Improving Performance Through Allocation and Competition: Evidence from a Patient Choice Reform*

Mika Kortelainen[a] ⓡ Liisa T. Laine[b] ⓡ Konsta Lavaste[c] ⓡ Tanja Saxell[d] ⓡ Luigi Siciliani[e]

[a] *University of Turku , Finnish Institute for Health and Welfare*
[b] *University of Missouri, IZA*
[c] *Finnish Institute for Health and Welfare, University of Jyvaskyla*
[d] *Aalto University, VATT Institute for Economic Research, Helsinki GSE*
[e] *University of York*

Abstract

We study the effects of introducing choice and non-price competition into public health care markets with heterogeneous producers. We use a difference-in-differences design based on a regional patient choice reform and comprehensive administrative data on orthopedic surgeries in Finland. We find that the reform led to a reallocation of patients towards large teaching hospitals, increasing concentration in their markets. Waiting times decreased and more patients received care in all the hospitals exposed to the reform, with little effect on clinical quality or surgical expenditure. Back-of-the-envelope calculations reveal that in aggregate, the introduction of the market-based mechanism has sizable social benefits. (*JEL* I11,L11, I18, L32, L38)

Acknowledgements: We thank Rubaiyat Alam, Cristina Bratu, Petri Böckerman, Heather Brown, David Cutler, Partha Deb, Michele Fioretti, Geir Godager, Martin Hackmann, Unto Häkkinen, Tor Iversen, Kaisa Kotakorpi, Laura Lasio, Timothy Layton, Teemu Lyytikäinen, Jennifer Mayo, Jaakko Pehkonen, Tuomas Pekkarinen, Carol Propper, Steve Puller, Anders Munk-Nielsen, Adam Sacarny, Henri Salokangas, Stephan Seiler, Markku Siikanen, Lauri Sääksvuori, Otto Toivanen, Janne Tukiainen, Benjamin Ukert, Anthony Yu, and Jouko Verho for their comments and suggestions. We also thank conference and seminar participants at the 13th Nordic Workshop on Industrial Organization, the 16th Berlin IO Day, AEA/ASSA 2024, ASHEcon 2023, Bank of Finland Research Seminar, CEPR Health Economics 2023, EARIE 2023, EuHEA 2024, FINHESS, IHEA 2023, IIOC 2024, Kansas Health Economics Seminar 2024, NHESG meeting 2022, University of Missouri, University of Turku, and VATT Institute for Economic Research. This research was supported by the Academy of Finland (Decision No. 325110), the Yrjö Jahnsson Foundation, the Instrumentarium Science Foundation, and Vakuutustiedon Kehittämissäätiö. Author order randomized using Ray and Robson (2018) ⓡ. This version: March 3, 2025. Corresponding author: Liisa T. Laine, 615 Locust Street Building, Columbia, MO 65211, USA. Tel: +1-573-882-1335, Email: lainel@missouri.edu.

I. Introduction

In many countries the provision of essential services such as health care and education has traditionally been managed through non-market-based allocation mechanisms in the public sector. During the past decades, many of these services have been deregulated to adopt market-based allocation mechanisms via introduction of choice and non-price competition among public producers. It remains unclear, however, what kinds of equilibrium effects the adoption of market-based allocation mechanisms may have on individual and producer outcomes and what the resulting aggregate effects in these markets are.

We provide quasi-experimental evidence on the effects of introducing a market-based allocation mechanism in public health care using a major policy change in Finland. We study a regional patient choice reform for planned surgeries that was designed to enhance choice and improve hospital performance via increased non-price competition. This reform was implemented in 2007 in the area of South-West Finland, which comprises approximately one-fifth of the Finnish population. Prior to this reform, choice was restricted and patients were allocated to the closest public hospital within their own health care district. After the reform, patients could choose any public hospital within the reform area, while leaving patients and hospitals outside the reform area unaffected. We employ nationwide hospital discharge data and a difference-in-differences (DiD) design that uses the reform area with a market-based allocation mechanism as a treatment group and the remaining area with a restricted choice allocation mechanism as a control group.

Our DiD design with a binary treatment based on the regional patient choice reform provides us with a novel empirical framework for studying how the introduction of choice and non-price competition (the market mechanism) affects the allocation of patients and hospital performance compared to an allocation mechanism with more restricted patient choice. Our research design enables us to estimate the reform's average treatment effects, as well as its aggregate and heterogeneous effects across producers in the entire reform area. Thus, we differ from previous research on patient choice that has estimated the *differential changes* in hospital performance *by market structure* after the nationwide adoption of patient choice reforms (Cooper et al., 2011; Gaynor, Moreno-Serra and Propper, 2013; Moscelli et al., 2018; Brekke et al., 2021; Moscelli, Gravelle and Siciliani, 2021, 2023). To our knowledge, no previous studies have examined patient choice reforms introduced in some areas of a country, while leaving other areas unaffected, as such choice policies have been introduced simultaneously at the national level.

We find that the regional patient choice reform affected patients' hospital choices and allocation across several commonly performed planned surgeries: hip replacements, knee replacements, and all orthopedic surgeries. Specifically, the reform led to a reallocation of patients towards large teaching hospitals, possibly due to their better resources and higher perceived quality (Newsweek and Statista, 2022). This reallocation towards teaching hospitals led to an increase in concentration, as measured by the Herfindahl-Hirschman Index, by up to 8 percent in their markets. Notably, there were no hospital entries or exits post-reform.

We then show that hospitals responded to the reform and the resulting change in the competitive environment by improving their performance. We find that waiting times decreased by up to 60 days or 40 percent and more patients received care in all the hospitals exposed to the reform on average; although the volume increase was much larger in teaching hospitals that have both larger capacity and better resources. Hospitals also shortened the length of stay for orthopedic surgery by 4 hours or 7 percent, with little impact on clinical quality (e.g., emergency readmissions and revision surgeries), patient mix, or surgical expenditure on average. To the extent that clinical quality did not change, shorter stays indicate that hospitals used their resources more efficiently (Cooper, Gibbons and Skellern, 2018). This might have enabled hospitals to increase healthcare production post-reform.

Finally, we evaluate the net social benefit of introducing choice and non-price competition for orthopedic surgery in aggregate, as implied by the estimated increase in hospital volume (in total, an increase of $1,628$ patients or 8 % annually), faster access to medical treatment, and cost savings from shorter stays. Our back-of-the-envelope calculation suggests a net benefit of approximately $3,400$ euros per patient or an aggregate of 80 million euros a year in the entire reform area.

Our study demonstrates how market-based choice reforms can improve efficiency in health care production and patient allocation to hospitals. Thus, we contribute to the literature on health care competition, as reviewed by Gaynor, Ho and Town (2015) and Handel and Ho (2021). Previous studies have used DiD designs to estimate the differential changes in hospital performance by market structure after the simultaneous adoption of nationwide choice reforms, with variation in the effects of the reforms by market structure (Cooper et al., 2011; Gaynor, Moreno-Serra and

Propper, 2013; Brekke et al., 2021; Moscelli, Gravelle and Siciliani, 2021, 2023).[1] In contrast, we estimate the market-wide average and aggregate effects of enhanced versus more restricted patient choice using a binary DiD approach. We also compare these two approaches in our application. We highlight that the DiD approach used in the previous literature estimates a different parameter of interest than the average effect of the choice reform and can be understood as a heterogeneity analysis between more and less concentrated markets for the average effect that we can identify using the binary DiD approach with the unaffected control group. Furthermore, our binary DiD approach allows us to use hospital choices and market concentration as outcomes.[2] This differs from the previous studies that have used market concentration as a right-hand side variable defining differential exposure to the choice reform, rather than as an outcome. Our results suggest that improvements in conditions to compete due to greater choice can promote market concentration towards larger producers.[3]

Our paper thus also contributes to the debates on the causes and consequences of large producers' market dominance (Gaynor, 2016; Autor et al., 2017; Bighelli et al., 2022). On the one hand, large producers' market dominance could result from reduced competition, which can lead to higher health care prices (Gowrisankaran, Nevo and Town, 2015; Cooper et al., 2019) or longer waiting times (Gaynor, Laudicella and Propper, 2012) without improvements in quality. On the other hand, it could result from enhanced competition and choice, enabling markets to reallocate economic activity towards larger producers due to their competitive advantages (Van Reenen, 2018; Berry, Gaynor and Scott Morton, 2019; De Loecker and Syverson, 2021; Bighelli et al., 2022). Our results are consistent with the latter conjecture and suggest that pro-competitive choice reforms can improve the functioning of public hospital markets while increasing large hospitals' market dominance.

[1] DiD designs using variation in policy effects or exposure, rather than assignment to treatment and control groups, are used in many settings, especially when nationwide reforms or shocks apply at the same time to all individuals or the population of interest (Duflo, 2001; Acemoglu, Autor and Lyle, 2004; Finkelstein, 2007).

[2] Our reduced form approach relying on a quasi-experiment complements Gaynor, Propper and Seiler (2016), who estimated a structural demand model to study changes in the quality elasticity of demand after a nationwide choice reform.

[3] We also contribute to the literature studying market-based mechanisms in public services. For example, Campos and Kearns (2023) document the positive effects of a centralized school choice program on educational outcomes and school quality based on a DiD design similar to ours. Our findings indicate that while excess demand may diminish the effectiveness of enhanced patient choice in improving quality (Chalkley and Malcomson, 1998), enhanced choice can improve performance by shortening waiting times and increasing hospital volumes.

Finally, our study links the empirical literature analyzing the impacts of market-based choice reforms with an extensive literature studying the allocation of economic activity among heterogeneous producers in a variety of industries (e.g., Syverson, 2011; De Loecker and Syverson, 2021). Chandra et al. (2016) find that higher-quality hospitals gain more patients and grow more over time, suggesting an important role for market forces in health care. Unlike Chandra et al. (2016), we employ a quasi-experiment to provide novel insights into how the introduction of non-price market mechanisms change allocation and performance among heterogeneous producers. Our findings indicate that the choice reform led to a reallocation of activity towards large teaching hospitals and incentivized all hospital types (teaching and non-teaching) to shorten waiting times. Our results are thus informative about the effectiveness of choice reforms in improving public producers' performance in the presence of regulated prices and potentially overloaded waiting lists.

The rest of the paper proceeds as follows. Section II describes the institutional setting. Section III presents the data and descriptive statistics. Section IV describes our binary DiD approach for estimating the reform's effects using patient-level data and Section V presents the corresponding results. Section VI presents hospital-level analyses and Section VII presents back-of-the-envelope cost-benefits calculations. Section VIII applies the DiD approach using variation in the effects of the choice reform by market structure and highlights the link between this approach and our binary DiD approach. The last section concludes.

II. Institutional Setting

II.A. The Finnish Health Care System

Finland has a publicly-financed, decentralized health care system, where all residents are entitled to public health care services via universal public insurance. Public primary care is organized and financed by the municipalities ($N = 326$ in 2010) for their residents by law and it is provided in municipality-owned health centers. Primary care physicians act as gatekeepers for publicly provided non-emergency (planned) specialized health care.

Specialized health care such as planned surgical treatment is provided by public hospitals that are governed by public authorities called hospital (i.e., health care) districts ($N = 20$ in 2010). Each municipality is a member of one of the hospital districts and participates in governance and financing of that district together with other member municipalities. The hospital districts are responsible for organizing specialized health care services in their region. The public sector provides the majority of hospital care in Finland and the private sector accounts for only 5 percent of hospital activity (Keskimäki et al., 2019).

The Finnish hospital sector consists of heterogeneous producers. There are large university-based teaching hospitals, medium-sized central hospitals, and small regional hospitals. Every hospital district has either a teaching hospital or a central hospital and possibly also one or more regional hospitals. Central hospitals provide services in most specialties while small regional hospitals only provide services for common medical conditions and specialties such as orthopedics. Teaching hospitals, on the other hand, have better resources and expertise to provide services in all specialties and for a broad range of medical conditions. Teaching hospitals also partner with medical schools to provide education and conduct research, contributing to knowledge production and innovation, and they tend to be the early adopters of new technologies (c.f. Skinner and Staiger, 2015). There are five teaching hospitals in Finland and they have better perceived quality (Newsweek and Statista, 2022), higher patient volumes and larger capacity compared to non-teaching (central or regional) hospitals ($N = 58$) (Karhunen, 2020). Their better perceived quality and resources give teaching hospitals a competitive advantage over non-teaching hospitals.

II.B. Regional Patient Choice Reform and Incentives

We study a regional patient choice reform that was introduced in four hospital districts in South-West Finland in October 2007, comprising approximately one-fifth of the Finnish population, or approximately one million citizens. The regional reform concerned planned (non-emergency) surgeries and it preceded a nationwide choice reform implemented in May 2011. Figure 1 depicts the reform area in shaded color, the hospital districts ($N = 20$, of which 4 were in the reform area and 16 in the control area), and the geographical distribution of hospitals ($N = 63$, of which 15 were in the reform area and 48 in the control area) in Finland in 2004–2010.[4]

Prior to the reform, patient choice was highly restricted, and patients were typically referred to the closest hospital within their own hospital district (Finnish Government Proposal 90/2010). After the reform, surgical patients who lived in the shaded reform area could choose any public

[4]Online Appendix Figure A.1 further shows the locations of different hospital types: teaching, central, and regional.

hospital within and across hospital districts in the reform area,[5] leaving patients and hospitals outside the reform area unaffected.[6]

The reform's central policy goals were to adopt a market-based allocation mechanism by introducing choice and to enhance hospital performance (e.g., clinical quality and waiting times) through increased non-price competition (Pirkanmaa Hospital District, 2007). According to our hospital expert interviews, policymakers in hospital districts in the reform area also hoped that patients would substitute to hospitals with shorter waiting times, consequently shifting demand to ease the pressure in overly crowded hospitals. Indeed, hospital waiting times for non-urgent care were long in Finland. For example, in 2007, approximately 13 percent of patients had to wait more than 6 months for hospital care, which is the national maximum waiting time target set by law in March 2005 (THL, 2012).

The patient choice reform was inspired by earlier market-based patient choice reforms implemented in other Nordic countries and the United Kingdom (Pirkanmaa Hospital District, 2007). However, unlike these national reforms, the Finnish reform was regional and left patient choice outside the reform area unaffected and more restricted. Thus, the reform provides a novel setting for studying the effects of enhanced versus more restricted patient choice on public hospital choice and performance.

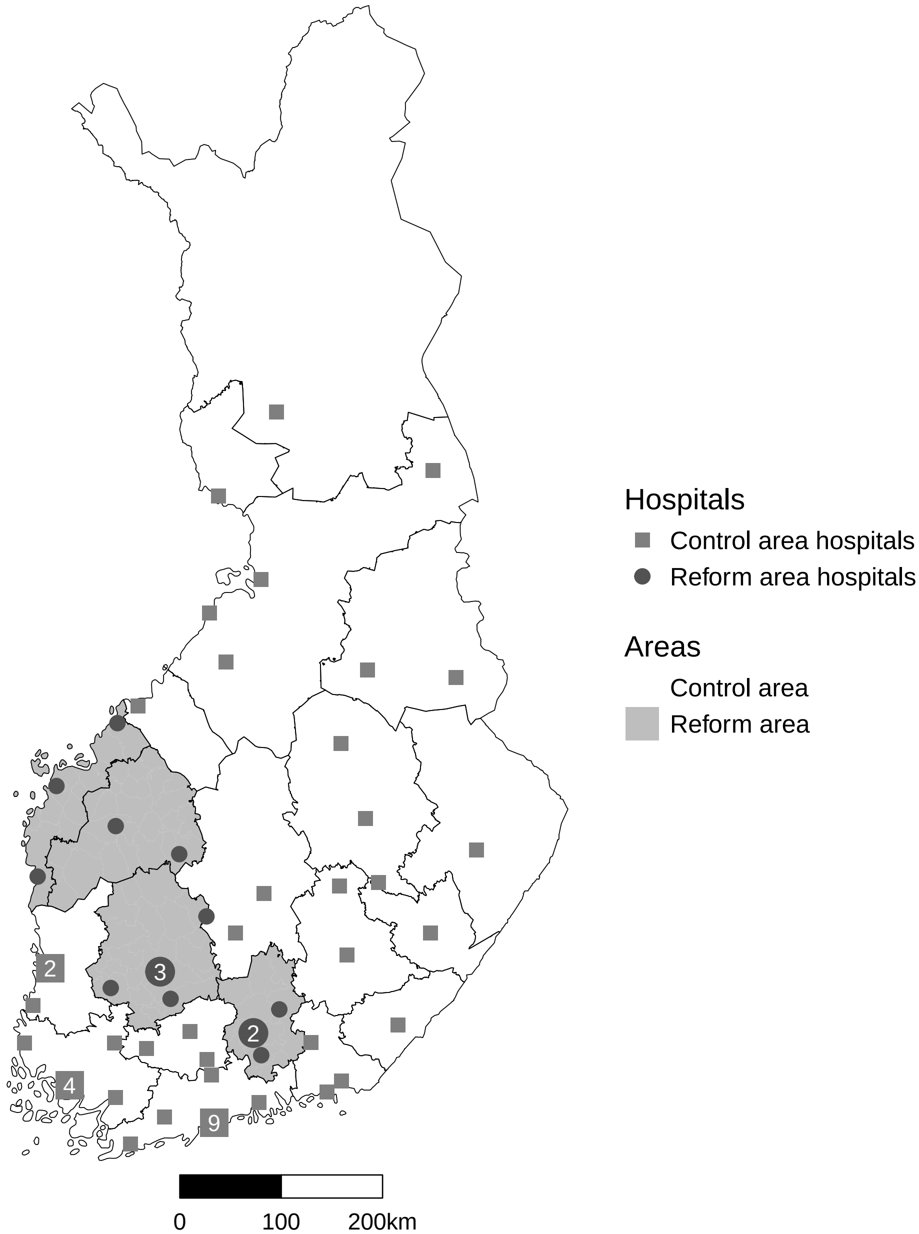
Patients access planned surgical treatment by a referral from their primary care physicians, and patients' hospital choices are guided by these physicians. The financial incentives for physicians associated with hospital referral decisions are minimal because public primary care physicians are salaried employees of municipalities rather than hospital districts. Private primary care physicians are also able to make referrals to public hospitals, but the receiving hospitals' specialists make the final decision on whether it is necessary for the patient to undergo the procedure or not.[7]

[5] Small municipality-owned hospitals were not officially part of the reform. We include them in the estimation samples, but have confirmed that their exclusion would not alter the results.

[6] In a few hospital districts in and outside the reform area, physicians were allowed to refer patients to hospitals located outside patients' hospital districts, but only when specific circumstances (such as long travel distance or waiting time) were met. Our results are robust to excluding such cases from the econometric analyses (Section V.C).

[7] A typical treatment episode for an orthopedic surgery proceeds as follows: (i) a primary care physician diagnoses the condition, helps in choosing a hospital, and writes a referral; (ii) the patient meets a specialist at the hospital, who makes the final treatment decision; (iii) the patient undergoes a surgery and in-patient care at the hospital; (iv) the patient receives post-operative care (in the same or different hospital) or is discharged; and (v) follow-up checks are performed.

