Alive and Kicking? Short-Term Health Effects of a Physician Strike in Germany

Daniel Avdic* Martin Karlsson† Lea Nassal‡ Nina Schwarz§

September 27, 2023

Abstract

We study the effects of a physician strike in German hospitals in 2006 on patient mortality. Leveraging a comprehensive dataset encompassing all hospital admissions in Germany and employing digitised records of strike participation, we estimate a difference-in-differences model to discern the causal effects of the strike. Our estimation results reveal a substantial decrease in hospital admissions during the strike period, whereas effects on hospital mortality are mostly driven by patient selection. To support this claim, we further show that emergency cases and more fragile patients, who were unable to substitute their immediate care needs, were more likely to be present in hospital during this period. Hence, in contrast to most other related studies, our results suggest that short term interruptions in access to healthcare may not have dramatic effects on healthcare quality provided that rationing of care by patient severity is carried out.

Keywords: physician strike, quality of care, rationing, Germany.

JEL Classifications: H75, I11, J52.

*Avdic: Deakin University

†Karlsson: CINCH, University of Duisburg-Essen, Essen, Mail: Weststadttürme Berliner Platz 6-8, 45127 Essen, Germany, Email: martin.karlsson@uni-due.de

‡Nassal: University of Duisburg-Essen

§Schwarz: University of Duisburg-Essen, Essen, Email: nina.schwarz@uni-due.de

1 Introduction

Labour disruptions by physicians are often seen as controversial because of physicians' obligation towards their patients (Metcalfe, Chowdhury and Salim, 2015). The International Labour Organisation (ILO) defines hospitals as an 'essential service' where a strike could in principle be prohibited due to a threat to life, personal safety or health (cf. International Labour Office, 2006). Even temporary disruptions to access to healthcare providers may have important effects on healthcare quality, such as emergency care. However, despite these concerns, physician strikes hit many countries, such as France in 2019, England in 2015, and Portugal in 2012 and 2019 and Germany in 2006 and 2020.

This paper analyses the short-term effects of a nationwide physician strike in Germany on inpatient mortality. The event took place in 2006 when physicians in tertiary hospitals took industrial action for three months between March and June.[1] It was one of the longest labour disruptions in German history and at the peak about 2/3 all physicians employed in tertiary hospitals participated. To empirically study the effects of the strike, we use rich administrative data on all hospital admissions and deaths in Germany for the years 2000-2008 to compare outcomes in strike- and non-strike impacted hospitals over time in a difference-in-differences model. The unprecedented and unanticipated nature of the strike provides a context where causal effects of the disruption on healthcare supply can be plausibly identified and quantified.

The extent of the adverse effects of a physician strike on patient health is not given a priori. On the one hand, labour disruptions may reduce quality of healthcare due to decreased geographical access, higher workload among healthcare staff and crowding effects. Avdic (2016) shows that an increased distance to hospitals due to closures of emergency departments resulted in a lower probability of surviving an acute myocardial infarction in the short run, due to an increase in the risk of out of hospital mortality. Piérard (2014), Lin (2014) and Aiken et al. (2002) show that the staff-to-

[1] The majority of German tertiary hospitals are organised on federal state level and thus provide medical care to all publicly as well as privately insured individuals. They incorporate medical schools and are also involved in research and teaching activities.

patient ratio is a determinant of patients' health outcomes. On the other hand, the adverse effects of a strike on healthcare quality can be mitigated through rationing care and triaging patients based on urgency. For example, hospitals can cancel or postpone elective surgeries, which could even lead to lower in-hospital mortality rates and concomitant complications in the short run. Cunningham et al. (2008) and Metcalfe, Chowdhury and Salim (2015) systematically review the medical literature on physician strikes and conclude that at least the withdrawal of medical services for a short period of time does not seem to increase mortality for the majority of labour disputes. A potential explanation is that in most cases some kind of emergency care is still available. However, most of the reviewed articles only present suggestive evidence and no causal analysis.

Our difference-in-difference estimates suggest that the strike led to a sharp reduction (12%) in hospital admissions in striking hospitals during the months the strike was ongoing compared to non-striking hospitals. Emergency admission rates increased by 5% and length of stay increased by 0.2 days. Estimates not adjusted for case-mix also indicate an increase in in-hospital mortality by 9-10%. Due to the potential endogeneity of patient composition because of care rationing, it is crucial to control for patient case-mix when studying strike effects in healthcare. To do this in the presence of a very large set of covariates representing case-mix, we implement a double selection Lasso model (Belloni, Chernozhukov and Hansen, 2014). Once we adjust for exogenous patient characteristics, the coefficient on mortality is reduced by half. Furthermore, our results indicate some spillover effects to nearby hospitals whereas there are no signs of a post-strike catch-up effects on admissions. We conclude that healthier patients at striking hospitals likely avoided care or were triaged, so that the patient composition during the strike had a higher share of 'frail' patients with higher underlying mortality risk.

We contribute to the scant literature on the causal effects of strikes in healthcare facilities. To our knowledge, only two papers have investigated the impact of physician strikes. Costa (2019) investigates labour disruptions caused by physicians in Portugal. He investigates the effects of nurses', physicians' and diagnostic and therapeutic technicians' (DTT) strikes on health outcomes. While no effects for nurses' and DTT strikes can be found, physician strikes seem to have increased in-hospital mortality by about 8%. Furthermore, he finds a shift from inpatient to outpatient services and an increase in complications around birth. Stoye and Warner (2023) analyse the impact of a number of short strikes of junior doctors in the United Kingdom in 2016. Even though our research question is related, our design is very different from those two papers. Whereas they consider short and repeated disruptions caused by strikes, we study the impact of a prolonged strike that affected entire hospitals over a period of three months. Therefore, the shock we consider represents a more serious challenge to the health care system, but it also entails more time for the system to respond to the challenge. In this sense, our analysis provides a very strong test of the resilience of the system in the presence of a major unexpected challenge.

In terms of strikes of other healthcare workers, Gruber and Kleiner (2012) analyse the health effects of nurses' strikes in New York between 1984-2004 and observe that strikes in hospitals lead to lower quality of medical care. They find an 18.3% increase in in-hospital mortality and an increase of 5.7% in 30-day readmission for patients admitted during a strike. Furthermore, Kronborg, Sievertsen and Wüst (2016) investigate a nurse strike in Denmark in 2008. They look at maternal and infant health and conclude that there was a decrease in prenatal midwife consultations, in length of stay after birth and in the number of home visits caused by the strike. In a follow-up study (Hirani, Sievertsen and Wüst, 2019) the same strike is used to investigate the impact of the timing of nurse home visits for newborns. Having missed earlier nurse visits due to the strike led to more regular and urgent general practitioner contacts in the first years of life compared to having missed later nurse visits. Furthermore, Friedman and Keats (2019) investigate effects on infant and neonatal health. Their conclusion is that babies born during a health worker strike in Kenya are more likely to die if they were born in a month with strike occurrence. An additional 15 deaths per 1,000 live births is observed one week and one month after birth which is an increase of 54-68%.

We show that the impact of strikes among physicians may generate different impacts on care quality compared to nurses and midwife strikes.

2 Background

2.1 The German Hospital Sector

The German hospital sector treats about 19 million patients every year and involves approximately 2,000 hospitals with 500,000 beds (Federal Statistical Office, 2018). Hospitals can be owned by public, private or non-profit organisations. Public hospitals can be separated further into hospitals run by the municipalities, the 16 federal states, the state and hospitals related to the statutory accident insurance (Goepfert and Conrad, 2013). The majority of the 36 university hospitals are run by the federal states. They are linked to the medical faculties of the universities and are thus also involved in research and teaching activities.

Roughly 10 percent of the 19 million patients treated in hospitals are admitted to university hospitals (Federal Statistical Office, 2018). Although they make up less than two percent of the hospitals, they treat around 10 percent of patients and hold about 20 percent of the physicians working in hospitals.[2] As there is free choice of healthcare providers, patients can freely choose the hospital they want to be treated in and physicians in the primary or ambulatory care sector are not obliged to transfer patients to a specific hospital (de Cruppé and Geraedts, 2017). Treatment costs are usually covered by a patient's health insurance, as health insurance is mandatory in Germany and all legal residents have universal coverage by either public or private health insurance (Blümel and Busse, 2015).[3]

[2] Table A.1 in Appendix A shows how the number of physicians and other health care workers in hospitals and university hospitals evolved between 2000 and 2010.

[3] The majority of patients (88%) are publicly insured and only about 11% are covered by a private health insurance (Bundesministerium für Gesundheit, 2019). The data used in this analysis covers privately and publicly insured patients. However, it is not possible to distinguish between privately and publicly insured.

2.2 Marburger Bund

The *Marburger Bund* is a professional organisation and the trade union for employed physicians in Germany, covering about 70 percent of all physicians employed in hospitals (Greef, 2012). Prior to the initiation of the physician strike in 2006, it has mainly acted as a professional organisation and was represented during collective bargaining by the *Vereinte Dienstleistungsgewerkschaft (ver.di)*, a multi-branch trade union (Greef, 2012).

Along with the negotiations on a new collective bargaining agreement for public services in 2005, the *Marburger Bund* demanded better working conditions for physicians working in hospitals (Marburger Bund Bundesverband, 2012). As *ver.di* was not willing to negotiate on an own collective bargaining agreement for physicians, the *Marburger Bund* decided to terminate the collective partnership with *ver.di* and become the first and only trade union for physicians.

The *Marburger Bund* and the employer's association officially started collective bargaining for hospitals on federal state level[4] on October 12, 2005 (Greef, 2012). As they were not able to reach a consent, the biggest-ever physician strike in German history started on March 16, 2006 and went on until June 16, 2006, when finally a collective bargaining agreement on federal state level was achieved. Table 1 presents a chronology of events starting in 2005. The table also presents information on the strike in hospitals on the municipality level. However, the focus in this paper will be on the strike at the federal state level, as university hospitals are organised on the federal state level.

[4]The *Marburger Bund* had to negotiate with two different employer's associations – the *Tarifgemeinschaft deutscher Länder* (TdL) which is responsible for public hospitals on the federal state level and the *Vereinigung kommunaler Arbeitgeberverbände* (VKA) which is responsible for public hospitals on the municipality level. The focus in this paper will be on the strike at the federal state level, as university hospitals are organised on federal state level.

