are between either of these groups and the hospitals without elections. By far, the estimate that is lowest in magnitude is the two-year window in the full sample, in which the coefficient drops to -0.032 and is not significant at conventional levels. All the other estimates, across the full sample and the election sample, are -0.049 or less and significant at least at the 10% level.

In Table 5, panel B, we focus on the length of the post-intervention period. Here we include only two, four, or six years of data from the post-election period in hospitals with successful unionization. As before, we continue to use the full panel of data from our control hospitals. We find that results for the aggregate outcome measure are broadly robust to the length of the post-unionization period. In the election sample, the coefficient estimate drops to slightly less than -0.05 in magnitude and is not significant at the 10% level for the two-year window. For the full sample, the drops are somewhat larger in magnitude and no monotonic relation exists between the number of years included and the size of the coefficient. The lowest estimated effect is -0.035 and is not significant at the 10% level; however, the estimates from the two- and six-year samples are both statistically significant at the 10% or smaller level. Overall, these results are consistent with our dynamic evidence in Table 3, which shows a clear fall in disease prevalence at the time of unionization, followed by some additional reduction subsequently.

In Table 6, we show that our results from the full sample and the election sample are robust to how we control for trends, how we control for the timing of elections, and whether we control for the outcome prevalence on admissions. Our first specification (columns (1) and (5)) estimates the impact of unionization, controlling for the effect of any election, successful or otherwise. In particular, we introduce a dummy that takes on the value 1 after a union election and 0 otherwise. Therefore, in this specification, our

Table 6. Robustness Checks

	Full sample				Election sample			
Variable	(1) With post-election dummy	(2) Quadratic trends	(3) Pre/post-election trends	(4) No POA controls	(5) With post-election dummy	(6) Quadratic trends	(7) Pre/post-election trends	(8) No POA controls
UTI	−0.749	−0.712	−0.551	−1.470*	−0.964	−0.870	−1.230	−1.730**
	(0.964)	(0.766)	(0.928)	(0.852)	(1.148)	(0.711)	(1.274)	(0.815)
PRU	−0.148	−0.134	−0.183	−0.222	0.006	−0.279	−0.015	−0.381*
	(0.238)	(0.201)	(0.262)	(0.210)	(0.289)	(0.208)	(0.282)	(0.217)
HAP	−2.036	−0.076	−1.747	0.156	−3.287***	−0.455	−3.085***	−0.572
	(1.463)	(0.706)	(1.532)	(0.819)	(1.031)	(0.633)	(1.144)	(0.962)
HAS	0.046	−0.012	0.034	−0.005	−0.052	−0.042	−0.049	−0.093
	(0.102)	(0.076)	(0.109)	(0.131)	(0.168)	(0.126)	(0.169)	(0.170)
SCA	−1.886***	−0.418	−1.674**	−0.343	−1.578**	−0.276	−1.414**	−0.145
	(0.728)	(0.391)	(0.739)	(0.416)	(0.637)	(0.428)	(0.638)	(0.504)
UGB	−0.044	0.011	0.092	−0.073	0.144	0.002	0.213	−0.045
	(0.290)	(0.266)	(0.376)	(0.267)	(0.382)	(0.372)	(0.428)	(0.401)
PNF	−1.036	−0.735	−1.103	−0.782	−1.219	−0.758	−1.434	−1.175
	(0.832)	(0.544)	(1.003)	(0.678)	(0.732)	(0.586)	(0.905)	(0.749)
MDB	−0.209	−0.154*	−0.175	−0.205**	−0.088	−0.128**	−0.038	−0.233**
	(0.156)	(0.081)	(0.169)	(0.098)	(0.166)	(0.059)	(0.190)	(0.104)
DVT	−0.466*	−0.179	−0.300	−0.350	−0.344	−0.104	−0.304	−0.359*
	(0.261)	(0.202)	(0.198)	(0.223)	(0.255)	(0.174)	(0.285)	(0.185)
CNS	−0.619	−0.602**	−0.196	−0.789**	−0.876*	−0.781***	−0.732	−1.046***
	(0.536)	(0.300)	(0.429)	(0.399)	(0.519)	(0.241)	(0.538)	(0.356)
WIN	−0.085	−0.204	−0.041	−0.218	0.001	−0.146	0.079	−0.216
	(0.213)	(0.126)	(0.202)	(0.168)	(0.205)	(0.114)	(0.218)	(0.176)
FTR	−0.589	−0.241	−0.490	−0.241	−0.439	−0.301	−0.408	−0.301
	(0.567)	(0.298)	(0.569)	(0.298)	(0.732)	(0.379)	(0.727)	(0.379)
IHD	−0.482	−0.395	−0.442	−0.395	−0.688	−0.478	−0.556	−0.478
	(1.357)	(0.530)	(1.357)	(0.530)	(1.354)	(0.640)	(1.434)	(0.640)
All	−0.091**	−0.055**	−0.069	−0.084***	−0.082**	−0.061**	−0.074*	−0.106***
	(0.045)	(0.025)	(0.044)	(0.031)	(0.037)	(0.024)	(0.043)	(0.035)
N	4,987	4,987	4,987	4,987	465	465	465	465

