

Physician Working Time Restrictions and Patient Mortality

Gonalo Lima Kathrine A. Lorentzen *

July 22, 2026

Abstract

We study how shortening physicians' work schedules affects healthcare provision and patient mortality. We exploit a 1982 reform in Denmark that strongly restricted hospital doctors' overtime, weekend, and nighttime shifts. Consistent with substitution from hours to workers, labor demand for physicians increased by 35% one year after implementation. Using a dose-response difference-in-differences strategy, we leverage variation in the intensity of the policy across hospital wards with different pre-reform reliance on overtime to document three main results. First, wards where labor demand changed the most substantially decreased in-hospital mortality rates, especially among emergency admissions. Second, while more physicians led to more planned admissions, better mortality outcomes are not fully explained by shifts in patient composition. Third, persistent post-discharge mortality declines and no change in readmissions or duration of stays suggest that gains were not at the expense of higher patient turnover. Our findings are consistent with overtime regulations correcting inefficiently long work hours in a frictional labor market.

JEL Classification: J22, J81, I11, I12, I18. **Keywords:** overtime; job sharing; physician productivity; healthcare policy; work time reduction.

*Lima: goncalo.lima@econ.ku.dk, Postdoc at the Department of Economics in the University of Copenhagen. Lorentzen: kalo@econ.ku.dk, Postdoc at Nova School of Business and Economics and the University of Copenhagen. We would like to thank Torben Heien Nielsen, Andrea Ichino, Thomas Høgholm Jørgensen, Sule Alan, Franziska Valder, Frederik Plesner Lyngse, Davide Dragone, Jonathan Zhang, as well as feedback from seminar attendees at the EUI, University of Bologna, Nova SBE and University of Copenhagen, for helpful discussions and feedback. Lima gratefully acknowledges funding from an EUI individual PhD grant. Lorentzen gratefully acknowledges funding from the Danish Independent Research Fund (Grant No. 9061-00035B) and the European Union (EQUALNovaERA). Both Lima and Lorentzen gratefully acknowledge funding from an EUI Early-Stage Research grant, as well as all the help and feedback from staff and researchers at CEBI, University of Copenhagen. The activities of CEBI are financed by the Danish National Research Foundation (Grant No. DNRF134). All views and errors are ours.

1. Introduction

How firms allocate total labor hours across workers is a fundamental organizational decision. The choice between getting the marginal hour from an additional hire or from an incumbent worker lies at the core of the firm’s labor demand problem (Hunt, 1999; Nadiri and Rosen, 1969; Oi, 1962). Depending on the marginal productivity of labor, fixed employment costs, and worker heterogeneity, it may be optimal to hire fewer workers on longer shifts rather than more workers on shorter ones. Yet, inefficiently long work hours may arise when firms’ and workers’ decisions are distorted—among other reasons—by search frictions (Germain, 2025), adverse selection (Landers et al., 1996; Akerlof, 1976), or incomplete internalization of fatigue-related productivity losses (Pencavel, 2015).

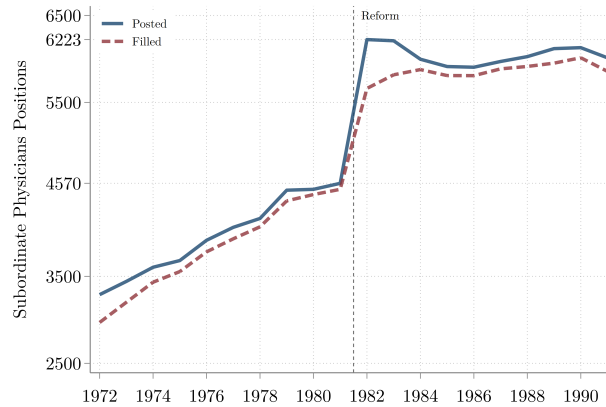
Hospitals provide an insightful setting for studying the allocation of hours across workers. Hospital physicians often work well beyond the general labor standards agreed for other economic sectors (Temple, 2014). Uncertain patient inflows, human capital specificity, and high coordination costs often justify physicians working long shifts. At the same time, relevant frictions may lead to socially inefficient outcomes. Hospitals’ monopsony power—especially over residents or junior doctors bound by training requirements to become specialists—may induce physicians to accept supplying more hours than they would want. As effort and fatigue are often hard to monitor, hospitals may also underestimate the drop in marginal productivity in the end of long shifts (Chan, 2018). In light of concerns with work-life balance and patient welfare, collective bargaining agreements across several countries have been setting increasingly stricter standards on physicians’ overtime, weekend, and nighttime hours (Maoz Breuer et al., 2023).¹ Yet, despite the prevalence of such regulations, there is limited if any causal evidence on the net effect of these types of policies on healthcare provision and patient outcomes.

In this paper, we address this gap by studying the labor market and healthcare effects of constraining physicians’ work hours. We exploit a 1982 reform in Denmark that strongly restricted early-career hospital doctors’ schedules, providing a natural experiment to examine the consequences of regulating work intensity in a high-skill, shift-based sector. The new collective bargaining agreement established three key provisions to the work of subordinate physicians hired by public hospitals: (i) a 40-hour workweek; (ii) at least half of those hours scheduled as daytime duty on weekdays; and (iii) restrictions on overtime, only to be used when hospitals could not meet staffing needs through new hires. These provisions effectively removed overtime as an adjustment margin and forced hospitals to substitute from hours-per-worker toward more workers. Consistent with this, we document a 35% surge in posted positions after the reform, with the number of employed subordinate physicians in Danish

¹ Subordinate doctors’ workweeks are limited to 40–50 hours across most European countries, 60 hours in Canada, and 80 hours in the United States (Maoz Breuer et al., 2023).

public hospitals increasing by 25% between 1981 and 1982 (Figure 1).

Figure 1. POSTED POSITIONS FOR SUBORDINATE PHYSICIANS IN DENMARK (1972-1991)



Notes. The figure presents the total number of posted and filled positions for subordinate physicians in public hospitals in Denmark, between 1972 and 1991. Subordinate physicians include both junior and senior medical residents, who were training for specialty recognition or have recently become specialists. The data comes from digitized statistical yearbooks produced by the Danish Health Authority. Appendix Figure G.6 presents analogous plots for the total number of physicians, and separately by physician's career stage.

The effect of this labor demand shock on patient outcomes is theoretically ambiguous. In a frictionless environment with optimal pre-reform labor allocation, overtime restrictions cannot possibly increase output. If a better labor allocation would have been possible without this constraint, hospitals would have already achieved it. However, several frictions may generate inefficiencies in the status quo, and lead to excessively high hours-per-worker. In these cases, restricting overtime could increase effective labor input and improve patient outcomes. Conversely, employment expansion could bring worse physicians, weaken continuity of care, or increase coordination costs across shifts, harming hospital performance. Which of these mechanisms dominates is ultimately an empirical question.

We build a novel dataset with all vacancies for physicians in public hospital wards in Denmark by digitizing a collection of statistical yearbooks published by the Danish Health Authority. This information allows us to document the first stage impact of overtime restrictions in the physician labor market, and characterize the extensive margin of labor demand across all hospital wards in the country. To measure the downstream effects on patient health, we access restricted-use administrative records on the universe of hospitalizations around the time of the reform. We complement this information with the death registries of the Danish population to measure mortality rates post-discharge. As our main unit of analysis, we focus on the hospital ward, as a hospital-specialty pair.

To evaluate the effects of the reform, we use a dose-response difference-in-differences design exploiting heterogeneity in pre-reform reliance on overtime across hospital wards. Although all wards were simultaneously affected by the reform, their intensity of exposure to the new regulations varied. Since the reform removed overtime as an adjustment margin, wards that relied more heavily on

overtime substituted relatively more hours for more workers to keep up with demand. While overtime itself is not consistently measured across workers in our setting, we exploit detailed information on predictors of labor demand change across reform periods. In particular, we use the share of unplanned (or acute) admissions—typically emergency cases whose treatment cannot be deferred—as a proxy for pre-reform overtime need. We exploit the fact that wards with higher concentrations of these acute cases were more likely to rely on resident physicians being on call for longer periods and to schedule more weekend and nighttime shifts before the reform. Using this pre-reform share of unplanned admissions as a continuous dose of treatment intensity, we compare outcomes across wards with different exposure levels over time in an event-study framework.

To validate the relevance of our exposure measure we show that wards in specialties with higher concentrations of unplanned admissions hire relatively more physicians post-reform. One standard deviation increase in the pre-reform share of unplanned admissions (23 p.p.) leads to an 18 percent rise in posted positions and 14 percent more subordinate-physician hires within three years. No differential pre-trends lend credibility to the interpretation that labor demand differences across specialties would have evolved similarly in the absence of this policy.

We find that the reform reduced in-hospital death rates. A one standard deviation increase in the pre-reform share of unplanned admissions (26 percentage points) reduces in-hospital deaths by 2.45 per thousand admissions (95% C.I.: -3.67; -1.24). Relative to a pre-reform mean of 34 deaths per thousand admitted patients, this corresponds to a 7 percent decline in yearly ward-level mortality for each standard deviation in pre-reform overtime intensity. Relying on an event-study type of specification and a strong version of parallel trends ([Callaway et al., 2024](#)), we show evidence that the divergence in mortality rate outcomes is likely attributable to the overtime restrictions reform. We show the effect is partially driven by a change in hospital admissions. Wards relatively more exposed to the reform admit more patients, with more of these being planned admissions. Our findings suggest that the reform allowed hospitals to lift healthcare supply constraints, enabling these wards to anticipate the treatment of these patients. Nevertheless, focusing only on the outcomes of unplanned admissions, a one standard deviation increase in overtime needs leads to 5.5 fewer deaths per thousand unplanned admissions (95% C.I.: -8.01,-3.05), or a 9.5% decrease relative to the pre-reform mean. Improvements in mortality rates are not confined to hospital stays, with effects persisting until at least one year after hospital discharge.

To address whether reductions in mortality were achieved through shifts in patient compositions we construct risk-adjusted measures of mortality. For each hospital ward we take the ratio between actual and expected death rates given the age distribution and case-mix of the ward’s admitted patients. Results using these risk-adjusted measures yield qualitatively similar interpretations, showing

that despite the relative increase in planned admissions, death rates declined beyond what would be expected.

To study whether the reform increased healthcare supply through higher patient turnover at the expense of quality of care, we examine its effects on post-discharge readmission rates and average length of stay. We find no detectable evidence consistent with shorter hospital stays or increased patient readmissions, even if in other hospital wards.

The interpretation of our results is robust to a series of changes in our study design. First, we assess the sensitivity to our overtime need proxy, by replicating the analysis with alternative proxies, such as the share of nighttime and weekend admissions. Second, results are not driven by differential attrition of hospital wards that close down or consolidate after the reform. Third, our specifications are robust to controlling for county-specific trends in the outcomes.

Related Literature.— Our paper relates to four main strands of work.

First, it contributes to the empirical literature on the effects of weekly working hours caps on labor market outcomes. Most existing research has focused on testing the impact on overall employment dynamics, with mixed evidence. While some studies of standard working hours reductions find positive effects on employment ([Germain, 2025](#); [Gunnemo and Hansen, 2025](#); [Fishback et al., 2024](#)), others find negative effects through increased hourly labor costs ([Asai et al., 2024](#); [Hunt, 1999](#)) and worker separations ([Crépon and Kramarz, 2002](#)), or no employment effects at all ([Batut et al., 2023](#); [Kawaguchi et al., 2017](#); [Sánchez, 2013](#); [Chemin and Wasmer, 2009](#); [Skuterud, 2007](#)). Our paper diverges from this literature in two important ways. Beyond a change in standard hours, the type of reform we study introduces considerable restrictions on the use of overtime and scheduling of working hours. Moreover, our paper not only examines the overall impacts on employment but also focuses on the effects on output in a sector where quality of provision is crucial to public policy.

Second, it specifically contributes to the emerging literature on time limitations to physicians' workweeks. Extant economic research on this topic has primarily focused on the upstream effects of these caps on the differential labor supply incentives across gender. Shorter workweeks encourage women to make less gender-stereotypical career choices, leading them to select medical specialties with greater reductions in working hours ([Lorentzen, 2024](#); [Wasserman, 2023b](#)). Long working hours have also been found to affect fertility timing and incentivize women with stronger preferences for children to sort into less time intensive specialties ([Wasserman, 2023a](#)). Similar to [Wasserman \(2023b\)](#), we exploit variation in reform intensity to identify our parameters of interest. However, beyond variation across specialties, we exploit variation at the hospital ward level. In addition, unlike previous research in this vein, we examine the downstream effects of workweek caps on the quality of health-

care provision, beyond the direct effects on the labor market.

Third, it contributes to understanding the productivity of healthcare professionals, and medical doctors in particular. Empirical evidence indicates that after long hours of emergency room work, physicians nearing the end of their shifts accept fewer patients, distort patient care, and use hospital resources less efficiently (Chan, 2018). Conversely, shorter hours might reduce the likelihood of repeated collaboration among practitioners, with potentially detrimental impacts on service quality. Indeed, evidence shows that past collaboration among team members creates team-specific human capital, improving patient outcomes (Chen, 2021). In our setting, fewer working hours also induced an inflow of new physicians, potentially filling additional vacancies with marginally less productive workers. Previous evidence highlights significant differences in physicians' value-added (Ginja et al., 2025), productivity (Chan and Chen, 2023; Doyle et al., 2010), and diagnostic skills (Chan et al., 2022; Currie and MacLeod, 2020; Currie and Macleod, 2017). Our paper most closely relates to recent empirical evidence on the effects of healthcare labor scarcity on mortality. Dodini et al. (2025) exploit variation in physician emigration from Sweden to Norway and find substantial mortality increases in Sweden, operating through higher hospitalization rates and premature discharges. Similarly, Schlenker (2025) finds increases in mortality from a negative labor supply shock of nurses in Germany. In contrast with these settings, we exploit a positive shock to labor demand for workers working fewer hours. To our knowledge, this is the first study to investigate the overall national-level effect of a reform on physicians' working hours on patient health outcomes.²

Finally, it speaks to the empirical literature on the efficacy of policy in improving public hospital performance. Greater competition in provision (Gaynor et al., 2013) and the professionalization of hospital management (Muñoz and Otero, 2025) have been shown to reduce hospital death rates. On the other hand, lack of competitive wages in the labor market for nurses leads to worse patient outcomes (Propper and Van Reenen, 2010). We provide novel evidence on how overtime restrictions in a frictional healthcare labor market can lead to better hospital performance and lower mortality.

The paper is organized as follows. In Section 2 we present the relevant institutional background, describe the reform we exploit, and outline expected effects of the reform. In Section 3 we describe the data sources, sample restrictions, and construction of the main outcomes used in the analysis. In Section 4 we detail our empirical strategy and main identification challenges. We report our main results in Sections 5 and 6, including robustness checks. In Section 7 we conclude and discuss further research.

² A meta-review of the medical literature identifies some empirical analyses of working hour limits on patient outcomes, with mixed evidence. Most reviewed papers are case studies on one or few hospital wards, and rely on self-reported data and pre-post analyses (Bolster and Rourke, 2015). We contribute to this type of literature by providing an empirical analysis across all hospital wards within a country.

2. The Physician Worktime Restrictions Reform in Denmark

In this section, we present the main features of the 1982 labor market reform that we exploit in the empirical analysis. We start by documenting the relevant institutional background. We then present the main features of the new collective bargaining agreement which set overtime restrictions with the goal of inducing a substitution from hours-per-worker to workers. Finally, we discuss the expected effects of the reform on the labor market and patient health. Appendix A provides further details on the setting and changes introduced by the reform. Appendix B formalizes our main theoretical claims.

2.1. Institutional Background

Hospital Care in Denmark.— The Danish healthcare sector is mostly publicly funded, with hospital care being free of charge to all residents. The system is organized into primary and secondary care. Hospitals fall under secondary care, including the provision of urgent healthcare services.³ Hospitals are then divided into hospital wards, according to specific medical specialties. In our period of analysis, the public hospital sector was characterized by county-level governance, with central government oversight. Patients in need of hospital care in Denmark were typically referred by their general practitioner to the public hospital within their area of residence. In acute or emergency situations, patients were taken to the nearest appropriate emergency department, regardless of residence.

Physicians' Career.— The career of hospital physicians was structured in three stages. Physicians started their careers as junior residents (*reservelæge*), hired on 1-year renewable contracts. At a second stage, physicians became senior residents (*1. reservelæge*), and were hired under five-year renewable contracts. Training in a particular medical specialty occurred during junior and senior residency, with the minimum duration being between 4.5 and seven years, but with an average duration of twelve years in our period of analysis. After specialty recognition, hospital doctors could move to become chief physicians, with responsibilities over hospital ward management, clinical training, and supervision of residents.

Job Matching and Wages.— The job market for physicians in Denmark combined regulated pay with a decentralized allocation of workers to available posts. Collective bargaining agreements (CBAs) established the terms of employment for public hospital physicians. For subordinate physicians—junior and senior residents—CBAs were negotiated between the Danish Association of Junior Hospital Doctors (FAYL) and public employers, represented by the Association of County Councils, two hospital municipalities (Copenhagen and Frederiksberg), and the Ministry of Finance. Collective bargaining took place every two years, where the parties determined whether to renew, amend, or—as in the

³ Primary care services, which include municipal health centers and general practitioners, can refer patients to hospital care.

rare instance of the reform we exploit—replace the existing agreement. Physicians were directly hired by hospitals according to publicly posted positions. In our main period of analysis, vacancy posting was centralized, but published by hospital ward, and dependent on their predicted healthcare demand. Subordinate physicians applied to their desired vacancy and were selected by hospitals on the basis of their application package. Importantly, this job matching mechanism was replaced by a lottery allocation for junior residents starting in 1984 (Fadlon et al., 2024).⁴ Despite decentralized competition for vacancies, physician hourly compensation, overtime premia, and working conditions were established according to the CBA. In our period of analysis, wage schedules included 15 tenure levels within each career stage, with specific provisions on overtime limits and pay for each level.

Work Schedules.— Before the reform, physicians’ working hours consisted of daytime duty (typically 8:00AM – 2:20PM) and shift work (outside the daytime interval or on weekends and holidays). Shifts were divided into three availability conditions: on-call on-site, on-call from home and on-call off-site. While daytime duty was characterized by collaborative work across seniority levels, with active supervision and training of junior and senior residents, on-call duty was characterized by multiple tiers of engagement. Junior residents carried most of the clinical workload outside of the daytime interval and staffed the front-line of hospital care, backed up by senior residents and chief physicians available on-call from the hospital or from home (Christy, 1984).⁵ Appendix A provides further details on the pre-reform regime.

2.2. The 1982 Reform

Starting in the 1960s, the expansion of higher education and no constraints to student intake led college enrollment in Denmark to grow exponentially (Hansen, 2018, 2015). By the 1970s, medical school graduates had more than doubled in just a few years (Lorentzen, 2024). With no commensurate increase in vacancies, an excess supply of early career physicians led to a slacking of the labor market. In 1981, hundreds of medical school graduates were unemployed or sub-employed in Denmark, while more than 800—equivalent to about 18% of available resident positions—had migrated to Sweden to work (Møller and Nexø, 2004). According to aggregate data on working hours, employed residents were working an average of 45.6 hours per week on regular and active on-call duties (Christy, 1983; Møller and Nexø, 2004).

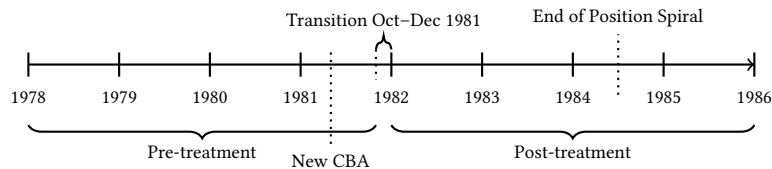
With the goal of increasing job sharing and normalizing physicians’ workweek, the Danish Association of Junior Hospital Doctors and the public employers entered into negotiations to sign a new

⁴ We report our difference-in-differences estimates up until 1984 to avoid the effect of this confounding policy in the labor market.

⁵ This multiple tier organization of physicians’ schedules across different stages was believed to ensure post-graduate training, while being cost-effective (Christy, 1984).

collective bargaining agreement. The new agreement was adopted on May 1981, against the will of 98% of the Danish Association of Junior Hospital Doctor's voting members. Following a strike and additional negotiations, the parties finally reached an interpretation agreement in December, 1981 (Lorentzen, 2024).

Figure 2. TIMELINE OF THE REFORM



Notes. The figure presents a stylized depiction of the timeline of the overtime restrictions reform in Denmark, between 1978 and 1986.

The new CBA and the interpretation agreement included three key provisions: (i) enforcement of a 40-hour workweek; (ii) at least half of those hours scheduled as daytime duty on weekdays; and (iii) restrictions on overtime, only to be used when hospitals could not fill new positions. All regulations applied to both junior and senior residents, excluding chief physicians.⁶ Although the new system had to be fully implemented in 1982, the original agreement established that each hospital ward should start planning the transition already by 1981, with a transition period in the last three months of the year. Figure 2 depicts a stylized timeline of the reform. Given the timing of the interpretation agreement and full implementation, we consider 1981 as the last pre-reform period, and 1982 as the first post-reform period.

Standard Workweek Hours.— The new CBA aimed at enforcing a 40-hour workweek, as subordinate physicians were not to work above this weekly standard, averaged over three months. Under the new rules, any overtime work in a given week had to be compensated by off-time in the following weeks, rather than overtime pay.⁷

Scheduling Restrictions.— The reform enforced new scheduling restrictions. To alleviate concerns with adequate post-graduate training, it guaranteed subordinate physicians a minimum of 20 hours on weekdays during daytime (Monday-Friday, 8AM-4PM), and consequently a 20-hour cap on nighttime duty. Moreover, mandatory rest periods were increased from a minimum of five to up to eleven hours for each 24 hours.⁸

⁶ The previous collective bargaining agreement had been signed in November 1977, and applied retroactively since April 1977.

⁷ Overtime could only be compensated with pay when compensatory leave was impossible to grant.

⁸ The new rest periods were implemented as part of new work environment agreement (*arbejdsmiljøaftale*) with the aim of ensuring compliance with a 1980 ministerial order on rest and weekly time off (*Bekendtgørelse nr. 372 af 15. august 1980 om hviletid og fridøgn*). The work environment agreement for subordinate physicians implemented rules on rest periods and weekly time off that were more lenient than general national rules—11 hours of rest per 24 hours and 1 weekly 24 hour rest period—but significantly more stringent than the pre-reform rules. Contrary to the other provisions

Overtime Restrictions.— Beyond the new compensatory leave rules, overtime was to be capped at five weekly hours per physician, or 20 weekly hours per duty team, averaged over a three-month period. The new regulations induced a so-called “position spiral” (*Normeringsspiral*), as new positions had now to be opened to fill these hours. Overtime restrictions were only active until a new CBA signed in June 1984, where they were dropped (Figure 2).⁹ To compensate for the fall in total pay through fewer hours, residents employed before the new CBA were bumped up two tenure levels in the wage table.¹⁰ Moreover, work during particular weekend and night hours was to be compensated at a higher inconvenience pay rate (*ulempegodtgørelse*).

2.3. The Expected Effects of the Reform

In Appendix B we develop a simple framework where we formalize the expected effects of overtime restrictions on the labor market. Here, we highlight the economic intuition that guides our empirical analysis, and draw our main testable hypotheses.

Environment.— We take hospital wards as our unit of analysis. Each ward differs by medical specialty, production technology, and local demand for healthcare. As public sector providers, we assume that wards maximize healthcare output—or minimize patient mortality—subject to a budget constraint. Each hospital ward chooses two types of labor inputs: (i) total number of workers (N); and (ii) hours-per-worker (h). Output has marginally decreasing returns relative to each input. As in our empirical setting, each hospital ward pays an hourly wage rate (w) according to a convex schedule. Physicians are paid the same wage rate until a standard hours threshold ($\bar{h}$); any hour worked above $\bar{h}$ is considered overtime and paid at a higher rate, $(1 + p)w$, with $p > 0$ as the overtime premium.

Overtime Allocations.— Hospital wards can vary in their use of overtime for mainly two reasons. First, hospital wards with different specialties and local healthcare demand vary in their reliance on overtime and shift work. According to the production technology of each ward, there is heterogeneity in the degree of substitutability between h and N —the extent to which one extra hour from an incumbent worker can be substituted by one hour supplied by an additional hire. For instance, if a surgeon’s shift ends mid-surgery, it may still be more productive to finish the case in overtime than swapping them for another equally productive surgeon. In other specialties, where work can be divided into smaller blocks with no negative consequences for productivity, wards can optimally hire more workers on shorter shifts. Likewise, when the arrival rate of patients is unpredictable and shifting patients’

in the CBA, the statutory rest hours were locally negotiated with hospitals, varied between 8 and 11 hours, but always higher than the pre-reform standard.

⁹ The 1984 CBA still required that a proposal to hire additional physicians had to be put forward if overtime could be converted into positions.

¹⁰ The special affordance made to the already employed (*paraplytillæg*) meant that, for a resident at the lowest tenure level, the change led to an increase of around 11% in base pay.

treatment constrained, the need for overtime increases as shifts require longer periods of availability. Second, if there is limited information on physicians' productivity, by underestimating the marginal productivity drop from additional hours, wards may use overtime hours at non-optimal levels.