Table 1. Chronology of events

Date	Event
14 Apr 2005	ver.di and employer's association start bargaining (federal)
9 Sep 2005	MB terminates collective partnership with ver.di
12 Oct 2005	MB and employer's association start bargaining (federal)
9 Mar 2006	MB and employer's association start bargaining (local)
16 Mar 2006	Strike begins (federal)
16 Jun 2006	MB and employer's association reach agreement for physicians (federal)
26 Jun 2006	Strike begins (local)
17 Aug 2006	MB and employer's association reach agreement for physicians (local)

Note.— MB = Marburger Bund; ver.di = Vereinte Dienstleistungsgewerkschaft; *Sources:* Greef (2012) and Martens (2008).

2.3 Physician Strike

The strike studied in this paper took place in the majority of German university and federal hospitals between March and June 2006[5]. It was one of the longest labour disruptions in German history (Martens, 2008). At first, these labour disruptions were restricted to a few university and federal hospitals and to one or two days per week, but soon the strikes expanded to other university hospitals and lasted longer and longer; sometimes even a whole week. At the peak of the strike, about 2/3 of all physicians in these hospitals participated.

In general, emergency, intensive care, and maternity units were generally excluded from the strike. This was ascertained by emergency agreements, signed by the university hospitals and the *Marburger Bund*. However, elective treatments and other non-urgent cases was affected by the strike.

3 Data

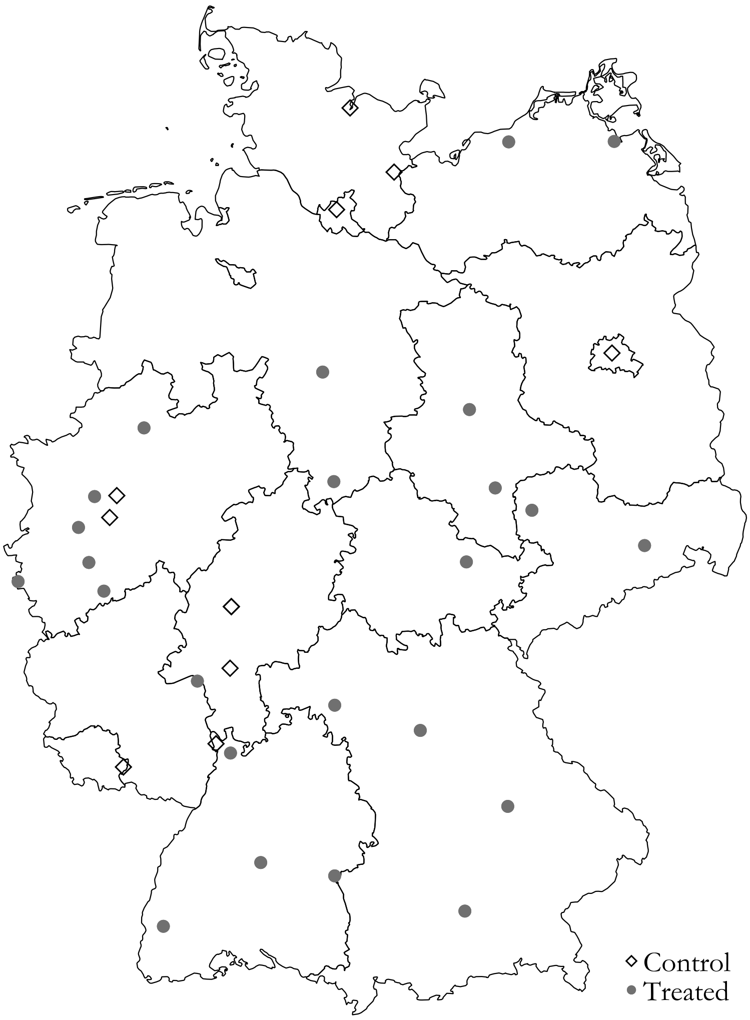
3.1 Strike Data

One challenge with studying the effects of the 2006 strike is that there exist no reliable administrative data which could be used for the analysis (Dribbusch, 2018). Therefore,

[5]There were 33 university hospitals in total, of which 24 participated in the strike. The university hospitals in Berlin, Frankfurt, Gießen and Marburg did not participate, as Berlin and Hesse are not part of the employer's association. Berlin has been excluded from the employer's association in 1994 and Hesse left in 2004 (Greef, 2012). Hamburg and Schleswig-Holstein (Kiel and Lübeck) did not take part as they had reached a collective agreement before the strike took place. The University hospitals in Homburg, Bochum, Mannheim and Wuppertal did not participate in the physician strike for other reasons.

we collected information on the strike from newspaper articles published in 2006 to obtain information on the exact date of the start and end of the labour disruptions for each university hospital. In total, physicians from 24 out of 33 university hospitals went on strike; nine of these participated from the first day onwards, while the other participating hospitals joined a few weeks later. The remaining nine university hospitals were not part of the collective bargaining agreement with the TdL in early 2006 and therefore did not participate in the labour disruptions. Figure 1 shows the location of treatment and control university hospitals in Germany.

FIGURE 1. Locations of treatment and control hospitals



3.2 Hospital Level Data

Hospital level data on admissions for the years 2000 to 2008 was obtained from the German Federal Statistical Office (FDZ der Statistischen Ämter des Bundes und der Länder, 2000-2008*a*). It contains individual level information on each hospital admission, including date of admission, length of stay, mortality and surgery indicators, 3-digit ICD-diagnosis and patient characteristics such as age, gender, and place of residence. Hospital level characteristics such as number of physicians and nurses are available on a yearly basis and used for balancing tests (see Section 4). Since German hospitals are required to provide this information by law, the data is complete and of high quality (see KHStatV, 1990).

The admission data is aggregated on day of admission and hospital level. For the main regressions on the hospital level, we keep only university hospitals as they provide a balanced sample whereas other hospitals could have been exposed to the strike on the municipality level starting in June 2006, for which we do not have strike data. The final sample consists of 143,744 observations. Table 2 presents descriptive statistics for the main outcome variables, treatment indicator and patient characteristics by treatment status of the hospital over all observation years.

TABLE 2. Descriptive statistics: Hospital data

	Treatment Hospital			Control Hospital		
	Mean	SD	Obs.	Mean	SD	Obs.
Treatment status	0.03	0.16	84,986	0.00	0.00	58,758
Female	0.49	0.05	84,986	0.49	0.07	58,758
Age	47.38	3.89	84,986	48.99	5.28	58,758
Hospital daily cases	159.28	69.88	84,986	156.27	111.23	58,758
Hospital log daily cases	4.95	0.54	84,986	4.79	0.81	58,758
Length of Stay	8.12	1.59	84,986	8.26	2.13	58,758
Surgery rate	37.06	22.17	84,986	33.57	22.03	58,758
Hospital log daily mortality	0.83	0.59	68,985	0.92	0.66	38,410
Mortality rate	1.67	1.45	84,986	1.94	2.17	58,758
10-Day mortality rate	0.91	1.11	84,986	1.08	1.64	58,758
Emergency rate	27.65	12.18	84,648	30.03	13.22	58,508
Urgency	0.33	0.09	84,648	0.35	0.09	58,508

NOTE.—Descriptive statistics by treatment status of the hospital. Variable definitions are available in Appendix B. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

The primary outcome variable in this analysis is the mortality rate (patient deaths per 100 admissions) in each hospital on each day. As suggested by Gruber and Kleiner (2012), we also investigate deaths occurring within 10 days after the day of admission as this outcome limits bias from strike-induced changes in length of stay. Furthermore, we estimate strike effects on length of stay, surgery rate (surgeries per 100 admissions), an ambulatory care sensitive (ACS) indicator based on Sundmacher et al. (2015)[6] and an emergency and urgency indicator based on the classification by Krämer, Schreyögg and Busse (2019)[7]. The latter outcome variables is informative of the extent to which the strike led to any changes of the patient composition in hospitals.

Exposure to the strike is defined as being admitted to a striking hospital during the strike period. The variable takes on the values zero or one. This variable does not capture patients that already were hospitalised when the strike began. We therefore use an alternative treatment indicator in a robustness check, defined as the share of patients on a specific day that were affected by the strike at any time during their hospital stay (see Section 6).

3.3 Mortality Census

Mortality data was also obtained from the German Federal Statistical Office (FDZ der Statistischen Ämter des Bundes und der Länder, 2000-2008*b*) and includes all deaths (both in and out of hospitals) and causes for death (3 digit ICD-Codes) between 2000 and 2008. We use district[8] of residence and date of death to link and merge the mortality data with the information on strikes we collected and then we collapse the data to the date of death and district level. Treatment assignment is based on district of residence, as the data does not allow for computing the exact distance between place of residence and striking hospitals.

Our final analysis data sample consists of approximately 1.3 million district-date

[6]The ACS indicator helps to identify hospitalisations which could have been avoided by effective and timely ambulatory care (Sundmacher et al., 2015). The analysis in this paper investigates whether the strike changed the ACS rate of a hospital by transferring patients to ambulatory care units outside of the hospital.

[7]Krämer, Schreyögg and Busse (2019) developed a diagnosis-based classification to determine whether a hospitalisation is elective or an emergency. They used machine learning and several predictor variables to classify the different diagnosis. A hospitalisation is defined as an emergency if their urgency indicator is above 0.5.

[8]The district (*Kreis*) is the intermediate administrative level between municipality (*Gemeinde*) and federal state (*Bundesland*). These units have responsibilities in the domains of public transport, road construction, and the construction of hospitals. There were in total 323 of them in 2006.

cells which we use as observations in our regressions. The district mortality data is useful to study effects on mortality outside of the hospital. For example, it could be that patients affected by the strike died outside of the hospital if they were discharged too early or failed to be admitted in time for a critical condition. As our data covers the post-strike period, we are also able to study any lagged effects of the strike to study, for example, net mortality effects of the strike. Descriptive statistics for the mortality data are presented in Table 3.

TABLE 3. Descriptive statistics: Mortality data

	Treatment district N = 77,477		Control district N = 32,880	
	Mean	SD	Mean	SD
Treatment Variable	0.03	0.16	0.00	0.00
Female	0.55	0.23	0.54	0.18
Married	0.38	0.23	0.38	0.18
Age at death	76.12	7.19	75.86	5.41
Log Deaths	1.92	0.90	2.46	1.01

NOTE.— Descriptive statistics by treatment status of the district. Variable definitions are available in Appendix B. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*b*), own calculations.