Notes: Dependent variables are measured as the incidence rate per 1,000 patients for the specified condition, except for the combined measure, All, which is measured in standard deviation units. All specifications contain hospital and year fixed effects as well as hospital-specific time trends. In addition, specifications 1 and 5 contain a dummy indicating the period after an election has taken place, specifications 2 and 6 contain a quadratic hospital-specific time trend, and specifications 3 and 7 contain time trends specific to the pre-election and post-election periods. Full sample results are estimated using all hospitals; election sample results are estimated using only hospitals that had either a winning or a losing unionization attempt. In addition, all specifications control for the proportion of patients who are female, white, black, Hispanic, and Asian; 8 age categories; 25 major diagnostic categories; and interactions among diagnostic categories, age, and gender. All models except for models 4 and 8 control for the proportion of patients whose specified condition was present on admission (POA) to the hospital. All models are weighted by the total number of patients. Robust standard errors, clustered by hospital, are in parentheses.

*$p < 0.10$; **$p < 0.5$; ***$p < 0.01$.

estimates are relative to election losses after the elections occur. The estimates of the effect of a union win are larger in magnitude in both the full sample and election sample. Although the standard errors in the election sample rise because of collinearity between the post-election dummy and the union-win dummy, the number of measures significant at the 10% level or less rises from two to four relative to the baseline estimates. Moreover, the coefficient of the impact on the overall patient outcome increases in magnitude by around one-third. The changes to the full sample are starker. Even though the standard errors are uniformly larger because of collinearity, the coefficients for two measures in particular strongly increase in magnitude and become highly significant: deep vein thrombosis, and shock or cardiac arrest. Although the two measures with significant coefficients in the baseline trends model in the full sample lose significance, in both cases, the coefficients rise in magnitude. The estimate for the overall patient outcome increases in magnitude by around two-thirds and is much more similar to the corresponding estimate from the election sample. This is sensible because we are now largely comparing election wins to election losses (by controlling for elections) even in the full sample.

In Table 6, columns (2) and (6), we control for quadratic trends to determine whether our results are robust to other ways of controlling for hospital-specific time trends. Our results are identical in both the full sample and the election sample to two decimal places and, in most cases, to three decimal places. That the variations in the parametric form of hospital-specific trends change the results to only a minimal extent is reassuring.

As we discussed in the context of dynamic specifications, with lagged treatment effects, estimates from the models with hospital-specific time trends may be affected by unionization itself. To ensure that hospital-specific time trends represent pre-existing trends rather than trends after unionization, Table 6, columns (3) and (7), include controls for time trends before and after unionization in addition to hospital-specific time trends. Allowing additional pre- and post-unionization trends increases the magnitude of our estimates from −0.056 to −0.059 in the full sample and −0.061 to −0.075 in the election sample. In both cases, the standard errors increase substantially because of greater collinearity of the treatment effect with the controls. The election-sample estimate remains significant at the 10% level, but the full-sample estimate drops slightly below the 10% significance level. Moreover, the reported coefficient on the union dummy does not account for the effect on the trend itself; it accounts only for the immediate impact in the election year. When we evaluate the average effect including the impact on trends using the mean length of the post-election period (3.2), we find that the estimate for the aggregate outcome is −0.106 in the full sample and −0.093 in the election sample (for more on interpreting coefficients with linear trend breaks, see McCrary 2007). In both samples, the average cumulative effects over the mean length of the post-election period are statistically significant at 5% levels. These estimates are similar to our evidence from the dynamic specifications. They indicate that the presence of lagged effects may, if anything, lead to a smaller estimate in specifications with hospital-specific time trends.