Hours-to-Workers Substitution.— With the introduction of overtime restrictions, hospital wards become severely restricted in making overtime allocations. Given the prevalence of overtime use prior to the reform (Christy, 1983), the new collective bargaining agreement must induce a substitution from hours (h) to workers (N), summarized in the following hypothesis:

Hypothesis 1. *Labor demand for physicians (N) shifts upwards after the reform.*

Figure 1 provides evidence in favor of Hypothesis 1, providing the motivation for our empirical analysis. Posted positions surge right after overtime restrictions are implemented. Crucially, this shock on the extensive margin of labor demand was based on a substitution from h to N , in line with the provisions in the new collective bargaining agreement. Figure 3 shows the change in the total number of available positions (N) for junior and senior resident physicians and the total number of active duty physician-hours demanded ($N \times h$) before 1982. Information for 1982 reflects the total hours forecast for that year based on the rules imposed by the new CBA (Christy, 1983).¹¹ While total labor-hours ($N \times h$) were projected to increase by 19%, the number of physicians was expected to increase by 36%, or almost double that proportion. The wedge between total hours and total workers implied a fall of about 12.5% in hours-per-physician (h), as triggered by the requirement to achieve a 40-hours workweek through elimination of overtime in 1982.

The second theoretical hypothesis is key to our empirical strategy, and provides a testable first stage to our empirical analysis:

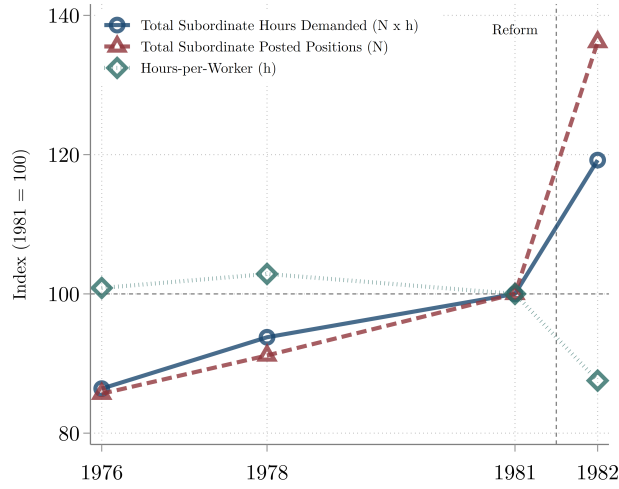
Hypothesis 2. *The change in labor demand for physicians is increasing on pre-reform overtime need.*

As discussed above, heterogeneity across hospital wards generates variation in overtime use. Wards with a relatively higher use of overtime pre-reform are relatively more affected by the new regulations. The reform induces a relatively stronger h for N substitution, with high-overtime wards having to hire relatively more workers as a consequence. In Section 4 we describe how we use this heterogeneity in overtime need to identify our reduced form effects. In Section 5, we show that higher pre-policy overtime need is empirically associated with larger percent changes in labor demand across reform periods.

Healthcare Quality and Supply.— How should this labor demand shock affect healthcare provision? Before the new CBA hospital wards were free to optimally adjust hours-per-worker (h). With the new

¹¹ The forecast is based on a survey-based official document produced by the Danish Association of Junior Hospital Doctors (*F.a.y.l.'s medlems-og overenskomstenueter 1972-1982*).

Figure 3. HOURS-TO-WORKERS SUBSTITUTION



Notes. The figure documents the trends in total hours, total number of physicians, and implied hours-by-physician demanded. Data points for 1982 are projections produced by the Danish Association of Junior Hospital Doctors according to data in a resource utilization survey (*F.a.y.l.'s medlems-og overenskomstenqueter 1972-1982*) (Christy, 1983). All values are normalized to be 100 in 1981, such that differences across years can be interpreted as percent changes relative to 1981.

constraints on h , optimal overtime allocations become infeasible. In a framework with no frictions in the labor market and perfect information on worker productivity, if a more productive labor allocation could have been possible before overtime restrictions, wards could have already achieved it. Under these conditions, health output could not possibly increase. Hiring additional physicians can translate into higher costs in terms of internal organization and quality of continued patient treatment. Conditional on the production function of each ward, shorter shifts may mean more hand-offs of cases across physicians. From an on-the-job training perspective, it may also be optimal for hospitals to expose physicians to long hours early in their careers as an investment on medical skills. With heterogeneous productivity, hospital wards may also optimally choose to assign longer hours to fewer physicians over potentially less productive new hires.

However, relevant frictions in the physician labor market in Denmark can lead to positive effects of overtime restrictions on health outcomes. Regulated wage schedules may not reflect the competitive equilibrium between labor supply and demand. The reform was driven by concerns about unemployment of recent medical graduates and work-life balance of early-career physicians. It could be that overtime premia were too low relative to what would be optimal if the price of overtime could freely adjust. In that case, pre-reform employment would be too low for optimal healthcare provision. Moreover, hospital wards may misperceive physicians' productivity. In workplaces with multiple overlapping tasks, unpredictable case arrival, and team work—typical of hospital care—individual productivity is hard to monitor. Empirical evidence shows that quality of medical decision making is highly sensitive to scheduling and declines near the end of shifts (Chan, 2018). With imperfect information on

effort, hospital wards may underestimate the marginal productivity drop of an extra hour, underplaying the effect of fatigue on quality of care. If fatigue costs were underestimated, this induced reallocation can raise output, depending on how far hospital wards' beliefs are from the true production function. Finally, even with heterogeneous productivity across physicians the effect on output can still be positive if productivity differences align with amenity preferences. For instance, recent evidence shows that female physicians sort into different specialties as a result of shorter workweeks (Lorentzen, 2024; Wasserman, 2023b). If the marginal entrants are of comparable or even higher productivity, the composition shift may increase output.

Given these different potential causal channels, the overtime restrictions introduced in the physician labor market in 1982 have theoretically ambiguous effects on health outcomes. In the following sections, we empirically estimate the net effect of the policy on patient health.

3. Data

In this section, we describe our main data sources and present relevant summary statistics related to the physician labor market and healthcare quality in Denmark around the reform. Appendix C presents the construction of our main datasets and outcomes in more detail.

3.1. Data Sources

Physician Labor Market.— We characterize physician labor demand and the extent to which it is met using digitized statistical yearbooks from the Danish Health Authority. We leverage an existing data set that builds on these sources by Lorentzen (2024) and supplement it with more granular-level data obtained from the same statistical yearbooks¹². The resulting dataset contains information on the number of posted positions (demand) and filled positions (met demand) at Danish hospitals. From this data, we construct datasets at different levels including the ward-by-year level and specialty-by-year level.¹³ Our main analyses on the first-stage impact on the physician labor market are on the specialty-by-year level, as this is the level at which we are able to join the physician labor market data with data on hospital admissions. This sample consists of 21 different medical specialties. See Appendix C for further details.

Healthcare Supply and Patient Outcomes.— We access a restricted-use administrative dataset—the *National Patient Register (LPR)*—covering the universe of hospitalizations in Denmark. We observe all inpatient hospitalizations at non-psychiatric hospital wards. For each patient contact, we observe the

¹² We scan, digitize, and clean the data from the following sources: *Medicinalstatistiske Meddelelser: Lægestillinger og Sengepladser på Institutioner (1972-1980)* and *Sygehusstatistik: Lægestillinger og Sengepladser på Institutioner (1981-1991)*. See Lorentzen (2024) for the full list of references to the archival material.

¹³ From the same sources, we also collect data on posted and filled positions on the county-by-year level and by career-stage as well as data on total bed capacity at the ward-by-year level

hour and date of admission, the date of discharge, the type of admission (planned or unplanned), hospital and hospital ward identifiers, the county in which the hospital is located, and the main specialty of the hospital ward. Relying on pseudonymized personal identifiers, we link hospitalizations to *The Cause of Death Register (DAR)* to obtain patients' date of death.

3.2. Sample Restrictions

In the empirical analysis we use an unbalanced panel of hospital wards by year, between 1978 and 1986.¹⁴ A ward is part of our sample if (i) it is not a psychiatric ward; (ii) admits patients; (iii) was operating before the beginning of our period of analysis; (iv) admits at least one acute and one non-acute patient; and (v) admits patients at least the year prior to the reform and in one of the years post-reform. Our panel contains a total of 424 wards in 106 hospitals. The sample includes over 7.3 million patient admissions, and covers 88% of the universe of hospitalizations in the period. See Appendix C for details on the construction of our main variables and on our sample restrictions.

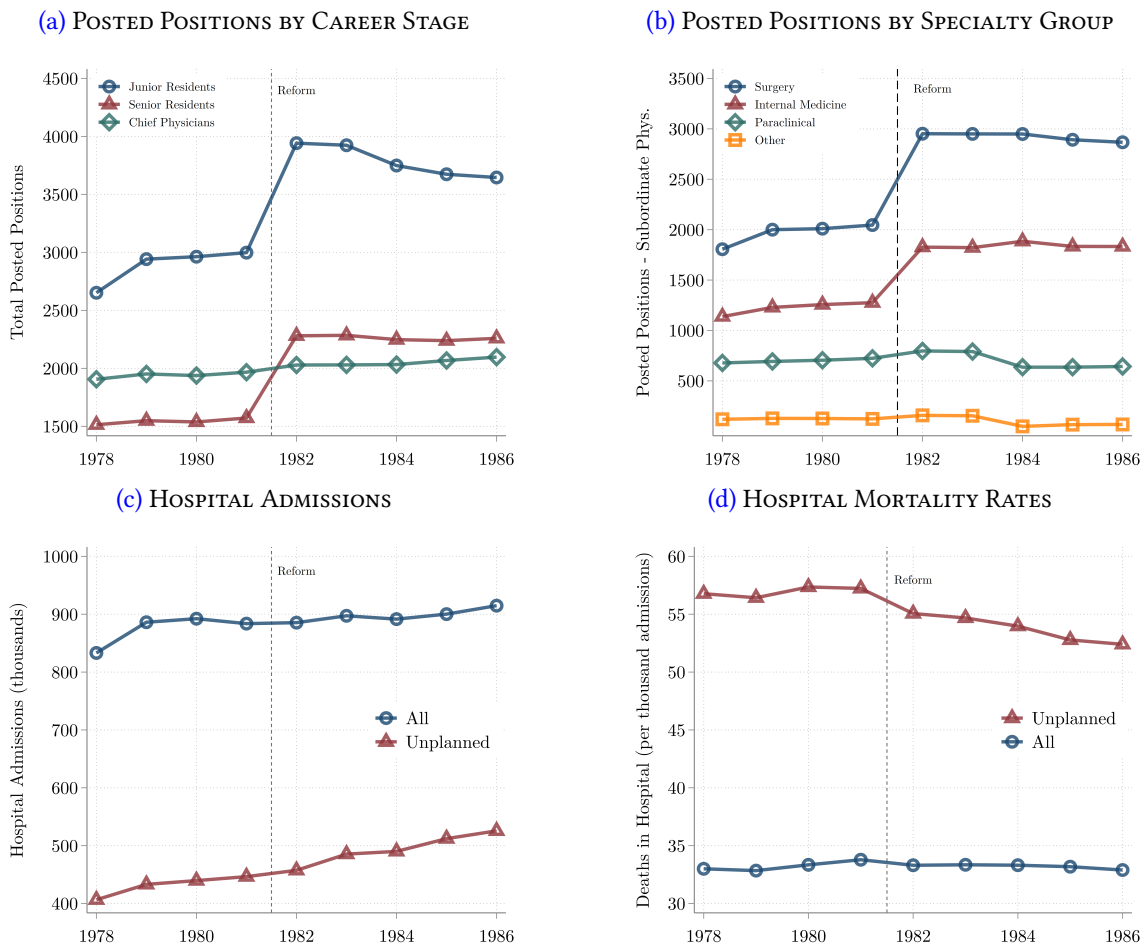
3.3. Main Outcomes

The primary measure of healthcare quality is the in-hospital all-cause mortality rate, a standard indicator of hospital performance (e.g., Muñoz and Otero, 2025; Chan et al., 2023; Doyle et al., 2015; Gaynor et al., 2013). To assess whether any effects extend beyond hospital discharge, we complement this outcome with mortality at longer horizons. Using linked patient admissions and death registry records, we construct mortality measures within 30 days, 60 days, and one year after discharge, including in-hospital deaths. As additional healthcare indicators, we examine patients' average length of stay (in days) and readmission rates into any hospital ward 14 days, 30 days, and 60 days after discharge. We also study heterogeneity by admission type, separately estimating all outcomes for planned and unplanned (acute) admissions. This distinction is central in our setting because unplanned admissions are typically non-deferrable emergencies, while planned admissions are more schedulable (Card et al., 2009).¹⁵ We compute all main outcomes separately for these two admission types. Finally, to address concerns about changes in patient composition across reform periods, we compute age-adjusted and risk-adjusted measures as the ratio between the observed rates and the predicted rates based on the age-composition and diagnostic of admitted patients in each ward (e.g., Finkelstein et al., 2025; Muñoz and Otero, 2025).

¹⁴ We focus on the period from 1978 to 1986 shield our analysis from other confounding policies, namely the introduction of the pre-reform collective bargaining agreement signed in April 1977 and the reduction of the standard workweek to 39 hours in January 1987.

¹⁵ See Section 4 and Appendix C for details on definitions and construction.

Figure 4. TRENDS AROUND THE REFORM



Notes. The figure documents relevant trends around the reform. Panel A presents the overall change in posted positions by physicians' career stage. Panel B depicts the trend in posted positions for subordinate physicians, separately by medical specialties grouped into broader categories. Appendix Table F.5 shows the assignment of specialties to each group. Data sources for Panels A and B are statistical books produced by the Danish Health Authority. Panel C documents trends in total hospital in-patient admissions, by type of admission. Panel D presents trends in hospital mortality rates by type of admission. Data sources for Panels C and D are restricted-use administrative registries of in-hospital patient admissions.

3.4. Descriptive Statistics

To motivate our empirical analysis, we here present descriptive evidence on the first-stage impact on the physician labor market and the downstream effects on healthcare during our period of analysis (1978-1986).

Physician Labor Market.— Figure 4 shows the trends in posted positions at Danish hospitals by career stage (Panel A) and specialty group (Panel B) from 1978 to 1986. First, consistent with the stated goals of the reform, Panel A shows that overtime restrictions are followed by a large, discontinuous increase in the number of posted positions. This surge in hiring is concentrated among subordinate physicians—junior and senior residents—with no impact on the chief physician labor market, regulated by a separate collective agreement. Between 1981 and 1982, the number of positions available for

subordinate physicians increased by 35% while filled positions increased by 25%.¹⁶ On the other hand, Panel B of Figure 4 shows that the surge in posted positions varied across broad specialty groups. The pattern is consistent with the idea that pre-reform overtime need and scheduling rigidity predicts labor demand responses through hours-to-workers substitution. While surgical specialties, in which long hours and non-daytime duty are common, increased their labor demand by more than 40%, para-clinical specialties, that typically rely very little on either, experienced labor demand growth of less than 10%. In Section 4 we describe how we rely on a related source of variation in our empirical analysis—pre-policy concentrations of acute patients, mostly driven by differences across wards’ specialties.

Healthcare Supply and Hospital Mortality.— In Figure 4 we also present trends in hospital admissions (Panel C) and in-hospital mortality (Panel D) throughout our period of analysis. Table 1 complements this information by showing summary statistics separately for the pre-reform (1978-1981) and the post-reform (1982-1986) periods. The table shows that total counts by ward, post-reform means are significantly higher than pre-reform means. For example, the mean number of yearly admissions increased by 230.97 (12.23%) across periods. Conversely, in-hospital mortality was significantly lower post-reform than pre-reform, with a decline of 2.69 fewer deaths per thousand admissions (-6.7%). Taken together, this shows that although more deaths occurred in hospital wards in the post-reform period, the in-hospital mortality *rate* was significantly lower, especially among unplanned admissions. Figure 4, Panel D, also shows a noticeable drop in the in-hospital mortality rate in the reform implementation year (1982) among unplanned admissions—a group of patients disproportionately treated by subordinate physicians. While this evidence descriptively documents significant differences across reform periods, in Section 4 we outline our empirical strategy to estimate the effects of the policy change.

¹⁶ Appendix E presents a more detailed analysis of labor demand change at the ward-level.

Table 1. SUMMARY STATISTICS: ANALYSIS SAMPLE (1978-1986)

	Pre-Reform		Post-Reform		Post - Pre	Post - Pre p-value
	Mean	SD	Mean	SD	Diff.	Diff.
Counts						
Hospital Admissions	1879.61	1326.88	2110.58	1400.74	230.97	0.00
Unplanned Admissions	945.12	785.75	1170.89	904.28	225.76	0.00
Planned Admissions	934.49	866.77	939.70	826.49	5.21	0.85
Deaths in Hospital	63.02	76.30	70.07	83.94	7.05	0.01
Rates (per thousand admissions)						
In-hospital Mortality	40.09	43.70	37.40	40.09	-2.69	0.05
In-hospital Mortality (among unplanned)	56.59	59.99	52.11	57.01	-4.48	0.02
In-hospital Mortality (among planned)	18.54	30.32	15.33	34.77	-3.21	0.00
30-day Mortality	57.20	57.60	55.78	54.51	-1.42	0.44
60-day Mortality	68.15	65.72	67.75	63.42	-0.40	0.85
1-year Mortality	127.19	111.14	131.19	110.31	4.00	0.28
30-day Readmissions	99.15	68.44	117.06	76.72	17.92	0.00
Other						
Length of Stay (Days)	12.72	13.66	10.46	9.55	-2.26	0.00
Share of Unplanned Admissions	0.48	0.25	0.53	0.25	0.04	0.00
Share of Night Admissions	0.24	0.14	0.25	0.14	0.01	0.12
Share of Weekend Admissions	0.11	0.06	0.12	0.06	0.00	0.27
Urban County	0.54	0.50	0.54	0.50	-0.00	0.98
Mean Patients' Age	45.40	16.80	47.03	16.45	1.64	0.00
Number of Observations	1691	1691	1970	1970	3661	3661
Number of Wards	424	424	424	424	424	424
Number of Hospitals	106	106	106	106	106	106

Notes. The table presents summary statistics for the analysis sample. The sample includes all non-psychiatric hospital wards with at least one observation in each reform period (see Section C.1 for more details). Rate variables refer to variables measured per 1000 admissions. Length of Stay (Days) is the average duration of hospitalizations in the ward, measured in days. "Share of Unplanned Admissions" refers to the share of a ward's admissions that are classified as "acute". "Share of Night Admissions" refers to the share that are admitted outside the interval 08:00AM-4:00PM. All variables are measured at the yearly level.

4. Empirical Strategy

In this section, we present our research design. To study the effect of physician overtime restrictions on patient outcomes, we rely on a dose-response difference-in-differences strategy and event studies. We present our source of identifying variation, as well as our preferred econometric specifications and estimators.

4.1. Identifying Variation: Differences in Pre-reform Overtime Need

To estimate the effect of overtime restrictions on patient outcomes, we exploit variation across hospital wards with different levels of pre-reform overtime need. We proxy wards' overtime need with pre-reform characteristics of their production function and local healthcare demand. Our approach has two main advantages. First, these demand features are predetermined and set constraints on wards' labor allocation decisions. Second, they capture the *need* for additional hours, while observed overtime could reflect strategic scheduling choices unrelated to true workload constraints.

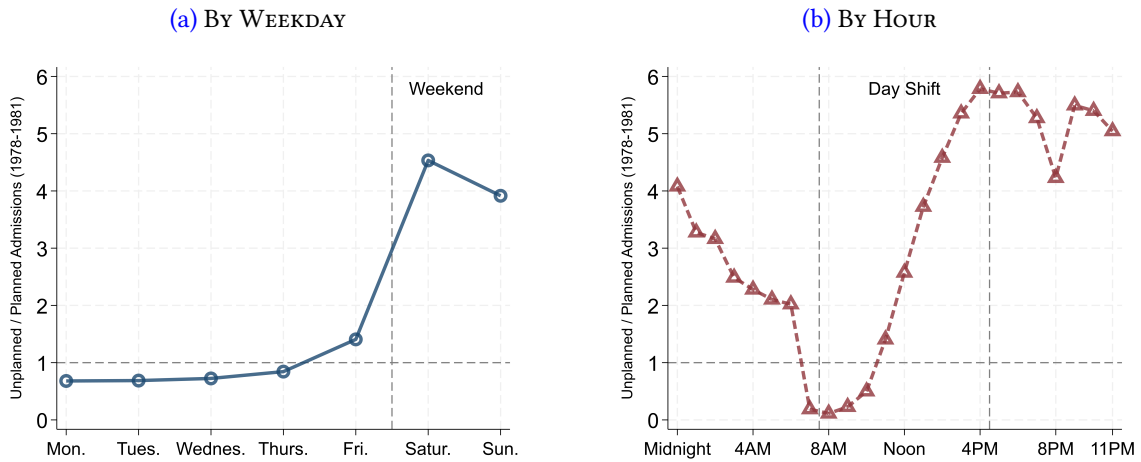
We use the pre-reform share of unplanned admissions in each hospital ward as our main measure of overtime need.¹⁷ We leverage three schedule-constraining properties of this type of cases. First, acute cases typically need immediate attention and treatment cannot be deferred. In contrast, the treatment of non-acute cases can be planned and adapted to physicians' daytime duty schedules. Second, the timing of arrival of unplanned admissions across the day is arguably exogenous to hospital ward decisions. Third, weekend and night work is more likely to be assigned to unplanned nondeferrable admissions.

In Figure 5 we show the pattern of unplanned to planned admissions across weekdays and hour of admission. Cases whose treatment can be planned are shifted towards workdays (Monday-Friday), and especially towards the beginning of the week. Patients admitted during the weekend are about 4.5 times more likely to be an unplanned rather than a planned admission (Panel A). This echoes the discussion in Card et al. (2009), that identifies patients with nondeferrable diagnosis by the relative likelihood of being admitted on weekends. Moreover, we show that within a day, admissions that can be planned are relatively more likely to be shifted to the beginning of the physicians' daytime shift. An admission during the night is almost 5 times more likely to be unplanned (Panel B).

All these properties predict relatively more overtime need. On the one hand, higher concentrations of unplanned admissions force more physicians to be available on-call, working longer shifts. On the other hand, by imposing restrictions on weekend and night work, the policy induces wards facing relatively more unpredictable demand to hire more physicians (see Section 2). Given these properties, we exploit variation across wards in the share of unplanned admissions pre-reform. By interacting a

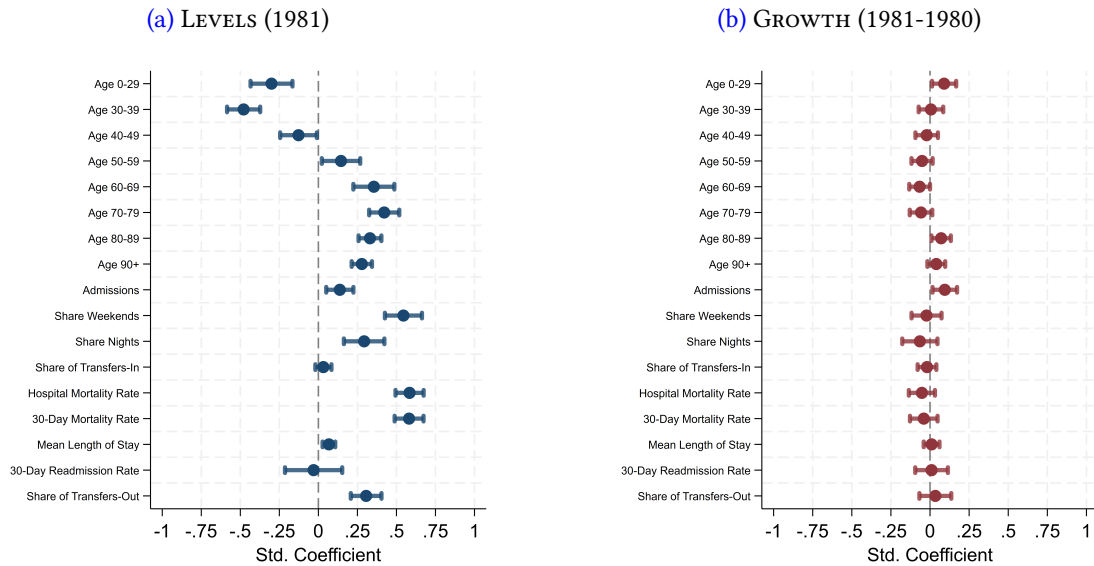
¹⁷ As detailed in Section C.2, unplanned admissions refer to acute cases that arrive without prior booking, typically through emergency departments.

Figure 5. UNPLANNED-TO-PLANNED ADMISSIONS (1978-1981)



Notes. The figure presents the ratio between the number of unplanned admissions over the number of planned admissions within different time periods. The sample includes all patient admissions between 1978 and 1981. Panel A shows the variable for which day of the week. Panel B shows the variable for each hour throughout the day. The vertical dash lines represent the start and end of the physician day shift. A day shift is between 8:00 and 16:00. The horizontal dash lines represent when the number of planned and unplanned admissions are the same.

Figure 6. BALANCE IN OUTCOMES AND COVARIATES



Notes. Panel A presents coefficients and 95% confidence intervals from OLS regressions of each variable on our dosage variable—share of unplanned admissions in 1980—one year prior to the onset of the reform. Panel B presents the same coefficients but with the dependent variable being the first difference change between 1980 and 1981. The dosage and dependent variables are standardized in both panels. Age distribution variables are the share of patients in those particular age bins. All regressions are weighted by the number of patient admissions in 1981, except for when the dependent variable is itself the number of admissions. Standard errors are clustered at the hospital ward level.

post-reform indicator with this proxy of overtime need pre-reform we aim to isolate variation in the intensity of overtime restrictions across wards.