Figure 1: Patient Choice Reform Area and Hospital Locations



Notes: Borders indicate hospital districts in 2007. The reform area includes four hospital districts, which comprise approximately one-fifth of the Finnish population, or approximately one million citizens. Finland is approximately the size of Germany, but with 1/15 of the population ($\sim$ 5.4 million citizens in 2010). The large squares and circles mark regions which had multiple (2–9) hospitals. In total, there were $N = 15$ hospitals in the reform area and $N = 48$ hospitals in the control area. The figure includes all public hospitals that performed planned orthopedic surgeries, although some of them did not perform hip and/or knee replacement surgeries.

Finnish public hospitals are reimbursed for the services produced from their patients' municipalities of residence and the money follows the patients. Each hospital district sets the reimbursement rates of their own hospitals administratively and bases these on administered prospective payments using diagnosis-related groups (DRGs) (Kautiainen, Häkkinen and Lauharanta, 2011).[8]

Health insurance reduces patients' financial consequences related to their hospital choice. In Finland, patients' co-payments for publicly provided health care services are moderate, capped by national legislation, and do not vary much between hospital districts (Hetemaa et al., 2018). For example, in 2008, the maximum co-payment for a surgery was 83.90 euros (Finnish Government Decree 464/2008), which at the time corresponded to 4 percent of the monthly disposable cash income of a median household.[9] Hence we do not expect hospital co-payments to affect patients' hospital choices to a great extent. However, more significant monetary costs can result for patients choosing a more distant hospital because (i) the Finnish population is spread out over a large geographical area, (ii) the distances between hospitals are long, and (iii) public insurance covers travel costs based on the cheapest mode of transport to the nearest hospital, regardless of the actual mode of transport or hospital choice (Paltta, 2008).

Although public hospitals are not profit maximizers in the same way as private for-profit hospitals are, they can face significant pressures to perform financially well and compete for patients by using non-price strategies (e.g., Gaynor, Moreno-Serra and Propper, 2013; Gaynor, Ho and Town, 2015).[10] Because money follows patients through municipality reimbursements in Finland, hospitals' can increase their revenues by attracting more patients if they have capacity to treat them.[11]

Given that administered co-payments are almost fixed, the way hospitals can attract more patients is by improving the quality of care and/or shortening waiting times. Moreover, to shorten waiting times, hospitals can increase supply by increasing resources or by making health care delivery more productive and efficient (OECD, 2020). Patient choice reform brought about a substantial shift in hospitals' ability to attract and compete for patients, with potential efficiency gains

[8]All of the reform area hospital districts reimbursed DRG tariffs either exclusively or combined with a fee-for-service model during our study period 2004–2010. We are not aware of any major changes in the hospital reimbursement systems during the study period.

[9]€83.90/(€28,400/12 months) $\approx$ 0.04 (Statistics Finland, 2010).

[10]Hospitals are assumed to be semi-altruistic with patient utility entering the hospital's objective function to some degree (Ellis and McGuire, 1986; Brekke, Siciliani and Straume, 2011, 2008).

[11]Hospitals can attract patients, for example, from outside their own hospital districts after the choice reform.

in health care production (Gaynor, Moreno-Serra and Propper, 2013; Longo et al., 2019).

Patients' ability to choose their hospital, and thereby the intensity of non-price competition, depends on the available information that the patients and their referring physicians can use when making their choice (Brown et al., 2023). There is publicly available information on hospital performance outcomes in Finland; although no specific patient review system similar to the one maintained by the English National Health Service is provided.

Hospital districts publish information on hospital-level waiting times by specialty on their individual websites. Nationwide statistics on hospital-level waiting times are also collected for common procedures such as hip and knee replacements (THL, 2012)—the surgeries we study—and the statistics were available throughout our study period. Moreover, the Finnish Institute for Health and Welfare (THL) publishes information on the clinical quality outcomes of hip and knee replacements, acute myocardial infarction (AMI), and stroke patients at hospital or hospital district level for example in 2007 (THL, 2021). We are not aware of any major changes in public information provision during our study period. In addition to clinical quality information, patients and their referring physicians can receive other information from unofficial sources such as colleagues, family, and peers. Patients can also learn about the choice reform from their referring physicians, unofficial sources, and the news (the Finnish Broadcasting Company, 2007).

III. Data

We use nationwide (de-identified) patient-level hospital discharge data that contains public hospital admissions and discharges in Finland in 2004–2010. Using the information on medical procedures produced by hospitals, we create three samples of orthopedic surgeries. Orthopedic surgeries are commonly studied procedures in the previous literature on hospital performance (Moscelli, Gravelle and Siciliani, 2021; Godøy et al., 2024). The first two samples include planned hip and knee replacement surgeries because we expect scope for choice and producer competition as a result of the choice reform: they were among the most common planned surgeries and were available in all types of hospitals. To get a more comprehensive picture of the choice reform, we also study the third sample containing all planned orthopedic surgeries. The sample also includes planned hip and knee replacement surgeries, which account for approximately 15 percent of the observations in the sample.[12]

In total, our samples contain 63 hospitals during the observation period, including in total 31,404 observations for the samples of hip replacement surgeries, 39,427 observations for knee replacement surgeries, and 473,617 observations for all orthopedic surgeries. We focus on patients aged 18–74 years at the time of hospital admission.[13] Moreover, we have excluded from the data a small portion (1–2 percent) of patients who obtained orthopedic surgery across reform and control area borders. Furthermore, one regional hospital closed down during the pre-reform period and we exclude all patients who received care in that hospital or lived close to it (see online Appendix Figure A.10). We match each observation with administrative data from Statistics Finland on the patient's date of death, demographics, and residence at the end of each year. Next we discuss the construction of the main variables, leaving the more detailed sample and variable description to online Appendix Section A1. We conclude this section by providing our descriptive statistics.

III.A. Measures of Hospital Choice

We construct four outcome variables related to hospital choice and allocation for surgical patients. Our first choice variable indicates if the patient received care in a teaching, rather than a non-teaching, hospital. Teaching hospitals are larger, have better resources and capacity and a better perceived quality compared to non-teaching hospitals (Section II.A; Silber et al. 2020; Newsweek and Statista 2022).

Our second choice variable is the distance traveled, which is the straight-line distance between each patient's residence and the location of the hospital where the patient received care. The third choice variable indicates if the patient's surgery was at the nearest hospital. And our fourth choice variable indicates if the patient's surgery was outside their hospital district of residence.

[12] Online Appendix Table A.1 shows the observation shares of hip and knee replacements and other common orthopedic surgeries. The most common orthopedic surgery is surgery for knee meniscus that requires small cuts and is less invasive than hip and knee replacements.

[13] We do not have detailed information on the residence of patients over 74 years of age, and our data contains only a few patients under 18 years of age. The previous literature on patient choice reforms also focuses on patients under 75 years of age (Gaynor, Moreno-Serra and Propper, 2013).

III.B. Measures of Hospital Performance

Clinical quality measures. We consider several indicators of care quality to measure hospital performance. First, we construct an indicator of whether the surgical patient had an emergency readmission within 30 days of discharge. Emergency readmissions are a commonly used metric of hospital quality (Varkevisser, van der Geest and Schut, 2012; Gupta, 2021; Moscelli, Gravelle and Siciliani, 2021). We also conduct additional analyses using more detailed measures of clinical quality for hip and knee replacements such as revision surgeries and mechanical complications and infections in the prosthesis (Section V.B).[14]

Waiting time. In addition to clinical quality, waiting time is another commonly studied aspect of hospital performance (Propper, Burgess and Gossage, 2008; Moscelli, Gravelle and Siciliani, 2021, 2023), and one of the targets of the choice reform (Section II.B). We measure waiting time by the number of days that the patient is on the waiting list for surgery.[15] Waiting times delay access to medical treatment and can depend on how efficiently hospitals organize their health care production (Section II.B).

Measures of efficiency and resource use. Measuring performance based on hospital efficiency (how well resources are used to achieve the output) is generally difficult because of the lack of high-quality data on clinical quality and costs (Cooper, Gibbons and Skellern, 2018). Even though our discharge data is detailed in many dimensions, it does not include comprehensive information on costs or the physical and human resources used. Therefore, we follow the previous literature and use length of stay as a proxy for hospital efficiency (Robinson et al., 1988; Gaynor, Moreno-Serra and Propper, 2013; Cooper, Gibbons and Skellern, 2018; Moscelli, Gravelle and Siciliani, 2021). To the extent that clinical quality does not change, a shorter stay indicates faster discharge, and thereby lower costs and less resources used for the same patient outcomes.

We also separately analyze a coarse measure of annual hospital operating expenditure (e.g., purchases of labor and material inputs) for all surgeries collected from all individual hospitals by

[14] We also study the quality of *emergency* care, which was not targeted by the reform, in order to evaluate the spillover effects of the reform (Section V.B).

[15] The patient is placed on the waiting list when a specialist has made the final decision on surgery. In most cases, waiting times are recorded comprehensively in the discharge data. However, some hospitals record waiting times less comprehensively (online Appendix Section A1), and our results are robust for excluding these hospitals from our samples (Section V.C).

the Finnish Institute for Health and Welfare (Section VI). In the discharge data, planned orthopedic surgeries represent more than 20 percent of all planned surgeries, although their share of hospital costs might be very different.

III.C. Measures of Hospital Volume and Market Structure

We measure hospital volume by the number of patients. We also measure market structure at the hospital level using the Herfindahl-Hirschman Index (HHI), which is based on patient market shares and allocations to each hospital (Gaynor, Moreno-Serra and Propper, 2013). In step one, we calculate a HHI value for each municipality by taking the sum of the squared patient market shares of hospitals using data on patients from that municipality only. In step two, we calculate the hospital-level HHI values by taking a weighted average of the values of the municipal-level HHI, where each municipality is weighted according to its share of the hospital's total patient volume. Because we use market shares in this calculation, the HHI varies between 0 and 1, and hospitals located in highly concentrated markets (high hospital-level HHI) also have high market shares (correlation 0.85–0.89). The hospital-level HHI measure captures the average degree of concentration in markets where the hospital operates, and allows large hospitals to operate in a greater market area, in comparison to smaller hospitals. We refer to this HHI calculated from observed hospital choices as the *actual* HHI and we calculate it separately for each estimation sample.

III.D. Patient and Surgery Type Covariates

Our main covariates are patient's age and sex at the time of admission in addition to the patient's surgery type, because many possible covariates (such as financial position or staffing) may be endogenous. We also estimate models using two additional covariates as a robustness check (Section V.C). The first is an indicator for a weekend admission (equal to one if admitted on Saturday or Sunday), as staff may be more limited on weekends. The second is the number of past emergency admissions each patient had within one year prior to surgery, as a proxy for severity and morbidity (case mix).

III.E. Descriptive Statistics

Table 1 reports the descriptive statistics of hospital choice outcomes for the samples of hip replacements, knee replacements, and all orthopedic surgeries (Panels A–C) in the pre-reform period (Q1/2004–Q3/2007). We present the statistics separately for the reform and control areas. In each sample, approximately one fifth of the patients resided in areas affected by the reform.

We find that patients in the reform and control areas had some differences in their hospital choice outcomes in the pre-reform period. For example, the hip and knee replacement patients were more likely to travel longer distances on average (31–32 vs 28 kilometers) and to receive care in a teaching hospital compared to those in the control area (56–59 vs 34–39 percent). For all orthopedic surgeries, the distance traveled was, however, shorter on average (23 versus 27 kilometers) and the probability of choosing a teaching hospital was lower in the reform than in the control area (23 versus 41 percent).

Table 1 also shows that hospital volumes were larger in the reform than in the control area in the samples of hip and knee replacements, whereas the reverse was true in the sample of all orthopedic surgeries. Moreover, the hospital-level means of actual HHI indicate a high degree of market concentration at the level of 0.63–0.87, with fairly high variation across hospitals and/or over time in the pre-reform period (SD 0.11–0.14).

Table 2 reports the descriptive statistics of our patient-level hospital performance outcomes in addition to the patient characteristics. Hospital performance was fairly similar between the reform and control areas in terms of the clinical quality outcome, the probability of 30-day emergency readmission (5–10 percent), and length of stay (2–8 days). In contrast, waiting times were longer in the reform area than in the control area (153–265 versus 144–212 days).

Table 2 shows that patient characteristics (age and sex) were similar in the reform and control areas in the pre-reform period. Only the average number of pre-surgery emergency admissions was smaller in the reform than in the control area, suggesting that hospitals in the reform area provided care to less severe patients. We address the differences between the reform and control areas and study the plausibility of the parallel trends assumption in our econometric approach in Section IV.

Hospitals were also heterogeneous in their volumes in the pre-reform period. Large teaching hospitals had much larger volumes than non-teaching hospitals, including medium-sized central

Table 1: Pre-Reform Descriptive Statistics of Choice Outcomes and Market Concentration

	Reform area			*Control area*		
	Mean	SD	N	Mean	SD	N
Panel A. Hip replacement surgeries						
Teaching hospital	0.56	0.50	3,432	0.39	0.49	12,565
Distance (km)	32.35	37.54	3,432	28.20	39.80	12,565
Nearest hospital	0.81	0.39	3,432	0.82	0.38	12,565
Different hospital district	0.06	0.23	3,432	0.03	0.17	12,565
Hospital volume	47.67	44.94	72	26.40	29.48	476
Actual HHI	0.86	0.11	72	0.83	0.14	476
Panel B. Knee replacement surgeries						
Teaching hospital	0.59	0.49	3,845	0.34	0.48	15,603
Distance (km)	30.88	35.13	3,845	27.79	40.27	15,603
Nearest hospital	0.79	0.41	3,845	0.83	0.38	15,603
Different hospital district	0.05	0.21	3,845	0.03	0.17	15,603
Hospital volume	48.06	54.71	80	31.27	31.01	499
Actual HHI	0.87	0.12	80	0.84	0.13	499
Panel C. All orthopedic surgeries						
Teaching hospital	0.23	0.42	45,359	0.41	0.49	209,319
Distance (km)	22.99	30.04	45,359	27.24	42.07	209,319
Nearest hospital	0.72	0.45	45,359	0.81	0.39	209,319
Different hospital district	0.03	0.17	45,359	0.04	0.20	209,319
Hospital volume	265.26	225.98	171	391.98	476.02	534
Actual HHI	0.63	0.14	171	0.71	0.13	534

Notes: The table reports descriptive statistics for 18–74-year-old patients undergoing an orthopedic surgery in the pre-reform period (Q1/2004–Q3/2007). Distance to the hospital (km) is continuous and the other choice outcomes, including teaching hospital, nearest hospital, and different hospital district, are binary (0/1). Hospital volume (number of patients) and actual HHI are calculated at the hospital-quarter level. The HHI is measured on a 0–1 scale, where a higher value indicates more market concentration.

Table 2: Pre-Reform Descriptive Statistics of Hospital Performance and Patient Characteristic Measures

	Reform area			*Control area*		
	Mean	SD	N	Mean	SD	N
Panel A. Hip replacement surgeries						
Emergency readmission in 30 days	0.07	0.25	3,432	0.08	0.28	12,565
Length of stay (days)	7.66	7.10	3,432	8.21	7.97	12,565
Waiting time (days)	213.56	171.38	3,325	170.35	160.44	9,052
Age	62.19	9.38	3,432	62.40	8.94	12,565
Female	0.50	0.50	3,432	0.53	0.50	12,565
Pre-surgery emergency admissions	0.31	0.95	3,432	0.45	1.28	12,565
Panel B. Knee replacement surgeries						
Emergency readmission in 30 days	0.10	0.30	3,845	0.10	0.30	15,603
Length of stay (days)	7.36	6.19	3,845	7.67	6.39	15,603
Waiting time (days)	264.88	217.33	3,660	212.22	201.82	11,028
Age	64.85	7.21	3,845	64.21	7.49	15,603
Female	0.67	0.47	3,845	0.66	0.47	15,603
Pre-surgery emergency admissions	0.34	0.95	3,845	0.42	1.15	15,603
Panel C. All orthopedic surgeries						
Emergency readmission in 30 days	0.05	0.22	45,359	0.06	0.24	209,319
Length of stay (days)	2.27	5.72	45,359	2.19	6.08	209,319
Waiting time (days)	152.66	169.11	32,280	143.94	175.08	131,815
Age	51.82	13.73	45,359	51.41	13.54	209,319
Female	0.51	0.50	45,359	0.53	0.50	209,319
Pre-surgery emergency admissions	0.47	1.22	45,359	0.61	1.46	209,319

Notes: The table reports descriptive statistics for 18–74-year-old patients undergoing an orthopedic surgery in the pre-reform period (Q1/2004–Q3/2007). Length of stay (days), waiting time (days), age (years) and pre-surgery emergency admissions (count) are continuous and the other variables, including emergency readmission within 30 days and female, are binary (0/1). Waiting time is missing for some patients, which is depicted as a smaller number of observations.

hospitals and small regional hospitals (91–1,281 versus 27–404 and 12–194 patients per quarter on average, respectively; online Appendix Table A.2).[16] Larger volume in teaching hospitals is likely to reflect their better capacity and resources. Although teaching hospitals are generally considered higher-quality and more expensive (Burke et al., 2019; Silber et al., 2020; Newsweek and Statista, 2022), their risk-adjusted readmission rates (clinical quality) and surgical expenditure per patient did not differ much from those of non-teaching hospitals (online Appendix Tables A.3 and A.4). This finding is consistent with the evidence for common medical and surgical conditions in other settings (Burke et al., 2019; Silber et al., 2020). For hip and knee replacements, mean waiting times were, however, shorter in teaching hospitals compared with medium-sized non-teaching hospitals (181–237 versus 202–250 days), whereas the reverse was true in for all orthopedic surgeries (167 versus 157 days).[17]

IV. Baseline Econometric Approach for Estimating Choice Reform Effects

The introduction of a market-based allocation mechanism may affect hospital choice and performance by increasing substitution and non-price competition across hospitals. Thus, we begin by estimating the effects of the regional patient choice reform on related patient outcomes, using our administrative patient-level data and a DiD approach with a binary treatment variable. We use the reform area as a treatment group and the remaining areas of the country as a control group. Specifically, we employ the following baseline specification:

$$y_{imht} = \beta_1 \text{Treated}_i \times \text{Post}_t + \mathbf{X}_{it}' \gamma + \lambda_t + \mu_m + \varepsilon_{imht}, \tag{1}$$

where y_{imht} is the outcome related to hospital choice or performance for patient i living in municipality m and received care at a hospital h in period (quarter) t. Treated_i is a binary indicator for the treatment group equal to one if patient i was located in the reform area.[18] Post_t is a binary post-reform indicator equal to one after the fourth quarter of 2007 (Q4/2007). We include quarter fixed effects λ_t to control for time-varying national-level shocks or trends, such as legislative

[16] Online Appendix Tables A.5 and A.6 further show that teaching hospitals were very similar in terms of their mean performance and patient composition, respectively, between reform and control areas.