Finally, in Table 6, columns (4) and (8), we assess the concern that the apparent reduction in disease incidence may be because of better screening of conditions during admission rather than to actual changes in outcomes during the hospital stay. For this reason, we remove from our control set the prevalence of these outcomes among patients during admission. Contrary to that hypothesis, we find that the magnitude of the coefficient rises in almost all the specifications, in both the full sample and the election sample. The coefficient for urinary tract infections becomes statistically significant at the 10% level or less in both samples. Moreover, the coefficients for pressure ulcers and deep vein thrombosis become statistically significant at the 10% or lower in the election sample. The coefficient for the overall measure of patient outcome rises in magnitude by more than 50% in both samples to −0.084 in the full sample and −0.108 in the election sample.

Discussion

Several explanations are possible for our findings of improved quality of outcomes following unionization. First, possibly our results do not reflect the causal effect of unions but, rather, endogeneity in the timing of unionization. In particular, a temporary drop in health care quality may have induced a union-organizing drive; subsequently, conditions may have improved, leading to a natural recovery in quality without unionization itself playing any role. Consistent with this endogeneity explanation, we do, in fact, see that unionization occurred in hospitals experiencing worsening patient outcomes. This is evident from the model in Table 3, column (1), in which the estimates use all hospitals and do not control for hospital-specific time trends. Contrary to the endogeneity explanation, however, we do not see a relative quality decline in our preferred specifications using the election sample (columns (3) and (4)) or in the full sample with hospital-specific time trends (column (2)). Moreover, we see improvement only after an election in the hospitals where the union won the election. Accounting for an "election effect" by adding a separate election dummy does not attenuate the findings; instead, the estimates from that specification are actually larger in magnitude. The findings from the election sample suggest that patient health outcomes would probably not have improved in hospitals with successful union elections in the absence of union formation.

A second possibility is that unionization led to a shift in patient selection. If the patients or doctors of patients knew about the unionization and thought it could negatively affect care, sicker patients could have been directed to other hospitals. Alternatively, unionized nurses could have transferred sicker patients to other hospitals at triage. This, in turn, could have led to healthier patients being admitted to the hospital, who were less likely to acquire diseases in the hospital. Although we do find a statistically insignificant but consistent reduction in number of patients per case load after unionization, we do not find any evidence of a shift in patient demographics, including age. Thus, we do not think that selection of patients is likely to explain our results on patient outcomes.

A third possible mechanism for a reduction in the patient-outcome measures is a change in reporting. We estimate the impact of unionization on outcomes that were not present on admission. Possibly, unionization leads to a change in the probability of detecting these conditions during admission. One variant of this argument is that hospitals start screening more diligently when admitting patients. In that case, an increase in the reported (but not actual) presence of these conditions on admission would be apparent. To the contrary, as we show in Table 6, our qualitative results hold with or without controls for the prevalence of these conditions being present on admission. Therefore, we do not think that our results merely reflect changes in reporting standards after unionization.

A fourth possible explanation is that, after unionization, conditions increasingly go unreported, thus leading to an apparent downturn in disease prevalence for conditions not present on admission. Because this reporting is largely done through doctors' diagnoses, however, this would be unlikely to happen without collusion between doctors and nurses; we think this is unlikely to have happened on any noticeable scale.

Because our findings appear to show the causal effects of unionization (as opposed to statistical artifacts), considering the various channels through which they may occur is useful. We note at the outset that we are not able to test for the importance of these specific potential channels, although doing so would be valuable in future work with suitable data. A useful way to divide potential mechanisms is into those that improve patient health outcomes through changes in the characteristics or behavior of nurses and those that do so through changes in the behavior of management, although some potential mechanisms could reasonably be placed in both categories. The former includes changes in turnover, morale, or effort in response to improved wages or working conditions, and the latter includes a reorganization of management practices, increased collaboration between management and labor, greater capital intensity, a reduction in shift lengths, and reduced staffing ratios.