Balance.— To characterize wards that were relatively more exposed to the reform, we examine how the share of unplanned admissions—our dosage variable—correlates with a set of ward-level covariates

and outcomes in the year prior to the policy change. Figure 6, Panel A, reports standardized OLS coefficients from regressing these 1980 characteristics on the share of unplanned admissions. Wards with higher exposure tend to be larger, admit older patients, and exhibit higher in-hospital and post-discharge mortality rates. Consistent with Figure 5, they also experience a higher incidence of weekend and nighttime admissions. Thus, more exposed wards differ systematically in observable characteristics from less exposed ones.

Panel B of Figure 6 shows that year-on-year changes in these characteristics are much less correlated with the dosage measure. Wards with higher exposure were not trending differently prior to the reform, a finding further supported by the event-study estimates. Together, these patterns suggest that while levels differ across wards, pre-reform trends do not, lending credibility to our identification strategy.

In Appendix D we further characterize the variation in the share of unplanned admissions across time, prior to the reform. Through a variance decomposition exercise, we start by showing that most of the pre-reform variation stems from differences across specialties and wards, rather than across years. We then discuss alternative proxies of overtime need and how these relate to our preferred variable. In particular, we characterize differences across wards in the share of weekend admissions, night admissions, as well as uncertainty in unplanned, weekend, and night admissions throughout the year. In Section 6.5 we show whether the interpretation of our main results changes with alternative definitions of our dosage variable.

4.2. Econometric Specifications and Identifying Assumptions

We exploit ward-level heterogeneity in pre-reform overtime need using the following event-study specification:

$$Y_{it} = \alpha + \gamma_i + \lambda_t + \sum_{\tau=1978}^{1980} \delta_{\tau} \mathbb{1}\{\tau = t\} D_i + \sum_{\tau=1982}^{1986} \beta_{\tau} \mathbb{1}\{\tau = t\} D_i + \varepsilon_{it}, \quad (1)$$

where Y_{it} is the outcome for hospital ward i and year t . D_i is the pre-reform share of unplanned admissions (in $t = 1980$), our main regressor, standardized to unit variance so that coefficients are read as effects of a one standard deviation increase in exposure to overtime restrictions. Ward fixed effects (γ_i) control for time-invariant characteristics of each hospital ward, whereas year fixed effects (λ_t) absorb common shocks to all wards within-year. We omit $t = 1981$, making the final pre-reform year the reference category. Pre-reform coefficients δ_{τ} trace differential pre-trends across wards, and serve as a visual parallel trends test. Our coefficients of interest are β_{τ} , which trace the dynamic response to the reform. To have a per-patient interpretation of our estimates and account for heterogeneity in ward size, we weight observations by the number of admitted patients in 1981 (except when admissions are

the dependent variable). To aid with the interpretation of our effects, we run separate specifications by type of admission, depending on whether patients' treatment could be deferred. To address misspecification concerns, we also estimate the effect of the reform using a log transformation of the dependent variable. To account for the cases of wards with no deaths, we also estimate the model through Poisson quasi-maximum likelihood estimation (QMLE) (Chen and Roth, 2024; Mullahy and Norton, 2024). In additional analyses, we also present stricter versions of our preferred specification in which we control for trends within counties or within broad specialty groups, by replacing year fixed effects by county-by-year (λ_{ct}) or specialty-by-year fixed effects (λ_{st}). We cluster the standard errors at the ward level, but explore other possible structures of error correlation in robustness analyses.

To have a summary estimate of the effect of the reform in the three years after the onset of the reform, we replace the event-time coefficients in Equation 1 by a post-reform indicator, using the following specification:

$$Y_{it} = \alpha + \gamma_i + \lambda_t + \beta \mathbb{1}\{t \geq 1982\} D_i + \varepsilon_{it}. \quad (2)$$

The coefficient β captures the average relative effect of overtime restrictions across differentially exposed wards after the reform is implemented. Our dosage definition and other variables remain defined as above. In the case of our simplified DiD specification, we restrict the treatment period to three years post reform (1982 – 1984), before the partial lifting of overtime restrictions may translate into reemergence of overtime hours that were previously eliminated (see Section 2).¹⁸ Finally, we run identical specifications to those in Equation 2 at the patient-level.

Relevance.— To interpret our results as effects of the reform, D_i must predict reform-induced hours-to-workers substitution. In Section 5 we run a similar event study as the one in Equation 1, using labor demand and hired physicians as the dependent variable of interest. Because we cannot link information on hired physicians at the hospital ward level to our dose variable obtained from data from the National Patient Register, we run our first stage analysis at the specialty (s) level. Yet, we show a strong increase in labor demand the higher the pre-reform share of unplanned admissions by specialty (Section 5).

Parallel Trends.— Interpreting our estimates as causal requires that outcomes for wards with different levels of exposure would have trended similarly absent the reform. We test differential pre-trends by examining the pre-reform coefficients δ_τ , in Equation 1. In Sections 5 and 6 we show no differential pre-trends across dose levels. Still, in the context of our continuous DiD design, without a stronger version of parallel trends, the interpretation of the two-way fixed effects (TWFE) estimates is ill-defined

¹⁸ We note that the removal of the position spiral implies that previously eliminated overtime may reemerge if the demand for health care supplied by subordinate physicians rises more than supply.

if there are heterogeneous treatment effects by dose (Callaway et al., 2024). Given this, we require a strong version of parallel trends where the average change in outcomes if all units experience dose d is equal to the observed path of outcomes of dose group d (Callaway et al., 2024):

$$\mathbb{E} [Y_{post}(d) - Y_{pre}(0)] = \mathbb{E} [Y_{post}(d) - Y_{pre}(0)|D = d], \quad \forall d \in \mathcal{D}, \quad (3)$$

where $\mathcal{D}$ is the support of our dose variable. Relying on this identifying assumption, we interpret our TWFE estimates as an average causal response from a marginal increase in D .

No Concurrent Shocks.— In Section 2 we carefully examine the main policy changes and find no other major changes in the healthcare institutional landscape in Denmark for our main period of analysis. Negotiations around the reform concluded in late-1981, with a three-month adjustment window. In our event-study analyses we show that there is no evidence consistent with anticipation in hiring (Section 5) or in our outcome variables (Section 6.1). In 1984, there were two other major changes in the physicians labor market in Denmark. First, overtime restrictions were partially lifted in 1984, as discussed above and in Section 2. Second, the job allocation mechanism for early-career physicians fresh out of medical school changed into a lottery-based system (Fadlon et al., 2024). Therefore, we restrict the interpretation of our estimates to a treatment period three years after the onset of the reform (1982-1984).

5. Working Time Restrictions and the Physicians' Labor Market

In this section, we provide empirical evidence on the first-stage impact of the reform on the physician labor market. Consistent with Hypothesis 2 and the arguments for our identifying source of variation in Section 4.1, we show that specialties with higher concentrations of unplanned admissions changed their labor demand the most. Appendix E presents additional descriptive evidence on overall changes across reform periods including at the hospital ward level.

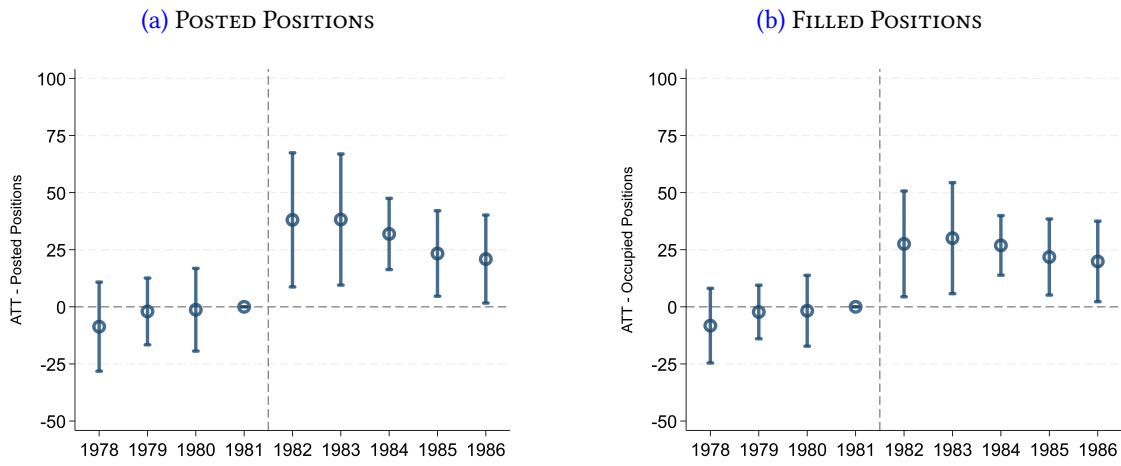
Extensive Margin of Labor Demand.— To study the first stage of the reform by overtime need, we start by merging the share of unplanned admissions in $t = 1980$ with information on the number of posted and filled subordinate physician positions at the specialty-by-year level (see Section 3). We then run event-study (Equation 1) and simple difference-in-differences (Equation 2) specifications (Appendix E). We include a total of 21 medical specialties for which we have information.

Figure 7, Panel A, shows dynamic effects on labor demand across reform periods, with no differential pre-trends. Post-reform point estimates peak in 1982 – 1983 and then gradually decrease in magnitude, consistent with wards front-loading demand with the introduction of the overtime restrictions reform and with reductions in posted positions occurring with the end of the position spiral in the

middle of 1984.¹⁹ Over a baseline of 177 open positions by specialty pre-reform, the policy led to a 17% change in labor demand per standard deviation of exposure in the three years after its implementation.

Hired Physicians.— Dynamic effects on hired physicians show an identical pattern (Figure 7, Panel B), although with a relatively smaller magnitude. Consistent with the wedge between physicians’ labor supply and demand right after the introduction of overtime restrictions (Figure 1), occupied positions increase only by 23.5 new positions (Appendix Table E.4, Column 2, C.I.: 10.9; 36) for a one standard deviation difference in our proxy of overtime need.

Figure 7. LABOR MARKET EFFECTS: EVENT-STUDY AT THE SPECIALTY LEVEL



Notes. The sample includes all specialties that can be merged between the archival data on labor market positions and admission registries. The figure documents coefficient of a regression with the same specification as Equation 1, but at the specialty level and using labor outcomes as dependent variables. Panel A presents estimated effects for the change in posted positions, while Panel B presents the estimated effects for filled positions. All estimates are relative to the baseline year 1981. The dashed vertical lines indicate onset of the new working-hour system. The capped bars depict 95% confidence interval using robust standard errors.

In Appendix E we further show that the share of unplanned admissions pre-reform is highly correlated with ($\rho = 0.56$) and explains almost a third of the variation in the percent change in labor demand across specialties. In contrast to our main specification, we use the percent change in labor demand between 1983 and 1981 as the dependent variable, to allay concerns with the effect being driven by baseline specialty size before the reform. We show that a one standard deviation difference in the share of unplanned admissions pre-reform is associated with a 5.5 p.p. increase (p -value = .008) in the demand for new subordinate physicians, weighted by the number of posted positions at baseline (Figure E.5). Overall, our first-stage findings, even if at the specialty-level, are consistent with overtime need predicting a stronger hours-per-worker to worker substitution after the implementation of overtime restrictions (Hypothesis 2). Next we turn to the effect on patient outcomes.

¹⁹ With the removal of the position spiral, the public employers issued a guarantee running until April 1st 1987, stating that the number of posted positions would be reduced by at most 360 positions cf. [CIS nr 6380 af 19/07/1984](#).

6. Working Time Restrictions and Health Outcomes

In this section we provide reduced-form evidence of the effect of physician working time restrictions in Denmark on patient outcomes. First, we show that the policy led to a significant reduction in in-hospital mortality in wards relatively more affected by working time restrictions, especially among unplanned cases. Second, we find that after the implementation of the new working hour system, the most affected wards were able to admit relatively more planned cases. Third, we find that these changes in healthcare production among relatively more affected wards led to no changes in readmission rates or the average length of stay. Finally, we show that the results are robust to a series of alternative treatment dosage definitions, specifications, and sample restrictions.

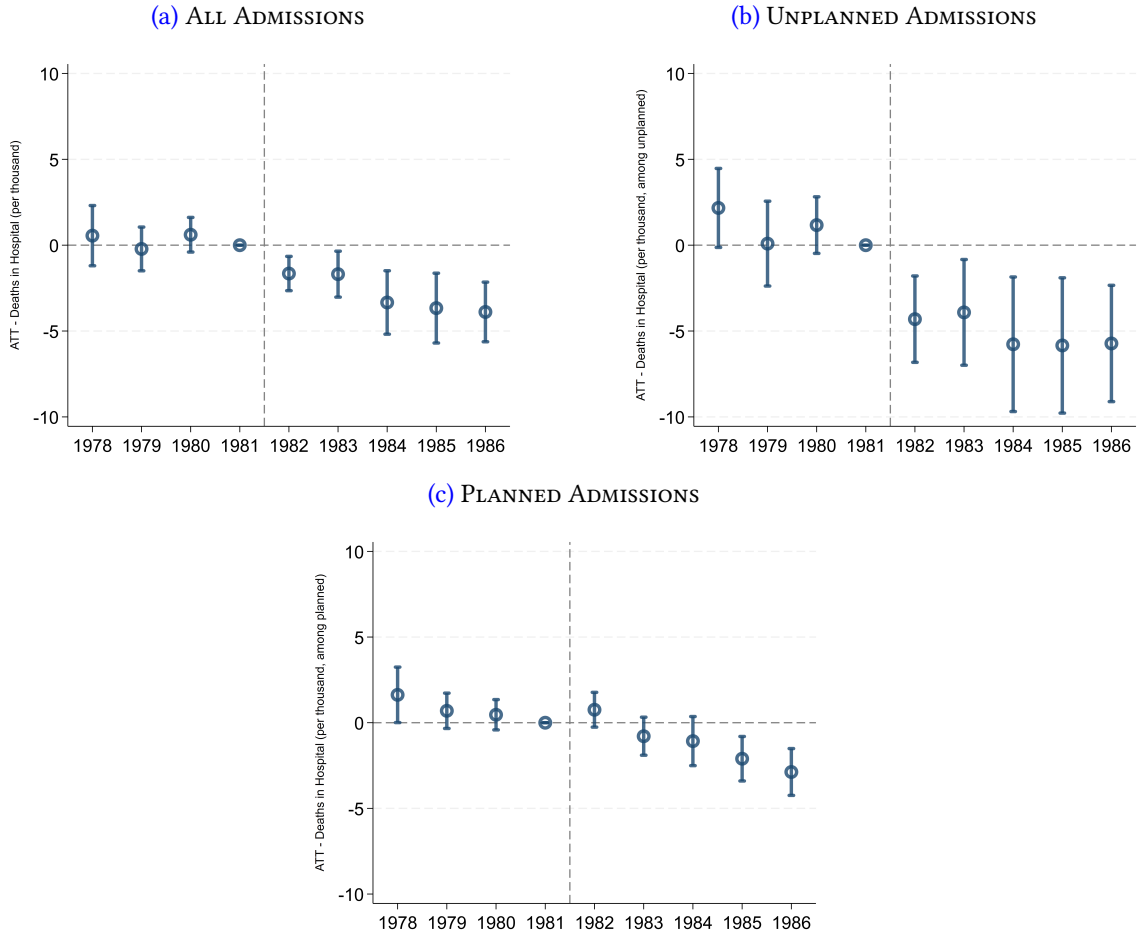
6.1. In-hospital Mortality

We start by showing that the working time restriction reform in Denmark reduced in-hospital mortality. Figure 8, Panel A, plots the event-study coefficients of the specification outlined in Equation 1, for our 1978-1986 sample of hospital wards in Denmark. We observe that from 1982—the first year of implementation—in-hospital mortality rates gradually decrease relative to 1981 and are statistically significant ($p < 0.01$). To summarize the estimates, we aggregate pre- and post-reform periods until 1984 in the dose-response difference-in-differences (DiD) specification outlined in Equation 2. As Table 2, Panel A, Column 1 reports, we find that in the three years after the onset of the reform, one standard deviation difference in pre-reform overtime need—25 percentage points in the share of unplanned admissions in 1980—reduces in-hospital mortality by 2.45 deaths per thousand admissions (95% CI: -3.67; -1.24). Relative to the weighted pre-reform mean of 34.08 deaths per thousand admissions, this represents a 7.1% decline in in-hospital mortality for each standard deviation in overtime need. In Appendix Figure G.7 we show that the mortality decline is driven by wards where hours-for-workers substitution induced by overtime restrictions is larger.

The main concern is that wards were already in diverging trends pre-reform. However, consistent with no differential pre-trends across different treatment dosages, the dynamic effects show that pre-reform estimates are small and statistically indistinguishable from zero (Figure 8, Panel A). Therefore, it does not seem that in-hospital mortality differences across wards would have similarly changed without the reform.

Another concern with our main specification is that the dependent variable is in levels, which may not capture proportional changes in mortality rates across wards with different baseline levels of admissions. Appendix Table F.8 reports difference-in-differences estimates from a log specification. Three years after the onset of the reform, a one-standard-deviation increase in the dose variable is associated with about a 4.9% decline in in-hospital mortality (95% C.I.: -0.091; -0.007) for all admissions and a 7.1%

Figure 8. IN-HOSPITAL MORTALITY: EVENT-STUDY AT THE WARD LEVEL



Notes. The figure documents coefficients δ_t and β_t of the event-study specification as in Equation 1. Panel A presents estimated effects for the change in deaths in hospital per thousand admissions, among all admissions. Panel B presents the estimated effects for deaths in hospital originating from unplanned admissions per thousand unplanned admissions. Panel C presents the estimated effects for deaths in hospital originating from planned admissions per thousand planned admissions. All estimates are relative to the baseline year 1981, and results are weighted by the number of admitted patients in 1981. The dashed vertical lines indicate onset of the reform. The capped bars depict 95% confidence interval using robust standard errors clustered at the hospital ward level.

decline (95% C.I.: -0.111; -0.031) for unplanned admissions. However, the log specification excludes about 6% of ward-year observations with zero recorded deaths. As shown by [Chen and Roth \(2024\)](#) and [Mullahy and Norton \(2024\)](#), estimates can be sensitive in the presence of extensive-margin effects. We therefore also estimate a Poisson QMLE specification in levels, whose coefficients are interpretable as log point changes. Appendix Table F.9, Column 1, yields the same qualitative pattern as our main specification, but with a smaller proportional effect for all admissions (3.7%; 95% C.I.: -0.07; -0.004) and a similar magnitude for unplanned admissions.

We also run the same difference-in-differences specification at the level of patient admissions. Appendix Table F.10 shows that the decrease in the individual's probability of dying in hospital is consistent with our ward-level results. We find that, at the patient-level, one standard deviation increase in

the dose variable leads to 2.15 fewer deaths per thousand admissions, or 6.4% over the mean (Appendix Table F.10, Panel A, Column 1).

Table 2. EFFECTS OF THE REFORM ON MORTALITY

	Mortality Rates			
	In-hospital (1)	30-days (2)	60-days (3)	1-year (4)
<i>Panel A: All Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-2.45 (0.617) [-3.67,-1.24]	-2.27 (0.690) [-3.63,-0.91]	-1.93 (0.748) [-3.40,-0.46]	-1.27 (0.998) [-3.23,0.70]
R-Squared	0.962	0.968	0.972	0.981
Mean Dep. Var.	34.08	48.44	57.72	106.80
<i>Panel B: Unplanned Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-5.53 (1.262) [-8.01,-3.05]	-6.69 (1.566) [-9.77,-3.62]	-6.64 (1.634) [-9.85,-3.43]	-6.52 (1.752) [-9.96,-3.07]
R-Squared	0.944	0.952	0.957	0.972
Mean Dep. Var.	58.00	79.53	91.90	150.03
<i>Panel C: Planned Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-1.02 (0.520) [-2.05,-0.00]	-0.83 (0.710) [-2.23,0.56]	-0.88 (0.868) [-2.59,0.82]	0.27 (1.245) [-2.18,2.72]
R-Squared	0.807	0.845	0.863	0.957
Mean Dep. Var.	9.87	16.96	23.13	62.87
Observations	2,911	2,911	2,911	2,911
Nr. of Wards	424	424	424	424
Mean Exposure Var.	0.48	0.48	0.48	0.48
S.D. Exposure Var.	0.25	0.25	0.25	0.25

Notes. The table presents estimates of β in Equation 2 for the impact of the reform on mortality rates, using the sample of hospital wards between 1978 and 1984. Panel A presents results for all admissions, while Panels B and C show estimates only among unplanned and planned admissions, respectively. D_i is the standardized share of unplanned admissions in each ward i , in 1980. Each column refers to a different measure of patient mortality, measured per thousand admissions. In-hospital refers to the total number of deaths in the hospital ward per thousand admitted patients (Column 1). Columns 2 through 4 present mortality rates at different horizons post-discharge: 30 days, 60 days, and 1 year. All observations are weighted by the total number of admissions (Panel A), unplanned admissions (Panel B), or planned admissions (Panel C) in 1981. The mean of the dependent variable is computed pre-reform and similarly weighted. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

By Type of Admission.— Next, we examine the effects of the reform on different types of patients. We start by focussing on unplanned admissions—patients arriving through the emergency room or admitted on a non-elective basis. Acute patients have a stronger link to the margin affected by overtime restrictions. Acute cases arrive at relatively more unpredictable hours and generate most of weekend and nighttime work, covered by residents’ relatively longer hours pre-reform (see Section 4). If overtime restrictions have an effect through a fatigue channel, the effect should be relatively stronger among this type of patients. Moreover, contrary to planned admissions, unplanned admissions are typ-

ically nondeferrable. The effects on overall mortality could then partly reflect endogenous healthcare supply decisions, where planned admissions are deferred to later dates.

Figure 8, Panel B, and Table 2, Panel B, Column 1, replicate the event-study and DiD specifications on the set of unplanned admissions. We find that in-hospital mortality rates among unplanned cases fall by 5.53 for each thousand admissions, for one standard deviation increase in overtime need pre-reform (95% C.I.: -8.01, -3.05), or 9.5% relative to the mean. Among planned admissions, mortality rates also fall by about 1 less death for each thousand admissions (95% C.I.: -2.05, -0.00), or 10% relative to the pre-reform mean. As for all patient types, we find no evidence of differential pre-trends (Figure 8). We interpret the larger point estimate for acute cases—relative to the drop among all patients—as suggestive evidence that the reform operates through a fatigue channel, which is most relevant for acute care that takes place around the clock.

6.2. Post-discharge Mortality

Table 2 shows that improvements in patient survival extend well beyond the hospital stay. Columns 2–4 report deaths per thousand admissions at 30 days, 60 days, and one year after discharge. Tracking mortality beyond discharge is informative because in-hospital deaths may be affected by discharge timing or recording practices, while post-discharge mortality reflects patients’ underlying health after treatment and continuity of care. For all admissions combined, mortality declines remain visible up to one year post-discharge, though they gradually attenuate over time. Among unplanned admissions—the group most exposed to the reform—the reduction in mortality is more pronounced and persistent: a one standard deviation increase in exposure reduces 1-year mortality by 6.5 per thousand admissions, or 4.3% over the mean (Panel B, Column 4). By contrast, planned admissions show relatively short-lived changes (Panel C), suggesting that the gains are concentrated among emergency cases where timely availability of physicians is most critical. Since post-discharge deaths are recorded independently of hospital data, these patterns cannot be explained by changes in coding practices or discharge policies. Event-study estimates confirm the absence of differential pre-trends, reinforcing a causal interpretation of these sustained mortality improvements (Appendix Figure G.10).

6.3. Interpreting Changes in Mortality

A decline in mortality rates may reflect different causes. On the one hand, relatively more exposed hospital wards may prevent more deaths overall, conditional on the same intake of patients. On the other hand, wards may admit more patients without a commensurate rise in mortality. Moreover, a systematic shift in the type of patients admitted may not necessarily reflect an improvement in quality of care. Distinguishing between these sources of mortality decline is therefore essential for interpreting the effects of the reform.

Healthcare Supply.— To separate expanded capacity from improved survival conditional on admission, we re-estimate our main specifications using admissions and death counts—rather than mortality rates—as dependent variables.

Appendix Figure G.11 and Table F.11 show that the reform eased pre-existing supply constraints. A one standard deviation increase in overtime need led to roughly 58 additional yearly admissions per ward, corresponding to a 3 percent increase relative to the pre-reform mean of 1,879 (Table F.11, Panel A, Column 1). This expansion was driven by planned admissions, consistent with Figure G.12, which documents a sizeable reduction in the share of unplanned cases following the reform. We interpret this pattern as evidence that the reform expanded effective healthcare: hospitals were able to handle a larger caseload without a commensurate rise in deaths in hospital.

Patient Composition.— Mortality improvements could also arise if wards treated systematically healthier patients after the reform. To evaluate this possibility, we construct mortality measures that adjust for changes in patient characteristics.

We first compute age-adjusted mortality rates, dividing observed deaths by the expected number of deaths implied by each ward’s age distribution. Following Finkelstein et al. (2025), we use the predicted death rate in different age bins²⁰ in the first year of our sample (1978), to compute the expected death rate in every ward-year cell based only on the age distribution. The second measure accounts for differences in diagnostic mix. Using data on over 790 thousand hospitalizations in 1978, we estimate a logit model of mortality on age bins and diagnostic categories to predict each patient’s conditional probability of dying in hospital (as in Muñoz and Otero, 2025). Diagnoses are grouped into 17 Charlson comorbidity categories (Charlson et al., 1987) and one residual low-risk group.²¹ Our risk-adjusted mortality measure is the ratio of observed to predicted deaths in each ward-year, rescaled by the overall 1978 death rate to yield coefficients that can be interpreted as deaths per thousand admissions.