[17] Regional hospitals had the shortest mean waiting times in every sample. This might reflect lower demand rather than greater supply in regional hospitals.

[18] The results remain intact if we define the treatment group indicator based on the hospital's location: Treated_h.

changes, shifts in physician referral practices, and evolving clinical guidelines. We also include fixed effects for the patient's municipality of residence, μ_m, to control for persistent differences between municipalities (and also between the reform and control areas), for example, in patients' location, preferences, or morbidity.[19]

$\mathbf{X}_{it}$ includes patient-specific control variables: age, sex, and fixed effects for the patient's surgery type. We use a minimal set of patient characteristics (age and sex) as control variables in our baseline analysis to avoid the bad control problem (Angrist and Pischke, 2009). We also show the robustness of the results to controlling for the patient's pre-existing health status or morbidity (Section V.C). We cluster standard errors at the level of the patient's municipality of residence ($N = 319$) to account for the possibility of within-area correlation in unobservables among patients located in the same geographical area.

The coefficient of interest β_1 identifies the average treatment effect of the choice reform on patient outcome, using variation across regions in the adoption of the patient choice reform and the assignment to separate treatment and control groups. The statistical inference is based on standard p-values but we also compute sharpened q-values, which adjust inference for the probability of false discoveries (Type I errors) when testing multiple hypotheses, following Anderson (2008).[20]

Our first key identifying assumption is that the reform did not have indirect spillover effects to the untreated control area. Such spillover effects could arise if the reform led to the reallocation of patients from the control area to the reform area. In our setting, such reallocation is unlikely, however. Choice opportunities remained unchanged and more restricted in the control area post-reform (Section II.B). Moreover, obtaining care across reform and control area borders was very uncommon (1–2 percent of all observations) throughout our observation period and we excluded these observations from our samples (Section III).

Our second key identifying assumption is that in the absence of the patient choice reform, patient outcomes in the reform and control areas would have followed parallel trends. Online

[19] We include fixed effects for patients' home municipalities rather than for hospitals in the baseline specification. Municipality fixed effects are more relevant for choice outcomes because they reflect patients' location. For example, given their location, patients can decide whether to choose a hospital in another hospital district. Hospital fixed effects are, however, more relevant for hospital performance outcomes because hospitals determine performance for their own patients. Our results for hospital performance outcomes are robust for including hospital fixed effects (Section V.C).

[20] In cases where many hypotheses are rejected, sharpened q-values can be smaller than p-values.

Appendix Figures A.2–A.5 show the trends in outcome means, residualized by only municipality fixed effects, in the reform and control areas. We calculate the (relative) half-year time periods from the adoption of the reform in October 2007, $l \in \{-7,...,5\}$, and normalize the outcome in "one period before reform adoption" ($l = -1$) to zero, similar to the common practice in event study frameworks. We find only few systematic differences in the mean outcome trends in the pre-reform period between the reform and control areas. Some of the reform area outcome means are, however, noisy in the samples of hip and knee replacements, while the noise is much less pronounced in the larger sample of all orthopedic surgeries. Furthermore, the changes in the outcome means in the post-reform period appear to come from the reform, rather than the control area, supporting the credibility of our research design.

To further examine potential pre-existing trends and the dynamic effects of the choice reform graphically, we estimate the following binary treatment event study specification:

$$y_{imht} = \text{Treated}_i \times \sum_{l=-7}^{5} \delta_{1,l} \mathbb{1}[l = t] + \mathbf{X}'_{it}\tau + \lambda_t + \mu_m + \varepsilon_{imht}. \tag{2}$$

The coefficients for the pre-reform period $\delta_{1,l}$, $l < -1$ capture a possible pre-existing trend in the outcome variable, whereas the coefficients $\delta_{1,l}$, $l > -1$ for the post-adoption periods capture the dynamic effect of the choice reform in each of these periods. We use the same set of controls and fixed effects as in specification (1) and again normalize the coefficients for the indicators "one period before reform adoption" to zero, $\delta_{1,-1} = 0$.

V. Baseline Results

V.A. Effects on Hospital Choice and Allocation

We first investigate how the regional patient choice reform affected patients' hospital choices and allocation. We present the results from estimating the baseline binary DiD specification (1) using our patient-level data for hip replacements, knee replacements, and all orthopedic surgeries in Panels A–C of Table 3, respectively.

We find that the reform led to the reallocation of patients towards large teaching hospitals (Column 1 of Table 3). Specifically, the probability of choosing a teaching, instead of a non-teaching, hospital increased by 4–5 percentage points (10–12 percent in comparison to the pre-reform mean) for hip and knee replacements. We also find a slightly smaller and statistically significant increase of 3 percentage points (7 percent) for all orthopedic surgeries.[21]

Figure 2 presents the event study estimates for the probability of choosing a teaching hospital obtained by estimating specification (2). We find that this probability began to increase half a year after the implementation of the reform in every sample. The lag in the effects may result from waiting times for planned surgery and patients or physicians adjusting to the new choice system. The event study estimates reveal little evidence of pre-trends in the choice outcome, providing supporting evidence for the credibility of our research design.[22]

We also find that the reform, which expanded patients' choice of hospitals within and across hospital districts, induced patients to choose more distant hospitals (Column 2 of Table 3, online Appendix Figure A.6).[23] Depending on the sample, the travel distance increased by approximately 2–6 kilometers or 8–20 percent compared to the pre-reform mean distance. The DiD point estimate is statistically significant at the 5 percent level for hip replacements and all orthopedic surgeries, while not statistically significant for knee replacements.

We also studied whether the reform changed the probability to receive care in the nearest hospital. The DiD parameter estimates are statistically significant for knee replacements (Column 3 of Table 3 and online Appendix Figure A.7), indicating a 4 percentage point (5 percent) increase.[24] In contrast, we find that a larger share of patients chose a hospital beyond their own hospital district (Column 4 of Table 3 and online Appendix Figure A.8). The magnitude of the DiD point estimates (Table 3) ranges from 1 to 3 percentage points (23–89 percent at the sample mean), although the coefficient estimate for knee replacements is not statistically significant. In sum, the reform impacted hospital choice and patient reallocation by inducing patients to travel longer distances and across hospital district borders to obtain orthopedic surgery at large teaching hospitals.

[21] All orthopedic surgeries also include a large number of less-invasive surgeries (see Table 2), for which patients might prefer convenience over expected quality and thus might be more willing to choose the nearest (non-teaching) hospital, instead of a more distant teaching hospital. Consistent with this, we find that the point estimate for the choice of teaching hospital is no longer statistically significant and the probability of choosing the nearest hospital increases when we exclude hip and knee replacements from the sample of all orthopedic surgeries (online Appendix Table A.7).

[22] The pre-treatment coefficients are also quite precisely estimated, especially for hip and knee replacements. Thus, it is unlikely that pre-existing trends that produce meaningful bias in the treatment effects estimates would not be undetected with substantial probability (c.f. Roth, 2022).

[23] However, we find that the travel distance decreased for patients living in some hospital district border areas.

[24] The nearest hospital could also be located outside the patient's own hospital district, especially if they live near the district's border area.

Table 3: The Effects of the Reform on Hospital Choice and Allocation Outcomes

	Teaching hospital[a]	Distance (km)[b]	Nearest hospital[c]	Different hospital district[d]
	(1)	(2)	(3)	(4)
Panel A. Hip replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.051***	5.922**	-0.015	0.031**
	(0.017)	(2.845)	(0.019)	(0.013)
p-value	[0.002]	[0.038]	[0.444]	[0.018]
Sharpened q-value	{0.010}	{0.038}	{0.125}	{0.028}
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.429	29.091	0.818	0.035
N	31,404	31,404	31,404	31,404
Panel B. Knee replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.039***	2.387	0.039**	0.016
	(0.014)	(1.946)	(0.019)	(0.011)
p-value	[0.004]	[0.221]	[0.034]	[0.141]
Sharpened q-value	{0.016}	{0.125}	{0.055}	{0.104}
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.393	28.399	0.819	0.034
N	39,427	39,427	39,427	39,427
Panel C. All orthopedic surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.026***	2.413**	-0.006	0.009*
	(0.007)	(1.174)	(0.009)	(0.005)
p-value	[0.000]	[0.041]	[0.456]	[0.092]
Sharpened q-value	{0.001}	{0.066}	{0.140}	{0.089}
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.377	26.481	0.794	0.039
N	473,617	473,617	473,617	473,617
Surgery type FEs	✓	✓	✓	✓
Municipal FEs	✓	✓	✓	✓
Age & sex	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 319$). Sharpened q-values calculated sample-wise.

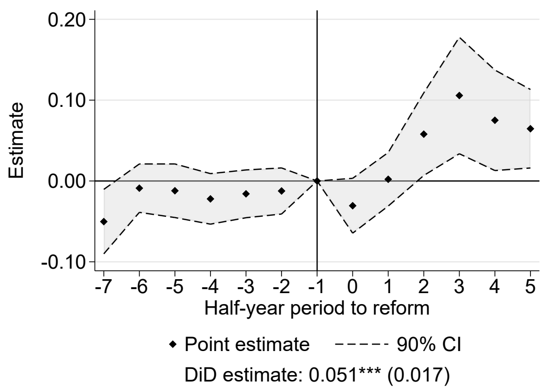
[a] Equals one if the patient received care at a teaching (university) hospital.

[b] Distance from the patient's residence to the hospital in kilometers.

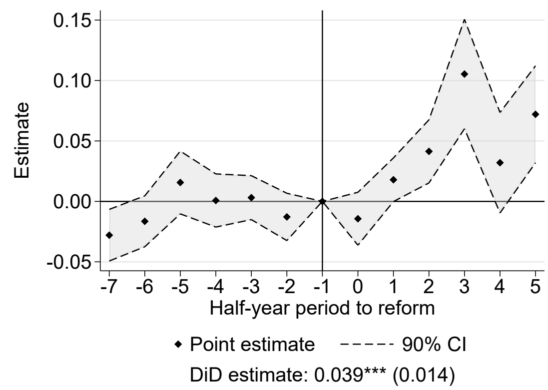
[c] Equals one if the patient received care at the nearest hospital.

[d] Equals one if the patient received care outside their hospital district of residence.

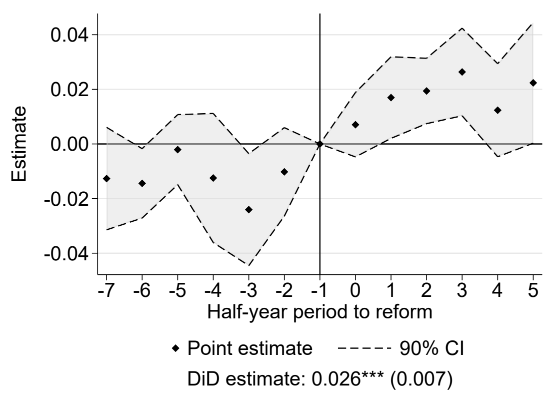
Figure 2: The Effect of the Reform on the Probability of Surgical Patients Receiving Care at a Teaching Hospital



Panel A. Hip replacements



Panel B. Knee replacements



Panel C. All orthopedic surgeries

Notes: Estimated using equation (2). Includes the DiD estimates (in percentage points) shown in Column 1 of Table 3.

V.B. Effects on Hospital Performance

Table 4 displays the results on the effects of the choice reform on hospital performance, as measured by the quality of care, waiting time, and length of stay. Column 1 shows that the reform had little effect on quality of care, as measured by the 30-day emergency readmission probability, for hip replacement (Panel A), knee replacement (Panel B), and all orthopedic surgery patients (Panel C). The point estimates are small in magnitude and not statistically significant. The 95 percent confidence intervals allow us to rule out effects larger than 1–2 percentage points (13–17 percent). The corresponding event study estimates are shown in online Appendix Figure A.9 and they confirm that almost none of the post-reform point estimates are statistically different from zero.

Column 2 in Table 4 and Figure 3 show the estimated effects on waiting times. We find that in every sample, waiting times decreased dramatically post-reform. The waiting time estimates are precise, and similarly to the results for other outcomes and samples, there is only little evidence of pre-trends for hip replacements and all orthopedic surgeries. However, in the sample of knee replacements, waiting times started to decrease more quickly in the reform area than in the control area already in the pre-reform period. Based on the outcome trends (online Appendix Figure A.5), this is caused by a national law change that introduced a maximum waiting time target of 6 months at the beginning of our observation period (Q1/2005). Even though our econometric approach captures national-level shocks such as law changes through time fixed effects, the introduction of the maximum waiting time target may have disproportionately affected long waiting times for knee replacements in the reform area.

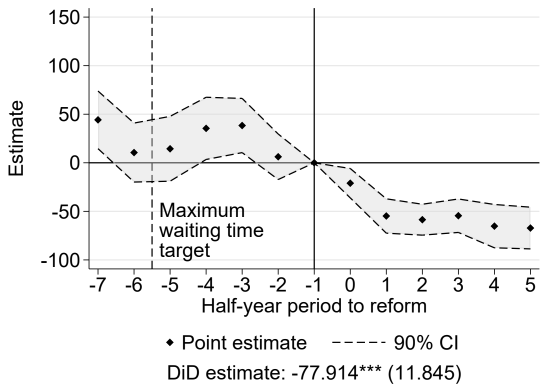
To isolate the effect of the patient choice reform from that of the maximum waiting time target, we use data from periods when waiting times were already adjusted to this law change, Q4/2006–Q4/2010 (Table 5). We find that the post-reform decrease in waiting times was substantial even if the period affected by the waiting time target is excluded: 55–58 days (36–39 percent) for hip and knee replacements, and 27 days (22 percent) for all orthopedic surgeries. Based on the DiD results and outcome trends, the patient choice reform was effective in reducing waiting times below the 6-month maximum for waiting times set by national legislation. Our results suggest that the two policies—maximum waiting time target and choice reform—can act as complementary policy tools in addressing overloaded waiting lists in public hospitals. In Section VI, we show that the

Table 4: The Effects of the Reform on Quality of Care, Waiting Time, and Length of Stay

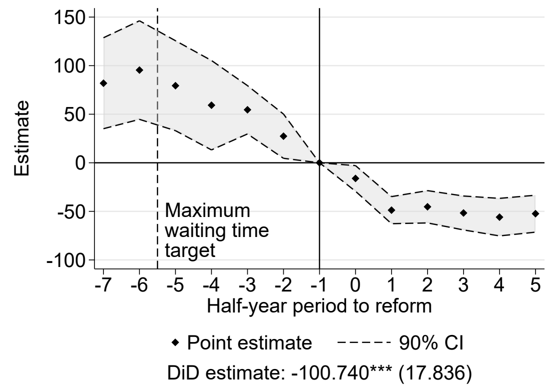
	Readmission[a] (1)	Waiting time[b] (2)	Length of stay[c] (3)
Panel A. Hip replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	-0.000	-77.914***	-0.478
	(0.007)	(11.845)	(0.298)
p-value	[0.995]	[0.000]	[0.110]
Sharpened q-value	{0.497}	{0.001}	{0.124}
$\text{mean}(y_{imht}\|\text{Post}_t = 0)$	0.080	181.991	8.091
N	31,404	24,222	31,404
Panel B. Knee replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.002	-100.740***	-0.501*
	(0.008)	(17.836)	(0.273)
p-value	[0.839]	[0.000]	[0.068]
Sharpened q-value	{0.388}	{0.001}	{0.073}
$\text{mean}(y_{imht}\|\text{Post}_t = 0)$	0.102	225.373	7.609
N	39,427	29,759	39,427
Panel C. All orthopedic surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.002	-20.723***	-0.161***
	(0.003)	(6.709)	(0.050)
p-value	[0.407]	[0.002]	[0.002]
Sharpened q-value	{0.158}	{0.004}	{0.004}
$\text{mean}(y_{imht}\|\text{Post}_t = 0)$	0.060	145.636	2.205
N	473,617	316,454	473,617
Surgery type FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of patient's home municipality ($N = 319$). Sharpened q-values calculated sample-wise.
[a] Equals one if the patient had an emergency readmission (to any hospital) within 30 days after discharge.
[b] Number of days between the specialist's final treatment decision and the surgery. Some of the values are missing, which results in a smaller N compared to the other columns (see online Appendix Section A1.7 for more details).
[c] Number of days between admission and discharge.

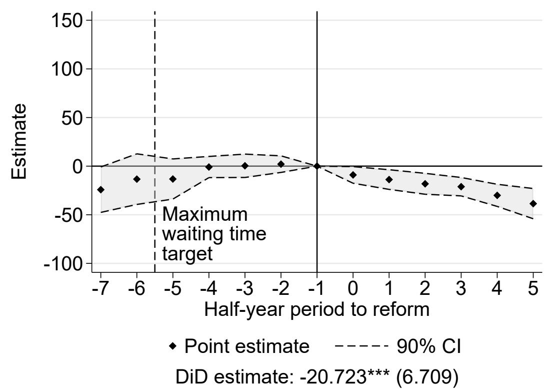
Figure 3: The Effect of the Reform on Waiting Times of Planned Surgery Patients



Panel A. Hip replacements



Panel B. Knee replacements



Panel C. All orthopedic surgeries

Notes: Estimated using equation (2). Includes the DiD estimates (in number of days) shown in Column 2 of Table 4. The dashed line indicates the introduction of the maximum waiting time target that guarantees medical treatment within 6 months.

reduction in waiting times coincides with a large increase in hospital volumes.

Hospitals may reduce the length of stay in order to improve efficiency and to free up resources to care more patients with shorter waiting times. We find that the estimates for length of stay are less precise for hip and knee replacements (Column 3 of Table 4, Figure 4). On the other hand, for all orthopedic surgeries, the length of stay decreased by 4 hours (0.161×24 hours ≈ 4) or 7 percent from the pre-reform mean of 2 days and the DiD estimate is statistically significant at the 1 percent level. The reform might have incentivized hospitals to shorten the length of stay for patients needing relatively less invasive orthopedic surgery, because their health is less likely to deteriorate due to shorter stays compared to patients in need of major surgery such as a hip or knee

Table 5: The Effect of the Choice Reform on Waiting Time Estimates After Adjusting to the Introduction of the Maximum Waiting Time Target

	Hip replacements		Knee replacements		All orthopedic surgeries	
	Whole time period (1)	After target (2)	Whole time period (3)	After target (4)	Whole time period (5)	After target (6)
$\text{Treated}_i \times \text{Post}_t$	-77.914***	-55.161***	-100.740***	-58.376***	-20.723***	-26.522***
	(11.845)	(7.479)	(17.836)	(9.699)	(6.709)	(5.212)
p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]	[0.000]
Sharpened q-value	{0.001}	{0.001}	{0.001}	{0.001}	{0.002}	{0.001}
mean(y_{imht}|$\text{Post}_t = 0$)	181.991	140.993	225.373	163.066	145.636	118.062
N	24,222	15,266	29,759	19,243	316,454	199,492
Surgery type FEs	✓	✓	✓	✓	✓	✓
Municipal FEs	✓	✓	✓	✓	✓	✓
Age & sex	✓	✓	✓	✓	✓	✓

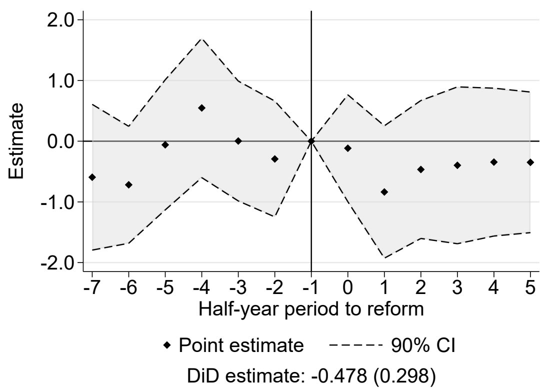
Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Outcome in all columns is waiting time in days. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 319$). Columns 1, 3, and 5 re-display baseline estimates from Column 2 of Table 4. Columns 2, 4, and 6 display estimates using data from Q4/2006–Q4/2010 only. Sharpened q-values calculated sample-wise (i.e., jointly for (1)–(2), (3)–(4), as well as (5)–(6)).

replacement. To the extent that clinical quality did not change, as indicated by our earlier results on emergency readmissions, shorter stays indicate improved efficiency post-reform.