4 Econometric Strategy

We apply a difference-in-differences (DID) empirical design to estimate the causal effects of the 2006 strike. To this end, we use the nine non-striking university hospitals as a control group for the 24 striking university hospitals to estimate group-specific changes in mortality and other outcomes of interest before, during and after the strike occurred. Whereas the two groups may exhibit systematic differences in size and staffing levels, they are likely to be exposed to the same general trends in terms of admissions and case-mix. Therefore, assuming that in the absence of the strike, the two groups would follow a parallel trend seems reasonable, and in this case the DID specification would identify the average treatment effect on the treated. Below we carefully assess the plausibility of this assumption and choose our empirical specification based on the balance of some key hospital characteristics. We also consider an alternative control group in robustness checks. Subsequent analyses, described in detail below, will explore underlying channels for our main results, including changes in patient composition and spillover effects in non-striking hospitals.

4.1 Empirical Model

Our main regression specification is defined by the following equation:

$$y_{it} = \beta_0 + \beta_1 S_{it} + \mu_i + \nu_t + H_i \cdot t + \epsilon_{it} \quad (1)$$

where y_{it} is one of the outcome variables listed in Table 2 for hospital i on day t, S_{it} is a strike indicator which is 1 for strike-hit hospitals during the strike period and 0 otherwise, μ_i are hospital fixed effects, ν_t are date of admission fixed effects and $H_i \cdot t$ are hospital-specific linear time trends. Date of admission fixed effects include fixed effects for year, week, and day of week. The analysis exploits regional and temporal variation in strike exposure and hospital fixed effects can control for time-invariant hospital heterogeneity, such as average case volume, number of beds and share of specialists. Under the assumption that treatment and control hospitals would have followed a common trend in the absence of the strike, conditional on the included control variables, the parameter β_1 measures the average treatment effect on the treated (ATT) of the strike on the outcome of interest. All regressions are weighted by the baseline average admissions per day in each hospital and standard errors are clustered on hospital level.[9]

In addition to in-hospital mortality, we also investigate aggregate district mortality effects by estimating the following regression model:

$$y_{jt} = \beta_0 + \beta_1 S_{jt} + \mu_j + \nu_t + C_j \cdot t + \gamma X_{jt} + \epsilon_{jt} \quad (2)$$

where y_{jt} is the logarithmised number of deaths occurring on day t in district j, S_{jt} is a strike indicator variable equal to one for districts with a strike-hit hospital during the

[9] Although there are only 33 university hospitals in total, some of these are based in multiple locations. We treat each location as a separate entity meaning that there are 44 clusters in total. As number of admissions might be affected by the treatment, we also present unweighted estimates for the main outcome variables in Table A.2 in Appendix A. Point estimates are similar to the weighted estimates. Furthermore, we implement case-mix adjustment in Section 4.2

strike period and zero otherwise, μ_i are district fixed effects, ν_t are date fixed effects and $C_j \cdot t$ are district specific time trends. Date fixed effects include year fixed effects, week fixed effects and day of week fixed effects. Finally, X_{jt} is a vector of time-varying district-level patient characteristics, including the share of females, and the share of married individuals. The coefficient of interest is β_1, measuring the effect of the strike on the log number of deaths in districts with a striking hospital.

Table 4 presents hospital balancing tests to study the comparability between treated and untreated hospitals in our sample.[10] To study this, we regress each of the variables reported in the table on a set of control variables and the treatment variable indicator. The point estimates from column (1), where we only include year fixed effects as controls, suggest that treated hospitals are on average larger with greater number of beds, more staff and a higher share of intensive care unit beds. However, once we additionally control for hospital fixed effects and hospital specific trends, the test suggests that the two groups are well balanced on observable characteristics.

TABLE 4. Covariate balancing tests

	Observations	Mean	(1)	(2)	(3)
N Beds	395	1,130.48	432.597**	-22.986	-5.064
			(191.028)	(27.176)	(15.843)
S ICU Beds	395	0.08	0.020**	0.001	0.002
			(0.009)	(0.002)	(0.003)
N Physicians	395	573.50	337.988***	37.294	8.034
			(106.905)	(23.928)	(20.568)
S Female Physicians	395	0.33	-0.003	0.001	-0.004
			(0.017)	(0.008)	(0.011)
N Physicians in Charge	395	40.74	15.316	-2.302	-3.242
			(12.246)	(4.399)	(3.192)
N Assistant Physicians	395	433.12	227.567**	20.984	10.834
			(87.075)	(17.948)	(11.049)
N Non-Physicians	395	3,294.67	1972.984***	202.625*	44.069
			(560.346)	(108.468)	(104.998)
S Female Non-Physicians	395	0.77	-0.010	-0.008*	-0.002
			(0.013)	(0.004)	(0.003)
N Trainee Positions	395	352.22	134.706**	-20.573	-22.718
			(66.874)	(19.103)	(17.253)
N Childbirth	395	1,167.57	508.146	51.139	9.083
			(413.760)	(50.383)	(38.120)
S C-sections	306	0.36	0.000	-0.021	-0.008
			(0.037)	(0.025)	(0.019)
Year FE			✓	✓	✓
Hospital FE				✓	✓
Hospital Specific Trends					✓

NOTE.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions successively include year fixed effects, hospital fixed effects and hospital specific time trends. In comparison to the main specification the treatment period is defined as 2006 as data was only available on a yearly level. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

[10] As the data is only available yearly, we coded the entire year 2006 as treatment period.

4.2 Patient Selection

There are at least two challenges to the identification of causal effects in our setting which might affect the validity or interpretation of any results derived from our regression model. First, it could be that there are underlying trends in the outcome variables that might have caused the strike. Due to the unprecedented nature of the strike, we argue that it is unlikely that patients were anticipating such radical labour disruptions. Nevertheless, we address this concern by presenting event study estimates for the main outcome variables and test for linear pre-trends.

A second issue regarding the estimation of strike effects on mortality is that the strike might have changed the patient composition. While emergency care was still available, elective surgeries and other non-emergent cases might have been affected. If emergency care patients have worse health outcomes than non-emergency care patients, this change in the case-mix will contribute to our DID estimates. Whereas the estimated effect captures the causal effect of the strike on the hospital-level mortality rates, it is not the causal effect that is relevant from a policy perspective. In order to infer whether the quality of care was compromised during the strike, we need to base estimates of mortality effects on a comparable population of patients. We address this concern in two ways: first, we seek to control for patient composition in the empirical analysis, and second, we estimate effects of the strike on district mortality rates.

Combining the various patient and case characteristics that are available in the data – age, gender, ICD codes, indicators of urgency, emergency and deferability – there are 337 specific covariates that may be considered. In order to avoid overfitting, we rely on a double selection Lasso model, where covariates that predict either our main treatment indicator, or the mortality outcome, are included in the specification (Belloni, Chernozhukov and Hansen, 2014). Our Lasso selection specification eventually selected 163 covariates for the analysis. We add them as control variables to the main regression specification.[11]

[11] In Appendix C, we instead consider the covariate selection approach suggested by Hirano and Imbens (2001) and show that results are robust to this alternative variable selection approach.

4.3 Analysis of Spillover Effects

If admissions in striking hospitals decrease and only more fragile patients are admitted, healthier patients may decide to either delay hospitalisation, seek treatment at an other hospital or substitute other types of care for hospital care. Our estimates of post-strike admissions and mortality will give an indication of whether postponement is a common strategy. In what follows, we also analyse the extent to which surrounding non-university hospitals were affected by the strike.

We use the year 2005 as our benchmark to calculate, for each district, the share of patients who receive treatment in any strike-hit university hospital in that year. We argue that this variable is a good proxy to study strike spillover effects, since the pressure on surrounding non-treated hospitals should be greatest in districts where a high share of patients treated in strike-hit hospitals at baseline. The baseline patient shares is thus used as a regressor in the following regression model:

$$y_{it} = \beta_0 + \beta_1 S_{it} + \mu_i + \nu_t + C_i \cdot t + \epsilon_{it} \quad (3)$$

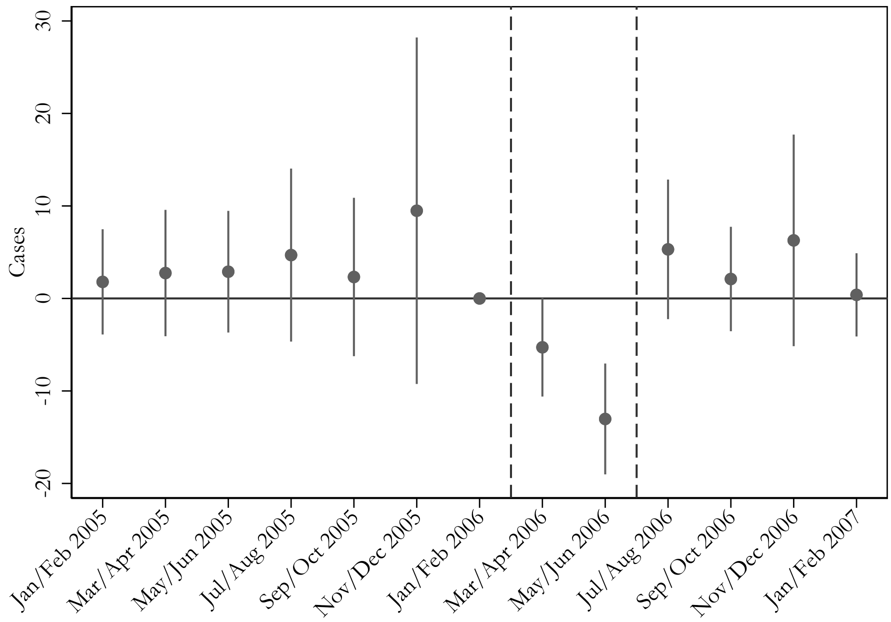
where y_{it} is one of the hospital outcome variables for district i on day t. S_{it} is an interaction between the treatment indicator and the pre-strike share of patients treated in university hospitals. Remaining variables are defined as above.

5 Results

5.1 Hospital Admissions and Patient Mortality

We first investigate whether the strike affected the number of daily hospital admissions as a consequence of the 2006 physician strike. Figure 2 illustrates bi-monthly event study estimates for the relative change in hospital admissions for striking and non-striking hospitals. The absence of a trend for the point estimates in the time periods leading up to the strike suggests that the critical common trend assumption required for a causal interpretation of our DID estimator is valid. Considering the estimates for the strike period itself, indicated by the vertical dashed lines, we see a significant relative drop in admissions in striking hospitals in May and June 2006 by about 15 cases per day when the labour disruptions were the most intense. Finally, the coefficient estimates once again turn statistically indistinguishable from zero once the strike ends, indicating that striking hospitals experienced a permanent drop in their admission rates as a consequence of the strike.