Appendix Table F.13 reports age-adjusted mortality results for in-hospital, 30-day, and 1-year mortality. The estimates are qualitatively similar to our baseline mortality results, with generally negative effects across horizons and larger declines among unplanned admissions, suggesting that changes in age composition alone do not explain the mortality improvements in more exposed wards. Appendix Figure G.8 shows that these results are not confounded by differential pre-trends.

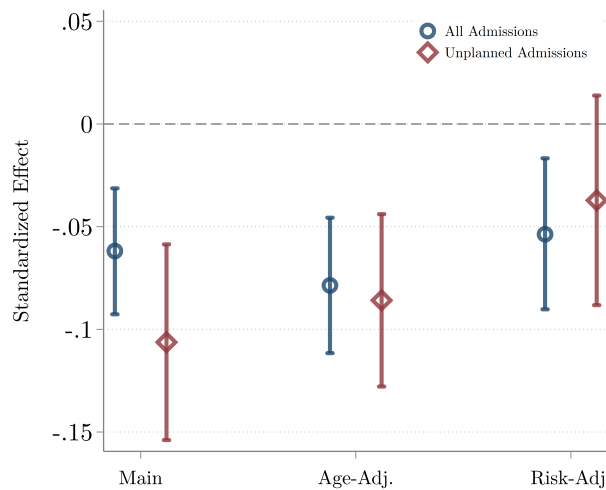
Appendix Table F.14 shows the results for the risk-adjusted in-hospital mortality rates. We find that a one standard deviation increase in the proxy of overtime need leads to a drop in 1.19 deaths per thousand admissions (95% C.I.: -2.00; -0.37), or about 4.2% over the pre-reform risk-adjusted mean

²⁰ The age bins are 0 for new-borns and then 5 year intervals up until 89. The last bin includes all patients aged 90 and above.

²¹ Charlson comorbidity groups are identified from ICD-8 diagnostic codes in Danish registry data.

(Column 1). These results show that in relatively more exposed wards deaths decreased more than what would be predicted by their patient death risk. Restricting to unplanned admissions we also see a decline of 1.85 risk-adjusted deaths per thousand admissions, though with limited statistical precision (95% C.I.: -4.39; 0.69). Appendix Figure G.9 reports the coefficients in an event-study specification, showing no evidence of differential pre-trends.

Figure 9. STANDARDIZED DiD ESTIMATES FOR ADJUSTED IN-HOSPITAL MORTALITY RATES



Notes. The figure summarizes the difference-in-differences estimates of Equation 2 for adjusted mortality outcomes, comparing the baseline mortality specification to age-adjusted and risk-adjusted, constructed as described in the main text. Age-adjusted mortality is computed by dividing observed deaths by the expected number of deaths implied by each ward's age distribution, using predicted death rates in different age bins from 1978. Risk-adjusted mortality is computed as the ratio of observed to predicted deaths in each ward-year, where predictions are based on a logit model of mortality on age bins and diagnostic categories estimated on 1978 data, rescaled by the overall 1978 death rate. All observations are weighted by the total number of admissions in 1981. For comparability, we standardize effect sizes by dividing by the weighted standard deviation of the dependent variable in the pre-reform period, with weights given by the number of admissions in the ward in 1981. Point estimates are shown with 95% confidence intervals computed from robust standard errors clustered at the hospital ward level.

Figure 9 summarizes these results. It plots the standardized difference-in-differences estimates for in-hospital mortality under the baseline, age-adjusted, and risk-adjusted outcome definitions. The point estimates are generally smaller in magnitude for the adjusted mortality measures, but they remain negative. This suggests that while changes in patient composition may account for part of the observed mortality decline, they do not fully explain it. The reform led to improvements in patient survival beyond what would be expected based on changes in age or diagnostic mix alone.

In sum, these results indicate that the decline in mortality rates following the reform cannot be fully explained by changes in patient composition or shifts in healthcare supply. Wards more affected by overtime restrictions expanded their capacity to admit patients without experiencing a proportional rise in deaths, and mortality improvements persist even after adjusting for patient age and diagnostic mix. The evidence therefore points to gains in healthcare efficiency. Hospitals treated more patients with fewer deaths per admission, consistent with improvements in quality of care rather than just

shifts towards lower-risk patients.

6.4. Other Outcomes: Readmissions and Length of Stay

An important concern is that lower in-hospital mortality rates and additional admissions may reflect earlier patient discharges or other changes that merely postpone deaths to post-discharge periods. To assess whether overtime restrictions altered the production function of hospital wards or compromised overall quality of care, we examine two outcomes: (i) readmission rates 14, 30, and 60 days after discharge, and (ii) the average length of stay (LOS) per admission.

Table 3. EFFECTS ON READMISSION RATES AND LENGTH OF STAY

	Readmission Rates			Length of Stay
	14-days (1)	30-days (2)	60-days (3)	(4)
<i>Panel A: All Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.60 (1.164) [-2.89, 1.69]	-2.73 (2.304) [-7.26, 1.80]	-1.89 (2.219) [-6.25, 2.47]	-0.14 (0.103) [-0.35, 0.06]
R-Squared	0.846	0.922	0.938	0.916
Mean Dep. Var. Pre-Reform	59.01	97.77	142.87	9.31
<i>Panel B: Unplanned Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	0.15 (1.198) [-2.20, 2.51]	1.27 (1.729) [-2.13, 4.67]	2.39 (1.984) [-1.51, 6.29]	-0.15 (0.127) [-0.40, 0.10]
R-Squared	0.698	0.797	0.846	0.851
Mean Dep. Var. Pre-Reform	59.10	96.35	142.15	10.52
<i>Panel C: Planned Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	1.20 (1.730) [-2.20, 4.60]	-2.04 (4.391) [-10.67, 6.59]	0.10 (4.391) [-8.54, 8.73]	-0.24 (0.163) [-0.56, 0.08]
R-Squared	0.878	0.930	0.939	0.866
Mean Dep. Var. Pre-Reform	58.41	98.81	143.05	8.16
Observations	2,911	2,911	2,911	2,911
Nr. of Wards	424	424	424	424
Mean Exposure Var.	0.48	0.48	0.48	0.48
S.D. Exposure Var.	0.25	0.25	0.25	0.25

Notes. The table presents estimates of β in Equation 2. The sample includes all hospital wards with at least one observation in each reform period (see Section C.1 for more details). D_i is the standardized share of unplanned admissions in each ward i , in 1980. Readmissions refer to the total number of admitted patients that are readmitted in the same or any other hospital ward within 14, 30, or 60 days after discharge from hospital ward i per thousand admitted patients (Columns 1-3). The dependent variable in Column 4 is the number of days the average admitted patient spends in ward i . All observations are weighted by the number of total patient admissions in 1981. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

Table 3 reports the difference-in-differences estimates, using the same specification as in Equation 2. Across all time horizons, the point estimates for readmission rates are negative but statistically imprecise, with 95 percent confidence intervals that include zero (Columns 1–3). Likewise, the estimated change in average LOS is small and not statistically significant (-0.14; 95% C.I.: -0.35; 0.06). Event-study

results show no evidence of differential pre-trends for either outcome (Appendix Figure G.13).

When disaggregating by admission type, the results have a qualitatively identical interpretation. Among unplanned (acute) admissions, where treatment urgency is particularly important, there is no indication that physicians shortened hospital stays or discharged patients earlier (Table 3, Panel B). Among planned admissions, where hospitals could have potentially used shorter stays to expand capacity, we also find no evidence of reduced LOS or higher readmission risk, with noisy estimates statistically indistinguishable from zero (Table 3, Panel C).

In sum, these results suggest that the supply expansion and mortality reductions documented above are not driven by shorter stays, premature discharges, or higher patient turnover. Instead, the evidence points to hospital wards maintaining at least similar lengths of stay and readmission patterns while achieving better patient survival. These findings are consistent with efficiency gains from the reform rather than changes in discharge behavior.

6.5. Robustness and Sensitivity

To ensure the robustness of our main results we examine the sensitivity of our findings to alternative: (i) proxies of overtime need; (ii) sample restrictions; (iii) specifications; and (iv) permutation tests.

Alternative Dosages.— As argued in Section 4.1, our preferred measure of treatment intensity—hospital wards’ pre-reform share of unplanned admissions—captures their exposure to overtime restrictions. Nonetheless, one may worry that it is not the only or most relevant source of identifying variation. To address this concern, we re-estimate our main specifications using alternative proxies for *overtime need*: the pre-reform shares of weekend and nighttime admissions, which also reflect exposure to the scheduling constraints imposed by the reform (Section 2).

Appendix Figure G.14 presents the event-study estimates based on these alternative measures. Across all definitions, we find qualitatively similar patterns: mortality rates decline more in wards with greater exposure to overtime restrictions. Estimated magnitudes are somewhat smaller when exposure is measured by nighttime admissions, particularly among unplanned cases, where estimates are less precise.

Finally, to capture the multidimensional nature of scheduling intensity, we construct a composite exposure index. We compute, for each ward, the mean and within-year variation of unplanned, weekend, and nighttime admissions, and extract their common component through principal-component analysis (see Appendices C and D). Using this latent measure yields effect estimates that are quantitatively and qualitatively similar to our baseline specification (Appendix Figure G.15).

These findings show that the estimated mortality effects are robust to alternative definitions of

exposure and persist across multiple dimensions of scheduling intensity, reinforcing the interpretation that the reform’s impact operates through restrictions on overtime work, as well as weekend and night shifts.

Alternative Samples.— We test robustness using different analysis samples. First, we re-run our difference-in-differences specification in a balanced sample of hospital wards. Because our main analysis uses an unbalanced sample, it could be that effects are driven by hospital units that close down or are consolidated across the period, and affect the composition of the observations post-reform. To allay concerns with this type of attrition, we replicate our findings using a balanced sample of hospital wards, restricting the analysis only to those wards that admit patients in every year throughout the period between 1978 and 1984. The restricted sample thus excludes observations from 36 wards in our main sample. Appendix Table F.15 reports the results for this balanced sample, showing that the estimates remain similar, even if slightly less precise.

Second, we re-estimate our main specifications excluding gynecology and obstetrics (GynOb) wards. These wards treat substantially larger number of patients per year than those in other specialties, giving them a higher influence in our estimates. To ensure that our results are not driven by this differences in scale across wards, we repeat the analysis after removing GynOb wards from the sample. Appendix Table F.16 shows that the qualitative interpretation of our main findings is unchanged.

Other Specifications.— Variation in our dosage variable arises primarily from differences across medical specialties and in the type of local healthcare demand faced by each hospital ward. One concern is that post-reform differences could instead reflect unobserved location- or specialty-specific trends. To address this, we augment our main specification by adding year-by-county (λ_{tc}) or year-by-specialty-group fixed effects (λ_{ts}). The former restricts within-year comparisons to wards operating in the same county, while the latter limits comparisons to wards within the same specialty group.²² Tables F.7–F.12 report the corresponding difference-in-differences estimates across different columns.

Results including year-by-county fixed effects are quantitatively similar to our baseline estimates, suggesting that spatially correlated shocks are unlikely to drive our findings. When we instead include year-by-specialty fixed effects, the estimated effects become smaller and statistically less precise. This attenuation is expected: adding λ_{ts} absorbs much of the relevant identifying variation, since pre-reform differences in unplanned admissions are largely explained by variation across specialties (Appendix D, Table D.1). By comparing only wards within the same specialty in a given year, this specification removes meaningful cross-specialty variation, such as between high-intensity acute specialties and

²² Specialties are grouped into General Surgery, Specialized Surgery, Gynecology and Obstetrics, Internal Medicine (General), Internal Medicine (Specialties), and Other Specialties. Appendix Table F.6 lists the specific specialties in each group.

lower-intensity ones. Once this between-specialty variation is differenced out, the remaining within-specialty variation is limited. Nevertheless, the point estimates remain negative, although with confidence intervals that include zero.

Permutation Tests.—To assess whether our findings could result from spurious correlations between wards’ pre-reform overtime exposure and subsequent mortality changes, we perform a permutation test. Specifically, we randomly reassign the pre-reform share of unplanned admissions across hospital wards and re-estimate our main specification using these placebo exposure values. Repeating this procedure 1001 times yields a distribution of simulated treatment effects. Appendix Figure G.16 compares this placebo distribution to our baseline estimates, and by type of admissions. Our estimated effects lie far in the lower tail of the simulated distributions, indicating that the observed mortality decline is unlikely to have arisen by chance.

Alternative Inference.—We re-estimate our difference-in-differences specification (Equation 2) using alternative assumptions about the correlation structure of the error term. Appendix Figure G.17 reports 95% confidence intervals based on standard errors clustered at the hospital-ward, hospital, county, and specialty-group levels. Across all clustering choices, the estimated effects on in-hospital and 1-year mortality remain qualitatively unchanged, confirming that our main results are not sensitive to different clustering of the standard errors.

7. Conclusion

In this paper, we exploit a strong restriction on physicians’ overtime hours and scheduling to study how the allocation of hours across workers shapes productivity in public hospitals. Leveraging ward-level differences in pre-reform overtime need, we show that the 1982 Danish overtime restriction reform triggered a sizeable substitution from hours to workers. For each standard deviation in our intensity measure of exposure to the policy, wards hired 14% more resident physicians. We find that in-hospital mortality rates fell by an average of 7% for a one standard deviation increase in exposure to the reform. The labor demand shock seems to have lifted previous capacity constraints, allowing hospital wards to admit relatively more planned, non-acute patients. Nevertheless, larger impacts among unplanned admissions and reductions in risk-adjusted mortality rates suggest that improvements in hospital performance were not fully driven by changes in patient composition post-reform. Importantly, the effects are not short-lived and extend to out-of-hospital mortality at least until one year after discharge. Readmission rates and the average length of stay did not change with the new policy indicating that improvements in hospital performance were not offset by greater post-discharge healthcare utilization.

Our findings suggest that binding overtime restrictions can induce efficiency gains in public hospi-

tals. By forcing a reallocation of work toward additional personnel, such policies may relieve fatigue-related inefficiencies and enhance patient safety. The results illustrate how targeted labor market regulations can correct misaligned incentives in sectors where productivity is difficult to monitor or price. From a policy perspective, our findings contribute to the long-standing debate about resident physician work hours. While the weekly working hours of the average worker in the economy have been steadily declining over the last decades, physicians' workweeks have significantly lagged behind this trend. Our results show that the job-sharing policy in Denmark was effective in bringing more physicians into the profession without harming patients' health. Crucially, the new work hours arrangement was implemented in a context of excess labor supply prior to the reform, and would likely be less effective when shortages of professionals are an important concern.

In ongoing work, we are investigating why hospital performance gains were not achieved before the reform. Even if shorter-hours allocations were feasible, they may not have been privately optimal given prevailing frictions. Overtime is a highly flexible margin for covering demand spikes, helping to maintain continuity of care and deal with non-deferrable work. But overtime can also reduce performance through fatigue, with costs that are difficult to measure and therefore not fully internalized in scheduling decisions. In addition, wage rigidity, hiring and training frictions, or coordination problems can sustain an overtime-heavy equilibrium even when adding physicians would raise performance. By constraining overtime and forcing substitution toward the extensive margin, the reform may have shifted hospitals to a higher-performance labor allocation that was previously hard to implement or not incentive-compatible. Clarifying these mechanisms can inform when overtime restrictions are likely to raise productivity in other long-hours sectors facing similar distortions.

References

- Akerlof, George (1976) “The economics of caste and of the rat race and other woeful tales,” *Quarterly Journal of Economics*, 90 (4), 599–617, [10.2307/1885324](#). 2
- Asai, Kentaro, Marta C Lopes, and Alessandro Tondini (2024) “Firm-Level Effects of Reductions in Working Hours,” *Job Market Paper*. 5
- Batut, Cyprien, Andrea Garnero, and Alessandro Tondini (2023) “The employment effects of working time reductions: Sector-level evidence from European reforms,” *Industrial Relations*, 62 (3), 217–232, [10.1111/irel.12323](#). 5
- Bolster, Lauren and Liam Rourke (2015) “The Effect of Restricting Residents’ Duty Hours on Patient Safety, Resident Well-Being, and Resident Education: An Updated Systematic Review,” *Journal of Graduate Medical Education*, 7 (3), 349–363, [10.4300/JGME-D-14-00612](#). 1. 6
- Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant’Anna;Anna (2024) “Difference-in-Differences with a Continuous Treatment,” *arXiv:2107.02637v4*, [10.2139/ssrn.4716682](#). 4, 22
- Card, David, Carlos Dobkin, and Nicole Maestas (2009) “Does medicare save lives?,” *Quarterly Journal of Economics*, 124 (2), 597–636, [10.1162/qjec.2009.124.2.597](#). 14, 18
- Chan, David C. (2018) “The Efficiency of Slacking off: Evidence From the Emergency Department,” *Econometrica*, 86 (3), 997–1030, [10.3982/ecta13565](#). 2, 6, 12
- Chan, David C, David Card, and Lowell Taylor (2023) “Is There a VA Advantage? Evidence from Dually Eligible Veterans,” *American Economic Review*, 113 (11), 3003–3043, [10.1257/aer.20211638](#). 14
- Chan, David C, Matthew Gentzkow, and Chuan Yu (2022) “Selection with Variation in Diagnostic Skill: Evidence from Radiologists,” *Quarterly Journal of Economics*, 137 (2), 729–783, [10.1093/qje/qjab048](#). 6
- Chan, David and Yiqun Chen (2023) “The Productivity of Professions: Evidence from the Emergency Department,” *Working Paper*. 6
- Charlson, Mary E., Peter Pompei, Kathy L. Ales, and C. Ronald MacKenzie (1987) “A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation,” *Journal of Chronic Diseases*, 40 (5), 373–383, [10.1016/0021-9681\(87\)90171-8](#). 28, 47
- Chemin, Matthieu and Etienne Wasmer (2009) “Using Alsace-Moselle local laws to build a difference-in-differences estimation strategy of the employment effects of the 35-hour workweek regulation in France,” *Journal of Labor Economics*, 27 (4), 487–524, [10.1086/605426](#). 5

- Chen, Jiafeng and Jonathan Roth (2024) “Logs with Zeros? Some Problems and Solutions,” *Quarterly Journal of Economics*, 139 (2), 891–936, [10.1093/qje/qjad054](#). 21, 25
- Chen, Yiqun (2021) “Team-Specific Human Capital and Team Performance: Evidence from Doctors,” *American Economic Review*, 111 (12), 3923–3962, [10.1257/aer.20201238](#). 6
- Christy, Morten (1983) “Sygehusenes forbrug af lægetimer og lægestillinger 1972-1982,” *Ugeskrift for læger*, 145 (28), 2199–2201. 8, 11, 12
- (1984) “Vending af Vagtpyramiden,” *Ugeskrift for Læger*, 146 (4), 331–332. 8
- Crépon, Bruno and Francis Kramarz (2002) “Employed 40 hours or not employed 39: Lessons from the 1982 mandatory reduction of the workweek,” *Journal of Political Economy*, 110 (6), 1355–1389, [10.1086/342807](#). 5
- Currie, Janet M and W. Bentley MacLeod (2020) “Understanding Doctor Decision Making: The Case of Depression Treatment,” *Econometrica*, 88 (3), 847–878, [10.3982/ecta16591](#). 6
- Currie, Janet and W Bentley Macleod (2017) “Diagnosing Expertise: Human Capital, Decision Making, and Performance among Physicians,” *Journal of Labor Economics*, 35 (1), 1–43, [10.2307/26553249](#). 6
- Dodini, Samuel, Petter Lundborg, Katrine Løken, and Alexander Willén (2025) “The Fatal Consequences of Brain Drain,” *IZA DP No. 17819*, [www.iza.org](#). 6
- Doyle, Joseph J, Steven M Ewer, and Todd H Wagner (2010) “Returns to physician human capital: Evidence from patients randomized to physician teams,” *Journal of Health Economics*, 29 (6), 866–882, [10.1016/j.jhealeco.2010.08.004](#). 6
- Doyle, Joseph J., John A. Graves, Jonathan Gruber, and Samuel A. Kleiner (2015) “Measuring returns to hospital care: Evidence from ambulance referral patterns,” *Journal of Political Economy*, 123 (1), 170–214, [10.1086/677756](#). 14
- Fadlon, Itzik, Frederik Plesner Lyngse, and Torben Heien Nielsen (2024) “Early Career Setbacks and Women’s Career-Family Trade-Off,” *CEBI Working Paper 06/22*, [www.cebi.ku.dk](#). 8, 22
- Finkelstein, Amy, Matthew Notowidigdo, Frank Schilbach, and Jonathan Zhang (2025) “Lives vs. Livelihoods: The Impact of the Great Recession on Mortality and Welfare,” *Quarterly Journal of Economics*, 140 (3), 2269–2328, [10.2139/ssrn.4717244](#). 14, 28, 46
- Fishback, Price, Chris Vickers, and Nicolas L Ziebarth (2024) “Labor Market Effects of Workweek Restrictions: Evidence from the Great Depression,” *American Economic Journal: Macroeconomics*, 2024 (4), 77–113, [10.1257/mac.20220188](#). 5
- Gaynor, Martin, Rodrigo Moreno-Serra, and Carol Propper (2013) “Death by Market Power: Reform, Competition, and Patient Outcomes in the National Health Service,” *American Economic Journal: Economic Policy*, 5 (4), 134–166, [10.1257/pol.5.4.134](#). 6, 14

- Germain, Antoine (2025) “Working time reductions and monopsony power,” *Job Market Paper*. 2, 5
- Ginja, Rita, Julie Riise, Barton Willage, and Alexander Willén (2025) “Does Your Doctor Matter?,” *Journal of Political Economy: Microeconomics*, 3 (3), 497–538, [10.1086/732838](https://doi.org/10.1086/732838). 6
- Gunnesmo, Marius Fredagsvik and Casper Worm Hansen (2025) “Labor-Market Effects of Introducing the 8-Hour Workday,” *CEPR Discussion Paper No. DP20785*, www.cepr.org. 5
- Hansen, Else (2015) “Universiteter og idealer i velfærdsstaten,” *Uddannelseshistorie (year)*, 76–91. 8
- (2018) “Planlægningens (oversete) årti. Argumenter for adgangsbe­grænsning i 1970’erne,” *Historisk Tidsskrift*, 118 (1), 77–112, <https://tidsskrift.dk/historisktidsskrift/article/view/106531>. 8
- Huettner, Frank and Marco Sunder (2012) “Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values,” *Electronic Journal of Statistics*, 6, 1239–1250, [10.1214/12-EJS710](https://doi.org/10.1214/12-EJS710). 49
- Hunt, Jennifer (1999) “Has Work-Sharing Worked in Germany?,” *Quarterly Journal of Economics*, 114 (1), 117–148, <https://www.jstor.org/stable/2586949>. 2, 5, 41
- Israeli, Osnat (2007) “A Shapley-based decomposition of the R-Square of a linear regression,” *Journal of Economic Inequality*, 5 (2), 199–212, [10.1007/s10888-006-9036-6](https://doi.org/10.1007/s10888-006-9036-6). 49
- Kawaguchi, Daiji, Hisahiro Naito, and Izumi Yokoyama (2017) “Assessing the effects of reducing standard hours: Regression discontinuity evidence from Japan,” *Journal of the Japanese and International Economies*, 43, 59–76, [10.1016/j.jjie.2016.12.002](https://doi.org/10.1016/j.jjie.2016.12.002). 5
- Landers, Renée M, James B Rebitzer, and Lowell J Taylor (1996) “Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms,” *American Economic Review*, 86 (3), 329–348. 2
- Lorentzen, Kathrine Aaby (2024) “Family-friendly Jobs and Occupational Sorting across Gender : Evidence from Introduction of a Maximum 40-hour Workweek,” *Working Paper* (9061). 5, 8, 9, 13
- Maoz Breuer, Rina, Ruth Waitzberg, Adin Breuer et al. (2023) “Work like a Doc: A comparison of regulations on residents’ working hours in 14 high-income countries,” *Health Policy*, 130, 104753, [10.1016/j.healthpol.2023.104753](https://doi.org/10.1016/j.healthpol.2023.104753). 2
- Møller, Jes Fabricius and Sniff Andersen Nexø (2004) *Mellem kald og kamp - Yngre læger i 100 år*, København: Maanedsskrift for Praktisk Lægegering. 8
- Mullahy, John and Edward C Norton (2024) “Why Transform Y? The Pitfalls of Transformed Regressions with a Mass at Zero*,” *Oxford Bulletin of Economics and Statistics*, 86 (2), 417–447, [10.1111/obes.12583](https://doi.org/10.1111/obes.12583). 21, 25
- Muñoz, Pablo and Cristóbal Otero (2025) “Managers and Public Hospital Performance,” *American Economic Review*, 115 (11), 4040–4074, [10.1257/aer.20240615](https://doi.org/10.1257/aer.20240615). 6, 14, 28, 46

- Nadiri, MI and Sherwin Rosen (1969) “Interrelated factor demand functions,” *American Economic Review*, 59 (4), 457–471, <http://www.jstor.org/stable/10.2307/1813210>. 2
- Oi, Walter Y. (1962) “Labor as a Quasi-Fixed Factor,” *Journal of Political Economy*, 70 (6), 538–555, [10.1086/258715](https://doi.org/10.1086/258715). 2
- Pencavel, John (2015) “The Productivity of Working Hours,” *The Economic Journal*, 125 (589), 2052–2076, [10.1111/ECOJ.12166](https://doi.org/10.1111/ECOJ.12166). 2
- Propper, Carol and John Van Reenen (2010) “Can pay regulation kill? Panel data evidence on the effect of labor markets on hospital performance,” *Journal of Political Economy*, 118 (2), 222–273, [10.1086/653137](https://doi.org/10.1086/653137). 6
- Sánchez, Rafael (2013) “Do reductions of standard hours affect employment transitions?: Evidence from Chile,” *Labour Economics*, 20, 24–37, [10.1016/j.labeco.2012.10.001](https://doi.org/10.1016/j.labeco.2012.10.001). 5
- Schlenker, Oliver (2025) “The Deadly Consequences of Labor Scarcity : Evidence from Hospitals,” *EXC Working Paper No. 42*. 6
- Skuterud, Mikal (2007) “Identifying the potential of work-sharing as a job-creation strategy,” *Journal of Labor Economics*, 25 (2), 265–287, [10.1086/511379](https://doi.org/10.1086/511379). 5
- Temple, John (2014) “Resident duty hours around the globe: Where are we now?,” *BMC Medical Education*, 14 (1), [10.1186/1472-6920-14-S1-S8](https://doi.org/10.1186/1472-6920-14-S1-S8). 2
- Wasserman, Melanie (2023a) “Do Long Work Hours Constrain Women’s Fertility Choices?,” *Working Paper*. 5
- (2023b) “Hours Constraints, Occupational Choice, and Gender: Evidence from Medical Residents,” *Review of Economic Studies*, 90 (3), 1535–1568, [10.1093/restud/rdac042](https://doi.org/10.1093/restud/rdac042). 5, 13

A. Institutional Background on Working Hours and Shift Schedules

A.1. Pre-Reform Regime (1978–1981)

Before the 1982 working hour system, hospital physicians' weekly schedules comprised two main components: scheduled "daytime duty" and unscheduled "on-call duty," with overtime and rest rules layered on top.