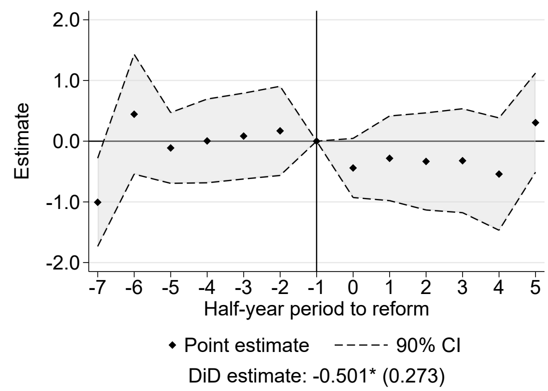
Additional quality outcomes and spillover effects to emergency care. Emergency readmissions may not, however, be sensitive enough to capture changes in all clinical quality attributes. To address this concern, we also estimated the effects of the reform on more detailed measures for hip and knee replacements: indicators for revision surgery, mechanical complication in the prosthesis, and infection or inflammation in the prosthesis (Urquhart et al., 2010; Mäkelä et al., 2011; Bayliss et al., 2017; Fleischman et al., 2019). The results in online Appendix Table A.8 show that the probability of mechanical complication decreased by 0.9 percentage points (22 percent) for knee replacement surgeries, indicating an improvement in quality post-reform in this sample. In contrast, the point estimates for the other quality outcomes and samples are not statistically significant, similar to the findings for emergency readmissions.

The choice reform for planned non-emergency surgery might still have spillover effects on the quality of *emergency* care that was not targeted by the reform. To test for this possibility, we follow the prior work (Kessler and McClellan, 2000; Cooper et al., 2011; Gaynor, Moreno-Serra and

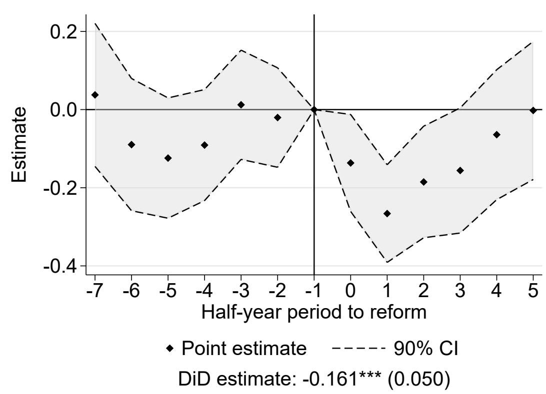
Figure 4: The Effect of the Reform on Length of Stay of Planned Surgery Patients



Panel A. Hip replacements



Panel B. Knee replacements



Panel C. All orthopedic surgeries

Notes: Estimated using equation (2). Includes the DiD estimates (in number of days) shown in Column 3 of Table 4.

Propper, 2013) and estimate the effects of the reform on emergency care clinical quality and length of stay for acute myocardial infarction (AMI), stroke, and hip fracture patients.[25] The DiD estimates in online Appendix Table A.9 show no other statistically significant effects than a decrease in stroke patients' emergency readmissions (4 percentage points or 21 percent). These findings complement our baseline results that show little statistically significant effect on the quality of care for orthopedic surgery.

V.C. Robustness Checks

We conduct a number of robustness checks for the baseline results (online Appendix Section A3). We begin by testing whether our results are sensitive to the inclusion of additional controls in the baseline specifications. For choice-related outcomes, we add controls for staffing and morbidity (indicator for weekend admission, and the number of pre-surgery emergency admissions). For the hospital performance outcomes, we also add hospital fixed effects to control for time-invariant hospital-level factors such as their average patient mix. The results do not change much (Tables A.10 and A.11).

We also study the robustness of our results regarding our baseline samples. First, although there was no hospital entry or exit post-reform, one small hospital in the reform area closed down during the pre-reform period (year 2006). To mitigate the possible bias in the estimates due to the closure and the resulting reallocation of patients to other hospitals, we excluded patients treated in the closed-down hospital or living in municipalities near it from our baseline estimations (see Figure A.10). The baseline results are, however, robust to the inclusion of these patients to the estimation sample (Tables A.12 and A.13). Second, in a few hospital districts, patients had the opportunity to obtain referrals to hospitals outside of their own hospital district under specific circumstances, such as long travel distance. Our results are robust to the exclusion of these hospital districts from the econometric analyses, and, in some cases, this exclusion improves the precision and increases the magnitude of our estimates compared to the baseline estimates (Tables A.14–A.17). Third, some hospitals had a joint hospital identifier in the data and cannot be distinguished from each other.

[25] In addition to revealing possible spillover effects, the benefit of analyzing the outcomes of emergency patients is that it mitigates patient selection into hospitals based on quality and/or morbidity (emergency patients usually receive care in the nearest hospital) (Kessler and McClellan, 2000; Moscelli et al., 2018).

This creates some measurement error in our outcomes related to travel distances and the indicator of choosing the nearest hospital, which are calculated based on the location of the largest hospital under the joint identifier. When we exclude hospitals with joint identifiers from the sample, we generally find larger effects compared to the baseline estimates (Tables A.18 and A.19).

To address potential bias from missing values in waiting times, we re-estimated the results regarding waiting times (i) without hospital districts in which more than 30 percent of waiting time values were missing, (ii) without surgeries for which more than 30 percent of waiting time values were missing, (iii) without one reform area hospital which did not report its waiting times in Q1/2008–Q4/2009, and (iv) using data for the years 2006 and 2010 only, when the share of missing values was generally low. The re-estimated results are similar or even larger than our baseline results (Table A.20).

VI. Hospital-Level Analyses and Heterogeneity

VI.A. Hospital Volume, Concentration, and Mean Performance

Next, we study the effects of introducing choice and non-price competition from the hospital's perspective, rather than the patient's. We estimate the effects of the reform on hospital volumes and market concentration. In addition, we study whether heterogeneous hospitals (teaching versus non-teaching) responded differently to the reform. We estimate the following specification using hospital-quarter-level data:

$$\begin{aligned} y_{ht} = &\beta_1 \text{Treated}_h \times \text{Post}_t + \beta_2 \text{Post}_t \times \text{Teaching}_h \\ &+ \beta_3 \text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h + \mathbf{X}'_{ht}\gamma + \lambda_t + \eta_h + \varepsilon_{ht}, \end{aligned} \tag{3}$$

where y_{ht} is the outcome for hospital h and period (quarter) t, Treated_h is the indicator of hospital h being located in the treatment group, and Teaching_h is the teaching hospital indicator. η_h are hospital fixed effects, which absorb time-invariant hospital-level factors such as Treated_h, Teaching_h, their interaction, hospital location, and average patient mix. The treatment and post-reform indicators (Treated_h and Post_t) are the same as in the patient-level specification (equation (1)). Covariates $\mathbf{X}_{ht}$ are also the same, but transformed to the hospital-level means in quarter t. The coefficient β_1 shows the effect of the reform in non-teaching hospitals and β_3 shows the differential effect in teaching hospitals compared to non-teaching hospitals. For the specification estimating the average hospital-level effects (without differential effect), we set $\beta_2, \beta_3 = 0$.[26] We cluster the standard errors at the hospital district level ($N = 20$).

Table 6 shows the effects of the patient choice reform on hospital volumes and market concentrations, as measured by the number of patients and actual hospital-level HHI. The results for the average effects are shown in Columns 1 and 3, while the differential effects are shown in Columns 2 and 4. We find that hospital volume increased by 8–38 percent on average after the reform (Column 1). The increase was statistically significant for knee replacements and all orthopedic surgeries and not statistically significant for hip replacements. In every sample, hospital volumes increased disproportionally in teaching hospitals (Column 2). For hip and knee replacements, this increase was 19–29 patients per quarter.[27] In relative terms, it corresponds to an increase of 65–87 percent compared with the mean volume of all hospitals (29–34 patients per quarter) and 21–29 percent compared with the mean volume of teaching hospitals (91–102 patients per quarter, see online Appendix Table A.2).

The DiD point estimates suggest a decrease in market concentration, as measured by the actual HHI, on average; the 95 percent confidence intervals allow us to rule out effects larger than 7–9 percent (Column 3). The average treatment effects, however, masks considerable heterogeneity in the effects across different types of hospitals and their markets (Column 4). The point estimates for non-teaching hospitals show that concentration decreased by 2–6 percent in their markets. The estimates are statistically significant for hip replacements and all orthopedic surgeries and not statistically significant for knee replacements. In contrast, reflecting the disproportionate increase in teaching hospitals' volumes, concentration in their markets increased by 1–9 percent after the choice reform and the estimates are statistically significant.

Table 7 shows the effects of the reform on hospital-level means of performance outcomes in teaching and non-teaching hospitals. The estimates for emergency readmission rates were generally small and statistically insignificant (Columns 1 and 2). In line with the increased hospital volumes and supply, the reform reduced mean waiting times in hospitals exposed to the reform on average (Column 3). The point estimates are negative for both teaching and non-teaching hospitals (Column 4), although they become smaller in magnitude when we exclude time periods affected by

[26] More specifically, we estimate the following specification: $y_{ht} = \beta_1 \text{Treated}_h \times \text{Post}_t + \mathbf{X}'_{ht}\gamma + \lambda_t + \eta_h + \varepsilon_{ht}$.

[27] The sum of the coefficients on $\text{Treated}_h \times \text{Post}_t$ and $\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$: $\beta_1 + \beta_3$.

Table 6: The Effects of the Reform on Hospital Volumes and Market Structure

	Hospital volume[a]		Actual HHI[b]	
	DiD (1)	Heterogeneity (2)	DiD (3)	Heterogeneity (4)
Panel A. Hip replacement surgeries				
$\text{Treated}_h \times \text{Post}_t$	5.707**	2.524*	-0.031	-0.049**
	(2.387)	(1.354)	(0.022)	(0.022)
	[0.027]	[0.078]	[0.174]	[0.041]
	{0.038}	{0.058}	{0.066}	{0.043}
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		16.551***		0.122***
		(3.832)		(0.039)
		[0.000]		[0.006]
		{0.003}		{0.015}
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	29.192	29.192	0.834	0.834
N	1,009	1,009	1,009	1,009
Panel B. Knee replacement surgeries				
$\text{Treated}_h \times \text{Post}_t$	12.755***	8.975***	-0.012	-0.021
	(3.884)	(1.601)	(0.024)	(0.027)
	[0.004]	[0.000]	[0.614]	[0.453]
	{0.006}	{0.001}	{0.258}	{0.222}
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		20.407***		0.084*
		(4.508)		(0.047)
		[0.000]		[0.092]
		{0.001}		{0.075}
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	33.589	33.589	0.842	0.842
N	1,070	1,070	1,070	1,070
Panel C. All orthopedic surgeries				
$\text{Treated}_h \times \text{Post}_t$	27.128***	16.467	-0.036**	-0.042**
	(7.826)	(11.037)	(0.015)	(0.015)
	[0.003]	[0.152]	[0.023]	[0.014]
	{0.016}	{0.038}	{0.024}	{0.020}
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		115.152***		0.051**
		(36.463)		(0.022)
		[0.005]		[0.030]
		{0.016}		{0.025}
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	361.245	361.245	0.693	0.693
N	1,295	1,295	1,295	1,295
Hospital and time FEs	✓	✓	✓	✓
Age & sex mix[c]	✓	✓	✓	✓
Surgery types[d]	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Includes hospital-quarter-level observations from Q1/2004 to Q4/2010. Standard errors clustered at the hospital district level ($N = 20$). Standard errors in brackets, standard p-values in square bracket, and sharpened q-values in curly brackets. Sharpened q-values calculated sample-wise, i.e. panel-wise.

[a] Number of patients.

[b] Observed market concentration measured on a 0–1 scale. See online Appendix Section A1.10 for more information.

[c] Shares of females, 18–29, 30–39, 40–49, 50–59, and 70–74-year-old patients of hospital's total patient volume. Baseline = share of 60–69-year-old male patients.

[d] Shares of different procedure codes among hospital's patients.

the implementation of the maximum waiting time target (online Appendix Table A.21). While most of these hospital-type specific estimates for mean waiting times are imprecise, they suggest that all hospital types improved their performance in response to increased choice and competition. Our statistically insignificant point estimates further suggest that the reform reduced the mean length of stays for all orthopedic surgeries in both teaching and non-teaching hospitals (Columns 5 and 6)—without major changes in clinical quality (Columns 1 and 2). [28]

VI.B. Case Mix and Resource Use

Because the reform changed hospitals' volumes, it could also have affected their patient composition and case mix. For example, if teaching hospitals provided care to sicker patients post-reform, it could explain the null effects on their patients' emergency readmissions (quality of care). To assess these changes, we apply the estimation strategy from equation (3) to hospital-level averages of patient demographics (age, sex) and morbidity (pre-surgery emergency admissions). The results in online Appendix Table A.22 show minimal change in patient composition across hospital types post-reform, except for hip replacements, where teaching and non-teaching hospitals treated sicker patients.

We also examine whether the reform affected hospitals' annual surgical expenditure and a coarse measure of productivity: expenditure per patient. Using a specification similar to equation (3), with hospital-year data and the case mix index in $\mathbf{X}_{ht}$, we find a statistically insignificant and very small decrease in surgical expenditure for all hospitals on average (online Appendix Table A.23).[29] Although, our data on resource use is coarse and limited, this finding suggests no major overall increase in hospitals' resource use due to the reform.[30] Surgical expenditure increased over 40 percent at teaching hospitals, likely due to patient reallocation and a disproportionate increase in volumes. There is no economically or statistically significant change on expenditure per patient, suggesting no additional resources were used for improved patient outcomes.

[28]For hip and knee replacements, the mean length of stay decreased by 12–16 percent in non-teaching hospitals, while the estimated effects in teaching hospitals were smaller and less conclusive.

[29]The case mix index is also part of the statistics on the costs and productivity of hospitals from the Finnish Institute for Health and Welfare.

[30]Using information in hospital districts' annual reports, we also confirmed that teaching hospitals did not operate at full capacity or increase capacity in terms of the number of beds post-reform.

Table 7: The Effects of the Reform on Hospital-level Means of Quality of Care, Waiting Times and Length of Stay

	Emergency readmission[a]		Waiting time[b]		Length of stay[c]	
	DiD	Hetero-geneity	DiD	Hetero-geneity	DiD	Hetero-geneity
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Hip replacement surgeries						
$\text{Treated}_h \times \text{Post}_t$	0.002	0.003	-83.428**	-68.589	-0.997*	-1.437***
	(0.009)	(0.011)	(36.518)	(46.117)	(0.502)	(0.289)
	[0.797]	[0.773]	[0.034]	[0.153]	[0.062]	[0.000]
	{0.549}	{0.549}	{0.087}	{0.182}	{0.114}	{0.001}
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		-0.008		-47.303		1.623***
		(0.021)		(46.136)		(0.472)
		[0.690]		[0.318]		[0.003]
		{0.549}		{0.361}		{0.012}
$\text{mean}(y_{ht}\|\text{Post}_t = 0)$	0.083	0.083	172.949	172.949	9.017	9.017
N	1,009	1,009	791	791	1,009	1,009
Panel B. Knee replacement surgeries						
$\text{Treated}_h \times \text{Post}_t$	0.007	0.007	-113.316**	-101.002*	-0.761*	-0.959**
	(0.014)	(0.017)	(42.295)	(57.356)	(0.413)	(0.425)
	[0.616]	[0.666]	[0.015]	[0.094]	[0.081]	[0.036]
	{0.500}	{0.500}	{0.155}	{0.198}	{0.198}	{0.169}
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		-0.005		-38.022		0.766
		(0.017)		(59.088)		(0.477)
		[0.765]		[0.528]		[0.124]
		{0.516}		{0.500}		{0.211}
$\text{mean}(y_{ht}\|\text{Post}_t = 0)$	0.109	0.109	208.711	208.711	8.313	8.313
N	1,070	1,070	812	812	1,070	1,070
Panel C. All orthopedic surgeries						
$\text{Treated}_h \times \text{Post}_t$	0.001	0.000	-14.984*	-10.619	-0.036	-0.017
	(0.004)	(0.005)	(8.308)	(11.679)	(0.095)	(0.111)
	[0.841]	[0.927]	[0.087]	[0.375]	[0.712]	[0.879]
	{1.000}	{1.000}	{0.536}	{0.890}	{1.000}	{1.000}
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		0.005		-33.429**		-0.193
		(0.004)		(12.385)		(0.131)
		[0.284]		[0.014]		[0.157]
		{0.890}		{0.147}		{0.646}
$\text{mean}(y_{ht}\|\text{Post}_t = 0)$	0.057	0.057	125.849	125.849	1.911	1.911
N	1,295	1,295	1,114	1,114	1,295	1,295
Hospital and time FEs	✓	✓	✓	✓	✓	✓
Age & sex mix[d]	✓	✓	✓	✓	✓	✓
Surgery types[e]	✓	✓	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Includes hospital-quarter-level observations from Q1/2004 to Q4/2010. Standard errors clustered at the hospital district level ($N = 20$). Standard errors in brackets, standard p-values in square bracket, and sharpened q-values in curly brackets. Sharpened q-values calculated sample-wise, i.e. panel-wise.

[a] 30-day emergency readmission rate (0–1).

[b] Mean waiting time in days. Some hospitals did not report waiting times in some quarters, which results in a smaller N compared to the other columns (see online Appendix Section A1.7 for more details). See also online Appendix Table A.21 for the waiting time estimates unaffected by the introduction of the maximum waiting time target.

[c] Mean length of stay in days.