FIGURE 2. Event study results: Hospital admissions



NOTE.— The figure shows event study estimates for admissions, controlling for day of week fixed effects and hospital fixed effects. The vertical lines represent the beginning and end of the strike. 95% confidence intervals included. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Formal DID estimates are reported in Table 5, confirming the event study result that the strike significantly decreased hospital admissions. In particular, when including time and hospital fixed effects and hospital-specific linear time trends, the estimates from Columns (1) and (2) suggest a decrease in daily admissions by 19-21 cases on average, or 12 percent. Columns (3) and (4) of the table reports results from applying the covariate adjustment approach described in Section 4.2. While this adjustment marginally attenuates the point estimate, it remains large and highly statistically significant.

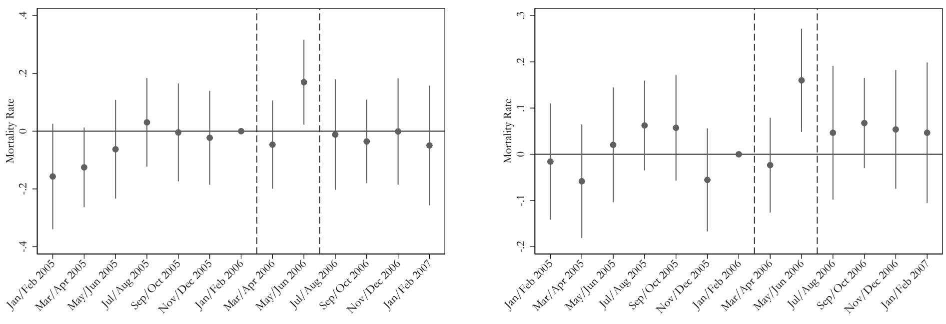
TABLE 5. Difference-in-differences estimates: Hospital admissions

	Observations	Mean	(1)	(2)	(3)	(4)
Cases	143,744	158.32	-21.460***	-19.041***	-15.583***	-13.217***
			(4.646)	(2.596)	(3.444)	(1.717)
Log Cases	143,744	4.90	-0.122***	-0.118***	-0.082***	-0.079***
			(0.019)	(0.014)	(0.015)	(0.010)
Covariate adjustment					✓	✓
Year FE			✓	✓	✓	✓
Hospital FE			✓	✓	✓	✓
Hospital Specific Trends				✓		✓

NOTE.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions include day of week fixed effects, week fixed effects, year fixed effects, hospital fixed effects and hospital specific time trends. Regressions are weighted by number of admissions. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Next, we analyse the consequences of the 2006 strike on in-hospital mortality rates. Figure 3 presents event study graphs for the overall mortality rate and 10-day mortality rate. Again we can observe that striking and non-striking hospitals had similar mortality trends in the lead-up to the strike, confirming that our empirical strategy appears sound in this context. We also see a reverse trend in the estimates for the strike months with a relative increase in both in-hospital and 10-day mortality rates for striking hospitals during the intensive strike period in May and June. These estimates return to being close to zero and statistically insignificant at the end of the strike period.

FIGURE 3. Event study results: Hospital mortality



(a) In-hospital Mortality Rate (b) 10-Day Mortality Rate

NOTE.— The figures show event study estimates for mortality rate and 10-day mortality rate, controlling for day of week fixed effects and hospital fixed effects. The vertical lines represent the beginning and end of the strike. 95% confidence intervals included. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

DID estimates corresponding to the event study plots for the mortality outcomes are presented in Columns (1) and (2) of Table 6. The point estimates imply that the mortality rate increased by roughly 0.16 percentage points, corresponding to an increase of around nine percent compared to the baseline mean of 1.75 deaths per 100 admissions. The estimates for 10-day mortality are somewhat lower at 0.07 percentage points, or 7.3 percent, but remain significant at the five percent level of statistical significance. Finally, reported estimates from the third row of the table indicate that the number of deaths in hospital drop by 0.07 from a baseline of 2.46; a reduction by 2.8 percent. This estimate is not significantly significant, yet consistent with the estimates for admissions (-12%) and mortality (+9%) since $0.88 \cdot 1.09 = 0.96$.

Columns (3) and (4) of the table reports results from applying the covariate adjusted approach described in Section 4.2. The results suggest that, conditional on the selected covariates, the strike had a moderate effect on case-specific mortality rates: the increase is now down to 0.8 percentage points or 4.6 percent. This suggests that a substantial part of the above estimated mortality effect is due to selection of sicker patients into striking hospitals during the strike period. This interpretation is confirmed by the fact that when we estimate the effect of the strike on predicted mortality, we obtain a treatment effect of 0.06 percentage points.

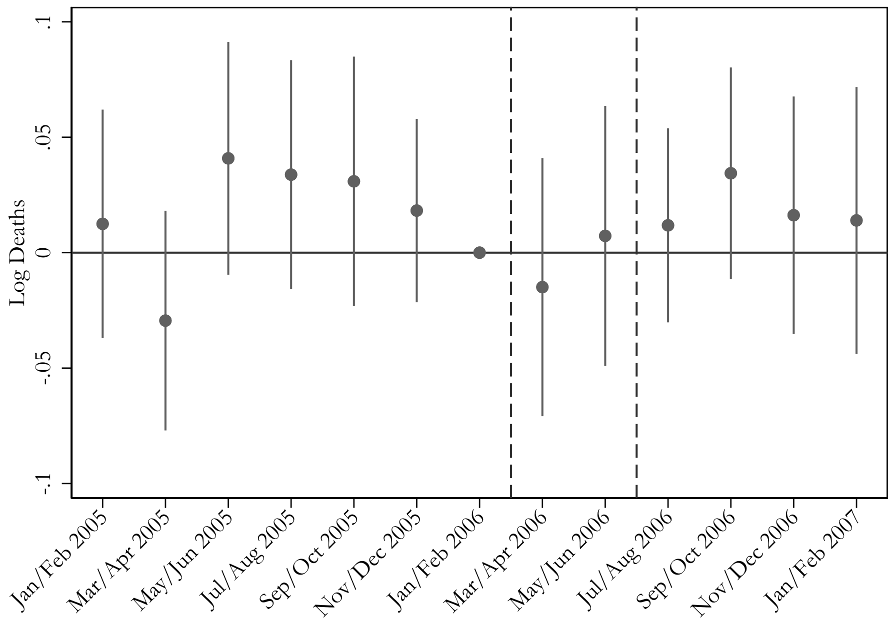
Table 6. Difference-in-differences estimates: Hospital mortality

			Baseline Specification		Covariate Adjusted	
	Observations	Mean	(1)	(2)	(3)	(4)
Mortality Rate	143,744	1.75	0.159***	0.165***	0.080**	0.085**
			(0.043)	(0.041)	(0.037)	(0.035)
10-Day Mortality Rate	143,744	0.96	0.074**	0.073**	0.022	0.021
			(0.036)	(0.035)	(0.030)	(0.030)
Number of Deaths	143,744	2.46	-0.099	-0.069	-0.088	-0.062
			(0.087)	(0.056)	(0.077)	(0.054)
Year FE			✓	✓	✓	✓
Hospital FE			✓	✓	✓	✓
Hospital Specific Trends				✓		✓
Covariate adjustment					✓	✓

Note.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions include day of week fixed effects, week fixed effects, year fixed effects, hospital fixed effects and hospital specific time trends. Regressions are weighted by number of admissions. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

If there were noteworthy and significant strike effects on mortality, these should also be visible for the district mortality. We therefore present event study estimates for treatment and control districts in Figure 4. The graph suggests no significant increase in mortality during the strike period or thereafter.

FIGURE 4. Event study results: District mortality



NOTE.— The figure shows event study estimates for log mortality, controlling for date fixed effects and district fixed effects. The vertical lines represent the beginning and end of the strike. 95% confidence intervals included. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*b*), own calculations.

5.2 Care Rationing and Hospital Spillovers

In this section, we explore how the physician strike affected the patient composition and potential spillover effects to other non-striking hospitals. We first study changes in the striking hospitals' patient compositions. Table 7 reports estimates for a set of outcomes associated with patient characteristics, including age, sex, rates of Ambulatory Case Sensitive (ACS) and emergency cases, an indicator for urgency (yes/no) and a deferability index, and hospital decisions, including length of stay and rate of surgery. We distinguish between these two sets of outcomes as the latter are more related to hospitals' choices in responding to the strike, whereas the former are exogenous to the hospital.

The estimates presented in the table suggest that the strike led to a significant reduction of patients' age by 0.41 years and significant increases in both emergency rates and the probability that a patient is classified as an urgent case. This is consistent with the fact that emergency care patients in general tend to be younger on average.[12]

[12]Figure A.1 in Appendix A shows the age distribution of non-emergency versus emergency care patients.

With respect to hospital decisions, we estimate a statistically significant increase in length of stay by 0.24 days. These findings together suggest that patients in relatively poorer health conditions were admitted to striking hospitals during the strike period. We do not find significant changes in the composition of gender or ACS conditions. For surgery rates, we find a considerable, albeit only marginally significant, reduction by about 4.5 percent, suggesting a possible postponement of scheduled surgeries during the strike period.

TABLE 7. Difference-in-differences estimates: Patient composition

	Observations	Mean	(1)	(2)
		PATIENT COMPOSITION		
Age	143,744	47.89	-0.369**	-0.411***
			(0.157)	(0.146)
Female	143,744	0.49	0.002	0.002
			(0.002)	(0.002)
ACS Rate	143,744	20.30	-0.188	-0.301
			(0.221)	(0.183)
Emergency Rate	143,156	28.41	1.558***	1.505***
			(0.349)	(0.348)
Urgency Indicator	143,156	0.34	0.012***	0.012***
			(0.003)	(0.002)
Deferability Index	143,549	0.12	0.003***	0.003***
			(0.001)	(0.001)
		HOSPITAL DECISIONS		
Length of Stay	143,744	8.17	0.255***	0.243***
			(0.055)	(0.052)
Surgery Rate	143,744	35.95	0.016	-1.629*
			(1.409)	(0.930)
Year FE			✓	✓
Hospital FE			✓	✓
Hospital Specific Trends				✓

NOTE.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions include day of week fixed effects, week fixed effects, year fixed effects, hospital fixed effects and hospital specific time trends. Regressions are weighted by number of admissions. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Next, we split our sample to compare results by emergency status to study the hypothesis in Gruber and Kleiner (2012) where stronger strike effects may be observed for non-emergency cases if healthier patients delay treatment or go to other health care facilities as a consequence of the strike. Table 8 presents separate DID estimates for emergency and non-emergency patients in our sample. Indeed, we see that the decrease in the number of admissions is considerably greater in magnitude for non-emergency cases (15.9%) compared to emergency cases (7.1%) suggesting a relative shift to more urgent cases. In addition, effects on mortality rates are stronger (relative to baseline means) and only significant for non-emergency admissions, and the impact on length of stay is also greater in magnitude for non-emergencies. Overall, the results in Table 8 suggest that healthier patients were less likely to receive care at striking hospitals compared to patients with poorer health status and more acute health conditions.