Daytime duty.— Physicians on wards with on-call obligations worked 6 hours and 20 minutes per weekday, beginning no earlier than 07:45 and no later than 10:00. In units without on-call responsibilities, the standard weekday shift was eight hours long.

On-call duty.— On-call covered the hours from the end of that weekday shift until the start of the next day-time duty (17 hours and 40 minutes). Hospitals classified on-call into three settings (off-site, at home, or on-site) and four "shift-classes" (A–D) according to expected active hours: Class A (0-2 hours); Class B (2-4 hours); Class C (4-6 hours); and Class D (6-8 hours).

Overtime.— Any duty beyond the scheduled daytime shift on a non-on-call day (at least half an hour) counted as overtime. In Class D on-call, hours worked beyond an eight-hour "effective" threshold were also treated as overtime.

Rest Periods.— After completing an on-call shift, a physician was guaranteed at least five hours of uninterrupted rest before the next daytime duty.

A.2. Post-Reform Regime

The 1982 collective agreement transformed duty classification into a single "normal duty" category plus on-call, tightened rest requirements, and raised overtime premiums.

Normal Duty.— Ward managers schedule planned work—daytime, evening, or night—within a new weekly hour cap.

- *Daytime:* Starts between 06:00 and 10:00 and must end by 18:00; at least twenty of the weekly hours must fall between 08:00 and 16:00 (twenty-hour rule).
- *Evening:* Begins no earlier than 14:00 and extends past 17:00.
- *Night:* Begins no earlier than 22:00.

On-call Duty.— Encompasses any hours outside normal duty, including optional twenty-four-hour weekend and holiday shifts. Availability factors—fractions of each on-call shift deemed "active" for pay—are set at one-third when doctors stay at home and two-thirds when they remain on-site; twenty-four-hour shifts carry higher factors (14 hours at home, 20 hours on-site).

Rest Periods.— The reform pledged compliance with national rest laws (eleven hours rest per twenty-four-hour period, one full day off per week²³), while allowing limited local waivers:

- *Daily rest:* May be negotiated down to eight hours: Once per twenty-four-hour period for normal duty, or up to ten such days per month for on-call—provided lost rest is compensated later.
- *Weekly rest:* May be rescheduled so that no more than twelve days separate two full rest periods.

Overtime Pay.— Effective on-call hours carry an additional one-third premium. From April 1982, this rose to one-half of the base pay.

²³ Bekendtgørelse nr. 372 af 15. august 1980 om hviletid og fridøgn.

B. Model

B.1. Setup

We consider an economy with a finite number of firms, H . In our model, firms are hospital wards that have no profit motive, but seek to maximize a healthcare output metric Y , subject to an exogenous budget B . Firms choose two types of labor inputs: hours-per-worker (h) and workers (N). We start by assuming that workers are homogeneous.

We also assume a simple Cobb-Douglas production function, given by:

$$Y(h, N) = hN^{1/\theta}, \quad \theta \in (0, 1). \quad (4)$$

This formulation makes the key trade-off transparent. Wards can raise output by extending shifts (h) or by hiring more workers (N). The θ parameter governs the marginal rate of technical substitution between h and N . As $\theta \rightarrow 1$, there are weaker returns to adding workers, so increasing hours is relatively more attractive. A lower θ implies stronger returns to adding workers, so hiring is relatively more attractive than longer shifts.

We assume that wards are heterogeneous in θ , such that there is a distribution $F(\theta)$ over hospital wards in the economy. Heterogeneity across hour productivity is a natural assumption in our context, as different hospital wards have different types of patients and medical conditions, which may require different levels of attention and care. For example, wards that treat more acute patients may have a higher θ , as they may require more hours of availability per patient.

B.2. Hospital Wards Problem Before the Reform

The problem faced by hospital wards before the reform is similar to the one in [Hunt \(1999\)](#), but we abstract from labor-capital substitution and wards do not maximize profit. Wards have a convex pay schedule provided by the triplet $\{w, p, \bar{h}\}$. Workers are compensated at rate w per hour up until the standard hours limit $\bar{h}$. For each additional hour worked above $\bar{h}$, workers are paid with an overtime premium $(1 + p)w$. Each worker has a fixed cost f . We assume that the fixed cost per worker is sufficiently large to justify the use of overtime hours. Furthermore, the overtime premium is such that $p < \frac{f}{w\bar{h}}$.²⁴ Contrary to the case of profit-maximizing firms, each hospital ward has a budget constraint B .

Each hospital ward faces the following problem:

²⁴ As we show later, this condition ensures that the optimal overtime solution is strictly positive.

$$\begin{aligned}
\max_{h,N} \quad & Y(h, N) - N [wh + wp\mathbb{1}\{h > \bar{h}\}(h - \bar{h}) + f], \\
\text{s.t.} \quad & N [wh + wp\mathbb{1}\{h > \bar{h}\}(h - \bar{h}) + f] \leq B, \\
& h \geq 0, N \geq 0.
\end{aligned} \tag{5}$$

The solution to the problem above leads us to the first results regarding the optimal scheduling of hours in terms of θ . Given the convex budget constraint, wards start by solving two separate interior problems, depending on whether h falls below or above standard hours $\bar{h}$.

Result 1. *The ward has two possible interior solutions to the optimal labor allocation across h and N :*

- (i) *Without overtime:* $h_0^*(\theta) = \frac{\theta}{1-\theta} \frac{f}{w}$; $N_0^*(\theta) = (1 - \theta) \frac{B}{f}$.
- (ii) *With overtime:* $h_1^*(\theta) = \frac{\theta}{1-\theta} \frac{f - wp\bar{h}}{w(1+p)}$; $N_1^*(\theta) = (1 - \theta) \frac{B}{f - wp\bar{h}}$.

Proof. Straightforwardly derived from the first order conditions (FOC) of the ward's problem in Equation 5. First, we derive the FOC for the case where $h \leq \bar{h}$ and solve for h to get h_0^* . We use the binding constraint to then derive N_0^* . Second, we derive the FOC for when $h > \bar{h}$ and get h_1^* and N_1^* . $\square$

Whether each of the solutions is implemented depends on the relationship between θ and the wage schedule parameters $\{w, p, \bar{h}, f\}$. We start by dividing the support of θ into two feasibility regions: (i) one where h_0^* is feasible, and (ii) one where h_1^* is feasible.

Result 2. *Overtime feasibility regions are such that:*

- (i) *For $\theta \in (0, \theta_0]$, where $\theta_0 \equiv \frac{w\bar{h}}{f + w\bar{h}}$: $h_0^*(\theta)$ is feasible.*
- (ii) *For $\theta \in (\theta_1, 1)$, where $\theta_1 \equiv \frac{w(1+p)\bar{h}}{f + w\bar{h}}$: $h_1^*(\theta)$ is feasible.*
- (iii) *For $\theta \in (\theta_0, \theta_1]$: $h^* = \bar{h}$.*

Proof. First, we show the derivation of θ_0 .

If the interior solution for the no-overtime hours allocation h_0^* falls in the overtime region such that $h_0^* > \bar{h}$, the hospital ward can only rely on the overtime solution h_1^* , as h_0^* is not feasible. It must then be that:

$$h_0^*(\theta) = \frac{\theta}{1-\theta} \frac{f}{w} \leq \bar{h} \iff \theta \leq \frac{w\bar{h}}{f + w\bar{h}} \equiv \theta_0. \tag{6}$$

So, for $h_0^*(\theta)$ to be implementable, $\theta \in (0, \theta_0]$, as in condition (i), Result 2.

On the other hand, if the interior solution for the overtime hours allocation h_1^* falls in the no-overtime region such that $h_1^* \leq \bar{h}$, the hospital ward can only rely on the no-overtime solution h_0^* , as h_1^* is not feasible. It must be that:

$$h_1^*(\theta) = \frac{\theta}{1-\theta} \frac{f - wp\bar{h}}{w(1+p)} > \bar{h} \iff \theta > \frac{w(1+p)\bar{h}}{f + w\bar{h}} \equiv \theta_1. \quad (7)$$

So, for $h_1^*(\theta)$ to be implementable, $\theta \in (\theta_1, 1)$, as in condition (ii), Result 2.

It is straightforward to show that these two feasibility regions do not overlap, as $\theta_1 > \theta_0$. By simplifying the inequality:

$$\frac{w(1+p)\bar{h}}{f + w\bar{h}} > \frac{w\bar{h}}{f + w\bar{h}} \iff p > 0, \quad (8)$$

which is true by assumption.

In the region where $\theta \in (\theta_0, \theta_1]$, neither of the interior solutions are feasible, and the hospital ward must rely on the corner solution $h^* = \bar{h}$, as in condition (iii), Result 2.

□

Given these conditions, the optimal allocation of hours and workers is summed up in Result 3. Figure B.1 illustrates the optimal allocation choices for each θ .

Result 3. *The hospital ward chooses the following allocation of hours, depending on its type θ :*

$$(h^*(\theta), N^*(\theta)) = \begin{cases} \left(\frac{\theta}{1-\theta} \frac{f}{w}, (1-\theta) \frac{B}{f} \right), & \theta \in (0, \theta_0], \\ \left(\bar{h}, \frac{B}{w\bar{h}+f} \right), & \theta \in (\theta_0, \theta_1], \\ \left(\frac{\theta}{1-\theta} \frac{f - wp\bar{h}}{w(1+p)}, (1-\theta) \frac{B}{f - wp\bar{h}} \right), & \theta \in (\theta_1, 1), \end{cases} \quad (9)$$

where $\theta_0 = \frac{w\bar{h}}{f + w\bar{h}}$ and $\theta_1 = \frac{w(1+p)\bar{h}}{f + w\bar{h}}$.

B.3. Overtime Restrictions

In our model we assume overtime restrictions ban any shift length above the standard hours cap ($\bar{h}$), equivalent to a hard constraint $h \leq \bar{h}$ on the choices of hospital wards. Overtime allocations become infeasible, making the cost schedule linear.

The implications are straightforward. For any ward that was using overtime ($\theta > \frac{w\bar{h}}{f + w(1+p)\bar{h}}$), the new feasible allocation is attained by substituting from hours to workers until the budget constraint binds. The change in the optimal policy of each θ ward is summarized in Result 4.

Result 4. *With overtime restrictions ($h \leq \bar{h}$), the change in demand for hours and workers is given by:*

$$(\Delta h(\theta), \Delta N(\theta)) = \begin{cases} (0, 0), & \theta \in (0, \theta_0], \\ (0, 0), & \theta \in (\theta_0, \theta_1], \\ \left(\bar{h} - \frac{\theta}{1-\theta} \frac{f-wp\bar{h}}{w(1+p)}, \frac{B[\theta(f+w\bar{h})-w(1+p)\bar{h}]}{(f+w\bar{h})(f-wp\bar{h})} \right), & \theta \in (\theta_1, 1), \end{cases} \quad (10)$$

where $\theta_0 = \frac{w\bar{h}}{f+w\bar{h}}$ and $\theta_1 = \frac{w(1+p)\bar{h}}{f+w\bar{h}}$.

Proof. For those where $\left(\theta \leq \frac{w(1+p)\bar{h}}{f+w\bar{h}}\right)$, the optimal choice of hours is still within the cap constraint, and wards have no incentive to change the optimal labor allocation. For those above the overtime threshold the new optimal hours are given by $h^* = \bar{h}$, according to the hard constraint on hours. Hours must fall ($\Delta h(\theta) < 0$) as these wards were choosing hours above $\bar{h}$ before overtime restrictions.

The new level of employment is such that the budget constraints binds, so:

$$N(f + w\bar{h}) = B \iff N^* = \frac{B}{f + w\bar{h}}, \quad (11)$$

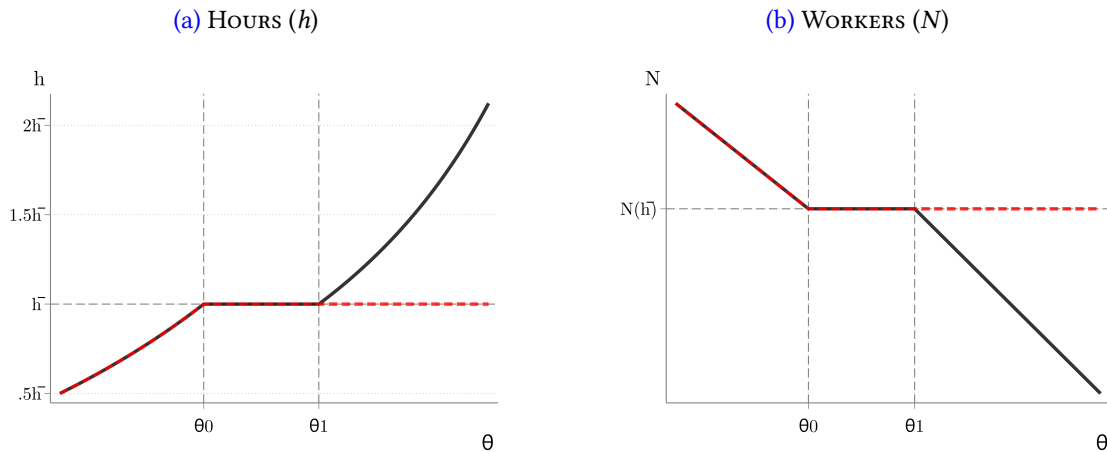
also independent of ward type θ . The change in employment is given by:

$$\frac{B}{f + w\bar{h}} - (1 - \theta) \frac{B}{f - wp\bar{h}} = \frac{B[\theta(f + w\bar{h}) - w(1+p)\bar{h}]}{(f + w\bar{h})(f - wp\bar{h})}. \quad (12)$$

Notice that because in the overtime region where $\theta > \frac{w(1+p)\bar{h}}{f+w\bar{h}}$, the change in employment after the reform is always positive.

□

Figure B.1. OVERTIME RESTRICTIONS ON OPTIMAL LABOR ALLOCATIONS BY θ



Notes. The figure illustrates the change in optimal choice of hours (Panel A) and workers (Panel B) for fixed parameters $\{w, p, \bar{h}, f\}$, along different values of θ , in the model. The solid black line represent optimal labor demand decisions for each level of θ without constraints on overtime allocations. The red dash line represents the optimal allocation after $h < \bar{h}$ is imposed.

C. Data

This section documents the restrictions on our main sample and how we construct our main variables using the data sources described in Section 3.1.

C.1. Sample Restrictions

Our main analysis dataset is an unbalanced panel of 424 hospital wards, 106 hospitals, covering the period 1978 – 1986. To construct the main analysis sample, we impose a series of restrictions. First, from the admission level data—with information on more than 8 million hospital admissions—we construct our main variables at the ward level. Second, we drop wards specialized in Psychiatry, as the complete information for these units are contained in the *Psychiatric Central Register* to which we have no access. Second, based on the flow of admissions, we define whether a given hospital ward is operating throughout the period. We define a ward to be always open if it is admitting patients between June 1977 and June 1986.²⁵ We only include in the sample wards that are observed before 1982, and at least once in the post-reform period (1982 – 1986). Therefore, our analysis sample includes wards that have at least one pre-reform and one post-reform year. Finally, we remove wards that are at the extreme of the distribution of our dosage variable (share of unplanned admissions). We only include hospital wards with a mix of planned and unplanned cases, thus excluding those that are only specialized in acute care or only admit patients in an elective basis. Table 1 in Section 3.4 presents summary statistics for this panel across time periods. To allay concerns with ward closures, we also construct a balanced panel of units open throughout the entire period of analysis.

Labor Outcomes Samples.— We use archival data on posted and filled positions by physician career stage at different levels of analysis (ward-by-year, specialty-by-year, and year). First, we exclude Psychiatry wards for consistency with the sample restrictions on the health outcomes sample. Second, we collapse this information at the specialty level and combine it with our main analysis sample, to recover information on exposure to the reform. We restrict this new sample at the ward-by-specialty level to include all specialties that can be identified in both data sources. The resulting sample includes a total of 21 different medical specialties, covering the period 1978 – 1986.²⁶ Our first stage analysis in Section 5 uses this sample.

²⁵ We allow the period of admission to start in 1977 to not capture wards that only start admitting patients in the beginning of our sample period.

²⁶ The specialties we include are: Anesthesiology, Clinical Oncology, Dermato-Venerology, Gastro Surgery, General Surgery, Gynecology and Obstetrics, Internal Medicine, Cardiology, Gastroenterology, Geriatrics, Nephrology, Pulmonary Diseases, Mixed Medicine, Neuro Surgery, Ophthalmology, Orthopedic Surgery, Otorhinolaryngology, Pediatrics, Plastic Surgery, Thoracic Surgery, and Urology.

C.2. Main Variables

Hospital Admissions ($\mathbf{Y}_{it}$).— We consider an hospital admission each registered case in the patient-level dataset. We identify two types of admissions. *Planned* admissions are booked in advance for a specific date, typically after referral from a general practitioner or outpatient specialist. By contrast, *unplanned* admissions refer to acute patients that arrive without prior booking, and are usually processed through the emergency department, nondeferrable and requiring immediate care. Based on the identification of the case as acute (*akut*) in the data, we categorize the admission as unplanned. During our pre-reform period of analysis, patients whose admissions are unplanned—52% of the cases—are almost 5 times more likely to die while in hospital and have 22% longer stays conditional on not dying. Consistent with acute patient cases being unplanned, weekend or night admissions are 5 times more likely to be acute (Section D, Figure 5). From the patient-level data, we then construct total hospital admissions, unplanned, and planned admissions at the ward-by-year level.

Mortality Outcomes ($\mathbf{Y}_{it}$).— From patient and death registry data, we construct multiple measures of mortality. We identify whether each patient dies in hospital or dies within d days of being discharged, with $d \in \{30, 60, 366\}$. From the patient-level data we then aggregate at the ward-by-year level. First, we construct total death counts, as well as by type of admission—planned or unplanned. Second, we construct mortality rates as the number of patients dying within each horizon per thousand admitted patients per year. We then compute mortality rates specific to each type of admission. Crucially, in post-discharge mortality includes deaths that occur within hospital. Therefore, they include both in-hospital and out-of-hospital deaths.

Age-Adjusted Outcomes ($\mathbf{Y}_{it}$).— Following [Finkelstein et al. \(2025\)](#), we adjust mortality outcomes by the age distribution of the patients in each ward. We proceed as follows. First, we divide patients into age bins, in intervals of 5 years, up until 89 years of age, with a last bin for those aged 90 or above. We then compute mortality rates for each age bin in 1978, as well as the overall mortality rate in that year. Then, for every ward in every year we compute the expected number of deaths given their patient composition. The age-adjusted measure is just the ratio between actual deaths and expected deaths. For consistency, we then multiply this by the overall mortality rate in 1978 to interpret results in terms of deaths per thousand admissions.

Risk-Adjusted Outcomes ($\mathbf{Y}_{it}$).— Following [Muñoz and Otero \(2025\)](#), we extend the age adjustment to take also into account the action diagnostic of each patient. We proceed as follows. As above, we distribute patients into age bins, in intervals of 5 years, up until 89 years of age, with a last bin for those aged 90 or above. We then distribute patients into one of the different Charlson comorbidity

index prognostic groups (Charlson et al., 1987) based on the ICD-8 codes found in the registry data, plus an additional residual group for all other medical conditions. Using data from 1978, we run logit models to predict the death risk of each patient based on their age and diagnostic group. We then aggregate these risk scores at the ward-year level to compute the number of expected deaths given the distribution of patients. Similarly to above, we finally compute the risk-adjusted measure as just the ratio between actual and expected deaths, multiplying it by the overall mortality rate in 1978.

Readmission Rates (Y_{it}).— We measure readmission rates as the number of patients that are readmitted into a hospital ward within d days after discharge by every thousand admitted patients, with $d \in \{14, 30, 60\}$. To be identified as a readmission, the patient should stay at least one day out of hospital care after the date of discharge. Patients that are admitted into another hospital ward on the date of discharge are coded as transfers and are not considered a readmission.

Length of Stay (Y_{it}).— The duration of each case is computed as the number of days passed between admission and discharge from the focal hospital ward. From the patient-level data, we then aggregate this variable to the ward-by-year level by taking the average length of stay over the number of admissions.

Post-Reform Period ($\mathbb{1}\{t \geq 1982\}$).— The policy we study was fully implemented in 1982, with a transition period in the last 3 months of 1981 (see Section 2). Throughout the main analysis, we identify all observations from 1982 onward as post-reform.

Physicians (L_{it}, L_{st}).— For our first-stage empirical analysis, we characterize the labor demand shock at different levels. In our main analysis in Section 5, we link our dosage variable (a ward’s unplanned admission share) with digitized archival data on the number of posted positions at the specialty level (specialty-specific labor demand in a year, L_{st}). We supplement our main analysis with ward-level descriptive evidence presented in Appendix E.²⁷

Exposure to the Reform (D_i).— In our research design, we exploit differences in reform intensity by hospital ward (see Section 4). We construct different versions of this variable at the ward level. For parsimony, our preferred exposure variable is the share of unplanned admissions pre-reform, measured in a single pre-reform year ($t = 1980$)²⁸. We also test our design by using alternative predictors for the change in labor demand. First, we use both the share of weekend and night admissions by hospital ward. Second, to capture medium to long-term scheduling uncertainty, we compute the mean of the standard deviation of these exposure variables across weeks, within $t = 1980$, for each hospital ward.

²⁷ As of now, we cannot link this source of unidentified ward-level data with the patient registers.

²⁸ Figure D.2a shows the distribution of this variable across hospital wards. We set $t = 1980$ to not be confounded by any anticipation. In Appendix D we show that the concentration of unplanned admissions varies substantially across wards but not across years, within wards.

We then use a latent variable from a combination of all these exposure variables, predictive of labor demand change across reform periods. Appendix D provides further detail on each of these variables and their relationship.

D. Identifying Variation

In this section we document the variation in the pre-reform share of unplanned admissions, our main proxy of overtime need. We also characterize the variation for alternative proxies for overtime need, used in robustness checks.

Variance Decomposition.—What explains variation in unplanned admissions across wards and across time? To what extent can these characteristics be attributed to production function characteristics of the wards relative to local demand characteristics? To shed light on this question we run a Shapley-Owen decomposition (Huettner and Sunder, 2012; Israeli, 2007) on the pre-reform share of unplanned admissions (D_{it}) varying at the ward-by-year level (1978-1981).

We start by running regressions of our overtime need proxies on different combinations of fixed effects. Each set of fixed effects relates to a specific level of variation. We consider four specific levels: ward (i), year (t), specialty (s), and county (c). Then, for each regression, we compute how much the adjusted R^2 changes when excluding the fixed effects for a specific level of variation j . Based on this information we can gauge the relative importance of each level of variation.

Table D.1 presents the partial R^2 for each level of variation j , and for three different proxies of overtime need: share of unplanned (Cols. 1-3), share of weekend (Cols. 4-6), and share of night admissions (Cols. 7-9). Columns 1, 4 and 7 present partial adjusted R^2 for specifications including only ward and year fixed effects. Columns 2, 5, and 8 include year and specialty fixed effects. Finally Columns 3, 6 and 9 include year, specialty, and county fixed effects.

Our Shapley-Owen decomposition exercise shows that most of the variation comes from differences across wards (i), rather than across years (t) (Table D.1, Cols. 1, 4, and 7). Ward fixed effects explain about 94% of the variation in the share of unplanned admissions, with this variable being relatively stable across years. A large portion of differences across wards is driven by differences in the type of healthcare provided. About 60% of the variation in our proxies of pre-reform overtime need are explained by the main specialty of the ward (s), rather than across years or across local healthcare demand (c).

Alternative Measures of Overtime Need.—In Section 6.5 we run robustness specifications where we use alternative definitions of overtime need. Instead of exploiting variation on the concentration of unplanned admissions, we use variation in the share of night and weekend admissions. Figure D.2 shows the distribution of each of these variables.