[d] Shares of females, shares of 18–29, 30–39, 40–49, 50–59, and 70–74-year-old patients of hospital's total patient volume. Baseline = share of 60–69-year-old male patients.

[e] Shares of different procedure codes among hospital's patients.

VII. Back-of-the-Envelope Calculation

We perform a simple back-of-the-envelope calculation to evaluate the aggregate social benefits and costs of the introduction of choice and non-price competition in the entire reform area. We outline the calculation and results for orthopedic surgery in this section and present further details in online Appendix Table A.24.

To calculate the aggregate benefits, we use the estimates of the health-related quality of life gained from orthopedic surgery (Jansson and Granath, 2011) and express their value in monetary terms (in 2010 euros) (Gyrd-Hansen, 2003). Based on these estimates, the estimated health gain per patient and year, denoted by g, would be $2,547$ euros.[31] Moreover, we assume that the number of years that each patient can benefit from these health gains is the discounted life expectancy, denoted by T, minus the waiting time (w) in *years*. For the waiting time, we use our results of $w_0 = 0.32$ years (118 days) in the pre-reform period and $w_1 = 0.25$ years ($118 - 27 = 91$ days) in the post-reform period on average (Table 5).[32] The discounted life expectancy at the average age of orthopedic patients (52 years) is $T = 19.8$ years assuming a discount factor of 3 percent and the expected age of death at 80 years (WHO, 2014).

The benefits of the reform arise from a shorter waiting time (a decrease of 27 days) and increased hospital volume. Specifically, we found that the reform increased hospital volume by 27 patients per hospital and quarter from the pre-reform average of 361 patients (Table 6), and there were 15 hospitals in the reform area. The approximated total volume per year would then be $y_0 = 361 \times 15 \times 4 \approx 21,680$ patients in the pre-reform period and $y_1 = (361+27) \times 15 \times 4 \approx 23,300$ patients in the post-reform period. A higher total volume increases the number of patients receiving orthopedic surgery with health gains (an increase of $y_1 - y_0 = 1,628$ patients or 8 percent annually). Shorter waiting times also enable patients to benefit from health gains for a longer period due to earlier treatment.

[31] For comparison, g would be $4,386$ and $3,113$ euros for hip and knee replacements, respectively (online Appendix Table A.24).

[32] Waiting time can also affect patient health g. For example, based on estimates by Nikolova, Harrison and Sutton (2016) and Gyrd-Hansen (2003), quality of life is reduced by approximately 0.009 percentage points ($14,149 \times 0.009/100 \approx 1.27$ euros) from an additional *day* of waiting for hip and knee replacements. Then, g would be reduced by $w_t/365.25 \times 1.27$ euros, depending on the waiting time in *years*, w_t (e.g., for hip replacements, $140.993 \times 1.27 \approx 179$ euros in the pre-reform period and $(140.993 - 55.161) \times 1.27 \approx 109$ euros in the post-reform period).

In sum, the aggregate benefits of the reform are driven by the increase in volume, shorter waiting times, and the health gains from orthopedic surgery: $g \times (T - w_1) \times y_1 - g \times (T - w_0) \times y_0$. Using the calculations above, we find that these benefits are substantial, approximately 85 million euros per year in the reform area, even with a conservative assumption that patient health g is not improved by shorter waiting after the reform.[33]

In terms of hospital costs, we found that the reform had little impact on hospitals' annual surgical expenditure on average (Section VI.B). However, here we use a conservative approach and assume that the aggregate costs increase with patient volume: $c \times (y_1 - y_0) > 0$, where c is the cost per patient.

The reform's annual net social benefit is approximately 74 million euros using a constant cost estimate of $c = 6,645$ euros for orthopedic surgery (Remes et al., 2007).[34] Additionally, the reform reduced the length of stay by 7 percent (4 hours) for orthopedic surgery (Table 4), which can lead to cost savings: : $c_1 < c_0$, where c_0 is the pre-reform cost and c_1 is the post-reform cost per patient. If we take this channel into account, we get cost estimates of $c_0 = 6,645$ and $c_1 = 6,398$ per patient. Then, the annual net social benefit of the reform is even larger, approximately 80 million euros ($3,400$ euros per patient).

VIII. Differential Changes by Market Structure

Next, we highlight the differences between our binary DiD approach and a DiD approach using variation in the effects of choice reform by pre-reform market structure, as used in the previous research on patient choice (Cooper et al., 2011; Gaynor, Moreno-Serra and Propper, 2013; Moscelli et al., 2018; Roos et al., 2020; Brekke et al., 2021; Moscelli, Gravelle and Siciliani, 2021, 2023). Our binary DiD approach, based on the regional reform, estimates the average treatment effect of introducing patient choice in comparison to the untreated control area with restricted patient choice. The binary DiD approach differs from the DiD approach used by the previous literature that estimates the *differential changes* in hospital performance by market structure after nationwide choice reforms. As such choice reforms have been introduced simultaneously at the national level

[33] $[g \times (T - w_1) \times y_1] - [g \times (T - w_0) \times y_0] = [2,547 \times (19.8 - 0.32) \times 21,675] - [2,547 \times (19.8 - 0.25) \times 23,302] \approx$ 85 million euros.

[34] 85 million euros $- (y_1 - y_0) \times c =$ 85 million euros $- (23,302 - 21,675) \times$ €6,645 $\approx$ 74 million euros.

(e.g., in the English National Health Service, the Netherlands, and Norway), all areas belong to the treatment group and there was no untreated control area.

In line with the previous literature, we use patient-level data from the treatment group only ($\text{Treated}_i = 1$) and estimate a DiD specification using variation across hospitals exposed to different market structures, rather than assignment to separate treatment and control groups as done in the binary DiD approach. We estimate the following specification, where we interact the post-reform indicator with a measure of market structure:

$$y_{imht} = \alpha_0 \widehat{\text{HHI}}_h + \alpha_1 \text{Post}_t + \alpha_2 \widehat{\text{HHI}}_h \times \text{Post}_t + \mathbf{X}'_{it}\nu + \mu_m + \hat{r}'_{it}\theta + \varepsilon_{imht}, \tag{4}$$

where $\widehat{\text{HHI}}_h$ is the market concentration faced by hospital h. To mitigate endogeneity in the market structure, we follow the standard practice (Kessler and McClellan, 2000; Gaynor, Moreno-Serra and Propper, 2013; Moscelli, Gravelle and Siciliani, 2021) and use the *predicted* rather than actual HHI from the pre-reform period. We control for the first-stage stage residuals ($\hat{r}_{it}$) from a hospital choice model to ensure that time-varying patient allocation to hospitals based on unobserved morbidity does not bias the estimates (Moscelli, Gravelle and Siciliani, 2021).[35] We bootstrap the clustered standard errors ($N = 67$ municipalities in the reform area) because $\widehat{\text{HHI}}_h$ is based on predicted choices. We base our inference on p-values and confidence intervals or their bounds because the bootstrap algorithm, a wild cluster bootstrap, does not produce standard errors. See online Appendix Section A2 for further details on the econometric approach.

The coefficient of interest, α_2, captures the differential change in hospital performance between more versus less concentrated markets after the choice reform (Cooper et al., 2011; Gaynor, Moreno-Serra and Propper, 2013). Note that if hospitals responded similarly to the reform in more versus less concentrated markets, then $\alpha_2 = 0$, even if the reform itself had a large effect on the outcome, which instead we are able to identify using our binary DiD approach in equation (1).

We present the results for the differential effects of the reform by market structure in Columns 1, 3 and 5 of online Appendix Table A.25. The coefficient estimates for the measure of clinical quality, the readmission probability, are not statistically significant. The coefficient estimates for

[35] The binary and continuous DiD results for hospital performance are not sensitive to the inclusion or exclusion of the control function residuals (online Appendix Tables A.26–A.27). Thus, changes in patient allocation to hospitals based on morbidity do not seem to drive our results.

waiting time are not statistically significant but strongly positive, suggesting that waiting times decreased more in *less* concentrated markets after the reform. For example, for hip replacements, a continuous DiD estimate of 68 indicates that a one standard deviation (0.13) decrease in the predicted HHI is associated with a $68 \times 0.13 \approx 9$ days decrease in waiting time post-reform. Moreover, the negative coefficient estimates for length of stay suggest a larger decrease in *more* concentrated markets after the choice reform.

Then, we expand our binary DiD approach to allow for the reform's effect to vary between more and less concentrated markets to illustrate how the DiD approach used by the previous hospital choice literature relates to our binary DiD approach. We use the following specification:

$$
\begin{aligned}
y_{ht} =& \kappa_0 \widehat{\text{HHI}}_h + \kappa_1 \text{Post}_t + \kappa_2 \text{Treated}_i \times \text{Post}_t + \kappa_3 \text{Treated}_i \times \widehat{\text{HHI}}_h \\
&+ \kappa_4 \widehat{\text{HHI}}_h \times \text{Post}_t + \kappa_5 \text{Treated}_i \times \widehat{\text{HHI}}_h \times \text{Post}_t \\
&+ \mathbf{X}'_{it} \boldsymbol{\nu_1} + \hat{r}'_{it} \theta_1 + \text{Treated}_i \times (\mathbf{X}'_{it} \boldsymbol{\nu_2} + \hat{r}'_{it} \theta_2) + \mu_m + \varepsilon_{imht}.
\end{aligned}
\tag{5}
$$

The results are presented in Columns 2, 4 and 6 of online Appendix Table A.25. The sum of the coefficients κ_4 and κ_5 equals to the coefficient of interest, α_2, in equation (4), which is shown by our results in Table A.25. However, α_2 does not capture the average treatment effect of the choice reform, which is instead captured by $\kappa_2 + (\kappa_5 \times \widehat{\text{HHI}}_h)$, where $\widehat{\text{HHI}}_h$ is evaluated at the sample mean. Taking waiting times as an example, the choice reform reduced on average waiting times for hip replacements by 84.3 days[36] (in line with Table 4 and Figure 3). The reduction was more pronounced in less concentrated markets (an additional 7.5 days[37] reduction for one standard deviation decrease in the predicted HHI). Similarly, for knee replacements, the reform reduced waiting times by 75.8 days[38], and the reduction was more pronounced in less concentrated markets (an additional 16.8 days[39] reduction for one standard deviation reduction in the predicted HHI).

In sum, the DiD approach used in the previous literature on patient choice estimates a different parameter of interest than the average effect of the choice reform. The approach can be interpreted as a heterogeneity analysis between more and less concentrated markets for the average reform

[36] $\kappa_3 + \kappa_5 \times \text{mean}(\widehat{\text{HHI}}_h) = -139 + 84 \times 0.653 \approx -84.3$
[37] $(\kappa_4 + \kappa_5) \times \text{SD}(\widehat{\text{HHI}}_h) = (-32 + 84) \times 0.146 \approx 7.5$
[38] $\kappa_3 + \kappa_5 \times \text{mean}(\widehat{\text{HHI}}_h) = -216 + 217 \times 0.644 \approx -75.8$
[39] $(\kappa_4 + \kappa_5) \times \text{SD}(\widehat{\text{HHI}}_h) = (-103 + 217) \times 0.147 \approx 16.8$

effect, which we identify using the unaffected control group and the binary DiD approach.

IX. Conclusions

We studied the effects of introducing a market-based allocation mechanism compared to an allocation mechanism with more restricted patient choice. We used a DiD approach based on a unique quasi-experiment: a regional patient choice reform in Finland that introduced choice and non-price competition among public hospitals. Using nationwide administrative data on orthopedic surgeries, we found that patients responded to the reform by traveling longer distances and across hospital districts to receive surgery at large teaching hospitals. The effects on patient behavior were considerable despite the country's low population and hospital densities, and despite the limited financial incentives for patients related to their hospital choice due to universal public insurance.

We found that the introduction of choice can promote market concentration by reallocating patients towards large teaching hospitals, potentially due to their better resources and higher perceived quality. However, this does not necessarily imply that the choice reform has increased clinical quality across providers.[40] We indeed show that the choice reform had little effect on clinical quality.

Our back-of-the-envelope calculation further suggests that in aggregate, the social benefits of the choice reform outweigh the costs by approximately 80 million euros a year ($3,400$ euros per patient) for orthopedic surgery. The reform had direct benefits for patients, as the total hospital volume increased by $1,628$ patients (8 percent) annually and waiting times (i.e., non-monetary costs) became much shorter.[41]

Our results suggest that the introduction of choice and non-price competition can improve hospital performance in the presence of long waiting times, which exist in many health care systems globally (OECD, 2020). Our results are directly relevant to health care systems relying on public production (including those of the United Kingdom, Sweden, and Norway, for example), but also

[40] This aligns with the theoretical literature showing that in the presence of excess demand, which translates into long waiting times, the effect of patient choice, measured by a higher demand responsiveness to quality, is diminished (Chalkley and Malcomson, 1998). This occurs because a marginal increase in quality would further increase the excess demand without affecting equilibrium volume.

[41] Hospitals could compete to avoid patients rather than attract them. On the other hand, if hospitals are altruistic, a marginal reduction in waiting time can increase volume. The introduction of patient choice reduces waiting times if the latter effect dominates (Brekke, Siciliani and Straume, 2008). Our results are consistent with this explanation.

to other systems with administratively set reimbursements to producers (such as the Medicare system in the United States). More generally, our results have relevance in public services, such as education, where market-based mechanisms have been adopted via the introduction of choice and non-price competition.

Our results based on the patient choice reform are informative about the role of market forces in public health care. The results tie greater competition and choice to higher concentration. The more competitive a market is, the more concentrated is the market towards large hospitals, potentially improving market allocation and performance. However, none of the smaller hospitals were closed post-reform. Instead, after the reform, patient volumes in smaller hospitals remain the same but waiting lists became shorter, with no entry or exit post-reform.

While our results indicate that greater choice can improve market allocation and performance, further research is needed to disentangle the demand and supply side responses to the enhanced patient choice. We also leave for future research to study the implications of choice reforms for redistribution as well as inequality in health care delivery and patient outcomes.

References

Acemoglu, Daron, David H. Autor, and David Lyle. 2004. "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury." *Journal of Political Economy*, 112(3): 497–551.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103(484): 1481–1495.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017. "Concentrating on the Fall of the Labor Share." *American Economic Review*, 107(5): 180–85.

Bayliss, Lee E., David Culliford, A. Paul Monk, Sion Glyn-Jones, Daniel Prieto-Alhambra, Andrew Judge, Cyrus Cooper, Andrew J. Carr, Nigel K. Arden, David J. Beard, and Andrew J. Price. 2017. "The Effect of Patient Age at Intervention on Risk of Implant Revision

After Total Replacement of the Hip or Knee: a Population-Based Cohort study." *The Lancet*, 389(10077): 1424–1430.

Berry, Steven, Martin Gaynor, and Fiona Scott Morton. 2019. "Do Increasing Markups Matter? Lessons from Empirical Industrial Organization." *Journal of Economic Perspectives*, 33(3): 44–68.

Bighelli, Tommaso, Filippo di Mauro, Marc J Melitz, and Matthias Mertens. 2022. "European Firm Concentration and Aggregate Productivity." *Journal of the European Economic Association*, 21(2): 455–483.

Brekke, Kurt R., Chiara Canta, Luigi Siciliani, and Odd Rune Straume. 2021. "Hospital Competition in a National Health Service: Evidence from a Patient Choice Reform." *Journal of Health Economics*, 79: 102509.

Brekke, Kurt R., Luigi Siciliani, and Odd Rune Straume. 2008. "Competition and Waiting Times in Hospital Markets." *Journal of Public Economics*, 92(7): 1607–1628.

Brekke, Kurt R., Luigi Siciliani, and Odd Rune Straume. 2011. "Hospital Competition and Quality with Regulated Prices." *The Scandinavian Journal of Economics*, 113(2): 444–469.

Brown, Zach Y., Christopher Hansman, Jordan Keener, and Andre F. Veiga. 2023. "Information and Disparities in Health Care Quality: Evidence from GP Choice in England." National Bureau of Economic Research Working Paper 31033.

Burke, Laura G., Dhruv Khullar, Jie Zheng, Austin B. Frakt, E. John Orav, and Ashish K. Jha. 2019. "Comparison of Costs of Care for Medicare Patients Hospitalized in Teaching and Nonteaching Hospitals." *JAMA Network Open*, 2(6): e195229–e195229.

Campos, Christopher, and Caitlin Kearns. 2023. "The Impact of Public School Choice: Evidence from Los Angeles's Zones of Choice." *The Quarterly Journal of Economics*, 139(2): 1051–1093.

Chalkley, Martin, and James M. Malcomson. 1998. "Contracting for Health Services with Unmonitored Quality." *The Economic Journal*, 108(449): 1093–1110.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson. 2016. "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector." *American Economic Review*, 106(8): 2110–44.

Cooper, Jack, Stuart V. Craig, Martin Gaynor, and John Van Reenen. 2019. "The Price Ain't

Right? Hospital Prices and Health Spending on the Privately Insured*." *The Quarterly Journal of Economics*, 134(1): 51–107.

Cooper, Zack, Stephen Gibbons, and Matthew Skellern. 2018. "Does Competition from Private Surgical Centres Improve Public Hospitals' Performance? Evidence from the English National Health Service." *Journal of Public Economics*, 166: 63–80.

Cooper, Zack, Stephen Gibbons, Simon Jones, and Alistair McGuire. 2011. "Does Hospital Competition Save Lives? Evidence From The English NHS Patient Choice Reforms." *Economic Journal*, 121(554): F228–F260.

De Loecker, Jan, and Chad Syverson. 2021. "An Industrial Organization Perspective on Productivity." In *Handbook of Industrial Organization, Vol. 4.* , ed. Kate Ho, Ali Hortaçsu and Alessandro Lizzeri, Chapter 3, 141–223. Amsterdam: Elsevier.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91(4): 795–813.

Ellis, Randall P., and Thomas G. McGuire. 1986. "Provider behavior under prospective reimbursement: Cost sharing and supply." *Journal of Health Economics*, 5(2): 129–151.

Finkelstein, Amy. 2007. "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare." *The Quarterly Journal of Economics*, 122(1): 1–37.

Finnish Government Decree 464/2008. 2008. "Valtioneuvoston asetus sosiaali- ja terveydenhuollon asiakasmaksuista annetun asetuksen muuttamisesta." *https://www.finlex.fi/fi/laki/alkup/2008/20080464*.

Finnish Government Proposal 90/2010. 2010. "Hallituksen esitys Eduskunnalle terveydenhuoltolaiksi sekä laeiksi kansanterveyslain ja erikoissairaanhoitolain muuttamiseksi sekä sosiaali- ja terveydenhuollon asiakasmaksuista annetun lain muuttamiseksi." *https://www.eduskunta.fi/FI/vaski/HallituksenEsitys/Documents/he_90+2010.pdf*.

Finnish Intitute for Health and Welfare (THL). 2012. "Trends in Access to Specialised Health Care in 2007–2012." THL Statistical Report 28/2012.