TABLE 8. Difference-in-differences estimates: Heterogeneity by emergency status

	EMERGENCY			NON-EMERGENCY		
	Mean	(1)	(2)	Mean	(1)	(2)
Cases	143.06	-18.516***	-16.631***	164.18	-21.976***	-19.793***
		(4.197)	(2.294)		(4.640)	(2.700)
Log Cases	3.48	-0.069***	-0.071***	4.57	-0.159***	-0.155***
		(0.016)	(0.010)		(0.023)	(0.020)
Mortality Rate	3.37	0.133	0.143	1.27	0.207***	0.207***
		(0.102)	(0.098)		(0.044)	(0.043)
10-day Mortality Rate	2.21	0.071	0.072	0.55	0.080***	0.074***
		(0.080)	(0.079)		(0.026)	(0.025)
Number of Deaths	1.19	-0.064	-0.055	1.30	-0.029	-0.015
		(0.049)	(0.038)		(0.046)	(0.033)
Length of Stay	9.31	0.241**	0.232**	8.00	0.368***	0.358***
		(0.107)	(0.092)		(0.061)	(0.057)
Surgery Rate	24.10	0.047	-1.256	40.31	0.074	-1.798*
		(1.220)	(0.786)		(1.591)	(1.063)
Year FE		✓	✓		✓	✓
Hospital FE		✓	✓		✓	✓
Hospital Specific Trends			✓			✓

NOTE.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions include day of week fixed effects, week fixed effects, year fixed effects, hospital fixed effects and hospital specific time trends. Regressions are weighted by number of admissions. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

To study spillover effects to other hospitals, we estimate Equation (3) for a set of outcomes. Table 9 reports the estimated β_1 parameter, interpreted as the average percentage change in the outcome variable for each percentage point of a hospital's pre-strike patient population that were treated in a striking hospital. The results indicate that the hospitals surrounding the treated university hospitals experience a surge in admissions by about eight percent. However, there is no indication that this increase in admissions worsens patient outcomes as the parameter estimates for mortality rates and length of stay are generally small and non-significant at conventional levels.

TABLE 9. Difference-in-differences estimates: Spillover analysis

	Observations	Mean	(1)	(2)
Log Cases	1,041,569	4.25	0.102***	0.078***
			(0.031)	(0.020)
Mortality Rate	1,041,569	2.57	-0.118	0.039
			(0.124)	(0.094)
10-Day Mortality Rate	1,041,569	1.62	-0.070	-0.014
			(0.090)	(0.081)
Length of Stay	1,041,736	8.19	0.483***	0.105
			(0.143)	(0.109)
Year FE			✓	✓
Hospital FE			✓	✓
Hospital Specific Trends				✓

NOTE.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions include day of week fixed effects, week fixed effects, year fixed effects, hospital fixed effects and hospital specific time trends. Except for log cases regressions are weighted by number of admissions. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

6 Robustness Checks

In this section, we examine a number of robustness checks to probe the stability of our empirical findings. First, to analyse whether seasonality is a confounding factor, we estimate hospital level placebo regressions for our outcomes dating back the strike to the same period of the year but in 2004. Table 10 presents the results showing small and statistically insignificant estimates across the board.

TABLE 10. Difference-in-differences estimates: Placebo regressions for 2004

	Observations	Mean	(1)	(2)
Cases	143,744	158.32	-1.032	-1.311
			(4.081)	(4.265)
Log Cases	143,744	4.90	0.009	0.007
			(0.014)	(0.015)
Mortality Rate	143,744	1.75	0.021	-0.006
			(0.058)	(0.038)
10-day Mortality Rate	143,744	0.96	0.002	-0.015
			(0.037)	(0.025)
Number of Deaths	143,744	2.46	0.002	-0.034
			(0.110)	(0.101)
Length of Stay	143,744	8.17	-0.021	-0.026
			(0.061)	(0.062)
Surgery Rate	143,744	35.95	2.610	2.395
			(1.853)	(1.865)
ACS Rate	143,744	20.30	0.056	0.077
			(0.178)	(0.180)
Emergency Rate	143,156	28.41	-0.121	-0.106
			(0.204)	(0.202)
Urgency Indicator	143,156	0.34	-0.002	-0.001
			(0.002)	(0.002)
Deferability Index	143,549	0.12	-0.000	-0.000
			(0.000)	(0.000)
Year FE			✓	✓
Hospital FE			✓	✓
Hospital Specific Trends				✓

NOTE.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions include day of week fixed effects, week fixed effects, year fixed effects, hospital fixed effects and hospital specific time trends. Except for log cases all regressions are weighted by the number of admissions. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Next, we test the stability of our findings to the choice of control group. To this end, we use all other hospitals in Hamburg and Berlin, who neither participated in the federal nor the municipality strike, as alternative controls.[13] Note that this alternative control group includes all hospitals in the two federal states, not only university hospitals, and is therefore potentially less comparable with respect to a range of factors, including size and scope of provided services. Nevertheless, Table 11 shows that our findings are largely robust to the choice of control group.

[13]Berlin did not participate in the strike, because it was not a member of the employers' association and Hamburg reached a collective agreement before the strike began.

Table 11. Difference-in-differences estimates: Alternative Control Group

	Observations	Mean	(1)	(2)
Cases	267,382	161.83	-21.977***	-19.536***
			(3.718)	(2.354)
Log Cases	267,382	4.72	-0.138***	-0.125***
			(0.016)	(0.013)
Mortality Rate	267,382	2.07	0.270***	0.199***
			(0.045)	(0.041)
10-day Mortality Rate	267,382	1.19	0.135***	0.095***
			(0.036)	(0.034)
Number of Deaths	267,382	3.07	0.006	-0.035
			(0.090)	(0.055)
Length of Stay	267,382	8.28	0.392***	0.261***
			(0.064)	(0.053)
Surgery Rate	267,382	37.35	-1.765	-3.004**
			(1.607)	(1.170)
ACS Rate	84,986	19.25	-0.483***	-0.462***
			(0.148)	(0.150)
Emergency Rate	84,648	27.65	1.494***	1.484***
			(0.350)	(0.353)
Urgency Indicator	84,648	0.33	0.012***	0.012***
			(0.002)	(0.003)
Deferability Index	84,865	0.12	0.003***	0.003***
			(0.001)	(0.001)
Year FE			✓	✓
Hospital FE			✓	✓
Hospital Specific Trends				✓

Note.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions include day of week fixed effects, week fixed effects, year fixed effects, hospital fixed effects and hospital specific time trends. Except for log cases all regressions are weighted by the number of admissions. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

We also test the robustness of our main findings by specifying an alternative treatment indicator, as the one we use in our main analysis does not capture patients that were already hospitalised when the strike started. The alternative treatment indicator is defined as the share of patients on a specific day that were affected by the strike at any time during their hospital stay. The reported estimates from Table 12 show that our results are robust to using this alternative definition of treatment group.

Table 12. Difference-in-differences estimates: Alternative Treatment Indicator

	Observations	Mean	(1)	(2)
Cases	143,744	158.32	-20.065***	-17.453***
			4.984	2.645
Log Cases	143,744	4.90	-0.107***	-0.103***
			0.019	0.014
Mortality Rate	143,744	1.75	0.128***	0.134***
			0.044	0.042
10-day Mortality Rate	143,744	0.96	0.055	0.054
			0.036	0.035
Number of Death	143,744	2.46	-0.097	-0.064
			0.095	0.062
Length of Stay	143,744	8.17	0.226***	0.217***
			0.058	0.054
Surgery Rate	143,744	35.95	-0.473	-1.856*
			1.510	0.945
ACS Rate	143,744	20.30	0.009	-0.134
			0.233	0.193
Emergency Rate	143,156	28.41	1.364***	1.312***
			0.334	0.329
Urgency Indicator	143,156	0.34	0.010***	0.010***
			0.002	0.002
Deferability Index	143,549	0.12	0.003***	0.003***
			0.001	0.001
Year FE			✓	✓
Hospital FE			✓	✓
Hospital Specific Trends				✓

Note.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions include day of week fixed effects, week fixed effects, year fixed effects, hospital fixed effects and hospital specific time trends. Except for log cases all regressions are weighted by the number of admissions. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Finally, we also test for linear pre-trends by estimating the following regression model on the pre-intervention sample:

$$y = \beta(trend \times treated) + \gamma treated + \delta trend + \epsilon, \tag{4}$$

where *trend* is a non-parametric time trend estimated for each $year \times month$ cell in the sample, and *treated* is a treatment indicator equal to one for striking hospitals and zero for non-striking hospitals. If the estimate β is significantly different from 0, we conclude the treated and control hospitals had significantly different trends leading up to the strike; thus rejecting the common trend assumption required for consistency of the DID estimator. The reported $\hat{\beta}$ in Table 13 indicate that the common trend assumption cannot be rejected at any conventional levels of statistical significance.[14] In addition, it should also be noted that our main specification includes hospital-specific

[14]The recent literature recommends caution regarding pre-trend tests as their power may be low (Roth et al., 2022). Nevertheless, the test does provide additional evidence concerning the plausibility of the common time trend assumption.

time trends which makes differential pre-trends less likely to be a problem.

Table 13. Pre-trend Test

	(1) Log Cases	(2) Mortality Rate	(3) 10-Day Mortality Rate	(4) Cases
Trend × Treated	-0.0000262 (-0.81)	-0.0000480 (-1.07)	-0.0000434 (-1.65)	-0.00972 (-0.92)

Note.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Except for log cases all regressions are weighted by the number of admissions. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

7 Conclusion

Strikes in essential industries, such as transport, healthcare and protection, have become more common in recent times due to budget pressures and worker dissatisfaction with rationalisation and retrenchment policies in the public sector. In the healthcare sector, longer working hours, less job security and higher workloads has increased the intensity of industrial action, which causes social and individual costs in the form of lower access and quality of care. Given the high-stake nature of the occupation, it is an important task to study the impact of strikes in the healthcare sector; in particular among specialists including physicians.