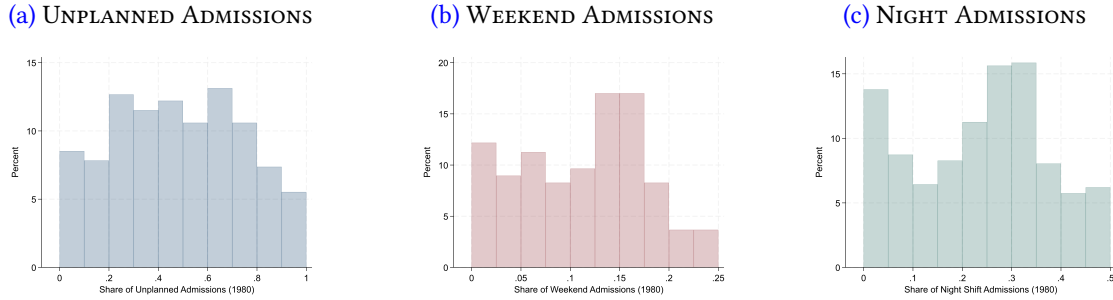
A relevant concern is whether wards with high concentrations of unplanned, weekend, or night admissions also have relatively lower variance in unplanned admissions across time. High-mean, low-variance in unplanned admissions could mean that some wards could prefer to employ more shifts to

Table D.1. SHAPLEY-OWEN VARIANCE DECOMPOSITION (1978-1981)

Level of Variation	Partial R^2								
	Share Unplanned			Share Weekend			Share Night		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ward (i)	0.939	–	–	0.953	–	–	0.922	–	–
Year (t)	0.001	0.001	0.001	0.00	0.00	0.00	0.00	0.00	0.00
Specialty (s)	–	0.6	0.599	–	0.616	0.616	–	0.53	0.524
County (c)	–	–	0.049	–	–	0.014	–	–	0.058
All	0.94	0.6	0.649	0.953	0.616	0.63	0.922	0.53	0.581

Notes. The table presents a Shapley-Owen variance decomposition of the adjusted R^2 of a regression of time-varying proxies of overtime need pre-reform into different sets of fixed effects. The method is detailed in the text of Appendix D. The dependent variables of each regression are the share of unplanned admission (Columns 1-3), the share of weekend admissions (Columns 4-6), and the share of night admissions (Columns 7-9). The sample includes all hospital wards with at least one observation in each reform period (see Section C.1 for more details), while only using information from the pre-reform period (1978-1981). All reports the adjusted R^2 of a regression that includes all sets of fixed effects in that specification.

Figure D.2. DISTRIBUTIONS OF OVERTIME NEED PROXIES



Notes. The figure presents histograms of our alternative overtime need proxies at the hospital ward level (D_i). For privacy protection reasons, the last bin of the histograms on the share of weekend admissions (Panel A) and the night admissions (Panel B) includes wards at or above that particular value of the variable, so that at least 5 wards are counted in each bin of the histogram.

deal with high yet predictable unplanned demand, rather than using overtime. However, in Section 5 we show evidence which is not consistent with this hypothesis.

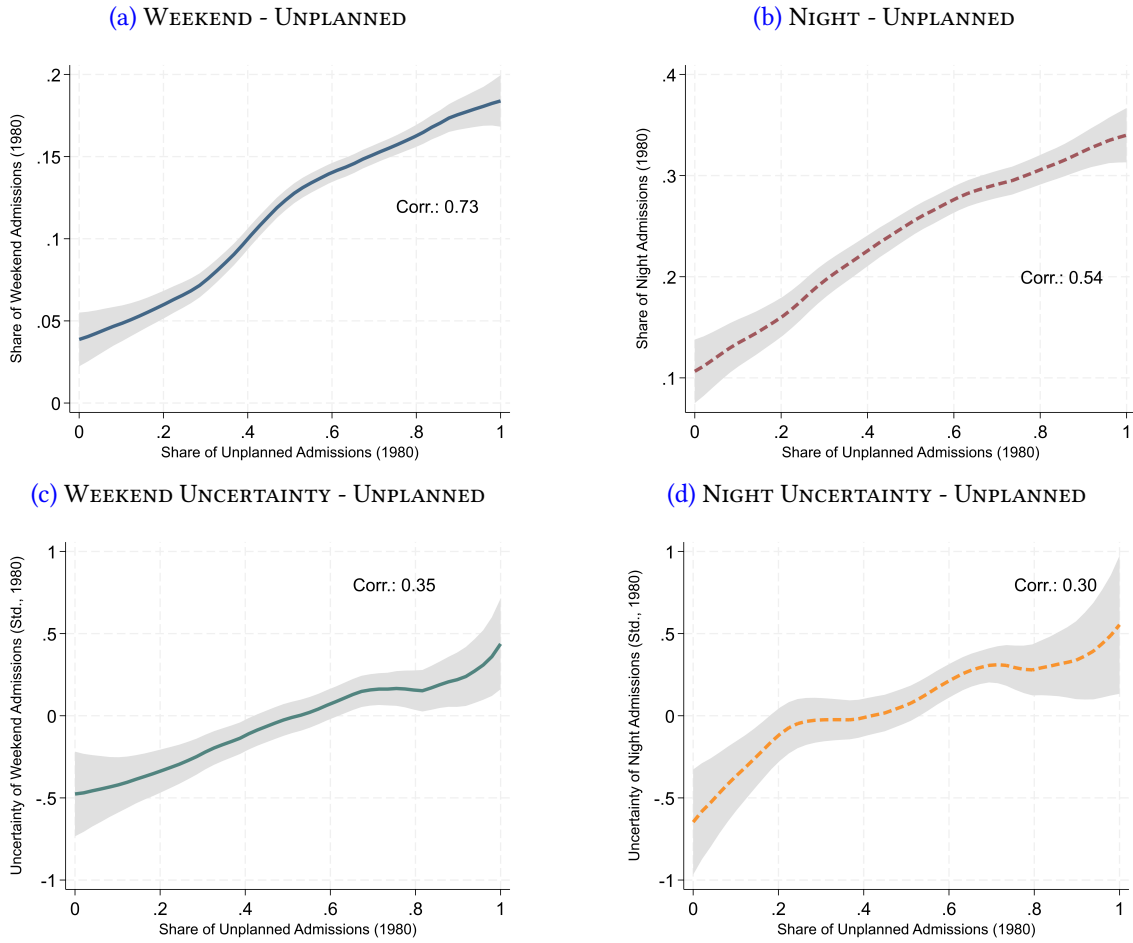
To further allay concerns, we compute a within-year standard deviation of week-level concentrations of unplanned, night, and weekend admissions for each ward. Table D.2 shows correlations between all these variables pre-reform. If anything, there seems to be a positive correlation between the share of unplanned admissions and uncertainty in weekend and night admissions. The correlation between the share of unplanned admissions and its weekly uncertainty across the year is slightly negative and close to zero ($\rho = -0.107$), as wards closer to the extreme of the distribution have no within-year variation in this variable. On the other hand, correlations between the concentration of unplanned admissions and both weekend ($\rho = 0.724$) and night admissions ($\rho = 0.519$) are strong (Figures D.3a-b). Likewise, wards with higher concentrations of unplanned admissions also have higher within-year uncertainty in the share of weekend admissions ($\rho = 0.304$) and night admissions ($\rho = 0.282$).

Table D.2. CORRELATION BETWEEN PROXIES OF OVERTIME NEED

<i>Variables</i>	Share Unplanned	Share Weekend	Share Night	SD Unplanned	SD Weekend	SD Night
Share Unplanned	1.000	–	–	–	–	–
Share Weekend	0.724	1.000	–	–	–	–
Share Night	0.519	0.769	1.000	–	–	–
SD Unplanned	-0.107	-0.306	-0.403	1.000	–	–
SD Weekend	0.304	0.420	0.164	0.336	1.000	–
SD Night	0.282	0.306	0.351	0.248	0.643	1.000

Notes. The table presents Pearson correlation estimates between six different proxies of overtime need, measured in $t = 1980$. The sample includes all hospital wards with at least one observation in each reform period (see Section C.1 for more details).

Figure D.3. CORRELATIONS BETWEEN OVERTIME NEED PROXIES



Notes. The figure documents local polynomial estimates of the relationship between different proxies of overtime need at the hospital ward level. The sample includes all hospital wards with at least one observation in each reform period (see Section C.1 for more details).

E. Overtime Restrictions and Labor Demand

E.1. Overall Change in Labor Demand

Here we document descriptive evidence on the increase in labor demand across time periods, at the ward level. Relying on archival data on the number of posted positions by hospital ward (see Section

3), we run the following pre-post specification:

$$L_{it} = \alpha + \lambda_i + \eta \mathbb{1}\{t \geq 1982\} + \varepsilon_{it}, \quad (13)$$

where L_{it} are different definitions of labor demand in levels, and hospital ward fixed effects (λ_i) absorb time-invariant differences across wards. We use all years for which we have information at the ward level (1979-1983).²⁹ The estimated $\hat{\eta}$ identifies the overall change across reform periods. We estimate Equation 13 through OLS and cluster the standard errors at the hospital level.

Table E.3 reports the main coefficients of this analysis. With the introduction of the overtime restrictions reform, hospital wards post an average of 3.68 new positions (Column 1, 95% C.I.: 3.39; 3.97). This corresponds to an increase of 36% over the pre-reform mean of 10.2 physician positions in each ward. The increase is driven by labor demand for subordinate physicians, for which there are 3.47 new positions post-reform (Column 2, 95% C.I.: 3.2; 3.74), a surge of 45%.

Table E.3. LABOR DEMAND AND BED CAPACITY (1979-1983)

	Total Physicians (1)	Sub. Physicians (2)	Sub. Physicians / 100 Beds (3)	Hospital Beds (4)
$\mathbb{1}\{t \geq 1982\}$	3.68 (0.148) [3.39,3.97]	3.47 (0.134) [3.20,3.74]	8.08 (1.378) [5.35,10.81]	-2.26 (0.769) [-3.79,-0.74]
Observations	2,180	2,180	2,180	2,180
Hospital Wards	465	465	465	465
Mean Dep. Var. Pre-Reform	10.17	7.75	16.11	66.76
Post-Reform Mean	14.04	11.39	24.72	65.23

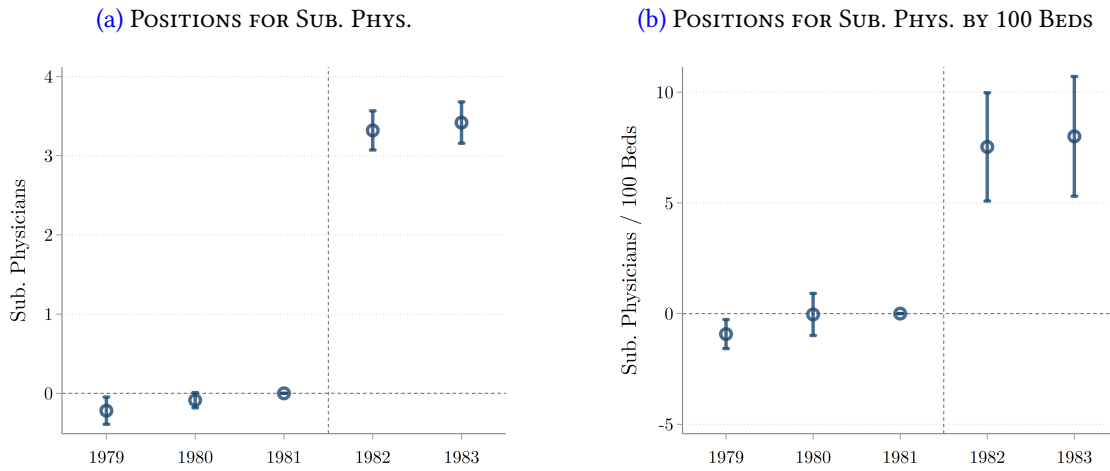
Notes. The data is drawn from statistical yearbooks, and includes all hospital wards in operating in the country between 1979 and 1983, excluding psychiatry wards. The table presents the change in the total number of positions open by type of physician (Columns 1 and 2) after the reform, as well as the change in subordinate physicians by 100 hospital beds (Column 3) and total bed capacity in hospital wards (Column 4). Regressions include hospital ward fixed effects. 95% confidence intervals for standard errors clustered at the hospital level are presented in between brackets.

In Columns 3 and 4 of Table E.3, we examine labor demand change in the context of each ward's healthcare supply capacity. Specifically, we use information on the number of beds in each hospital ward. Prior to the reform, there is an average of 16 subordinate physician positions for every 100 beds. After the reform, this increases by more than 8 (Column 3, 95% C.I.: 5.35; 10.81), to almost 25 such positions by 100 beds. The change in this ratio is partly driven by consolidation in the number of beds throughout the period, as wards decreased their bed capacity by 2.23 over a pre-reform baseline of 66.8 (Column 4, 95% C.I.: -3.79; -0.74).

Figure E.4 reports coefficients of event study versions of Equation 13, using 1981 as the reference

²⁹ In a future version of the paper, we will also include data for our entire period of analysis (1978-1986), as statistical books for the remaining years are still pending scanning and digitization.

Figure E.4. LABOR DEMAND CHANGE



Notes. The data is drawn from statistical yearbooks, and includes all hospital wards in operating in the country between 1979 and 1983, excluding psychiatry wards. The figure documents the change in labor demand relative to one year prior to the reform, using an event study version of Equation 13. Panel A presents the overall change in subordinate physicians at the hospital ward level. Panel B depicts the change in subordinate physicians by 100 hospital beds. The dashed vertical lines indicate onset of the new working-hour system.

period. Both panels highlight the discontinuous change in demand for subordinate physicians in the first post-reform year. Moreover, overtime restrictions were introduced after a relatively stable period in hiring. Overall, this evidence is consistent with other descriptive statistics in Section 3, and the prediction of our theoretical model (Hypothesis 1).

E.2. Effects on the Labor Market

Table E.4 shows that a one standard deviation difference in pre-reform unplanned admissions (23 p.p.) raises the number of posted positions by 31.1 (Column 1, C.I.: 16.1; 46) post-reform. Over a baseline of 176.5 open positions by specialty pre-reform, this corresponds to a 17% change per standard deviation across time periods.

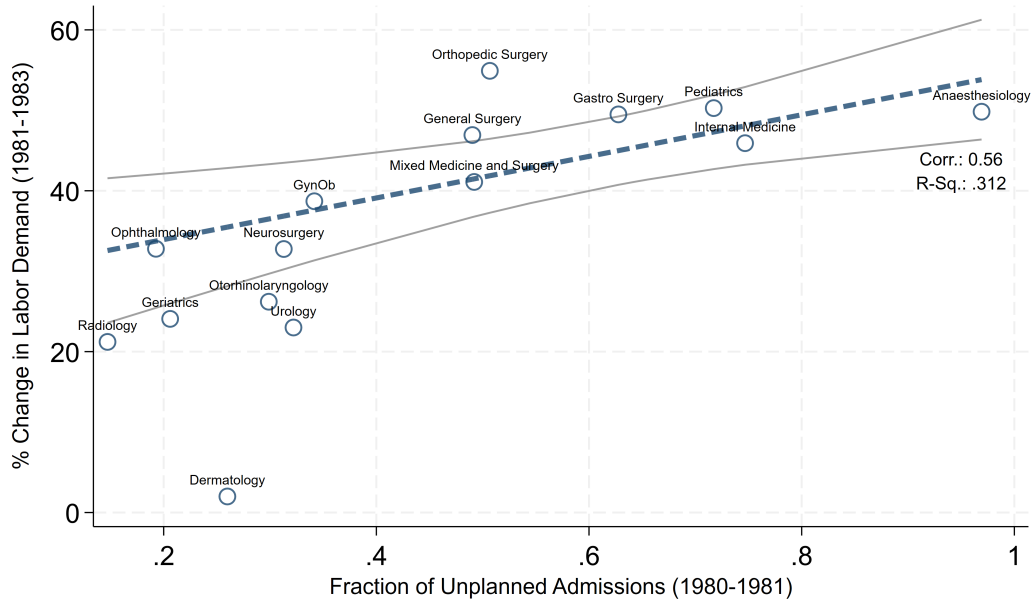
Figure E.5 further illustrates our findings, by plotting the relationship between our proxy of overtime need and the percent growth in posted positions between 1981 and 1983. It shows the fitted line from a weighted least squares regression, with weights given by the number of posted positions in 1981, as well as specialties with more than 50 positions pre-reform.

Table E.4. EFFECTS ON THE LABOR MARKET (1978-1984)

	Posted Positions (1)	Occupied Positions (2)
$\mathbb{1}\{t \geq 1982\} \times D_s$	31.05 (7.564) [16.10,45.99]	23.48 (6.358) [10.92,36.04]
R-Squared	0.968	0.974
Observations	181	181
Nr. of Specialties	21	21
Mean Dep. Var. Pre-Reform	176.5	174.3
Mean D_s	0.46	0.46
S.D. D_s	0.23	0.23
Specialty FE	Yes	Yes
Year FE	Yes	Yes

Notes. The table present estimates of β in Equation 2 for regressions at the specialty-by-year level. The sample includes all specialties that can be merged between the archival data on labor market positions and admission registries (see Section C.1 for more details). D_s is the standardized share of unplanned admissions in each specialty s , in 1980. 95% confidence intervals, computed from robust standard errors (in parentheses) are presented between brackets.

Figure E.5. CHANGE IN LABOR DEMAND (1981 – 1983) BY SHARE OF UNPLANNED ADMISSIONS



Notes. The figure shows the relationship between the share of unplanned admissions in each specialty (1980-1981) and the percent change in labor demand between 1981 and 1983. It plots and labels the position of each specialty. Only specialties with more than 50 available positions are presented. The dashed line represents the fitted line from a weighted least squares regression at the specialty level of the percent change in labor demand on share of unplanned admissions, with weights given by the number of positions in the specialty in 1980-1981. All specialties, even those not presented in the figure, are included in the regression. The figure also presents the Pearson correlation between the two variables and the R^2 of the regression.

F. Additional Tables

Table F.5. BROAD SPECIALTY GROUPINGS IN FIGURE 4

Specialty Group	Specialties Included
Surgery	Anesthesiology; General surgery; Gastro-Intestinal surgery; Hospital dentistry / oral surgery; Neurosurgery; Orthopaedic surgery; Plastic surgery; Thoracic surgery; Vascular surgery; Urology; Otorinolaryngologie; Ophthalmology; Gynecology and Obstetrics
Internal Medicine	Allergology; Cardiology; Endocrinology; Gastroenterology; Geriatrics; Haematology; Infectious diseases; Internal medicine; Long-term care medicine; Medical oncology; Nephrology; Neurology; Paediatrics; Pulmonology; Rheumatology; Tropical medicine
Paraclinical	Clinical biochemistry; Clinical immunology; Clinical microbiology; Clinical neurophysiology; Clinical physiology and nuclear medicine; Diagnostic radiology; Pathology; Physical medicine and rehabilitation; Radiation oncology; Transfusion medicine
Other	Dermatology & venereology; Public and occupational medicine; Public and social medicine

Notes. The table presents the four broad specialty groupings used to construct Figure 4 in Section 3.4. The table includes all the specialties in the digitized statistical yearbooks.

Table F.6. SPECIALTY GROUPINGS FOR REGRESSIONS

Specialty Group	Specialties Included
Internal Medicine (General)	Internal Medicine
Internal Medicine (Specialties)	Geriatrics; Hepatology; Hematology; Infectious Diseases; Cardiology; Endocrinology; Gastroenterology; Pulmonology; Nephrology; Rheumatology; Dermatology; Neurology; Oncology; Physiatry
General Surgery	General Surgery
Specialized Surgery	Vascular Surgery; Gastro Surgery; Plastic Surgery; Cardiothoracic Surgery; Urology; Neurosurgery; Orthopedic Surgery; Ophthalmology; Otorhinolaryngology; Maxillofacial Surgery
Gynecology and Obstetrics	Gynecology and Obstetrics
Other Specialties	Pediatrics; Anesthesiology; Mixed Medicine and Surgery; Unclassified

Notes. The table presents the six specialty groupings used to construct year-by-specialty-group fixed effects. The table includes all the specialties in the registry data of hospitalizations. For instance, paraclinical specialties are not included.

Table F.7. IN-HOSPITAL MORTALITY EFFECTS

	In-hospital Mortality Rate		
	(1)	(2)	(3)
<i>Panel A: All Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-2.45 (0.617) [-3.67,-1.24]	-2.39 (0.629) [-3.63,-1.16]	-1.32 (0.605) [-2.51,-0.13]
R-Squared	0.962	0.966	0.962
Mean Dep. Var.	34.08	34.08	34.08
<i>Panel B: Unplanned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-5.53 (1.262) [-8.01,-3.05]	-5.93 (1.187) [-8.26,-3.59]	-2.25 (1.954) [-6.09,1.59]
R-Squared	0.944	0.951	0.946
Mean Dep. Var.	58.00	58.00	58.00
<i>Panel C: Planned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-1.02 (0.520) [-2.05,-0.00]	-1.16 (0.589) [-2.32,-0.00]	-1.18 (0.558) [-2.28,-0.08]
R-Squared	0.807	0.814	0.813
Mean Dep. Var.	9.87	9.87	9.87
Observations	2,911	2,911	2,911
Nr. of Wards	424	424	424
Mean Exposure Var.	0.48	0.48	0.48
S.D. Exposure Var.	0.25	0.25	0.25
Hospital Ward FE	Yes	Yes	Yes
Year FE	Yes	No	No
Year X County FE	No	Yes	No
Year X Specialty FE	No	No	Yes

Notes. The table presents estimates of β in Equation 2 and variations. D_i is the standardized share of unplanned admissions in each ward i , in 1980. The dependent variable is in-hospital mortality rates measured in per thousand patient admissions. Each panel corresponds to a different set of patient admissions: all (Panel A), unplanned (Panel B), and planned (Panel C). Each column uses a different set of fixed effects, displayed in the table. All observations are weighted by the total number of admissions (Panel A), unplanned admissions (Panel B), and planned admissions (Panel C) in each ward in 1981. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

Table F.8. IN-HOSPITAL MORTALITY EFFECTS - LOG SPECIFICATION

	Ln In-hospital Mortality Rate		
	(1)	(2)	(3)
<i>Panel A: All Admission Types</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.049 (0.0213) [-0.091,-0.007]	-0.054 (0.0208) [-0.095,-0.013]	-0.030 (0.0251) [-0.080,0.019]
R-Squared	0.967	0.970	0.968
Observations	2,707	2,707	2,707
Nr. of Wards	412	412	412
Mean Dep. Var. Pre-Reform (No Log)	34.94	34.94	34.94
Mean Exposure Var.	0.50	0.50	0.50
S.D. Exposure Var.	0.25	0.25	0.25
<i>Panel B: Unplanned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.071 (0.0205) [-0.111,-0.031]	-0.082 (0.0222) [-0.126,-0.039]	-0.030 (0.0256) [-0.080,0.021]
R-Squared	0.964	0.967	0.966
Observations	2,526	2,526	2,526
Nr. of Wards	386	386	386
Mean Dep. Var. Pre-Reform (No Log)	60.42	60.42	60.42
Mean Exposure Var.	0.53	0.53	0.53
S.D. Exposure Var.	0.23	0.23	0.23
<i>Panel C: Planned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.060 (0.0395) [-0.138,0.018]	-0.080 (0.0410) [-0.161,0.001]	-0.041 (0.0408) [-0.121,0.039]
R-Squared	0.876	0.886	0.880
Observations	2,440	2,440	2,440
Nr. of Wards	396	396	396
Mean Dep. Var. Pre-Reform (No Log)	10.92	10.92	10.92
Mean Exposure Var.	0.50	0.50	0.50
S.D. Exposure Var.	0.24	0.24	0.24
Hospital Ward FE	Yes	Yes	Yes
Year FE	Yes	No	No
Year X County FE	No	Yes	No
Year X Specialty FE	No	No	Yes

Notes. The table presents estimates of β in Equation 2 and variations, with the dependent variable being in-hospital mortality rates in logs. D_i is the standardized share of unplanned admissions in each ward i , in 1980. The coefficients are interpreted in terms of log changes. Each panel corresponds to a different set of patient admissions: all (Panel A), unplanned (Panel B), and planned (Panel C). All observations are weighted by the total number of admissions (Panel A), unplanned admissions (Panel B), and planned admissions (Panel C) in each ward in 1981. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

Table F.9. IN-HOSPITAL MORTALITY EFFECTS WITH POISSON QMLE

	In-hospital Mortality Rate		
	(1)	(2)	(3)
<i>Panel A: All Admission Types</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.037 (0.0169) [-0.070,-0.004]	-0.031 (0.0188) [-0.068,0.006]	-0.027 (0.0192) [-0.064,0.011]
Mean Dep. Var.	34.08	34.08	34.08
<i>Panel B: Unplanned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.072 (0.0248) [-0.120,-0.023]	-0.082 (0.0236) [-0.128,-0.036]	-0.041 (0.0344) [-0.108,0.027]
Mean Dep. Var.	58.00	58.00	58.00
<i>Panel C: Planned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.017 (0.0336) [-0.082,0.049]	-0.001 (0.0316) [-0.063,0.060]	-0.095 (0.0449) [-0.183,-0.007]
Mean Dep. Var.	9.87	9.87	9.87
Observations	2,911	2,911	2,911
Nr. of Wards	424	424	424
Mean Exposure Var.	0.48	0.48	0.48
S.D. Exposure Var.	0.25	0.25	0.25
Hospital Ward FE	Yes	Yes	Yes
Year FE	Yes	No	No
Year X County FE	No	Yes	No
Year X Specialty FE	No	No	Yes

Notes. The table presents estimates of β in Equation 2 and variations, using Poisson quasi-maximum likelihood estimation. D_i is the standardized share of unplanned admissions in each ward i , in 1980. The dependent variable is in-hospital mortality rates. The coefficients are interpreted in terms of log changes. Each panel corresponds to a different set of patient admissions: all (Panel A), unplanned (Panel B), and planned (Panel C). All observations are weighted by the total number of admissions (Panel A), unplanned admissions (Panel B), and planned admissions (Panel C) in each ward in 1981. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