Fleischman, Andrew N., Majd Tarabichi, Zachary Magner, Javad Parvizi, and Richard H. Rothman. 2019. "Mechanical Complications Following Total Hip Arthroplasty Based on Surgical Approach: A Large, Single-Institution Cohort Study." *The Journal of Arthroplasty*, 34(6): 1255–1260.

Gaynor, Martin. 2016. "New Health Care Symposium: Consolidation And Competition In US Health Care." *Health Affairs blog.* *http://healthaffairs.org/blog/2016/03/01/new-health-care-symposium-consolidation-and-competition-in-us-health-care [accessed 16th May 2024].*

Gaynor, Martin, Carol Propper, and Stephan Seiler. 2016. "Free to Choose? Reform, Choice, and Consideration Sets in the English National Health Service." *American Economic Review*, 106(11): 3521–57.

Gaynor, Martin, Kate Ho, and Robert J. Town. 2015. "The Industrial Organization of Health-Care Markets." *Journal of Economic Literature*, 53(2): 235–84.

Gaynor, Martin, Mauro Laudicella, and Carol Propper. 2012. "Can Governments Do It Better? Merger Mania and Hospital Outcomes in the English NHS." *Journal of Health Economics*, 31(3): 528–543.

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): 134–66.

Godøy, Anna, Venke F. Haaland, Ingrid Huitfeldt, and Mark Votruba. 2024. "Hospital Queues, Patient Health, and Labor Supply." *American Economic Journal: Economic Policy*, 16(2): 150–81.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town. 2015. "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry." *American Economic Review*, 105(1): 172–203.

Gupta, Atul. 2021. "Impacts of Performance Pay for Hospitals: The Readmissions Reduction Program." *American Economic Review*, 111(4): 1241–83.

Gyrd-Hansen, Dorte. 2003. "Willingness to pay for a QALY." *Health Economics*, 12(12): 1049–1060.

Handel, Ben, and Kate Ho. 2021. "The Industrial Organization of Health Care Markets." In *Handbook of Industrial Organization, Vol. 5.* , ed. Kate Ho, Ali Hortaçsu and Alessandro Lizzeri, Chapter 16, 521–614. Amsterdam: Elsevier.

Hetemaa, Tiina, Katja Ilmarinen, Satu Kapiainen, Ilmo Keskimäki, Meri Koivusalo, Merja Korajoki, Ismo Linnosmaa, Eeva Liukko, Pasi Moisio, Susanna Mukkila, Kimmo

Parhiala, Mikko Peltola, Eeva Reissell, Liina-Kaisa Tynkkynen, and Maria Vaalavuo. 2018. "Social and Health Care Client Fees: Allocation, Effects and Fairness." Prime Minister's Office Publication Series 30/2018.

Jansson, Karl-Åke, and Fredrik Granath. 2011. "Health-related quality of life (EQ-5D) before and after orthopedic surgery." *Acta orthopaedica*, 82: 82–9.

Karhunen, Elina. 2020. "Study on the Number and Regional Distribution of Hospital Beds in Finland – Background Study for Assessing the Need for Construction Investments." Reports and Memorandums of the Ministry of Social Affairs and Health 2020:36.

Kautiainen, Kirsi, Unto Häkkinen, and Jorma Lauharanta. 2011. "Finland: DRGs in a Decentralized Health Care System." In *Diagnosis-Related Groups in Europe – Moving Towards transparency, Efficiency and Quality in Hospitals*. , ed. Reinhard Busse, Alexander Geissler and Wilm Quentin, Chapter 18, 321–338. New York, USA: McGraw-Hill Open University Press.

Keskimäki, Ilmo, Liina-Kaisa Tynkkynen, Eeva Reissell, Meri Koivusalo, Vesa Syrjä, Lauri Vuorenkoski, Rechel Bernd, and Marina Karanikolos. 2019. *Finland: Health System Review. Health Systems in Transition, vol. 21 (2)*, World Health Organization. Regional Office for Europe.

Kessler, Daniel P., and Mark B. McClellan. 2000. "Is Hospital Competition Socially Wasteful?" *The Quarterly Journal of Economics*, 115(2): 577–615.

Kortelainen, Mika, Liisa T Laine, Konsta Lavaste, Tanja Saxell, and Luigi Siciliani. 2023. "Improving Performance Through Allocation and Competition: Evidence from a Patient Choice Reform." VATT Working Paper 156.

Longo, Francesco, Luigi Siciliani, Giuseppe Moscelli, and Hugh Gravelle. 2019. "Does Hospital Competition Improve Efficiency? The Effect of the Patient Choice Reform in England." *Health Economics*, 28(5): 618–640.

Moscelli, Giuseppe, Hugh Gravelle, and Luigi Siciliani. 2021. "Hospital Competition and Quality for Non-Emergency Patients in the English NHS." *The RAND Journal of Economics*, 52(2): 382–414.

Moscelli, Giuseppe, Hugh Gravelle, and Luigi Siciliani. 2023. "The Effect of Hospital Choice and Competition on Inequalities in Waiting Times." *Journal of Economic Behavior & Organization*, 205: 169–201.

Moscelli, Giuseppe, Hugh Gravelle, Luigi Siciliani, and Rita Santos. 2018. “Heterogeneous Effects of Patient Choice and Hospital Competition on Mortality.” *Social Science & Medicine*, 216: 50–58.

Mäkelä, Keijo T, Mikko Peltola, Reijo Sund, Antti Malmivaara, Unto Häkkinen, and Ville Remes. 2011. “Regional and Hospital Variance in Performance of Total Hip and Knee Replacements: A National Population-Based Study.” *Annals of Medicine*, 43(Suppl. 1): S31–S38.

Newsweek and Statista. 2022. “World’s Best Hospitals 2022 – Finland.” *https://www.newsweek.com/worlds-best-hospitals-2022/finland [accessed 25 November 2024].*

Nikolova, Silviya, Mark Harrison, and Matt Sutton. 2016. “The Impact of Waiting Time on Health Gains from Surgery: Evidence from a National Patient-reported Outcome Dataset.” *Health Economics*, 25(8): 955–968.

OECD. 2020. *Waiting Times for Health Services: Next in Line. OECD Health Policy Studies*, Paris: OECD Publishing.

Osnes-Ringen, H., M. K. Kvamme, I. S. Kristiansen, M. Thingstad, J. E. Henriksen, T. K. Kvien, and H. Dagfinrud. 2011. “Cost-Effectiveness Analyses of Elective Orthopaedic Surgical Procedures in Patients with Inflammatory Arthropathies.” *Scandinavian Journal of Rheumatology*, 40(2): 108–115.

Paltta, Päivi. 2008. “Eligibility for Reimbursements of Health-Related Travel Costs in Finland and Selected Other Countries.” The Social Insurance Institution Studies in Social Security and Health 99.

Pirkanmaa Hospital District. 2007. “Miljoona suomalaista voi valita leikkaushoitopaikkansa.” *https://web.archive.org/web/20070826180414/http:/www.pshp.fi/default.aspx?contentid=1604&nodeid=9469&contentlan=1 [archived August 26th 2007, accessed 17th June 2024].*

Propper, Carol, Simon Burgess, and Denise Gossage. 2008. “Competition and Quality: Evidence from the NHS Internal Market 1991–9.” *The Economic Journal*, 118(525): 138–170.

Ray, Debraj, and Arthur Robson. 2018. “Certified Random: A New Order for Coauthorship.” *American Economic Review*, 108(2): 489–520.

Remes, Ville, Mikko Peltola, Unto Häkkinen, Heikki Kröger, Juhana Leppilahti, Miika

Linna, Antti Malmivaara, Keijo Mäkelä, Olavi Nelimarkka, Ilmo Parvinen, Seppo Seitsalo, and Jarmo Vuorinen. 2007. "PERFECT - Arthroplasty: Costs and Effectiveness of Hip and Knee Arthroplasty (in Finnish)." Stakes Working Paper 29/2007.

Robinson, James C., Harold S. Luft, Stephen J. McPhee, and Sandra S. Hunt. 1988. "Hospital Competition and Surgical Length of Stay." *JAMA*, 259(5): 696–700.

Roos, Anne-Fleur, Owen O'Donnell, Frederik T. Schut, Eddy Van Doorslaer, Raf Van Gestel, and Marco Varkevisser. 2020. "Does Price Deregulation in a Competitive Hospital Market Damage Quality?" *Journal of Health Economics*, 72: 102328.

Roth, Jonathan. 2022. "Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends." *American Economic Review: Insights*, 4(3): 305–22.

Silber, Jeffrey H., Paul R. Rosenbaum, Bijan A. Niknam, Richard N. Ross, Joseph G. Reiter, Alexander S. Hill, Lauren L. Hochman, Sydney E. Brown, Alexander F. Arriaga, Rachel R. Kelz, and Lee A. Fleisher. 2020. "Comparing Outcomes and Costs of Surgical Patients Treated at Major Teaching and Nonteaching Hospitals: A National Matched Analysis." *Annals of Surgery*, 271(3): 412–421.

Skinner, Jonathan, and Douglas Staiger. 2015. "Technology Diffusion and Productivity Growth in Health Care." *The Review of Economics and Statistics*, 97(5): 951–964.

Statistics Finland. 2010. "The first phase of the recession lowered real income of high and low-income households." *https://stat.fi/til/tjt/2008/tjt_2008_2010-05-20_tie_001_en.html [accessed 29th May 2024].*

Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature*, 49(2): 326–65.

the Finnish Broadcasting Company. 2007. "A Million Finns Will Soon Have More Freedom to Choose their Hospital." *https://yle.fi/a/3-5791182 (in Finnish).*

THL. 2021. "PERFECT – PERFormance, Effectiveness and Cost of Treatment episodes." *https://thl.fi/en/web/thlfi-en/research-and-development/research-and-projects/perfect-performance-effectiveness-and-cost-of-treatment-episodes [accessed June 23rd 2022].*

Torkki, Paulus. 2012. "Best Practice Processes - What Are the Reasons for Differences in Productivity Between Surgery Units." PhD diss. Aalto University School of Science.

Urquhart, Donna M., Fahad S. Hanna, Sharon L. Brennan, Anita E. Wluka, Karin Leder, Peter A. Cameron, Stephen E. Graves, and Flavia M. Cicuttini. 2010. "Incidence and Risk Factors for Deep Surgical Site Infection After Primary Total Hip Arthroplasty: A Systematic Review." *The Journal of Arthroplasty*, 25(8): 1216–1222.

Van Reenen, John. 2018. "Increasing Difference Between Firms: Market Power and the Macro Economy." *Changing Market Structures and Implications for Monetary Policy, Kansas City Federal Reserve: Jackson Hole Symposium*, 19–65.

Varkevisser, Marco, Stéphanie A. van der Geest, and Frederik T. Schut. 2012. "Do Patients Choose Hospitals With High Quality Ratings? Empirical Evidence from the Market for Angioplasty in the Netherlands." *Journal of Health Economics*, 31(2): 371–378.

World Health Organization (WHO). 2024. "Health Data Overview for the Republic of Finland – Life Expectancy." *https://data.who.int/countries/246 [accessed 21st May 2024].*

ONLINE APPENDIX

Improving Performance Through Allocation and Competition: Evidence from a Patient Choice Reform

Mika Kortelainen[a] ⓡ Liisa T. Laine[b] ⓡ Konsta Lavaste[c] ⓡ Tanja Saxell[d] ⓡ Luigi Siciliani[e]

[a] University of Turku, Finnish Institute for Health and Welfare

[b] University of Missouri, IZA

[c] Finnish Institute for Health and Welfare, University of Jyväskylä

[d] Aalto University, VATT Institute for Economic Research, Helsinki GSE

[e] University of York

A1. Data

We use de-identified, non-coded data on public hospital discharges and admissions in Finland in 2004–2010. The discharge data, officially known as the Care Register for Health Care, are from the Finnish Institute for Health and Welfare.

A1.1. Hospital Treatment Spell Construction

The discharge data are at the level of hospital treatment episodes and do not contain identifiers for hospital treatment spells.[42] Hence we use information on hospital admission and discharge dates to identify the spells. We made the following assumptions when assigning observations to the spells: hospital treatment episodes a and b of patient i belong to the same hospital treatment spell if

- admission date in episode $a \leq$ admission date in episode $b \leq$ discharge date in episode a, or
- admission date in episode a = admission date in episode b.

Treatment episodes in nursing homes and other long-term care facilities are excluded from the data before constructing the hospital treatment spells. We analyze treatment spells that began between January 1st 2004 and December 31st 2010 (our observation period).

A1.2. Sample Construction

For our baseline estimations, we construct three estimation samples of surgeries from the discharge data: planned (primary) hip replacement surgeries, planned (primary) knee replacement surgeries, and all planned orthopedic surgeries. We identified the surgeries based on the main procedure code and the type of hospital admission: non-emergency/planned.[43] In total, the hip replacement sample includes 8, the knee replacement sample includes 7, and the full orthopedic sample includes 658 detailed surgery types.

For our analyses on the quality spillover effect on emergency care, we also constructed three emergency care samples based on the main diagnosis code (ICD-10): AMI, stroke, and hip fracture

[42] Hospital treatment spell is a collection of treatment episodes (often just one episode) denoting the entirety of a patient's hospital stay from admission to the initial hospital to discharge from the final hospital. For example, a surgery and the following post-operative care are often two separate episodes but comprise one spell. Note that a patient might be transferred to another hospital for post-operative care, hence one hospital treatment spell can include care in several hospitals.

[43] Primary surgery means the initial rather than revision surgery. Our sample of all orthopedic surgeries includes both primary and revision surgeries because they are separately identifiable only for some surgeries such as hip and knee replacements.

admissions. The main inclusion criteria for different samples are shown in the table below. We also identified all planned surgeries for the purpose of calculating the predicted HHI that we use in our continuous DiD approach.

Main Inclusion Criteria of Different Estimation Samples

Sample/description	Procedure codes[a]	Diagnosis codes (ICD-10)[b]
Planned surgeries:		
Primary hip replacement surgeries	NFB*	
Primary knee replacement surgeries	NGB*	
Orthopedic surgeries	N* excluding the ones in which third character is a number[c]	
All surgeries	A*–Q* excluding the ones in which third character is a number[c]	
Emergency admissions:		
AMI admissions		I21.0, I21.1, I21.2, I21.3, I21.4, I21.9, I22.0, I22.1, I22.8, I22.9
Stroke admissions		I60*, I61*, I62*, I63* I16.4, I16.6, G46*, I67.2, I69.8, R47.0
Hip fracture admissions		S72.0, S72.1, S72.2

[a] According to the Finnish version of the Nordic Classification of Surgical Procedures (NCSP), see (in Finnish): https://urn.fi/URN:ISBN:978-952-245-858-2. The code may continue with any set of characters after the symbol *, indicating a more disaggregated classification.
[b] According to ICD-10 codes, see: https://icd.who.int/browse10/2010/en#/.
[c] Identification of surgical procedures follows the official statistics by the Finnish Institute for Health and Welfare: https://urn.fi/URN:NBN:fi-fe201205085475.

One reform area hospital closed-down at the end of the pre-reform period (2006). Its inclusion would bias our results and, hence, we exclude all patients who were treated in that hospital. This creates another source of bias, as patients who lived close to that hospital would appear in the samples in the post-reform period (and not in the pre-reform period). Hence, we exclude all patients who lived close to the closed-down hospital (see Figure A.10).

A1.3. Combining Discharge Data with Additional Data Sources

After constructing the samples from the discharge data, we combined them with additional de-identified, non-coded administrative data. First, patients' dates of death are from Statistics Finland's Causes of Death Registry. Second, the proxy of patients' residences at the end of each year are from Grid Database by Statistics Finland (grid locations). The grid locations are recorded on December 31st each year and thus the variable is not available for those who died or emigrated from Finland mid-year. We assume that their grid location at the time of their admission was as at the end of the previous year. Grid location data is de-identified by Statistics Finland and the grid locations are recorded missing if grid cells had too few inhabitants in a given year. In cases when the grid locations are recorded missing (≈ 0.5 percent), we use the coordinates of the center of patient's home municipality. Throughout the paper, we use the 2010 municipality classification. Third, hospitals' locations are determined by the centers of the municipalities where they are located (we do not have more detailed information on hospital locations). We link the discharge data to the provider registry (TOPI) from the Finnish Institute for Health and Welfare to obtain information on hospitals' municipalities. We also link municipalities with the coordinates of their centers. Fourth, we linked patients' and hospitals' municipalities to hospital districts by data from the Association of Finnish Municipalities. We use the hospital district classification of 2010.

A1.4. Measures of Hospital Performance

We provide details on the measures of hospital performance (for more details on the construction of other variables, see the working paper version of this paper: Kortelainen et al. (2023)). We calculate the 30-day follow-up period of emergency readmission from discharge from the last, rather than the initial, hospital in the treatment spell to account for transfers to another hospital after surgery (Torkki, 2012). The practice is similar to transferring patients to post-acute care facilities in the U.S. In our data, 4–18 percent of patients (depending on the sample) were transferred to another hospital for post-acute care. Similarly, the length of stay includes days in multiple hospitals in case there were hospital transfers.

Waiting time is readily available in the discharge data, but some hospitals or hospital districts have reported it less consistently than others. In the hip replacement sample, 23 percent of waiting time values are missing, while the same is true for 25 percent of observations in the knee replacement sample and for more than 33 percent of the sample of all orthopedic surgeries.[44] We implement comprehensive robustness checks in Section VIII, and confirm that the missing values do not bias the estimates of the

[44]Compared to hip and knee replacements, it is more likely that hospitals do not record waiting times for minor procedures included in the broad class of all orthopedic surgeries.

A1.5. Measures of Hospital Choice

	Outcome	Source	Description
(1)	Teaching hospital	Discharge data	An indicator equal to one if the patient received care at a teaching (university-based) hospital.
(2)	Distance	Discharge data, patient location (grid) data, & hospital location[a]	A straight line distance between patient i's residence and hospital h's location in kilometers. Some neighboring hospitals are recorded under a joint identifier, and in these cases we use the location of the largest hospital under the joint identifier.
(3)	Nearest hospital	Discharge data, patient location (grid) data, & hospital location[a]	An indicator equal to one if the patient received care atthe hospital nearest to their residence. Created based on outcome (2).
(4)	Different hospital district	Discharge data, patient location (grid) data, & hospital location[a]	An indicator equal to one if the patient i's hospital district was not the same as hospital h's hospital district.

[a] See variable (7) in online Appendix Section A1.8.

A1.6. Measures of Hospital Volume and Market Structure

	Outcome	Source	Description
(1)	Hospital volume	Discharge data	Number of patients in hospital h and quarter t.
(2)	Actual HHI	Discharge data	Concentration of hospital h's market in quarter t. Ranges from 0 (minimal concentration) to 1 (monopoly). For more information, see online Appendix Section A1.10.

A1.7. Hospital Performance: Clinical Quality, Waiting Time, and Length of Stay

We calculate the 30-day follow-up period of emergency readmission from discharge from the last, rather than the initial, hospital in the treatment spell to account for transfers to another hospital after surgery (Torkki, 2012). The practice is similar to transferring patients to post-acute care facilities in the U.S. In our data, 4–19 percent of patients (depending on the sample) were transferred to another hospital for post-acute care. Similarly, the length of stay includes days in multiple hospitals in case there were hospital transfers.