We explore the short-term impacts of a nationwide physician strike in German university hospitals in 2006 on patient mortality risk. We compare changes in outcomes over time in hospitals that were subject to striking physicians to other hospitals in a difference-differences empirical design. Our results show that hospital admissions in striking hospitals dropped by an estimated 12 percent during the strike period compared to non-striking hospitals. Moreover, we find that in-hospital mortality rates increased by nine percent over the same period. However, the effect on mortality is attenuated after adjusting for patient case-mix, suggesting that the effect is also due to a change in the patient composition during the strike period. This result is indicative of a triaging strategy where more urgent or frailer cases gain precedence over more healthy patients. Finally, we find no indications of neither temporal (post-strike period) or spatial (other nearby hospitals) spillover admission effects from which we conclude that the drop in admissions in striking hospitals were most likely permanent.

Our results contrast sharply with the related strike literature in Economics. For example, Gruber and Kleiner (2012), find a 26 percent decrease in admissions during the nurse strikes and an increase of 18 percent in in-hospital mortality. We interpret this glaring difference as an indication that strikes in healthcare (and elsewhere) can have very different impacts on organisational efficacy depending on the extent and range of services affected by the disruption. One drawback of the analysis in this paper is that death is a rare and serious outcome (Cunningham et al., 2008). It is still possible, and even likely, that other quality or patient care indicators, such as readmission rates or patient satisfaction, were affected. Furthermore, there might be other economic consequences of strikes, such as financial losses for the hospitals or an increased workload for non-striking physicians that we do not explore further in this paper. Establishing the total impact of strikes on a fuller set of social, economic, and health indicators should therefore constitute an important avenue for further research.

References

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." *Jama*, 288(16): 1987–1993.

Austin, Peter C. 2009. "Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples." *Statistics in medicine*, 28(25): 3083–3107.

Avdic, Daniel. 2016. "Improving efficiency or impairing access? Health care consolidation and quality of care: Evidence from emergency hospital closures in Sweden." *Journal of health economics*, 48: 44–60.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." *The Review of Economic Studies*, 81(2): 608–650.

Blümel, Miriam, and Reinhard Busse. 2015. "The German health care system, 2015." *International profiles of health care systems*, 69–76.

Bundesministerium für Gesundheit. 2019. "Daten des Gesundheitswesens 2019." *Bundesministerium für Gesundheit.* https://www.bundesgesundheitsministerium.de/themen/krankenversicherung/zahlen-und-fakten-zur-krankenversicherung/kennzahlen-daten-bekanntmachungen.html Retrieved on April 21, 2020.

Costa, Eduardo. 2019. "License to Kill? The Impact of Hospital Strikes." *The Impact of Hospital Strikes (May 20, 2019).*

Cunningham, Solveig Argeseanu, Kristina Mitchell, KM Venkat Narayan, and Salim Yusuf. 2008. "Doctors' strikes and mortality: a review." *Social Science & Medicine*, 67(11): 1784–1788.

de Cruppé, Werner, and Max Geraedts. 2017. "Hospital choice in Germany from the patient's perspective: a cross-sectional study." *BMC health services research*, 17(1): 720.

Dribbusch, Heiner. 2018. "Das Einfache, das so schwer zu zählen ist: Probleme der Streikstatistik in der Bundesrepublik Deutschland." *Industrielle Beziehungen*, 25(3).

FDZ der Statistischen Ämter des Bundes und der Länder. 2000-2008*a*. "Krankenhausstatistik der Jahre 2000-2008."

FDZ der Statistischen Ämter des Bundes und der Länder. 2000-2008*b*. "Todesursachenstatistik der Jahre 2000-2008."

Federal Statistical Office. 2018. "Grunddaten der Krankenhäuser 2017, Fachserie 12 Reihe 6.1.1." *Statistisches Bundesamt (Destatis).*

Friedman, Willa, and Anthony Keats. 2019. "Disruptions to health care quality and early child health outcomes: Evidence from health worker strikes in Kenya." mimeo.

Goepfert, Andreas, and Claudia B Conrad. 2013. *Unternehmen Krankenhaus.* Georg Thieme Verlag.

Greef, Samuel. 2012. *Die Transformation des Marburger Bundes: Vom Berufsverband zur Berufsgewerkschaft.* Springer-Verlag.

Gruber, Jonathan, and Samuel A Kleiner. 2012. "Do strikes kill? Evidence from New York state." *American Economic Journal: Economic Policy*, 4(1): 127–57.

Handouyahia, Andy, Tony Haddad, and Frank Eaton. 2013. "Kernel matching versus inverse probability weighting: a comparative study." *International Journal of Mathematical and Computational Sciences*, 7(8): 1218–1233.

Hirani, Jonas, Hans Henrik Sievertsen, and Miriam Wüst. 2019. "Beyond Treatment Exposure: The Timing of Early Interventions and Children's Health." mimeo.

Hirano, Keisuke, and Guido W Imbens. 2001. "Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization." *Health Services and Outcomes research methodology*, 2(3-4): 259–278.

International Labour Office. 2006. "Freedom of Association and others. Digest of Decisions and Principles of the Freedom of Association Committee of the Governing Body of the ILO." *International Labour Office, Geneva.*

KHStatV. 1990. "Krankenhausstatistik-Verordnung vom 10. April 1990 (BGBl. I S. 730), die zuletzt durch Artikel 1 der Verordnung vom 10. Juli 2017 (BGBl. I S. 2300) geändert worden ist." *Bundesamt für Justiz.*

Krämer, Jonas, Jonas Schreyögg, and Reinhard Busse. 2019. "Classification of hospital admissions into emergency and elective care: a machine learning approach." *Health care management science*, 22(1): 85–105.

Kronborg, Hanne, Hans Henrik Sievertsen, and Miriam Wüst. 2016. "Care around birth, infant and mother health and maternal health investments–Evidence from a nurse strike." *Social Science & Medicine*, 150: 201–211.

Lin, Haizhen. 2014. "Revisiting the relationship between nurse staffing and quality of care in nursing homes: An instrumental variables approach." *Journal of health economics*, 37: 13–24.

Marburger Bund Bundesverband. 2012. "Marburger Bund Bundesverband Chronologie 2005-2007." Marburger Bund Bundesverband.

Martens, Helmut. 2008. "Primäre Arbeitspolitik und Gewerkschaften im Gesundheitswesen: Der Ärztestreik 2006 als Beispiel primärer Arbeitspolitik in Zeiten tiefgreifender gesellschaftlicher Umbrüche." Arbeitspapier 143, Hans-Böckler Stiftung.

Metcalfe, David, Ritam Chowdhury, and Ali Salim. 2015. "What are the consequences when doctors strike." *BMJ*, 351: h6231.

Piérard, Emmanuelle. 2014. "The effect of physician supply on health status: Canadian evidence." *Health Policy*, 118(1): 56–65.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. 2022. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *arXiv preprint arXiv:2201.01194*.

Stoye, George, and Max Warner. 2023. "The effects of doctor strikes on patient outcomes: Evidence from the English NHS." *Journal of Economic Behavior & Organization*, 212: 689–707.

Sundmacher, Leonie, Diana Fischbach, Wiebke Schuettig, Christoph Naumann, Uta Augustin, and Cristina Faisst. 2015. "Which hospitalisations are ambulatory care-sensitive, to what degree, and how could the rates be reduced? Results of a group consensus study in Germany." *Health Policy*, 119(11): 1415–1423.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

Appendix A Additional tables and figures

Table A.1. Physicians and Health Care Workers in Hospitals

	All Hospitals			University Hospitals		
Year	Total	Physicians	Other Medical	Total	Physicians	Other Medical
2000	1,100,471	122,062	978,409	172,867	24,398	148,469
2001	1,101,356	123,819	977,537	173,114	24,758	148,356
2002	1,112,421	126,047	986,374	174,850	25,084	149,766
2003	1,096,420	128,853	967,567	173,091	25,154	147,937
2004	1,071,846	129,817	942,029	168,980	25,171	143,809
2005	1,063,154	131,115	932,039	170,160	25,435	144,725
2006	1,064,377	133,649	930,728	171,895	25,781	146,114
2007	1,067,287	136,267	931,020	172,782	26,241	146,541
2008	1,078,212	139,294	938,918	173,182	26,488	146,694
2009	1,096,520	143,967	952,553	177,535	27,632	149,903
2010	1,112,959	148,696	964,263	181,954	28,443	153,511

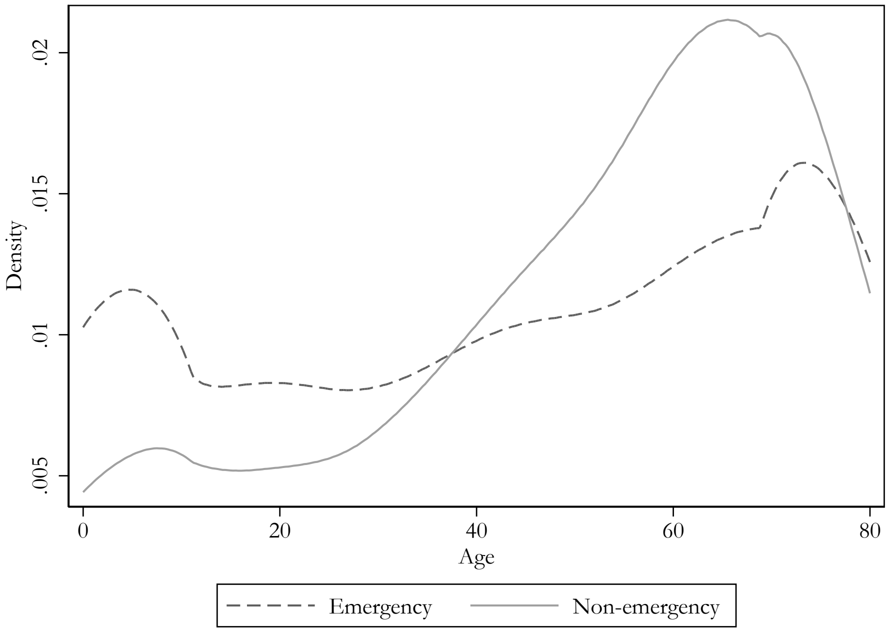
Note.— Information System of the Federal Health Monitoring.