Table F.10. PATIENT-LEVEL ANALYSIS: IN-HOSPITAL MORTALITY

	$\mathbb{1}\{\text{Dies in Hospital}\} \times 1000$			
	(1)	(2)	(3)	(4)
<i>Panel A: All Admission Types</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-2.15 (0.629) [-3.39,-0.92]	-2.19 (0.652) [-3.47,-0.91]	-0.90 (0.593) [-2.07,0.26]	-2.41 (0.578) [-3.55,-1.27]
R-Squared	0.043	0.043	0.043	0.071
Observations	5,652,711	5,652,711	5,652,711	5,652,711
Mean Dep. Var.	33.5	33.5	33.5	33.5
Mean Exposure Var.	0.51	0.51	0.51	0.51
S.D. Exposure Var.	0.23	0.23	0.23	0.23
<i>Panel B: Unplanned Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-4.59 (1.084) [-6.72,-2.46]	-5.06 (1.047) [-7.11,-3.00]	-0.67 (1.291) [-3.21,1.87]	-4.61 (1.017) [-6.61,-2.61]
R-Squared	0.044	0.044	0.044	0.079
Observations	2,935,468	2,935,468	2,935,468	2,935,468
Nr. of Wards	424	424	424	424
Mean Dep. Var.	57.0	57.0	57.0	57.0
Mean Exposure Var.	0.60	0.60	0.60	0.60
S.D. Exposure Var.	0.21	0.21	0.21	0.21
<i>Panel C: Planned Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.63 (0.457) [-1.53,0.26]	-0.87 (0.489) [-1.83,0.09]	-0.36 (0.443) [-1.23,0.51]	-0.74 (0.423) [-1.57,0.09]
R-Squared	0.023	0.023	0.023	0.033
Observations	2,717,243	2,717,243	2,717,243	2,717,243
Nr. of Wards	424	424	424	424
Mean Dep. Var.	9.8	9.8	9.8	9.8
Mean Exposure Var.	0.41	0.41	0.41	0.41
S.D. Exposure Var.	0.20	0.20	0.20	0.20
Hospital Ward FE	Yes	Yes	Yes	Yes
Year FE	Yes	No	No	Yes
Year X County FE	No	Yes	No	No
Year X Specialty FE	No	No	Yes	No
Age Bin FE	No	No	No	Yes

Notes. The sample includes all admissions in hospital wards with at least one observation in each reform period (see Section C.1 for more details). D_i is the standardized share of unplanned admissions in each ward i , in 1980. The dependent variable is an indicator of whether the admission results in death in the hospital, multiplied by 1000 so that results are interpreted per thousand admissions. Specifications in each of the columns use a different set of fixed effects. Column 1 reproduces the estimates of β in Equation 2 at the level of individual admissions. Columns 2 and 3 use year-by-county and year-by-specialty fixed effects, respectively. Column 4 uses our preferred specification with added age bin fixed effects. Each panel correspond to regressions in different types of admissions: all (Panel A), unplanned (Panel B), and planned (Panel C). We present the mean of the dependent variable pre-reform. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

Table F.11. EFFECTS OF THE REFORM ON TOTAL PATIENT ADMISSIONS

	Total Admissions		
	(1)	(2)	(3)
<i>Panel A: All Admission Types</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	57.63 (17.846) [22.55,92.71]	50.23 (19.256) [12.38,88.08]	45.26 (23.727) [-1.38,91.89]
R-Squared	0.958	0.961	0.959
Mean Dep. Var.	1879.61	1879.61	1879.61
<i>Panel B: Unplanned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	23.41 (18.095) [-12.16,58.98]	23.32 (17.468) [-11.01,57.66]	1.90 (17.663) [-32.81,36.62]
R-Squared	0.937	0.942	0.942
Mean Dep. Var.	945.12	945.12	945.12
<i>Panel C: Planned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	34.22 (18.260) [-1.68,70.11]	26.91 (18.057) [-8.59,62.40]	43.35 (19.208) [5.60,81.11]
R-Squared	0.942	0.946	0.944
Mean Dep. Var.	934.49	934.49	934.49
Mean Exposure Var.	0.48	0.48	0.48
S.D. Exposure Var.	0.25	0.25	0.25
Hospital Ward FE	Yes	Yes	Yes
Year FE	Yes	No	No
Year X County FE	No	Yes	No
Year X Specialty FE	No	No	Yes

Notes. The table presents estimates of β in Equation 2. The sample includes all hospital wards with at least one observation in each reform period (see Section C.1 for more details). D_i is the standardized share of unplanned admissions in each ward i , in 1980. The dependent variable is total admissions in each ward. Each column presents a specification with different sets of fixed effects. Each panel correspond to regressions in different types of admissions: all (Panel A), unplanned (Panel B), and planned (Panel C). We present the mean of the dependent variable pre-reform. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

Table F.12. EFFECTS OF THE REFORM ON TOTAL DEATHS IN HOSPITAL

	Total Deaths		
	(1)	(2)	(3)
<i>Panel A: All Admission Types</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	2.45 (1.120) [0.25,4.65]	2.31 (1.098) [0.15,4.47]	0.40 (1.036) [-1.63,2.44]
R-Squared	0.964	0.966	0.965
Mean Dep. Var. Pre-Reform	63.02	63.02	63.02
<i>Panel B: Unplanned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	2.53 (1.099) [0.37,4.69]	2.54 (1.064) [0.45,4.63]	0.16 (0.959) [-1.72,2.05]
R-Squared	0.963	0.965	0.965
Mean Dep. Var. Pre-Reform	53.84	53.84	53.84
<i>Panel C: Planned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.08 (0.265) [-0.61,0.44]	-0.23 (0.296) [-0.81,0.35]	0.24 (0.312) [-0.38,0.85]
R-Squared	0.848	0.861	0.851
Mean Dep. Var. Pre-Reform	9.19	9.19	9.19
Mean Exposure Var.	0.48	0.48	0.48
S.D. Exposure Var.	0.25	0.25	0.25
Hospital Ward FE	Yes	Yes	Yes
Year FE	Yes	No	No
Year X County FE	No	Yes	No
Year X Specialty FE	No	No	Yes

Notes. The table presents estimates of β in Equation 2. The sample includes all hospital wards with at least one observation in each reform period (see Section C.1 for more details). D_i is the standardized share of unplanned admissions in each ward i , in 1980. The dependent variable is total deaths in each ward. Each column presents a specification with different sets of fixed effects. Each panel correspond to regressions in different types of admissions: all (Panel A), unplanned (Panel B), and planned (Panel C). We present the mean of the dependent variable pre-reform. We present the mean of the dependent variable pre-reform. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

Table F.13. EFFECTS OF THE REFORM ON AGE-ADJUSTED MORTALITY RATES

	Age-adjusted Mortality Rates		
	In-hospital (1)	30-days (2)	1-year (3)
<i>Panel A: All Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-2.05 (0.439) [-2.91,-1.19]	-1.87 (0.574) [-3.00,-0.75]	-2.64 (1.268) [-5.13,-0.14]
R-Squared	0.922	0.927	0.929
Mean Dep. Var. Pre-Reform	27.92	40.76	93.95
<i>Panel B: Unplanned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-4.03 (1.001) [-6.00,-2.06]	-3.93 (1.188) [-6.26,-1.60]	-3.02 (2.594) [-8.12,2.08]
R-Squared	0.914	0.920	0.917
Mean Dep. Var. Pre-Reform	58.59	80.11	146.12
<i>Panel C: Planned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.59 (0.150) [-0.88,-0.29]	-0.56 (0.249) [-1.05,-0.07]	-1.91 (0.718) [-3.32,-0.50]
R-Squared	0.922	0.927	0.936
Mean Dep. Var. Pre-Reform	6.88	12.45	51.88
Observations	2,911	2,911	2,911
Nr. of Wards	424	424	424
Mean Exposure Var.	0.48	0.48	0.48
S.D. Exposure Var.	0.25	0.25	0.25

Notes. The table presents estimates of β in Equation 2 for age-adjusted mortality rates. The sample includes all hospital wards with at least one observation in each reform period (see Section C.1 for more details). D_i is the standardized share of unplanned admissions in each ward i , in 1980. Each column refers to a different measure of patient mortality, measured per thousand admissions: in-hospital (Column 1), 30-day (Column 2), and 1-year (Column 3). Mortality rates are age-adjusted by scaling observed deaths by the expected number of deaths implied by the ward-specific age composition of admitted patients, and rescaled by the overall mortality rate in 1978. Panel A presents results for all admissions, while Panels B and C show estimates among unplanned and planned admissions, respectively. All observations are weighted by the total number of admissions (Panel A), unplanned admissions (Panel B), or planned admissions (Panel C) in 1981. The mean of the dependent variable is computed pre-reform and similarly weighted. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

Table F.14. EFFECTS OF THE REFORM ON RISK-ADJUSTED MORTALITY RATES

	Risk-Adjusted In-hospital Mortality Rate		
	(1)	(2)	(3)
<i>Panel A: All Admission Types</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-1.19 (0.415) [-2.00,-0.37]	-1.35 (0.445) [-2.23,-0.48]	-0.61 (0.565) [-1.72,0.50]
R-Squared	0.898	0.907	0.902
Mean Dep. Var. Pre-Reform	27.68	27.68	27.68
<i>Panel B: Unplanned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-1.85 (1.292) [-4.39,0.69]	-2.53 (1.240) [-4.97,-0.09]	2.24 (2.142) [-1.97,6.45]
R-Squared	0.856	0.867	0.862
Mean Dep. Var. Pre-Reform	75.12	75.12	75.12
<i>Panel C: Planned Admissions</i>			
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.15 (0.214) [-0.57,0.27]	-0.13 (0.299) [-0.72,0.46]	-0.48 (0.294) [-1.06,0.09]
R-Squared	0.468	0.477	0.478
Mean Dep. Var. Pre-Reform	3.27	3.27	3.27
Observations	2,911	2,911	2,911
Nr. of Wards	424	424	424
Mean Exposure Var.	0.48	0.48	0.48
S.D. Exposure Var.	0.25	0.25	0.25
Hospital Ward FE	Yes	Yes	Yes
Year FE	Yes	No	No
Year X County FE	No	Yes	No
Year X Specialty FE	No	No	Yes

Notes. The table presents estimates of β in Equation 2. The sample includes all hospital wards with at least one observation in each reform period (see Section C.1 for more details). D_i is the standardized share of unplanned admissions in each ward i , in 1980. The dependent variable is the risk-adjusted in-hospital mortality rate, as described in the main text. It is the ratio between actual deaths and expected deaths given the patient case-mix, multiplied by the overall in-hospital mortality rate in 1978. Each column presents a specification with different sets of fixed effects. Each panel correspond to regressions in different types of admissions: all (Panel A), unplanned (Panel B), and planned (Panel C). We present the mean of the dependent variable pre-reform. We present the mean of the dependent variable pre-reform. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

Table F.15. EFFECTS ON PATIENT MORTALITY AND READMISSION RATES, BALANCED SAMPLE

	Mortality Rate			Readmission Rate
	In-hospital (1)	30-days (2)	1-year (3)	30-days (4)
<i>Panel A: All Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-2.45 (0.651) [-3.73, -1.17]	-2.21 (0.723) [-3.63, -0.79]	-1.19 (1.040) [-3.24, 0.85]	-3.01 (2.424) [-7.78, 1.75]
R-Squared	0.963	0.969	0.982	0.920
Mean Dep. Var.	34.16	48.44	106.89	98.02
<i>Panel B: Unplanned Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-5.65 (1.349) [-8.30, -2.99]	-6.76 (1.662) [-10.02, -3.49]	-6.58 (1.851) [-10.22, -2.94]	0.99 (1.838) [-2.63, 4.60]
R-Squared	0.946	0.953	0.972	0.796
Mean Dep. Var.	58.58	79.53	150.03	96.42
<i>Panel C: Planned Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-1.35 (0.493) [-2.32, -0.38]	-1.18 (0.699) [-2.55, 0.20]	-0.00 (1.259) [-2.48, 2.47]	-2.06 (4.534) [-10.98, 6.85]
R-Squared	0.847	0.960	0.962	0.928
Mean Dep. Var.	9.74	16.82	62.97	99.18
Observations	2,702	2,702	2,702	2,702
Nr. of Wards	386	386	386	386
Mean Exposure Var.	0.49	0.49	0.49	0.49
S.D. Exposure Var.	0.24	0.24	0.24	0.24

Notes. The table presents estimates of β in Equation 2. The sample includes all hospital wards admitting patients throughout the entire period (1978-1984). D_i is the standardized share of unplanned admissions in each ward i , in 1980. Each column refers to a different dependent variable, measured per thousand admissions. Columns 1-3 present estimates for regressions where in-hospital, 30-day post-discharge, and 1-year post-discharge mortality are the dependent variables. Column 4 reports results for 30-day readmission rates after discharge. All observations are weighted by the total number of admissions (Panel A), unplanned admissions (Panel B), and planned admissions (Panel C) in each ward in 1981. We present the weighted mean of the dependent variable pre-reform. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

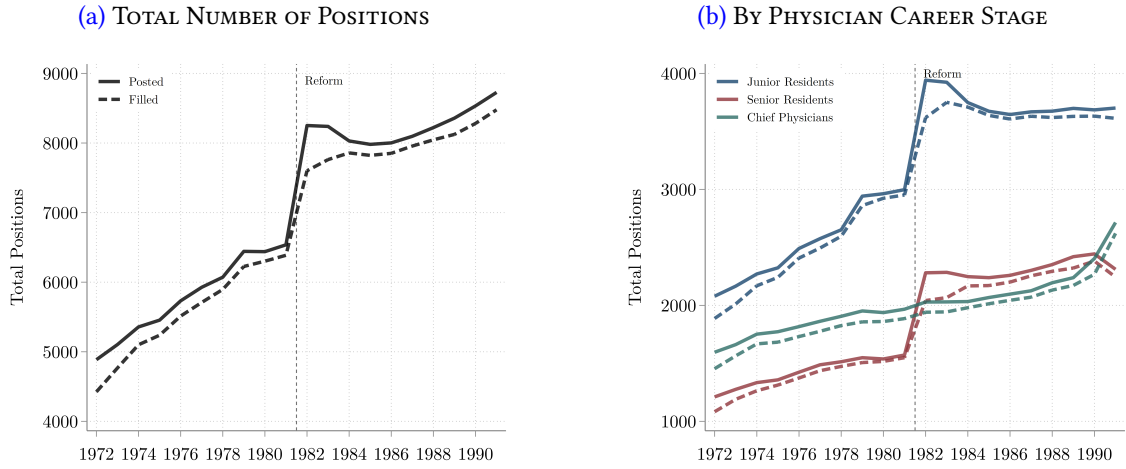
Table F.16. EFFECTS ON PATIENT MORTALITY AND READMISSION RATES, WITHOUT GYNOb WARDS

	Mortality Rate			Readmission Rate
	In-hospital (1)	30-days (2)	1-year (3)	30-days (4)
<i>Panel A: All Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-2.74 (0.739) [-4.19, -1.28]	-2.78 (0.833) [-4.41, -1.14]	-2.45 (1.246) [-4.90, -0.00]	-4.37 (3.014) [-10.30, 1.55]
R-Squared	0.955	0.962	0.976	0.924
Mean Dep. Var.	41.42	58.71	129.05	100.64
<i>Panel B: Unplanned Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-6.23 (1.530) [-9.23, -3.22]	-7.99 (1.952) [-11.83, -4.15]	-9.24 (2.112) [-13.39, -5.09]	2.70 (1.900) [-1.03, 6.44]
R-Squared	0.941	0.942	0.965	0.803
Mean Dep. Var.	65.76	90.16	169.69	95.78
<i>Panel C: Planned Admissions</i>				
$\mathbb{1}\{t \geq 1982\} \times D_i$	-0.91 (0.653) [-2.20, 0.37]	-0.61 (0.897) [-2.38, 1.15]	0.46 (1.594) [-2.68, 3.59]	-7.23 (5.354) [-17.76, 3.30]
R-Squared	0.792	0.829	0.949	0.935
Mean Dep. Var.	12.75	21.66	81.15	105.90
Observations	2,628	2,628	2,628	2,628
Nr. of Wards	383	383	383	383
Mean Exposure Var.	0.50	0.50	0.50	0.50
S.D. Exposure Var.	0.25	0.25	0.25	0.25

Notes. The table presents estimates of β in Equation 2. The sample excludes gynecology and obstetrics wards. D_i is the standardized share of unplanned admissions in each ward i , in 1980. Each column refers to a different dependent variable, measured per thousand admissions. Columns 1-3 present estimates for regressions where in-hospital, 30-day post-discharge, and 1-year post-discharge mortality are the dependent variables. Column 4 reports results for 30-day readmission rates after discharge. All observations are weighted by the total number of admissions (Panel A), unplanned admissions (Panel B), and planned admissions (Panel C) in each ward in 1981. We present the weighted mean of the dependent variable pre-reform. 95% confidence intervals, computed from robust standard errors clustered at the hospital ward level (in parentheses) are presented between brackets.

G. Additional Figures

Figure G.6. PHYSICIANS' LABOR DEMAND IN DENMARK (1972-1991)



Notes. Panel (A) displays the total number of posted and filled positions for all hospital physicians in public hospitals in Denmark, between 1972 and 1991. Hospital physicians during this period include chief physicians, junior residents and senior residents. Panel (B) displays the total number of posted and filled positions disaggregated by career stage. The data comes from digitized statistical yearbooks produced by the Danish Health Authority.

G.1. Mortality Changed Along the Dose

The event–study and DiD estimates impose a linear relationship between pre-reform overtime need and the post-reform change in mortality. To test for heterogeneity along D_i and visualize the shape of the relationship, we start by computing, for every ward i , the change in mortality rate:

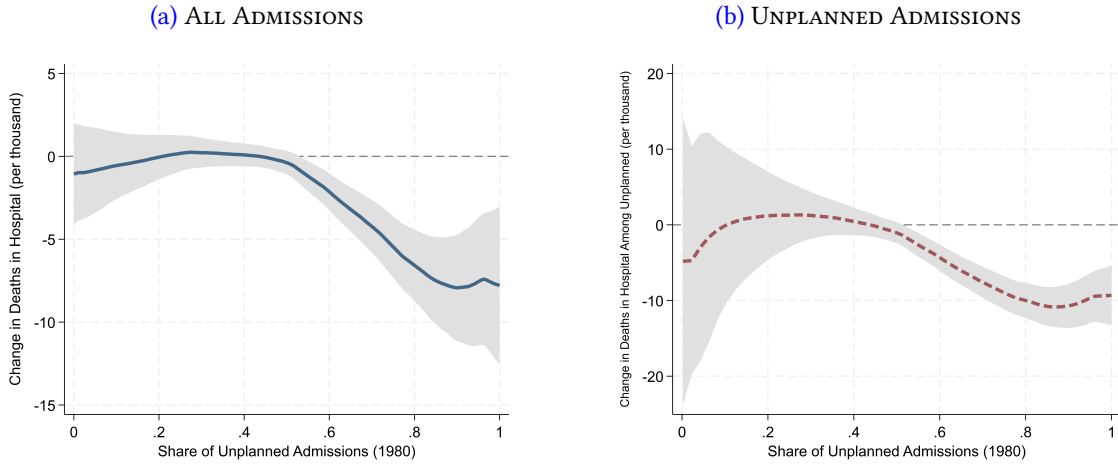
$$\Delta Y_i = \bar{Y}_{i,\text{post}} - \bar{Y}_{i,\text{pre}} = \frac{1}{T_{i,\text{post}}} \sum_{t=1982}^{1984} Y_{it} - \frac{1}{T_{i,\text{pre}}} \sum_{t=1978}^{1981} Y_{it}. \quad (14)$$

We then estimate $\mathbb{E} [\Delta Y_i | D_i = d]$ through a local-linear regression, with a bandwidth of 0.1, and weighting observations by the baseline number of admissions in 1981.

Figure G.7, Panel A shows the fitted curve for all admissions. Figure G.7, Panel B repeats the exercise for mortality among unplanned admissions only. The gray bands illustrate the 95% confidence intervals. The relationship between the share of unplanned admissions and the change in mortality rates across reform periods is essentially flat until about $D_i = 0.5$, with the slope turning negative from then on.

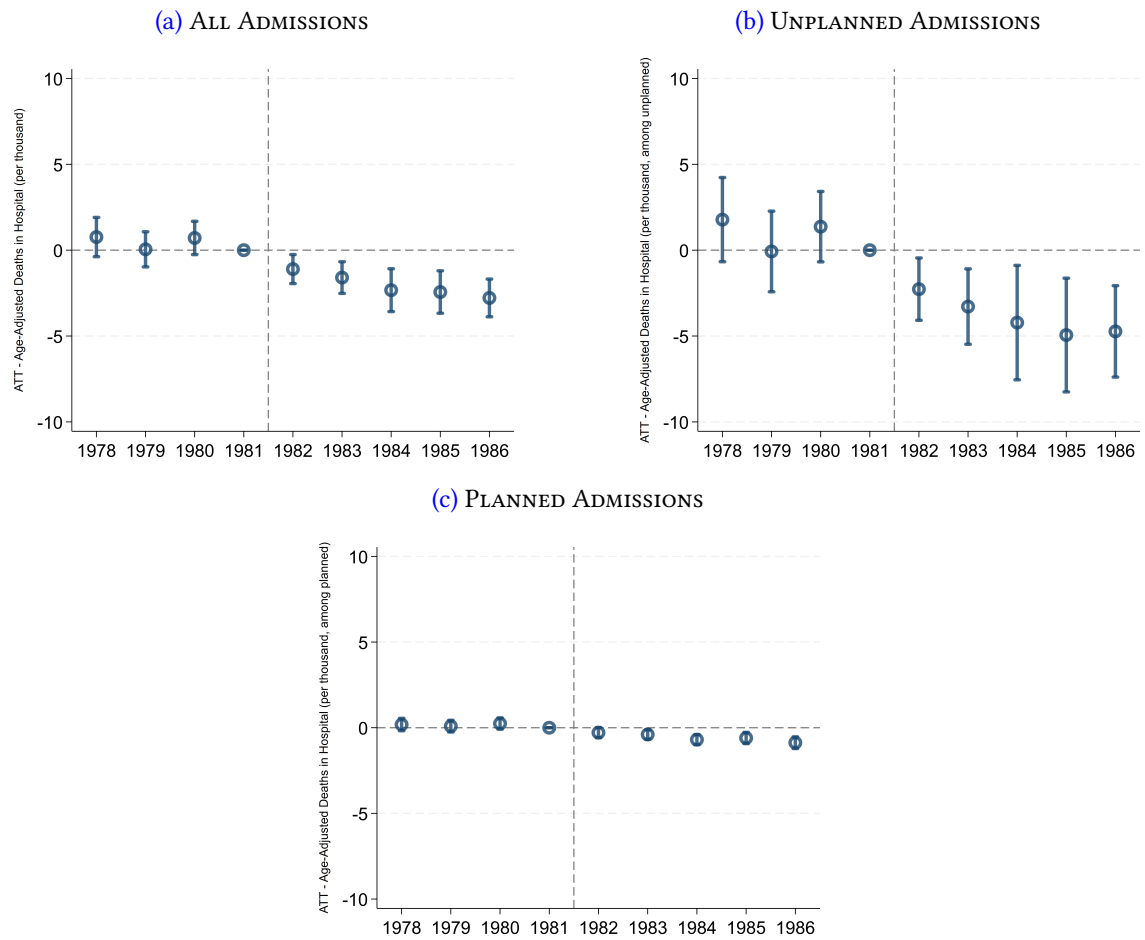
Overall, Figure G.7 reinforces the conclusions from our TWFE estimates. There is a stronger effect on wards where h -to- N substitution induced by overtime restrictions is larger. The effect is heterogeneous in the continuous exposure variable, becoming negative and significant only when the pre-reform share of unplanned admissions is above 50%.