To the extent that unions are able to raise nurses' wages, this may improve quality of care through increasing effort, reducing turnover, and possibly increasing morale. The nursing labor market has often been characterized as being monopsonistic, which highlights the possible role of easing recruitment and retention of nurses through better compensation. We note, however, that recent evidence from Matsudaira (2014) raised some questions about the simple monopsony explanation. Changes in worker behavior in response to increased wages is one of the main mechanisms through which unionization has been hypothesized to improve production quality in previous research (e.g., Mas 2006). Numerous cross-sectional studies of the nursing sector have found a positive union wage premium (Hirsch and Schumacher 1998; Spetz, Ash, Konstantinidis, and Herrera 2011). In manufacturing, however, some studies noted that marginal unionization seems not to have had an impact on average wages (e.g., DiNardo and Lee 2004); by contrast, and relevant to our findings here, other recent work has suggested that a victory by unions with strong support may in fact have had sizable impacts (Lee and Mas 2012).

Assuming that the unionization of nurses does indeed raise compensation, evidence is available that such increases frequently translate into improved nurse retention and patient outcomes. For instance, Schumacher (1997) and Seago et al. (2011) both documented higher nurse retention in higher-paying and unionized hospitals, and work by Propper and Van Reenen (2010) found that low wages for nurses relative to the local labor markets significantly increased patient mortality in British hospitals.

Improvements in nurse retention, morale, and, ultimately, effectiveness could also come from nonwage improvements in the nurses' work environments. As previously noted, health policy researchers have documented improvements in patient outcomes associated with aspects of the work environment, such as nurses' involvement in hospital governance, ongoing professional development opportunities, collegial nurse–physician relations, and accepted procedures for resolving nurse–supervisor disputes (Aiken et al. 2002; Kuokkanen et al. 2003; Friese et al. 2008). Because unionization is often associated with increased worker voice in determining working conditions (Freeman and Medoff 1984), improvements in the work environment are another clear mechanism through which unionization could impact patient outcomes.

Unionization may also lead to changes in the behavior of management that, in turn, affect patient outcomes. Studies have shown that management often reorganizes following the formation of a union to increase efficiency and quality, sometimes referred to as a shock effect (Slichter et al. 1960; Clark 1980b). If these reorganizations occur in the hospital sector and are effective, they could plausibly improve patient outcomes. Management may also respond to unionization by seeking collaboration with their newly organized labor force (Kochan 1986, 1994) and, in doing so, may improve employee efficiency or employee morale even in the absence of negotiated contracts with explicit favorable provisions regarding employee wages or working conditions.

Another potential mechanism driven by management reactions to unionization is that hospitals may move toward more capital-intensive production techniques, which may in turn improve patient outcomes. We note that Sojourner et al. (2015) reported evidence from nursing homes that staffing actually declined (relative to case load) following unionization, at least in the case of union victories. This reduction in staffing in nursing homes appears to be more consistent with increased capital intensity of production.

But, as previously noted, the acute care hospitals we study are much more capital intensive to begin with than are nursing homes and have quite different patterns of staffing, limiting our ability to extrapolate across the two studies.

Finally, a potential mechanism involving responses by both nurses and managers are changes in staffing levels. Nurses' unions bargain to a greater degree than most unions on staffing and work load, and in particular, many nurses' unions try to cap staffing ratios and set limits on the number of hours worked per shift. Lower patient-to-nurse staffing ratios can be seen as a core component of the work environment already discussed and will tend to reduce burnout and help nurses focus on and be more attentive to the patients under their care. Lower staffing ratios have consistently been shown to improve patient outcomes, including in studies that relied on quasi-random sources of variation in staffing ratios (Bell and Redelmeier 2001; Aiken et al. 2002; Evans and Kim 2006). Notably, beginning in 2004 California implemented AB394, which mandated minimum nurse staffing ratios statewide (for a detailed description and an analysis of the law's impact on patient outcomes, see Cook, Gaynor, Stephens, and Taylor 2012). Although the implementation and enforcement of AB394 occurred only toward the end of our sample period, the law was widely anticipated, and hospitals may have begun increasing nurse-to-patient staffing ratios prior to enforcement to ensure compliance. This unique institutional feature may reduce the likelihood that union-negotiated reductions in nurse staffing ratios are the primary mechanism underlying our findings.