Table A.2. Unweighted mortality regressions

	Observations	Mean	Estimate
Mortality Rate	143,744	2.15	0.117*
			(0.065)
10 Day Mortality Rate	143,744	1.23	0.084
			(0.050)
Year FE			✓
Hospital FE			✓
Hospital Specific Trends			✓

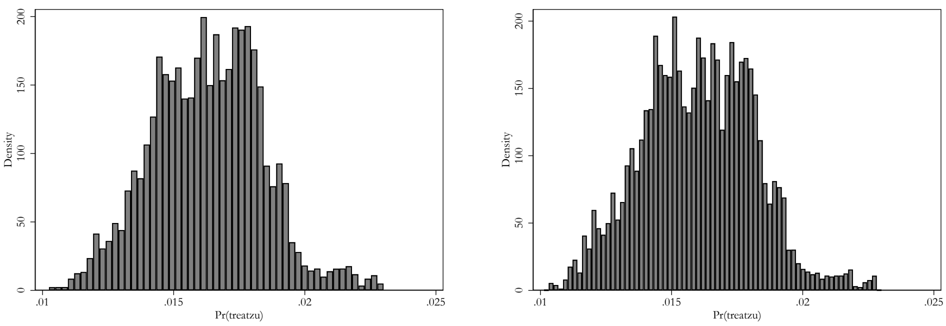
Note.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions include day of week fixed effects, week fixed effects, year fixed effects, hospital fixed effects and hospital specific time trends. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

FIGURE A.1. Age Distribution by Emergency Status



NOTE.— The figure shows the age distribution by emergency status. Based on individual level hospital data. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

FIGURE A.2. Propensity Scores Treatment and Control

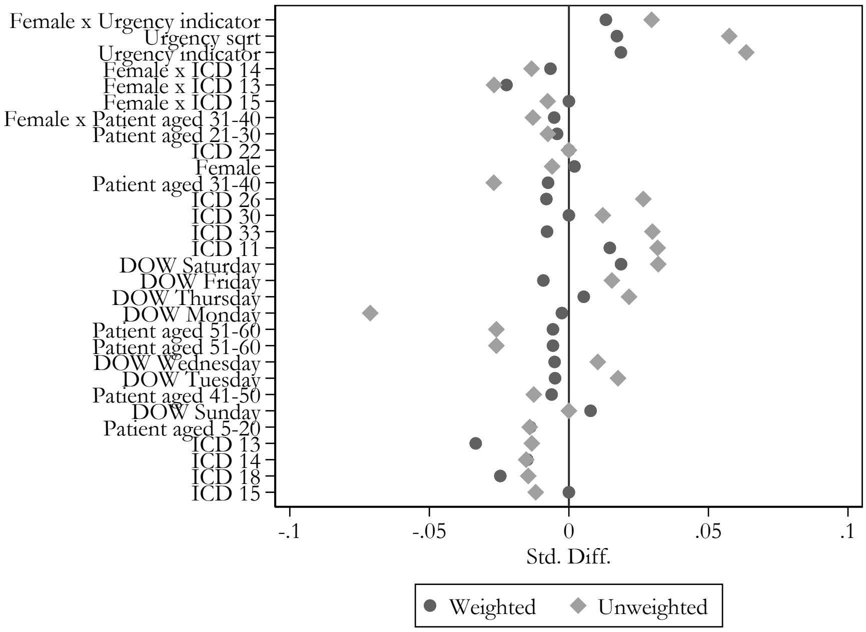


(a) Treatment

(b) Control

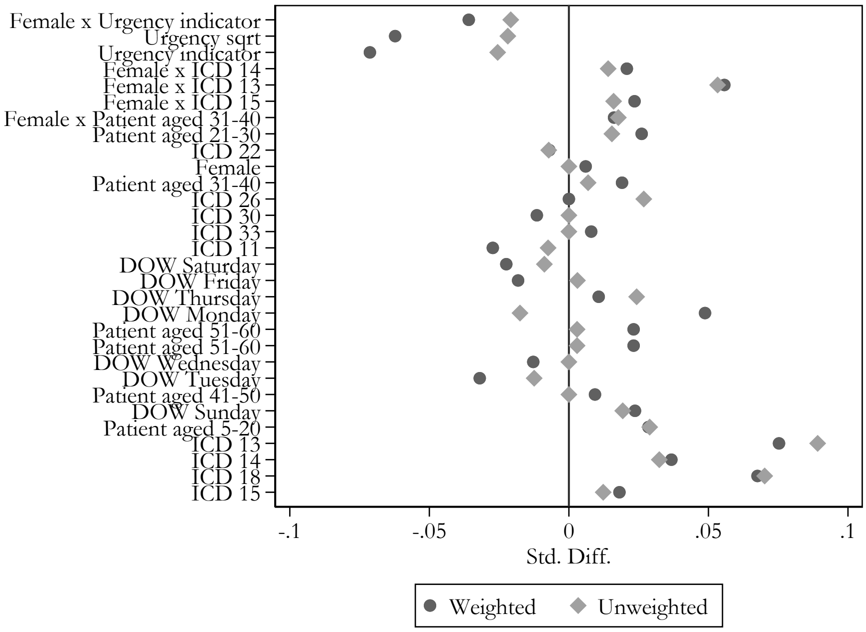
NOTE.— The figures show the estimated propensity scores, i.e. the probability of visiting a strike-hit hospital during the strike period conditional on the selected covariates, by applying a logistic regression. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Figure A.3. Standardised Differences TT vs. TC



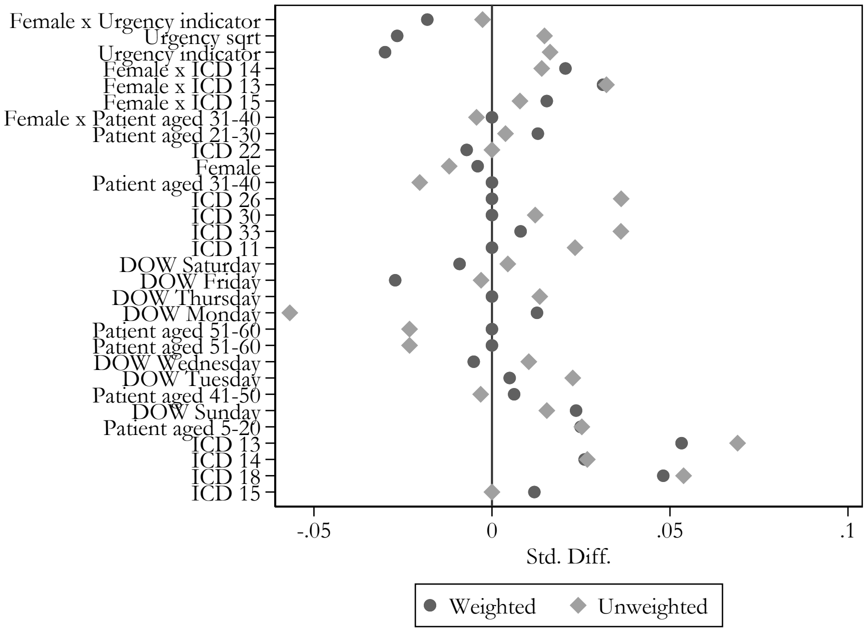
Note.— The figure shows the standardised difference before and after weighting between treatment hospitals and treatment period vs. treatment hospitals and control period. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Figure A.4. Standardised Differences TT vs. CT



Note.—The figure shows the standardised difference before and after weighting between treatment hospitals and treatment period vs. control hospitals and treatment period. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Figure A.5. Standardised Differences TT vs. CC



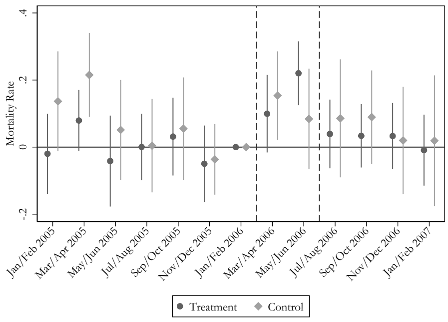
Note.— The figure shows the standardised difference before and after weighting between treatment hospitals and treatment period vs. control hospitals and control period. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Table A.3. Admissions by cause

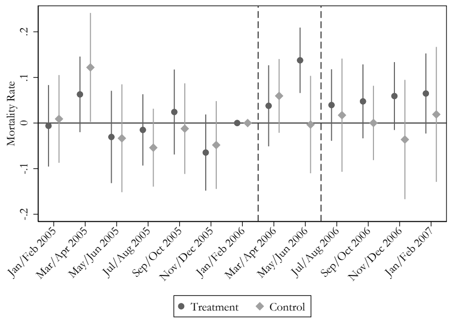
	Observations	Mean	(1)	(2)
Cardio	143,744	20.75	-2.465***	-2.323***
			(0.816)	(0.464)
Respiratory	143,744	7.17	-1.093***	-0.914***
			(0.320)	(0.221)
Infectious	143,744	3.40	-0.287**	-0.183**
			(0.139)	(0.082)
Metabolic	143,744	4.68	-0.636***	-0.544***
			(0.125)	(0.089)
Neoplastic	143,744	33.54	-3.622***	-2.844***
			(0.846)	(0.511)
Birth	143,744	7.96	-0.026	-0.036
			(0.146)	(0.139)
Medical Complication	143,744	0.62	-0.114**	-0.111**
			(0.056)	(0.045)
Year FE			✓	✓
Hospital FE			✓	✓
Hospital Specific Trends				✓

Note.— Robust standard errors clustered at the hospital level are reported in parenthesis. Significance levels: $*$ 0.10 $**$ 0.05 $***$ 0.01. Regressions include day of week fixed effects, week fixed effects, year fixed effects, hospital fixed effects and hospital specific time trends. Except for log cases regressions are weighted by number of admissions. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Figure A.6. Descriptives Mortality Rate