	Outcome	Source	Description
Panel A. Main clinical quality measure			
(1)	Emergency readmission within 30 days	Discharge data	An indicator equal to one if the patient had an emergency admission to *any* hospital for *any* reason within 30 days of being discharged from the last hospital in the treatment spell.
Panel B. Additional planned clinical care quality measures			
(2)	Revision surgery within 2 years[a]	Discharge data	An indicator equal to one if the patient had planned admission with main procedure code NFC or NGC to *any* hospital within 730 days of the initial surgery.
(3)	Mechanical complication of prosthesis within 2 years[a]	Discharge data	An indicator equal to one if the patient had admission with ICD-10 diagnosis code T84.0 to *any* hospital within 730 days of the initial surgery.
(4)	Infection or inflammation in prosthesis within 2 years[a]	Discharge data	An indicator equal to one if the patient had admission with ICD-10 diagnosis code T84.5 to *any* hospital within 730 days of the initial surgery.
Panel C. Additional emergency care quality measure			
(5)	Death within 30 days	Causes of death registry	An indicator equal to one if the patient died within 30 days after the initial treatment episode began (includes deaths in and out of the hospital).
Panel D. Length of stay			
(6)	Length of stay	Discharge data	Number of days between the admission and the departure in the treatment spell. Includes days in multiple hospitals in case there were hospital transfers.
Panel E. Waiting time			
(7)	Waiting time	Discharge data	Number of days between the specialist's final treatment decision and surgery.

Notes: See the classification of surgical procedures (in Finnish) in: `https://urn.fi/URN:ISBN:978-952-245-858-2`. See ICD-10 codes in: `https://icd.who.int/browse10/2010/en#/`.
[a] The discharge data do not allow us to identify whether the revision surgery, mechanical complication, infection, or inflammation concerned the same prosthesis as the initial surgery. It is possible that we identify that, for example, the patient underwent revision surgery within 2 years, even if the initial surgery concerned the left knee and the infection concerned the right knee. We assume that the magnitude of the bias resulting from these false positives is of minor magnitude.

Waiting time is readily available in the discharge data, but some hospitals or hospital districts have reported it less consistently than others. In the hip replacement sample, 21 percent of waiting time values are missing, while the same is true for 22 percent of observations in the knee replacement sample and for more than 30 percent of the sample of all orthopedic surgeries.[45] We implement comprehensive robustness checks in Section VIII, and confirm that the missing values do not bias the estimates of the effects on waiting times.

A1.8. Other Variables

[45]Compared to hip and knee replacements, it is more likely that hospitals do not record waiting times for minor procedures included in the broad class of all orthopedic surgeries.

	Variable	Source	Description
Panel A. Covariates			
(1)	Patient age	Discharge data	Patient age at the the time of admission.
(2)	Patient sex	Discharge data	An indicator equal to one if the patient is female.
(3)	N of emergency admissions within 1 year	Discharge Data	Number of emergency admissions that the patient had to *any* hospital within 365 days before the surgery or emergency admission under consideration. May include multiple admissions from one treatment spell.
(4)	Weekend admission	Discharge data	An indicator equal to one if the admission date was Saturday or Sunday.
Panel B. Other variables			
(5)	Patient residence	Location (grid) data	Finland was divided into 1x1 kilometer squares. The residence is the easting and northing coordinates (ETRS-TM95FIN) of the square in which the patient's residence was located.
(6)	Patient's municipality of residence	Location (grid) data	The municipality in which the patient was resident at the end of the year.
(7)	Hospital location	Provider registry (TOPI) & municipality center locations (Google Maps)	Easting and northing coordinates of the municipality in which hospital was located (ETRS-TM95FIN). Constructed based on hospital's municipality in TOPI and municipality center locations.
(8)	Hospital districts of patients and hospitals	(6), provider registry (TOPI), & hospital district data from Association of Finnish Municipalities	The hospital district to which the patient's municipality of residence or the hospital's municipality belonged. We use publicly available information on the hospital district of each municipality (patient's or hospital's) from the Association of Finnish Municipalities.
(9)	Reform area	Discharge data & (8)	An indicator equal to one if the hospital in which the patient was treated belonged to Vaasa, Etelä-Pohjanmaa, Pirkanmaa, or Päijät-Häme hospital districts.

A1.9. Hospital Expenditure

The discharge data do not include hospital expenditures, hence we use the Hospital Benchmarking Data from the Finnish Institute for Health and Welfare.[46] The data provide hospitals' operating expenditures by specialty, allowing us to measure hospitals' surgery-related expenditures, and a case mix index for controlling patients' severity.

	Outcome	Source	Description
Panel A. Expenditure outcomes			
(1)	Total expenditure	Hospital benchmarking data	Hospital h's annual surgery-related operating expenditures (millions of €[a]) in year t.
(2)	Expenditure per patient (treatment episode)	Hospital benchmarking data	Hospital h's annual surgery-related operating expenditures (€[a]) divided by hospital's DRG-weighted number of surgical patients (treatment episodes) in year t.
Panel B. Covariates			
(3)	Case mix index	Hospital benchmarking data	Number of DRG-weighted surgical patients of hospital h in year t is divided by absolute number of surgical patients (treatment episodes) of hospital h in year t and then transformed into an index by setting the whole country equal to 1.

[a] Deflated using Statistics Finland's price index of public health care services (see `https://stat.fi/en/statistics/jmhi`). Base year = 2000.

A1.10. Actual HHI

We constructed the hospital-level Herfindahl-Hirschman Index (HHI) in two steps. First, we calculated a municipality-level HHI value for each municipality m in each quarter t as a sum of squared market shares:

$$HHI_{mt}^{MUN} = \sum_{h=1}^{H} \left(\frac{n_{hmt}}{N_{mt}} \right)^2, \tag{A.1}$$

[46]See `https://thl.fi/en/web/thlfi-en/statistics-and-data/statistics-by-topic/health-care-services/hospital-benchmarking`

where n_{hmt} is the number of patients from municipality m who underwent surgery in hospital h in quarter t. N_{mt} is the total number of surgical patients from municipality m.

Second, we calculated the hospital-level HHI as a weighted average of the values of municipality-level HHI, using each municipality's share of the hospital's total patient volume as weights:

$$HHI_{ht}^{HOSP} = \sum_{m=1}^{M} \left(\frac{n_{mht}}{N_{ht}} \times HHI_{mt}^{MUN} \right), \tag{A.2}$$

where n_{mht} is the number of patients from municipality m who underwent surgery in hospital h in quarter t. N_{ht} is the total number of patients (from any municipality) who underwent surgery in hospital h. We refer to HHI_{ht}^{HOSP} as the *actual* HHI. We calculated it separately for samples of hip replacement surgeries, knee replacement surgeries, and all orthopedic surgeries.

A2. Continuous DiD and Control Function Approaches: Further Details

A2.1. Predicted HHI

Following Kessler and McClellan (2000) and the patient choice reform literature (Gaynor, Moreno-Serra and Propper, 2013), we construct a predicted version of the HHI and use it as the treatment intensity variable in the continuous DiD specification (4). We calculate the predicted HHI in four steps. First, we estimate the predicted hospital volumes using data on *all* planned surgery patients. We restrict the sample to patients aged 18–74 and allow them to choose any hospital in Finland. We estimate how travel distance and other observable patient and hospital characteristics affect the probability of patient i choosing hospital h in quarter t. Specifically, we estimate the following conditional logit model separately for each quarter during Q1/2004–Q4/2010:

$$\begin{aligned} U_{iht} &= V_{iht} + \varepsilon_{iht} \\ &= \alpha_{0t} + \alpha_{1t}\mathbf{X}_{iht} + \alpha_{2t}\mathrm{km}_{iht} + \alpha_{3t}\mathrm{km}_{iht}^2 \\ &\quad + \alpha_{4t}(\mathbf{X}_{iht} \times \mathrm{km}_{iht}) + \alpha_{5t}(\mathbf{X}_{iht} \times \mathrm{km}_{iht}^2) + \varepsilon_{iht}, \end{aligned} \tag{B.1}$$

where km_{iht} is the distance between patient i's residence and hospital h in kilometers and $\mathbf{X}_{iht}$ are the hospital characteristics: an indicator for hospital h being located in the same hospital district as patient i, an indicator for teaching hospital, and an indicator for regional hospital.[47] Patient i

[47] The specification for predicting hospital choices is similar to Moscelli et al. (2018) and Moscelli, Gravelle and Siciliani (2021). An alternative specification used by Gaynor, Moreno-Serra and Propper (2013) includes differential distances and interactions between hospital and patient characteristics, and we confirmed that it yields hospital volumes

chooses hospital h with the probability of:

$$P_{iht} = \exp(V_{iht})\Big[\sum_{h' \in S_i} \exp(V_{ih't})\Big]^{-1}, \tag{B.2}$$

Second, we calculate the predicted volume of each hospital's patients from a given municipality m as a sum of probabilities over all these patients.[48] Third, we calculate the predicted municipality-level HHI as the sum of the squares of the predicted market shares (based on the predicted volumes):

$$\widehat{HHI}_{mt}^{MUN} = \sum_{h=1}^{H} \left(\frac{\widehat{n}_{hmt}}{\widehat{N}_{mt}}\right)^2 \tag{B.3}$$

$\widehat{n}_{hmt}$ is the predicted volume from municipality m, that is the number of patients from that municipality who received care at a hospital h in quarter t. $\widehat{N}_{mt}$ is the predicted number of patients from municipality m (received care at any hospital).

Fourth, the predicted hospital-level HHI is calculated as a weighted average of the values of the predicted municipality-level HHI, using each municipality's share of the hospital's predicted total patient volume as weights:

$$\widehat{HHI}_{ht}^{HOSP} = \sum_{m=1}^{M} \left(\frac{\widehat{n}_{mht}}{\widehat{N}_{ht}} \times \widehat{HHI}_{mt}^{MUN}\right) \tag{B.4}$$

$\widehat{n}_{mht}$ is the predicted number of patients from municipality m who received care at a hospital h in quarter t. $\widehat{N}_{ht}$ is the total predicted number of patients (from any municipality) received care at a hospital h.

Our model predicts approximately 79 percent of patients' choices correctly in Q1/2004–Q4/2010. The share is higher than what Gaynor, Moreno-Serra and Propper (2013) predict in a United Kingdom setting (approximately 75 percent). The correlation between the actual and predicted HHI is 0.76. The correlation is similar than what Moscelli et al. (2018) predict in the United Kingdom setting (0.65–0.80).

Finally, we fix the predicted HHI to its pre-reform mean, as follows:[49]

$$\widehat{HHI}_{h}^{HOSP} = \frac{\sum_{t=2004Q1}^{2007Q3} \widehat{HHI}_{ht}^{HOSP}}{15} \tag{B.5}$$

that result in similar HHI values to the specification we use.

[48] An alternative way is to determine the option with the highest probability and regard it as the predicted choice. The predicted volume would be counted as sums of these choices.

[49] Note that the conditional logit model (equation (B.1)) and its parameters are estimated separately for each period. In practice, we only use estimates from the pre-reform period in constructing the predicted HHI.

We refer to $\widehat{HHI}_h^{HOSP}$ used in our continuous DiD analyses as the *predicted* HHI.

A2.2. Control Function Approach

The treatment intensity variable in the continuous DiD approach (predicted HHI) is affected by patients' choices, which may correlate with unobserved quality. We follow Moscelli, Gravelle and Siciliani (2021) to control for time-varying patient selection to hospitals using a control function approach. In the first stage, we estimate the following conditional logit model separately for each quarter in Q1/2004–Q4/2010:

$$\begin{aligned} U_{iht} &= V_{iht} + \xi_{iht} \\ &= \lambda_{1t}\text{km}_{ih} + \lambda_{2t}\text{km}_{ih}^2 + \lambda_{3t}\text{km}_{ih}^3 + \lambda_{4t}\text{nearest}_{ih} \\ &\quad + \lambda_{5t}\text{sameHD}_{ih} + \lambda_{6t}\text{teaching}_h + \lambda_{7t}\text{regional}_h + \xi_{iht}, \end{aligned} \tag{B.6}$$

where km_{ih} is the distance between the patient and the hospital, nearest_{ih} is an indicator for h being the geographically nearest hospital for patient i, sameHD_{ih} is an indicator for patient i and hospital h being located in the same hospital district, and teaching_h and regional_h are indicators for teaching and regional hospitals.[50] The chosen covariates reflect geographical access (km_{ih}, nearest_{ih}) and factors potentially relevant to hospital choice post-reform (sameHD_{ih}, teaching_h, regional_h).[51] Patient i chooses hospital h with a probability of

$$P_{iht} = \exp(V_{iht})\Big[\sum_{h' \in S_i} \exp(V_{ih't})\Big]^{-1}, \tag{B.7}$$

After calculating the choice probabilities, we can derive the set of residuals (one for each hospital h):

$$\widehat{r}'_{it} = [\widehat{r}_{i1t}, \widehat{r}_{i2t}, \ldots, \widehat{r}_{iHt}] = C_{iht} - \widehat{P}_{iht} \tag{B.8}$$

All these residuals ($H = 45$–52 depending on the sample) are then added to the second-stage specification (4).

[50]Baseline = central hospitals. We let patients choose among all hospitals in Finland ($N = 45$–52 hospital IDs depending on the sample, where some neighboring hospitals are recorded under a joint identifier). This is essentially the same as in Moscelli, Gravelle and Siciliani (2021), who restrict patients' choice sets to the closest 50 hospitals in the country.

[51]The covariates differ from the choice model in section A2.1, because the predicted HHI concerns only the more restricted *pre-reform* choices, for which these factors were much less relevant, while the control function also targets the less restricted *post-reform* choices in the reform area.

A3. Tables

Table A.1: Top-20 Orthopedic Surgeries in 2004–2010

Rank	Group[a]	Description	Share[b]
1	NGD*	Operations on meniscus of knee	12%
2	NGF*	Operations on synovia and joint surfaces of knee	8%
3	NGB*	Primary prosthetic replacement of knee joint	8%
4	NFB*	Primary prosthetic replacement of hip joint	7%
5	NBG*	Excision, reconstruction and fusion of joints of shoulder	7%
6	NHG*	Excision, reconstruction and fusion of joints of ankle and foot	4%
7	NDM*	Operations on fascia, ganglia, synovial sheaths and bursae of wrist and hand	4%
8	NHK*	Operations on bone of ankle and foot	4%
9	NGA*	Exploratory procedures on knee and lower leg	4%
10	NBL*	Operations on muscles and tendons of shoulder and upper arm	3%
11	NHU*	Removal of implants and external fixation devices from ankle and foot	3%
12	NDU*	Removal of implants and external fixation devices from wrist and hand	2%
13	NAG*	Excision, reconstruction and fusion of joints of spine	2%
14	NBE*	Operations on capsules and ligaments of joints of shoulder	2%
15	NGE*	Operations on capsules and ligaments of knee joint	2%
16	NDG*	Excision, reconstruction and fusion of joint of wrist and hand	2%
17	NBA*	Exploratory procedures on shoulder and upper arm	1%
18	NHL*	Operations on muscles and tendons of ankle and foot	1%
19	NBT*	Miscellaneous operations on shoulder or upper arm	1%
20	NDL*	Operations on muscles and tendons of wrist and hand	1%
		Other orthopedic surgeries	19%

[a] According to the Finnish version of the Nordic Classification of Surgical Procedures (NCSP). The grouping is based on the procedure code's first three characters (the code may continue with any set of characters after the symbol *, indicating a more disaggregated classification).
[b] The number of observations for each group is divided by the number of observations for all orthopedic surgeries in 2004–2010 ($N = 473,617$).

Table A.2: Pre-Reform Descriptive Statistics: Hospital Volume and Actual HHI by Hospital Type

	Hospital volume[a]				Actual HHI[b]			
	All	Teaching	Central	Regional	All	Teaching	Central	Regional
Panel A. Hip replacement surgeries								
Mean	29.19	91.47	27.42	12.09	0.83	0.91	0.88	0.77
SD	(32.69)	(46.27)	(15.47)	(5.73)	(0.13)	(0.06)	(0.11)	(0.14)
N	548	75	223	250	548	75	223	250
Panel B. Knee replacement surgeries								
Mean	35.25	102.00	33.22	16.67	0.85	0.91	0.90	0.78
SD	(36.20)	(50.26)	(19.92)	(8.00)	(0.13)	(0.07)	(0.10)	(0.14)
N	543	75	223	245	543	75	223	245
Panel C. All orthopedic surgeries								
Mean	428.33	1,281.28	403.87	194.26	0.73	0.80	0.81	0.64
SD	(465.98)	(741.21)	(187.94)	(86.92)	(0.12)	(0.14)	(0.07)	(0.08)
N	548	75	223	250	548	75	223	250

Notes: Values calculated from hospital-quarter level data spanning from Q1/2004 to Q3/2007 (pre-reform period).
[a] Number of patients.
[b] Hospital-quarter level actual HHI (0–1).

Table A.3: Pre-Reform Descriptive Statistics: Quality, Waiting Time, and Length of Stay, by Hospital Type

	Risk-adjusted readmission rate[a]				Mean waiting time[b]				Mean length of stay[c]			
	All	Teaching	Central	Regional	All	Teaching	Central	Regional	All	Teaching	Central	Regional
Panel A. Hip replacement surgeries												
Mean	0.08	0.08	0.08	0.08	172.95	181.39	201.98	141.11	9.02	7.18	9.43	9.20
SD	(0.01)	(0.01)	(0.01)	(0.02)	(86.40)	(59.44)	(88.35)	(83.26)	(2.82)	(1.96)	(2.94)	(2.71)
N	548	75	223	250	432	74	177	181	548	75	223	250
Panel B. Knee replacement surgeries												
Mean	0.10	0.10	0.11	0.10	210.12	236.94	249.82	160.61	8.26	7.08	8.51	8.39
SD	(0.02)	(0.01)	(0.02)	(0.02)	(110.90)	(95.11)	(115.55)	(91.79)	(2.22)	(1.70)	(2.19)	(2.28)
N	543	75	223	245	433	74	177	182	543	75	223	245
Panel C. All orthopedic surgeries												
Mean	0.06	0.07	0.06	0.06	133.59	167.19	156.71	97.43	2.24	2.89	2.29	2.00
SD	(0.01)	(0.01)	(0.01)	(0.01)	(65.31)	(43.95)	(73.39)	(40.94)	(0.75)	(1.00)	(0.60)	(0.66)
N	548	75	223	250	507	75	221	211	548	75	223	250

Notes: Values calculated from hospital-quarter-level data spanning from Q1/2004 to Q3/2007 (pre-reform period).

[a] Risk-adjusted 30-day readmission rate (0–1). Risk-adjusted by predicting patients' probability of readmission within 30 days with age, sex, number of emergency admissions within one year prior to the surgery, weekend indicator, time fixed effects, and surgery type fixed effects.