Figure G.7. CHANGE IN HOSPITAL MORTALITY RATES BY DOSAGE



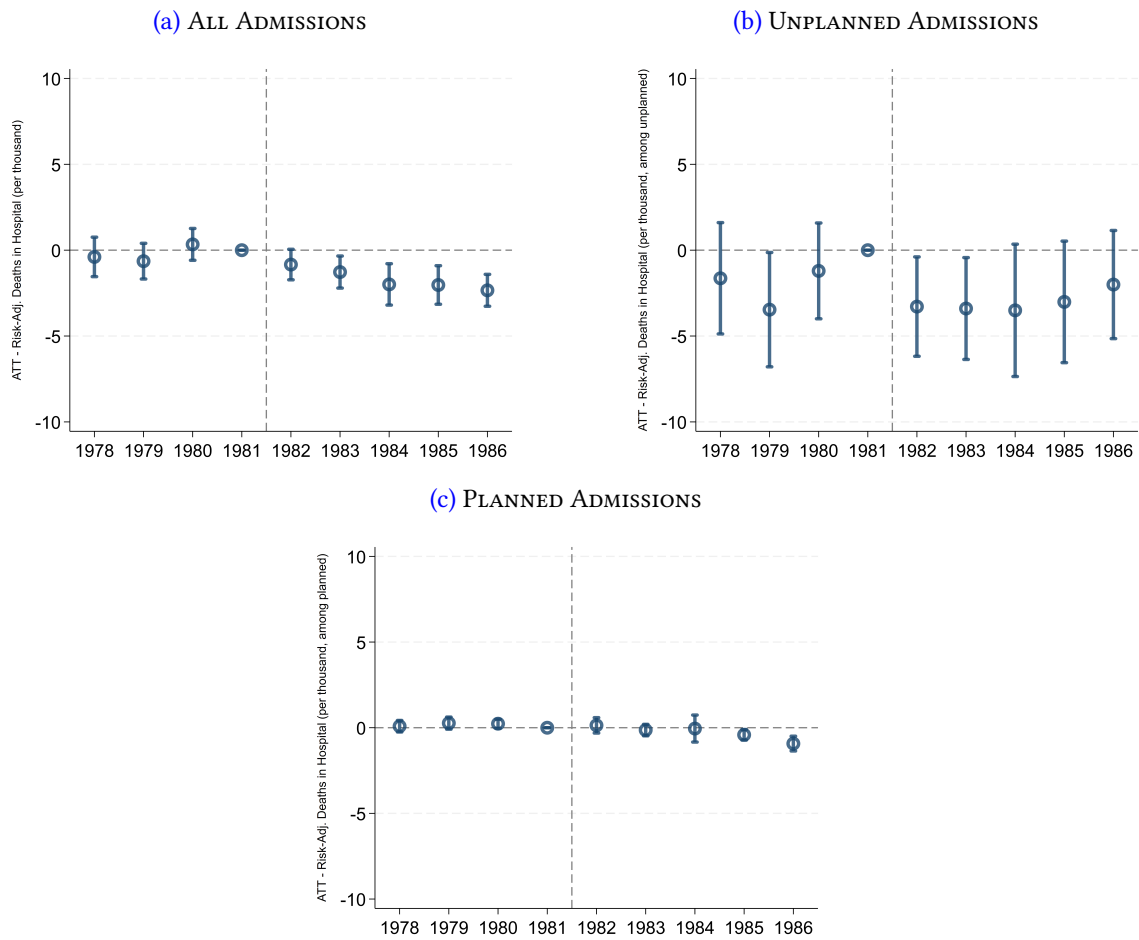
Notes. The figure documents the estimated curve of a local-linear polynomial regression, with bandwidth 0.10 of the post-pre change in hospital mortality (Panel A) or hospital mortality among unplanned admissions (Panel B) on the share of unplanned admissions in the hospital ward in 1980. The dependent variable is the ward-level change between post-reform (1982-1984) and pre-reform (1978-1981) mortality rates, defined as ΔY in the main text. 95% confidence bands are presented in gray.

Figure G.8. AGE-ADJUSTED IN-HOSPITAL MORTALITY: EVENT-STUDY AT THE WARD LEVEL



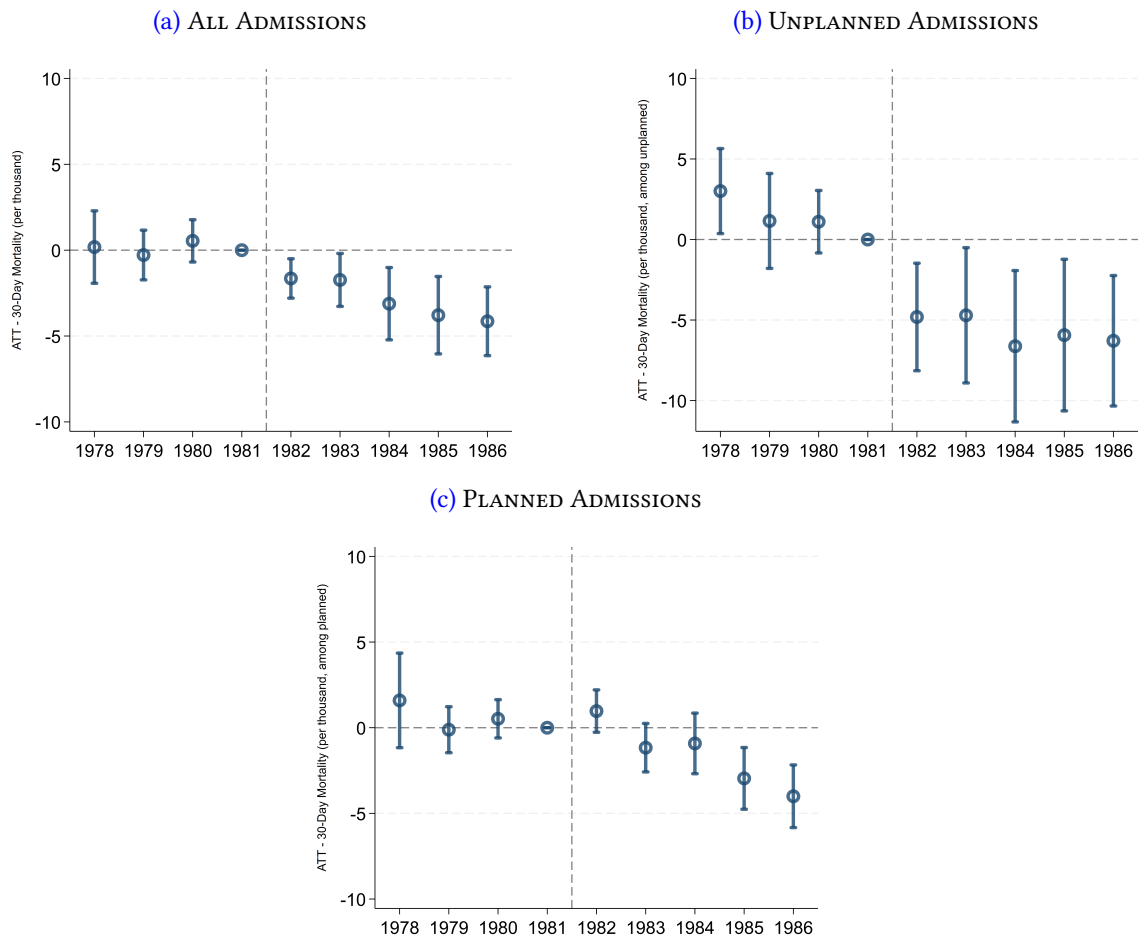
Notes. The figure documents coefficients δ_t and β_t of the event-study specification as in Equation 1 for age-adjusted mortality rates. Panel A presents estimated effects for the change in age-adjusted deaths in hospital per thousand admissions, among all admissions. Panels B and C present the same outcomes among unplanned and planned admissions, respectively. All estimates are relative to the baseline year 1981, and results are weighted by the number of admitted patients in 1981. The dashed vertical lines indicate onset of the reform. The capped bars depict 95% confidence interval using robust standard errors clustered at the hospital ward level.

Figure G.9. RISK-ADJUSTED IN-HOSPITAL MORTALITY: EVENT-STUDY AT THE WARD LEVEL



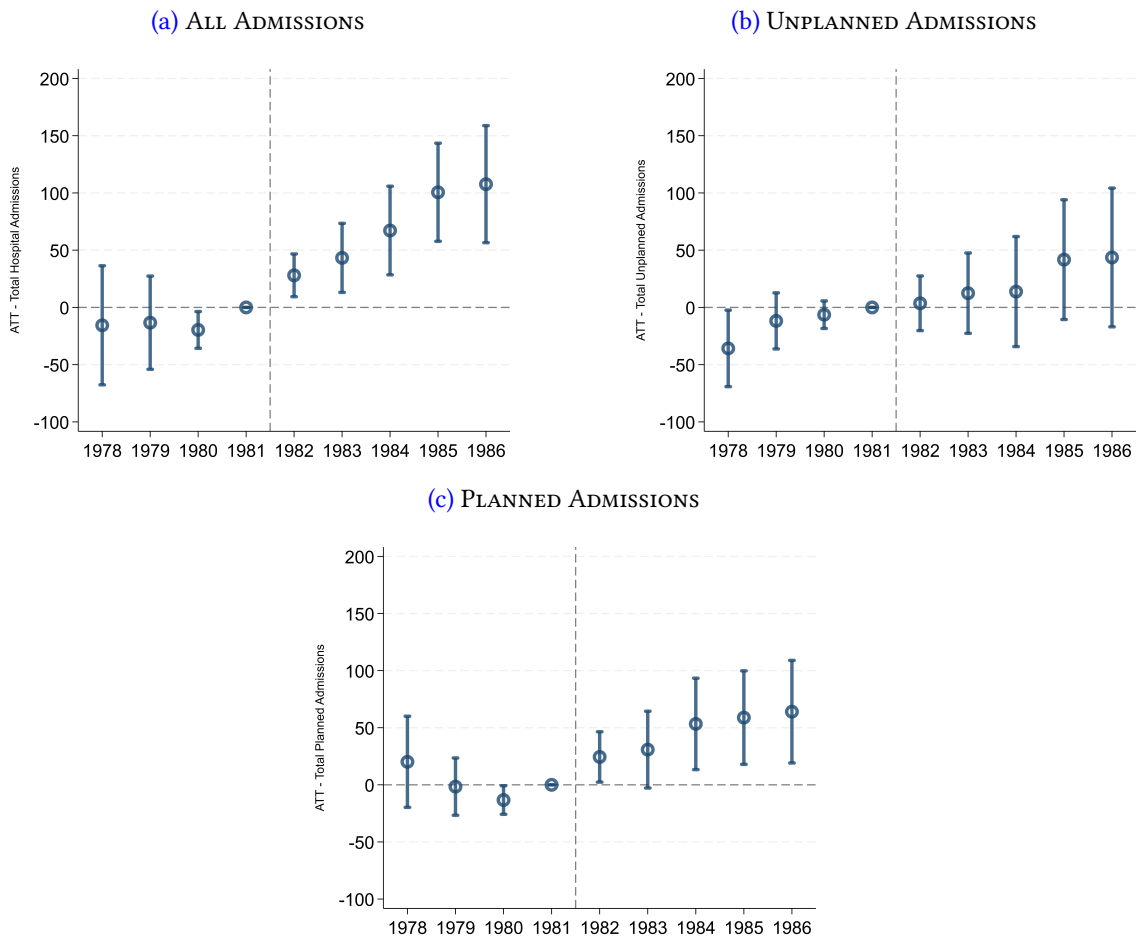
Notes. The figure documents coefficients δ_t and β_t of the event-study specification as in Equation 1 for risk-adjusted mortality rates. Risk-adjusted mortality rates are given by the ratio between actual deaths in hospital and expected deaths given the patient age distribution and case mix of the patients in each ward. Panel A presents estimated effects for the change in risk-adjusted deaths in hospital per thousand admissions, among all admissions. Panels B and C present the same outcomes among unplanned and planned admissions, respectively. All estimates are relative to the baseline year 1981, and results are weighted by the number of admitted patients in 1981. The dashed vertical lines indicate onset of the reform. The capped bars depict 95% confidence interval using robust standard errors clustered at the hospital ward level.

Figure G.10. 30-DAYS MORTALITY RATES: EVENT-STUDY AT THE WARD LEVEL



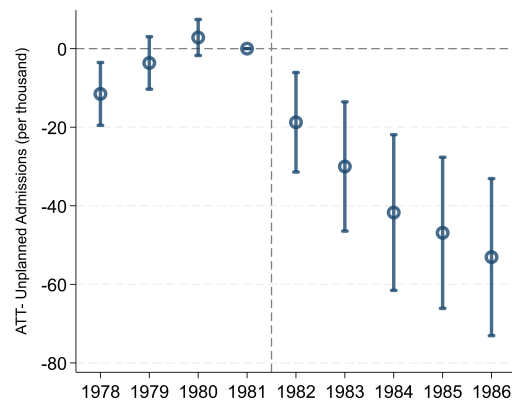
Notes. The figure documents coefficients δ_t and β_t of the event-study specification as in Equation 1 for mortality rates 30 days post-discharge. Panel A presents estimated effects for the change 30-days mortality rate per thousand admissions, among all admissions. Panels B and C present the same outcome among unplanned and planned admissions, respectively. All estimates are relative to the baseline year 1981, and results are weighted by the total number of admitted patients (Panel A), unplanned admissions (Panel B), and planned admissions (Panel C) in 1981. The dashed vertical lines indicate onset of the reform. The capped bars depict 95% confidence interval using robust standard errors clustered at the hospital ward level.

Figure G.11. HOSPITAL ADMISSIONS: EVENT-STUDY AT THE WARD LEVEL



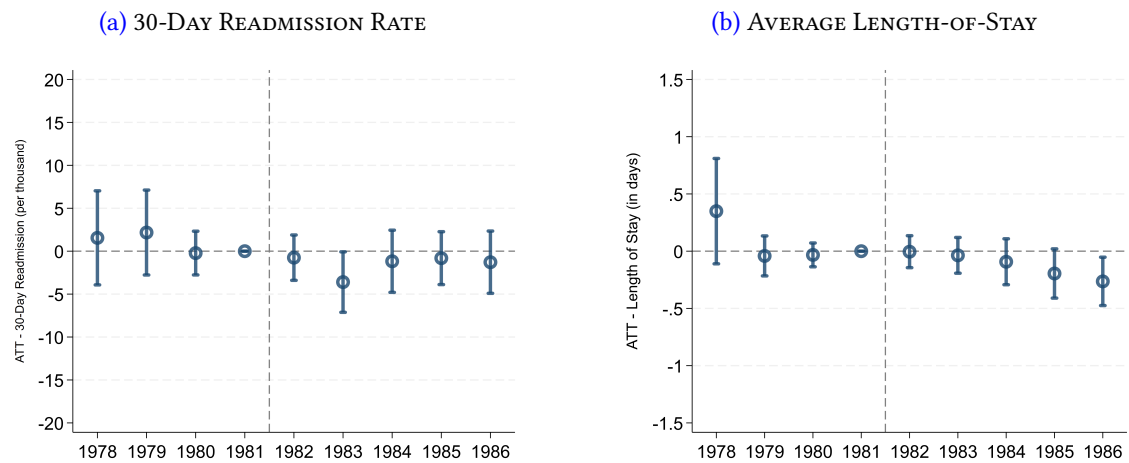
Notes. The figure documents the coefficients of a regression with the same specification as Equation 1. It shows estimated effects for the change in total hospital admissions (Panel A), unplanned admissions (Panel B), and planned admissions (Panel C) relative to the baseline year of 1981. The dashed vertical lines indicate onset of the new working-hour system. The capped bars depict 95% confidence interval using robust standard errors clustered at the hospital ward level.

Figure G.12. EFFECTS OF THE REFORM ON THE PROPORTION OF UNPLANNED ADMISSIONS: EVENT-STUDY AT THE WARD LEVEL



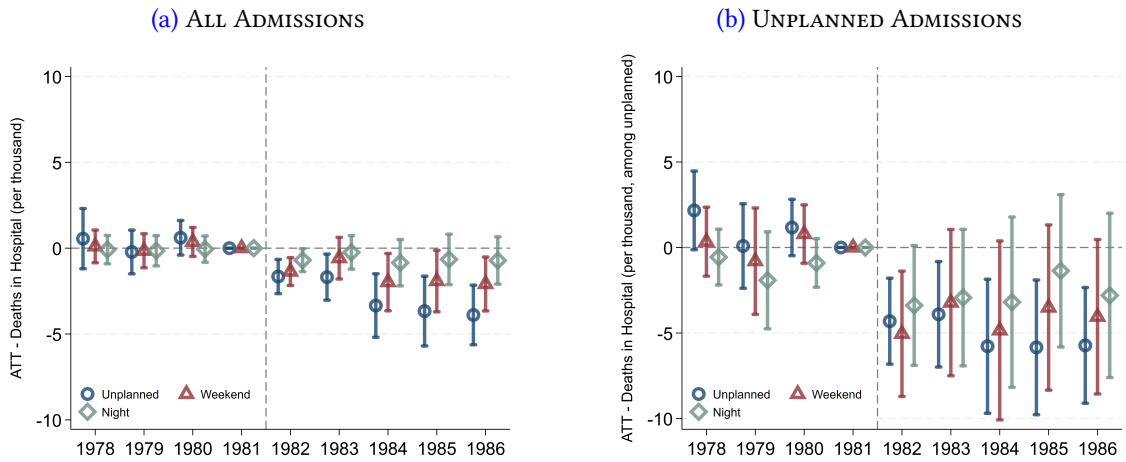
Notes. The figure documents the coefficients of a regression with the same specification as Equation 1. It shows estimated effects for the change in the share of unplanned admissions relative to the baseline year of 1981. The dashed vertical lines indicate onset of the new working-hour system. The capped bars depict 95% confidence interval using robust standard errors clustered at the hospital ward level.

Figure G.13. READMISSIONS AND AVERAGE LENGTH OF STAY: EVENT-STUDY AT THE WARD LEVEL



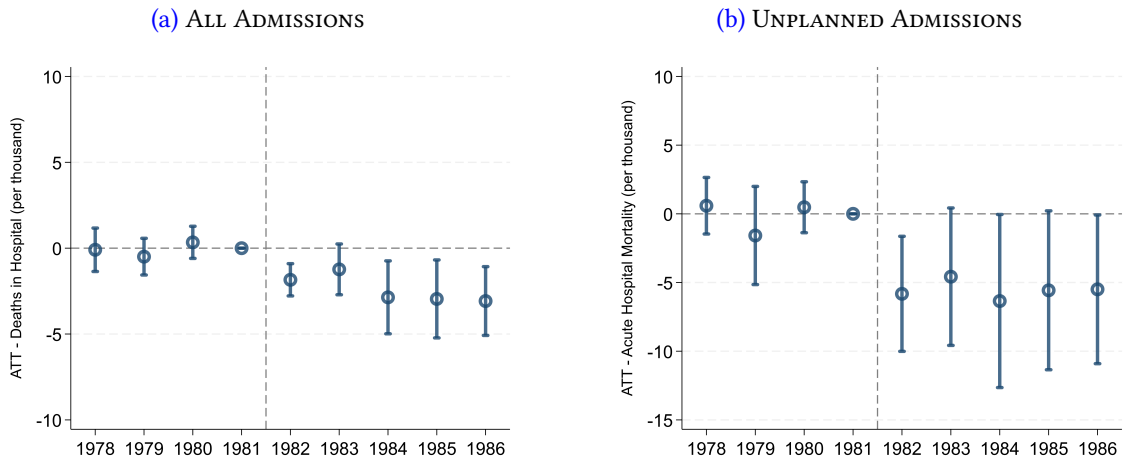
Notes. The figure documents the coefficients of a regression with the same specification as Equation 1. It shows estimated effects for the change in the rate of hospital readmission 30-days after discharge (Panel A), and the average number of days for each admitted patient (Panel B) relative to the baseline year of 1981. The dashed vertical lines indicate onset of the new working-hour system. The capped bars depict 95% confidence interval using robust standard errors clustered at the hospital ward level.

Figure G.14. HOSPITAL MORTALITY RATES - OTHER DOSAGE DEFINITIONS



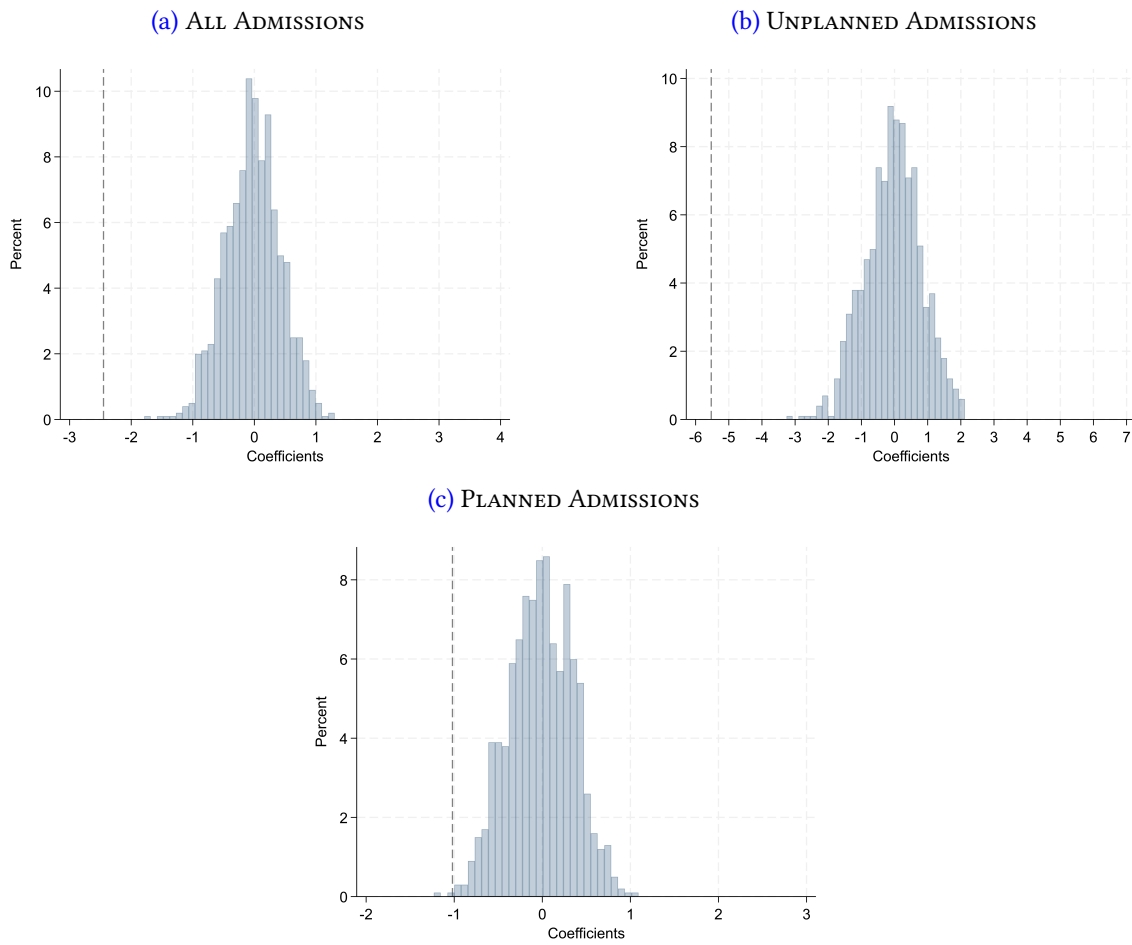
Notes. The figure documents coefficient of a regression with the same specification as Equation 1. It shows estimated effects for the change in deaths in hospital per thousand admissions, among all admissions (Panel A), and among unplanned admissions (Panel B). "Unplanned", "Weekend" and "Night" refer to different definitions of the treatment dosage D_i . Unplanned is the standardized share of unplanned admissions in each ward i , in 1980. The dashed vertical lines indicate onset of the new working-hour system. The capped bars depict 95% confidence interval using robust standard errors clustered at the hospital ward level.

Figure G.15. HOSPITAL MORTALITY: LATENT DOSAGE DEFINITION



Notes. The figure documents coefficient of a regression with the same specification as Equation 1. As the dosage variable uses a latent factor of several predictors of overtime use pre-reform (See Appendix D for more details). Each panel shows estimated effects for the change in deaths in hospital per thousand admissions, among all admissions (Panel A) and unplanned admissions (Panel B). The dashed vertical lines indicate onset of the new working-hour system. The capped bars depict 95% confidence interval using robust standard errors clustered at the hospital ward level.

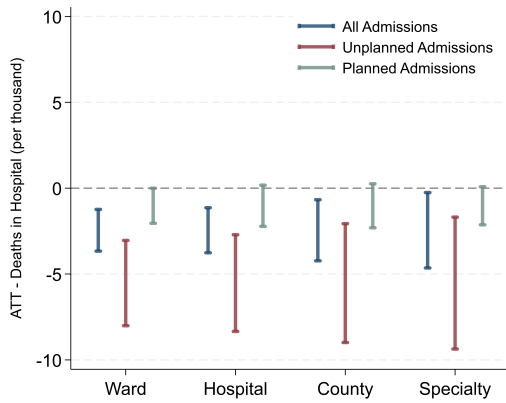
Figure G.16. PLACEBO DISTRIBUTIONS: RANDOM REASSIGNMENT OF DOSAGE VARIABLE



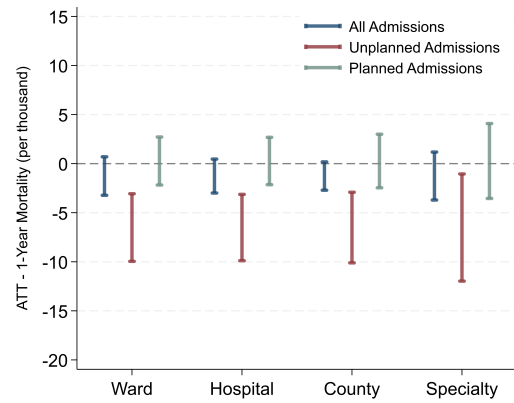
Notes. The figure presents coefficient distributions from 1001 sample permutations. In each sample, the dosage variable—the share of unplanned admissions in 1980—is randomly re-assigned across hospital wards. Each coefficient is the β coefficient in a specifications as in Equation 2. The dashed lines correspond to the estimates in our analysis sample, reported in F.7, Column 1.

Figure G.17. DIFFERENT CLUSTERING OF THE STANDARD ERRORS

(a) EFFECT ON IN-HOSPITAL MORTALITY RATE



(b) EFFECT ON 1-YEAR MORTALITY RATE



Notes. The figure depicts the 95% confidence intervals for different levels of clustering of the standard errors of the β coefficient of Equation 2. The figure presents these for in-hospital mortality rates (Panel A) and mortality rates 1-year post-discharge (Panel B). These include clustering at the ward (same as in the main specification), hospital, county and broad specialty group level. The 95% confidence intervals correspond to the coefficients presented in Panels A through C of Table 2, Columns 1 (for the case of the figure in Panel A) and 4 (for the case of the figure in Panel B).