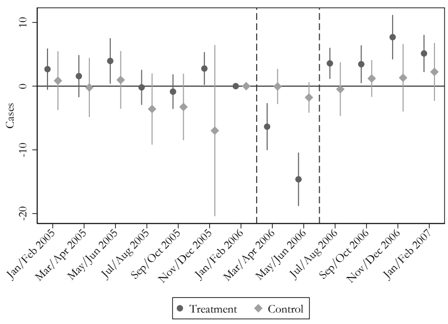
(a) In-hospital Mortality Rate



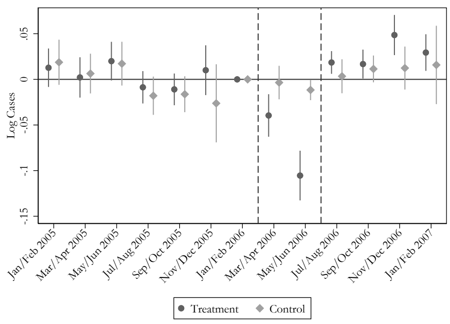
(b) 10-Day Mortality Rate

Note.— The figures show means for treatment and control hospitals. The vertical lines represent the beginning and end of the strike. 95% confidence intervals included. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Figure A.7. Descriptives Cases



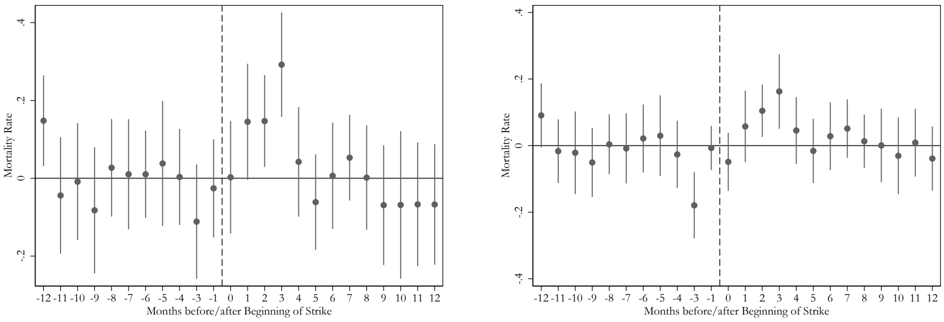
(a) Cases



(b) Log Cases

Note.— The figures show means for treatment and control hospitals. The vertical lines represent the beginning and end of the strike. 95% confidence intervals included. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

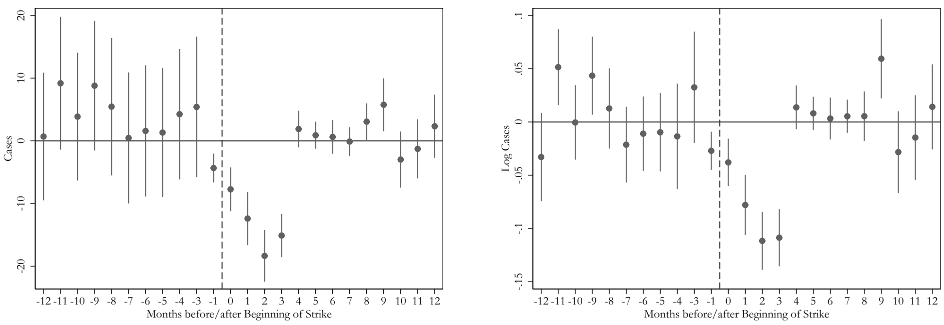
FIGURE A.8. Alternative Estimator - Mortality Rate



(a) In-hospital Mortality Rate (b) 10-Day Mortality Rate

NOTE.— The figures show event study estimates for mortality rate and 10-day mortality rate, controlling for day of week fixed effects and hospital fixed effects using the estimator of Sun and Abraham (2021). Reference period $t-2$. The vertical lines represent the beginning and end of the strike. 95% confidence intervals included. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

FIGURE A.9. Alternative Estimator - Cases



(a) Cases (b) Log Cases

NOTE.— The figures show event study estimates for cases and log cases, controlling for day of week fixed effects and hospital fixed effects using the estimator of Sun and Abraham (2021). Reference period $t-2$. The vertical lines represent the beginning and end of the strike. 95% confidence intervals included. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Appendix B Variable Definitions Control and Outcome

Treatment Variable: Dummy variable taking on the value 1 for admissions to a striking hospital during the strike period.

Female: Dummy variable taking on the value 1 for females.

Age: Age in years, continuous variable.

Married: Dummy variable taking on the value 1 for married individuals.

Cases: Number of admissions per day per hospital.

Log Cases: Logarithmised number of admissions per day per hospital.

Length of Stay: Length of hospital stay in days.

Surgery Rate: Surgeries per 100 admissions.

Mortality Rate: Deaths per 100 admissions.

10 Day Mortality Rate: Deaths within 10 days after admission per 100 admissions.

ACS Rate: Ambulatory care sensitive admissions based on (Sundmacher et al., 2015) per 100 admissions.

Urgency Indicator: Urgency indicator based on Krämer, Schreyögg and Busse (2019), continuous variable taking on values between 0 and 1.

Emergency Rate: Admissions indicated as emergency based on Krämer, Schreyögg and Busse (2019) per 100 admissions. Derived from urgency indicator. Admission is defined as an emergency admission if urgency indicator is above 0.5.

Log Deaths: Logarithmised number of deaths per day per district.

Appendix C Alternative Covariate Selection Approach

To explore changes in patient composition, we use propensity score weighting in combination with a covariate selection approach as suggested by Hirano and Imbens (2001). We first conduct a variable selection on potential exogenous pre-treatment patient and admission characteristics, such as ICD code dummies for the major diagnosis, urgency of the admission, urgency squared, age group dummies, gender or admission day of week dummies. The covariate selection is executed on the individual level hospital data and all variables with a t-statistic of at least 3.2 that might predict selection into treatment are chosen. Although some of the covariates might be outcome variables themselves, they are determined prior to treatment and exogenous on the patient level. This results in unbiased estimates of the treatment effect. In a second iteration, all previously selected variables are interacted with a female dummy and the procedure is repeated. In total, 30 relevant covariates are identified. The complete list of selected variables is available in Figure A.3 of of Appendix A.

After having identified the relevant covariates, we estimate the propensity score, i.e. the probability of visiting a strike-hit hospital during the strike period conditional on the selected covariates, by applying a logistic regression. The distributions of the estimated propensity scores for the treatment and control group are presented in Figure A.2 of Appendix A. Here, the treatment group refers to patients visiting a strike-hit hospital during the strike period, while the control group refers to patients who do not[15]. In order to conduct the analysis on the hospital level, we aggregate the individual-level data set – including the selected variables and the propensity scores – on day of admission and hospital level. We then calculate propensity score weights applying inverse probability weighting (IPW) for the ATT. In comparison to kernel weighting, IPW is computationally less intensive, as it does not require a bandwidth choice and it has a smaller variance (Handouyahia, Haddad and Eaton, 2013). Following Hirano and Imbens (2001), the weights are defined as:

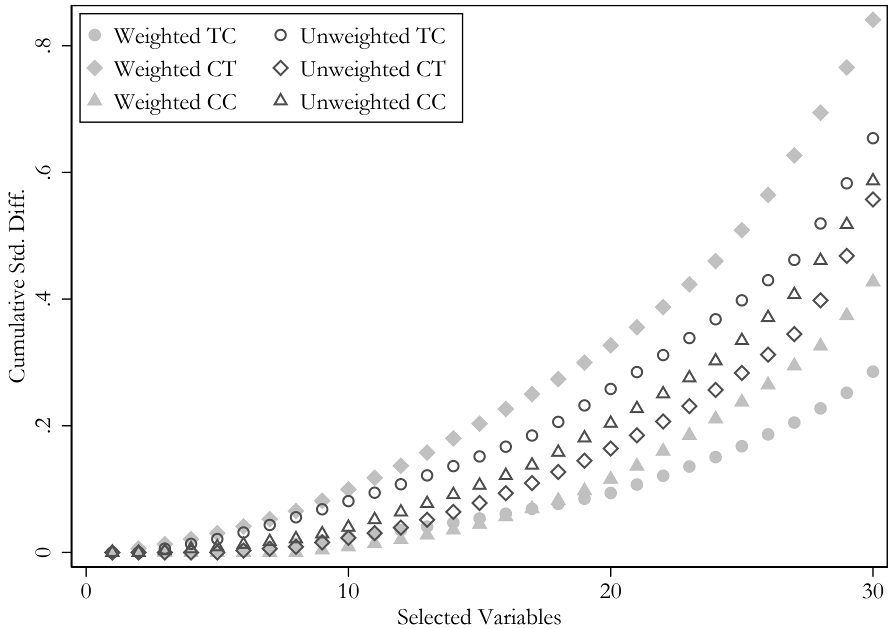
$$\omega(t,x) = \quad t + (1-t) \cdot \frac{\hat{p}(x)}{1-\hat{p}(x)} \tag{C.1}$$

where $\hat{p}(x)$ is the estimated propensity score and t is a treatment indicator which equals 1 for treated and 0 for control units. Thus, the weights are 1 for treated observations and $\frac{\hat{p}(x)}{1-\hat{p}(x)}$ for control observations.

To show that our weighted sample is sufficiently balanced with respect to the included covariates, Figure C.1 presents the accumulated standardised difference between treatment hospitals during the strike period (treatment group) and all possible combinations of control groups – before and after weighting – over all covariates. Except for the comparison between treatment hospitals during the treatment period and the group of control hospitals in the treatment period (CT), the standardised difference is always lower after weighting than before weighting. Detailed results for each of the 30 covariates are available in Figures A.3-A.5 of Appendix A. They confirm that covariates are more balanced after weighting than before weighting. Furthermore, the standardised difference is always below 0.25 after weighting, which indicates covariate balance across treatment and control groups (Austin, 2009).

[15] In particular, the control group consists of three different groups: (1) treated hospitals during the non-strike period; (2) control hospitals during the strike period and (3) control hospitals during the non-strike period.

FIGURE C.1. Cumulative Standardised Difference



NOTE.— The figure shows the accumulated standardised difference before and after weighting. *TC* indicates standardised difference between treatment hospital and treatment period vs. treatment hospital and control period, *CT* indicates standardised difference between treatment hospital and treatment period vs. control hospital and treatment period and *CC* indicates standardised difference between treatment hospital and treatment period vs. control hospital and control period. Source: FDZ der Statistischen Ämter des Bundes und der Länder (2000-2008*a*), own calculations.

Finally, we use a doubly robust estimator and estimate the following regression model using the above estimated propensity score weights in order to adjust our estimates for patient selection:

$$y_{it} = \beta_0 + \beta_1 S_{it} + \mu_i + \nu_t + H_i \cdot t + \gamma X_{it} + \epsilon_{it} \tag{C.2}$$

where all coefficients and variables are defined as in Section 4.1 and X_{it} is the mean of each selected covariate in a hospital i on day t.