Although our data do not allow us to implement tests that would distinguish among all these mechanisms directly, the timing of the changes we observed suggest some as being more likely than others. In particular, our dynamic estimates show the largest effects occur in the year of unionization, with modest additional improvements over the following four years. In many cases, however, an initial contract takes a substantial amount of time to negotiate; Ferguson (2008) reported that 62% of newly formed unions lacked a contract one year after an election victory and that 44% lacked a contract two years after the election. If this pattern holds for the unionization events in our sample, our findings have two possible interpretations. Under the first interpretation, the effects we find should be thought of as the effect of an "intent to treat," in which the actual treatment occurs only when a collective bargaining contract exists. In this case, the effect of the treatment on the treated is likely to be substantially larger than the estimates we report here. The alternative interpretation is that the wage improvements or formal changes in workplace rules that are negotiated through collective bargaining agreements are not the sole drivers of the observed improvements in patient outcomes. The sharp change after the election, followed by modest subsequent improvements, points toward this second interpretation. Likewise, changes in capital intensity are usually difficult to implement over a short time horizon and, therefore, seem unlikely to explain our main findings. In contrast, a reorganization of management structures or a more cooperative approach on the part of management may be implemented immediately following a successful unionization election. Such changes may immediately improve nurses' work environments along the dimensions we have discussed, making these mechanisms more consistent with our dynamic findings.

Conclusion

In this article, we examine the consequences of the unionization of nurses on health care quality. We find that hospitals with a successful unionization attempt experience a decline in the incidence of hospital-acquired illnesses compared to hospitals that experience a failed unionization attempt and compared to hospitals more broadly conditional on hospital-specific time trends. This holds true across a broad range of potentially nurse-sensitive medical outcomes, ranging from less serious illnesses such as urinary tract infections to critical ones such as in-hospital death. Our largest effects are for central nervous system complications, such as delirium and depression, and for metabolic derangement. These estimates show a decline of up to 58% for central nervous system disorders and 17% for metabolic derangement. The in-hospital death estimates, in particular, incorporate the effect of strikes and other disruptions on mortality (Gruber and Kleiner 2010), although they are small and statistically insignificant in our preferred specifications.

We find that our estimates are likely to represent a causal effect of unionization on the quality of care as opposed to a shift to reporting illnesses on admissions after unionization, greater selectivity over sicker patients, or unionization occurring during hospital decline followed by mean regression. But we do not currently have the data to separate out whether the effect of unionization is primarily through unionization itself, a change in pay, a change in staffing, or a change in work rules. Future research could disentangle these mechanisms, using data on union contracts, pay and staffing matched to hospitals, and patient transfers across hospitals. Moreover, with larger samples, regression discontinuity designs could be used to gain better identification. All these would provide substantial value-added for the understanding of medical labor markets and quality of hospital care more generally. They would also provide a better understanding of behavioral responses to labor market changes. Last, they would provide information for formulating labor market policy in the health care industry.

Appendix

Figure A.1. Distribution of Union Elections by Year

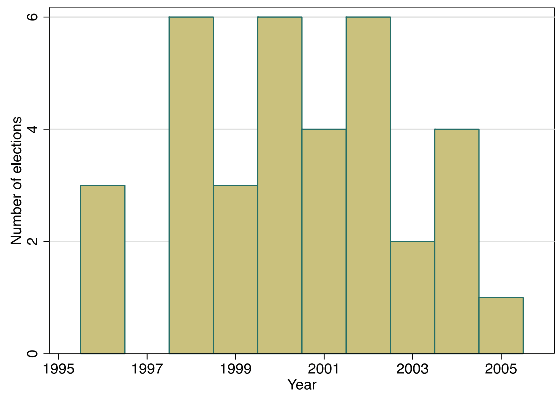
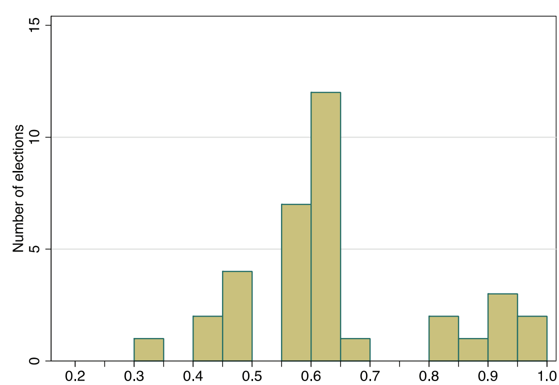


Figure A.2. Distribution of Union Elections by Vote Share



References

Aiken, Linda H., Jeannie P. Cimiotti, Douglas M. Sloane, Herbert L. Smith, Linda Flynn, and Donna F. Neff. 2011. The effects of nurse staffing and nurse education on patient deaths in hospitals with different nurse work environments. *Medical Care* 49(12): 1047–53.

Aiken, Linda H., Sean P. Clarke, Douglas M. Sloane, Julie Sochalski, and Jeffrey H. Silber. 2002. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *Journal of the American Medical Association* 288(16): 1987–93.

Ash, Michael, and Jean Ann Seago. 2004. The effect of registered nurses' unions on heart attack mortality. *ILR Review* 57(3): 422–42.

Bell, Chaim M., and Donald A. Redelmeier. 2001. Mortality among patients admitted to hospitals on weekends as compared with weekdays. *New England Journal of Medicine* 345(9): 663–68.

Blegen, Mary A., Colleen J. Goode, Shin H. Park, Thomas Vaughn, and Joanne Spetz. 2013. Baccalaureate education in nursing and patient outcomes. *Journal of Nursing Administration* 43(2): 89–94.
Card, David. 1996. The effect of unions on the structure of wages: A longitudinal analysis. *Econometrica* 64(4): 957–79.
Clark, Kim B. 1980a. Unionization and productivity: Micro-econometric evidence. *Quarterly Journal of Economics* 95(4): 613–39.
———. 1980b. The impact of unionization on productivity: A case study. *ILR Review* 33(4): 451–69.
Cook, Andrew, Martin Gaynor, Melvin Stephens, and Lowell Taylor. 2012. The effect of a nurse staffing mandate on patient health outcome: Evidence from California's minimum staffing regulation. *Journal of Health Economics* 31(1): 340–48.
DiNardo, John, and David S. Lee. 2004. Economic impacts of new unionization on private sector employers: 1984–2001. *Quarterly Journal of Economics* 119(4): 1383–441.
Evans, William N., and Beomsoo Kim. 2006. Patient outcomes when hospitals experience a surge in admissions. *Journal of Health Economics* 25(2): 365–88.
Ferguson, John-Paul. 2008. The eyes of the needles: A sequential model of union organizing drives, 1999–2004. *ILR Review* 62(1): 3–21.
Freeman, Richard, and James Medoff. 1984. *What Do Unions Do?* New York: Basic Books.
Friese, Christopher R., Eileen T. Lake, Linda H. Aiken, Jeffrey H. Silber, and Julie Sochalski. 2008. Hospital nurse practice environments and outcomes for surgical oncology patients. *Health Services Research* 43(4): 1145–63.
Gruber, Jonathan, and Samuel Kleiner. 2010. Do strikes kill? Evidence from New York State. *American Economic Journal: Economic Policy* 4(1): 127–57.
Hirsch, Barry T., and Edward J. Schumacher. 1998. Union wages, rents, and skills in health care labor markets. *Journal of Labor Research* 19(1): 125–47.
Kochan, Thomas A. 1986. *The Transformation of American Industrial Relations*. Ithaca, NY: Cornell University Press.
———. 1994. *The Mutual Gains Enterprise: Forging a Winning Partnership among Labor, Management, and Government*. Boston: Harvard Business Press.
Kochan, Thomas A., Adrienne E. Eaton, Robert B. McKersie, and Paul S. Adler. 2009. *Healing Together—the Labor-Management Partnership at Kaiser Permanente*. Ithaca, NY: Cornell University Press.
Krueger, Alan, and Alexandre Mas. 2004. Strikes, scabs and tread separations: Labor strife and the production of defective Bridgestone/Firestone tires. *Journal of Political Economy* 112(2): 253–89.
Kuokkanen, Liisa, Helena Leino-Kilpi, and Jouko Katajisto. 2003. Nurse empowerment, job-related satisfaction, and organizational commitment. *Journal of Nursing Care Quality* 18(3): 184–92.
Lake, Eileen T. 2002. Development of the practice environment scale of the nursing work index. *Research in Nursing & Health* 25(3): 176–88.
Laschinger, Heather K. Spence, and Joan Almost. 2003. Patient satisfaction as a nurse-sensitive outcome. In Diane Doran (Ed.), *Nursing Sensitive Outcomes: The State of the Science*, pp. 243–82. Sudbury, MA: Jones & Bartlett Learning.
Lee, David, and Alexandre Mas. 2012. Long-run impacts of unions on firms: New evidence from financial markets, 1961–1999. *Quarterly Journal of Economics* 127(1): 333–78.
Lee, Lung-Fei. 1978. Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables. *International Economic Review* 19(2): 415–33.
Lewis, H. Gregg. 1963. *Unionism and Relative Wages in the United States: An Empirical Inquiry*. Chicago: University of Chicago Press.
Litwin, Adam Seth. 2011. Technological change at work: The impact of employee involvement on the effectiveness of health information technology. *ILR Review* 64(5): 863–88.
Mas, Alexandre. 2006. Pay, reference points, police performance. *Quarterly Journal of Economics* 121(3): 783–821.
———. 2008. Labor unrest and the quality of production: Evidence from the construction equipment resale market. *Review of Economic Studies* 75(1): 229–58.

Matsudaira, Jordan. 2014. Monopsony in the low-wage labor market? Evidence from minimum nurse staffing regulations. *Review of Economics and Statistics* 96(1): 92–102.

McCrary, Justin. 2007. The effect of court-ordered hiring quotas on the composition and quality of police. *American Economic Review* 97(1): 318–53.

Needleman, Jack, Peter Buerhaus, Soren Mattke, Maureen Stewart, and Katya Zelevinsky. 2002. Nurse staffing levels and the quality of care in hospitals. *New England Journal of Medicine* 346(22): 1715–22.

Pappas, Sharon H. 2008. The cost of nurse-sensitive adverse events. *Journal of Nursing Administration* 38(5): 230–36.

Propper, Carol, and John Van Reenen. 2010. Can pay regulation kill? Panel data evidence on the effect of labor market outcomes on hospital performance. *Journal of Political Economy* 118(2): 222–73.

Schumacher, Edward J. 1997. Relative wages and exit behavior among registered nurses. *Journal of Labor Research* 18(4): 581–92.

Seago, Jean Ann, Joanne Spetz, Michael Ash, Carolina-Nicole Herrera, and Dennis Keane. 2011. Hospital RN job satisfaction and nurse unions. *Journal of Nursing Administration* 41(3): 109-114.

Slichter, Sumner H., James J. Healy, and Robert E. Livernash. 1960. *The Impact of Collective Bargaining on Management*. Washington, DC: Brookings Institution.

Sojourner, Aaron, Robert Town, David C. Grabowski, and Michelle Chen. 2015. Impacts of unionization on employment, product quality, and productivity: Regression discontinuity evidence from nursing homes. *ILR Review* 68(4): 771–806.

Spetz, Joanne, Michael Ash, Charalampos Konstantinidis, and Carolina Herrera. 2011. The effect of unions on the distribution of wages of hospital employed registered nurses in the United States. *Journal of Clinical Nursing* 20(1–2): 60–67.

Twigg, Diane, Christine Duffield, Alex Bremner, Pat Rapley, and Judith Finn. 2011. The impact of the nursing hours per patient day (NHPPD) staffing method on patient outcomes: A retrospective analysis of patient and staffing data. *International Journal of Nursing Studies* 48(5): 540–48.

Wessels, Walter. 1994. Do unionized firms hire better workers? *Economic Inquiry* 32(4): 616–29.