[b] Mean waiting time in days (no risk-adjustment). Some hospitals did not report any waiting times in some quarters, which is depicted as a smaller N. We analyze the robustness of the results regarding missing values in waiting times and summarize the results in Section V.C.

[c] Mean length of stay in days (no risk-adjustment).

Table A.4: Pre-Reform Descriptive Statistics: Hospital Expenditure for All Surgeries by Hospital Type

	No weights				DRG weights[a]			
	All	Teaching	Central	Regional	All	Teaching	Central	Regional
Panel A. Total expenditure (millions of €)[b]								
Mean	24.42	95.45	23.21	6.10				
SD	(36.97)	(66.39)	(8.66)	(3.09)				
N	118	15	48	55				
Panel B. Expenditure per patient (€)[c]								
Mean	586.30	950.73	610.67	465.64	411.88	452.75	425.43	388.92
SD	(174.75)	(118.42)	(91.04)	(55.42)	(42.65)	(40.63)	(38.31)	(32.72)
N	118	15	48	55	118	15	48	55

Notes: Values calculated from hospital-year-level data in the pre-reform period (2004–2006).
[a] Number of patients (treatment episodes) multiplied by DRG weights, which describe the relative average expenditure for operating on patients in a particular DRG category.
[b] Hospital's annual care-related expenditure in the surgical ward. (€, deflated using prices in 2000)
[c] Hospital's annual care-related expenditure in the surgical ward. (€, deflated using prices in 2000) divided by number of patients (treatment episodes) in the surgical ward.

Table A.5: Pre-Reform Descriptive Statistics: Quality, Waiting Time, and Length of Stay in Teaching Hospitals, by Area (Reform and Control)

	Risk-adjusted readmission rate[a]		Mean waiting time[b]		Mean length of stay[c]	
	Reform area	Control area	Reform area	Control area	Reform area	Control area
Panel A. Hip replacement surgeries						
Mean	0.08	0.08	220.87	171.35	6.19	7.43
SD	(0.01)	(0.01)	(45.87)	(58.60)	(0.74)	(2.10)
N	15	60	15	59	15	60
Panel B. Knee replacement surgeries						
Mean	0.10	0.10	302.14	220.37	6.34	7.26
SD	(0.01)	(0.01)	(92.74)	(89.05)	(0.60)	(1.83)
N	15	60	15	59	15	60
Panel C. All orthopedic surgeries						
Mean	0.08	0.06	186.69	162.32	4.58	2.47
SD	(0.00)	(0.01)	(32.24)	(45.34)	(0.65)	(0.51)
N	15	60	15	60	15	60

Notes: Values calculated from hospital-quarter level data spanning from Q1/2004 to Q3/2007 (pre-reform period).
[a] Risk-adjusted 30-day readmission rate (0–1). Risk-adjusted by predicting patients' probability of readmission within 30 days with age, sex, number of emergency admissions within one year prior to the surgery, weekend indicator, time fixed effects, and surgery type fixed effects.
[b] Mean waiting time in days (no risk-adjustment). Some hospitals did not report any waiting times in some quarters, which is depicted as a smaller N. We analyze the robustness of the results regarding missing values in waiting times and summarize the results in Section V.C.
[c] Mean length of stay in days (no risk-adjustment).

Table A.6: Pre-Reform Descriptive Statistics: Mean Age, Share of Females, and Mean Number of Previous Emergency Admissions in Teaching Hospitals, by Area (Reform and Control)

	Mean age[a]		Share of females[b]		Mean N of pre-surgery emergency admissions[c]	
	Reform area	Control area	Reform area	Control area	Reform area	Control area
Panel A. Hip replacement surgeries						
Mean	61.18	60.63	0.51	0.52	0.25	0.43
SD	(0.73)	(2.07)	(0.03)	(0.08)	(0.08)	(0.17)
N	15	60	15	60	15	60
Panel B. Knee replacement surgeries						
Mean	64.46	63.54	0.66	0.68	0.30	0.41
SD	(0.55)	(1.37)	(0.02)	(0.06)	(0.07)	(0.16)
N	15	60	15	60	15	60
Panel C. All orthopedic surgeries						
Mean	54.89	50.42	0.52	0.53	0.59	0.65
SD	(1.49)	(1.09)	(0.03)	(0.03)	(0.06)	(0.13)
N	15	60	15	60	15	60

Notes: Values calculated from hospital-quarter level data spanning from Q1/2004 to Q3/2007 (pre-reform period).
[a] Mean age at the time of admission (18–74).
[b] Share of females out of all patients.
[c] Mean number of emergency readmissions the hospital's patients had within 1 year before their surgery.

Table A.7: Robustness Check: All Orthopedic Surgeries Except Hip and Knee Replacements (Choice Outcomes)

	Teaching hospital[a] (1)	Distance (km)[b] (2)	Nearest hospital[c] (3)	Different hospital district[d] (4)
$\text{Treated}_i \times \text{Post}_t$	0.015^{***}	1.710^{*}	-0.004	0.005
	(0.005)	(1.020)	(0.007)	(0.005)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.372	26.128	0.799	0.040
N	402,422	402,422	402,422	402,422
Surgery type / Diagnosis code FEs	✓	✓	✓	✓
Municipal FEs	✓	✓	✓	✓
Age & sex	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 319$).
[a] Equals one if the patient received care at a teaching (university) hospital.
[b] Distance from the patient's residence to the hospital in kilometers.
[c] Equals one if the patient received care at the nearest hospital.
[d] Equals one if the patient received care outside their hospital district of residence.

Table A.8: The Effects of the Reform on Additional Quality Measures

	Revision[a] (1)	Infection[b] (2)	Complication[c] (3)
Panel A. Hip replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.008	-0.001	0.002
	(0.005)	(0.003)	(0.006)
$\text{mean}(y_{imht}\|\text{Post}_t = 0)$	0.034	0.012	0.056
N	31,404	31,404	31,404
Panel B. Knee replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.005	-0.000	-0.009*
	(0.004)	(0.003)	(0.005)
$\text{mean}(y_{imht}\|\text{Post}_t = 0)$	0.032	0.019	0.041
N	39,427	39,427	39,427
Surgery type FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 319$).
[a] Equals one if the patient had a revision surgery within 2 years of the initial surgery.
[b] Equals one if the patient had an infection or inflammation in the prosthesis within 2 years of the initial surgery.
[c] Equals one if the patient had a mechanical complication in the prosthesis within 2 years of the initial surgery.

Table A.9: The Effects of the Reform on Emergency Care Quality and Length of Stay

	Death within 30 days[a] (1)	Readmission[b] (2)	Length of stay[c] (3)
Panel A. AMI			
$\text{Treated}_i \times \text{Post}_t$	0.011	-0.001	-0.030
	(0.008)	(0.010)	(0.548)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.073	0.180	8.675
N	32,919	32,919	32,919
Panel B. Stroke			
$\text{Treated}_i \times \text{Post}_t$	0.001	-0.042***	-0.610
	(0.007)	(0.012)	(0.775)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.082	0.199	16.207
N	49,114	49,114	49,114
Panel C. Hip fracture			
$\text{Treated}_i \times \text{Post}_t$	-0.010	-0.025	-1.733
	(0.007)	(0.018)	(1.281)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.031	0.193	21.979
N	10,995	10,995	10,995
Diagnosis code FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 319$ depending on the sample).
[a] Equals one if the patient died (before or after discharge) within 30 days of admission.
[b] Equals one if the patient had an emergency readmission (to any hospital) within 30 days after discharge.
[c] Number of days between admission and discharge.

Table A.10: Robustness Check: Additional Controls (Choice Outcomes)

	Teaching hospital[a]	Distance (km)[b]	Nearest hospital[c]	Different hospital district[d]
	(1)	(2)	(3)	(4)
Panel A. Hip replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.051***	5.884**	-0.014	0.031**
	(0.016)	(2.845)	(0.019)	(0.013)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.429	29.091	0.818	0.035
N	31,404	31,404	31,404	31,404
Panel B. Knee replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.040***	2.360	0.039**	0.015
	(0.013)	(1.933)	(0.019)	(0.011)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.393	28.399	0.819	0.034
N	39,427	39,427	39,427	39,427
Panel C. All orthopedic surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.026***	2.413**	-0.006	0.009*
	(0.007)	(1.179)	(0.009)	(0.005)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.377	26.481	0.794	0.039
N	473,617	473,617	473,617	473,617
Surgery type / Diagnosis code FEs	✓	✓	✓	✓
Municipal FEs	✓	✓	✓	✓
Age & sex	✓	✓	✓	✓
Hospital FEs				
N of emergency admissions	✓	✓	✓	✓
Weekend	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 319$).
[a] Equals one if the patient received care at a teaching (university) hospital.
[b] Distance from the patient's residence to the hospital in kilometers.
[c] Equals one if the patient received care at the nearest hospital.
[d] Equals one if the patient received care outside their hospital district of residence.

Table A.11: Robustness Check: Additional Controls (Hospital Performance Outcomes)

	Readmission[a] (1)	Waiting time[b] (2)	Length of stay[c] (3)
Panel A. Hip replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	-0.003	-73.083***	-0.542*
	(0.007)	(12.429)	(0.293)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.080	181.991	8.091
N	31,404	24,222	31,404
Panel B. Knee replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.003	-100.056***	-0.601**
	(0.008)	(18.038)	(0.276)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.102	225.373	7.609
N	39,427	29,759	39,427
Panel C. All orthopedic surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.002	-20.717***	-0.165***
	(0.002)	(6.489)	(0.050)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.060	145.636	2.205
N	473,617	316,454	473,617
Surgery type / Diagnosis code FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓
Hospital FEs	✓	✓	✓
N of emergency admissions	✓	✓	✓
Weekend	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 319$).

[a] Equals one if the patient had an emergency readmission (to any hospital) within 30 days after discharge.

[b] Number of days between the specialist's final treatment decision and the surgery. Some of the values are missing, which results in smaller N compared to other columns (see online Appendix Section A1.7 for more details).

[c] Number of days between admission and discharge.

Table A.12: Robustness Check: Including Patients Affected by the Closed-Down Hospital (Choice Outcomes)

	Teaching hospital[a]	Distance (km)[b]	Nearest hospital[c]	Different hospital district[d]
	(1)	(2)	(3)	(4)
Panel A. Hip replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.055***	5.653**	0.003	0.034**
	(0.017)	(2.587)	(0.019)	(0.014)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.423	29.815	0.811	0.037
N	31,818	31,818	31,818	31,818
Panel B. Knee replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.043***	1.741	0.061***	0.018
	(0.014)	(1.826)	(0.021)	(0.011)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.388	29.103	0.812	0.035
N	39,920	39,920	39,920	39,920
Panel C. All orthopedic surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.026***	2.133*	0.015	0.007
	(0.007)	(1.201)	(0.014)	(0.006)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.374	26.861	0.789	0.040
N	477,475	477,475	477,475	477,475
Surgery type / Diagnosis code FEs	✓	✓	✓	✓
Municipal FEs	✓	✓	✓	✓
Age & sex	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 326$). All patients who received care in the closed-down reform area hospital included, as well as patients who lived in the municipalities neighboring the hospital. See Figure A.10 in Appendix A4 for a map.

[a] Equals one if the patient received care in a teaching (university) hospital.

[b] Distance from patient's residence to the hospital in kilometers.

[c] Equals one if the patient received care in the nearest hospital.

[d] Equals one if the patient received care outside their hospital district of residence.

Table A.13: Robustness Check: Including Patients Affected by the Closed-Down Hospital (Hospital Performance Outcomes)

	Readmission[a] (1)	Waiting time[b] (2)	Length of stay[c] (3)
Panel A. Hip replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	-0.003	-76.567***	-0.454
	(0.007)	(11.399)	(0.281)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.080	182.869	8.111
N	31,818	24,586	31,818
Panel B. Knee replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	-0.001	-96.056***	-0.438*
	(0.009)	(17.519)	(0.248)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.103	225.645	7.615
N	39,920	30,193	39,920
Panel C. All orthopedic surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.002	-17.408**	-0.164***
	(0.003)	(6.945)	(0.049)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.060	145.403	2.212
N	477,475	319,338	477,475
Surgery type / Diagnosis code FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 326$). All patients who received care in the closed-down reform area hospital included, as well as patients who lived in the municipalities neighboring the hospital. See Figure A.10 in Appendix A4 for a map.

[a] Equals one if the patient had an emergency readmission (to any hospital) within 30 days after discharge.

[b] Number of days between the specialist's final treatment decision and the surgery. Some of the values are missing, which results in smaller N compared to other columns (see online Appendix Section A1.7 for more details).

[c] Number of days between admission and discharge.

Table A.14: Robustness Check: Excluding Three Western Hospital Districts With Possibilities to Obtain Referral Outside Own Hospital District (Choice Outcomes)

	Teaching hospital[a]	Distance (km)[b]	Nearest hospital[c]	Different hospital district[d]
	(1)	(2)	(3)	(4)
Panel A. Hip replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.044^{***}	2.831^{**}	-0.013	0.033^{***}
	(0.013)	(1.365)	(0.016)	(0.010)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.451	27.462	0.826	0.026
N	29,220	29,220	29,220	29,220
Panel B. Knee replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.042^{***}	1.819	0.037^{**}	0.028^{***}
	(0.012)	(1.212)	(0.017)	(0.009)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.409	27.209	0.823	0.026
N	36,999	36,999	36,999	36,999
Panel C. All orthopedic surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.028^{***}	1.594	-0.004	0.011^{*}
	(0.009)	(1.018)	(0.010)	(0.006)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.397	25.936	0.793	0.037
N	446,243	446,243	446,243	446,243
Surgery type / Diagnosis code FEs	✓	✓	✓	✓
Municipal FEs	✓	✓	✓	✓
Age & sex	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 291$). Excludes patients who lived in the three Western hospital districts that some degree of patient choice.

[a] Equals one if the patient received care at a teaching (university) hospital.

[b] Distance from patient's residence to the hospital in kilometers.

[c] Equals one if the patient received care at the nearest hospital.

[d] Equals one if the patient received care outside their hospital district of residence.

Table A.15: Robustness Check: Excluding Three Western Hospital Districts Where Possible to Obtain Referral Outside Own Hospital District (Hospital Performance Outcomes)

	Readmission[a] (1)	Waiting time[b] (2)	Length of stay[c] (3)
Panel A. Hip replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.001	-79.616***	0.006
	(0.008)	(14.558)	(0.284)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.081	178.424	7.964
N	29,220	22,364	29,220
Panel B. Knee replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.009	-114.761***	-0.191
	(0.009)	(20.244)	(0.298)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.102	224.511	7.518
N	36,999	27,782	36,999
Panel C. All orthopedic surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.005	-28.846***	-0.108**
	(0.003)	(7.326)	(0.054)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.060	145.620	2.177
N	446,243	298,199	446,243
Surgery type / Diagnosis code FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 291$). Excludes patients who lived in the three Western hospital districts that allowed some degree of patient choice.

[a] Equals one if the patient had an emergency readmission (to any hospital) within 30 days after discharge.

[b] Number of days between the specialist's final treatment decision and the surgery. Part of the values are missing, which results in smaller N compared to other columns (see online Appendix Section A1.7 for more details).

[c] Number of days between admission and discharge.

Table A.16: Robustness Check: Excluding Four Northern Hospital Districts Where Possible to Obtain Referral Outside Own Hospital District (Choice Outcomes)

	Teaching hospital[a]	Distance (km)[b]	Nearest hospital[c]	Different hospital district[d]
	(1)	(2)	(3)	(4)
Panel A. Hip replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.052^{***}	5.825^{**}	-0.011	0.031^{**}
	(0.017)	(2.875)	(0.019)	(0.013)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.459	27.343	0.819	0.035
N	29,145	29,145	29,145	29,145
Panel B. Knee replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.042^{***}	2.319	0.043^{**}	0.016
	(0.014)	(1.994)	(0.019)	(0.011)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.422	26.779	0.819	0.034
N	36,538	36,538	36,538	36,538
Panel C. All orthopedic surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.026^{***}	2.535^{**}	-0.006	0.010^{*}
	(0.007)	(1.232)	(0.008)	(0.005)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.402	24.485	0.793	0.036
N	438,749	438,749	438,749	438,749
Surgery type / Diagnosis code FEs	✓	✓	✓	✓
Municipal FEs	✓	✓	✓	✓
Age & sex	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 281$). Excludes patients who lived in the four Western hospital districts that some degree of patient choice.

[a] Equals one if the patient received care at a teaching (university) hospital.

[b] Distance from patient's residence to the hospital in kilometers.

[c] Equals one if the patient received care at the nearest hospital.

[d] Equals one if the patient received care outside their hospital district of residence.

Table A.17: Robustness Check: Excluding Four Northern Hospital Districts Where Possible to Obtain Referral Outside Own Hospital District (Hospital Performance Outcomes)

	Readmission[a] (1)	Waiting time[b] (2)	Length of stay[c] (3)
Panel A. Hip replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	-0.003	-75.390***	-0.510*
	(0.007)	(11.931)	(0.300)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.077	183.282	7.935
N	29,145	22,885	29,145
Panel B. Knee replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.001	-94.949***	-0.528*
	(0.009)	(17.997)	(0.276)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.102	229.721	7.476
N	36,538	27,845	36,538
Panel C. All orthopedic surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.002	-17.366***	-0.173***
	(0.003)	(6.664)	(0.052)
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.059	148.135	2.174
N	438,749	293,532	438,749
Surgery type / Diagnosis code FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 281$). Excludes patients who lived in the three Western hospital districts that some degree of patient choice.

[a] Equals one if the patient had an emergency readmission (to any hospital) within 30 days after discharge.

[b] Number of days between the specialist's final treatment decision and the surgery. Some of the values are missing, which results in smaller N compared to other columns (see online Appendix Section A1.7 for more details).

[c] Number of days between admission and discharge.

Table A.18: Robustness Check: Excluding Hospitals That Used a Joint Hospital ID (Choice Outcomes)

	Teaching hospital[a]	Distance (km)[b]	Nearest hospital[c]	Different hospital district[d]
	(1)	(2)	(3)	(4)
Panel A. Hip replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.058***	8.830***	-0.060***	0.055***
	(0.017)	(3.377)	(0.021)	(0.017)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.315	31.952	0.805	0.041
N	21,785	21,785	21,785	21,785
Panel B. Knee replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.042***	4.357**	0.005	0.035***
	(0.011)	(1.966)	(0.018)	(0.010)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.290	30.239	0.812	0.038
N	28,093	28,093	28,093	28,093
Panel C. All orthopedic surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.032***	2.109**	-0.012	0.011**
	(0.007)	(1.033)	(0.010)	(0.005)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.220	27.563	0.788	0.042
N	322,286	322,286	322,286	322,286
Surgery type / Diagnosis code FEs	✓	✓	✓	✓
Municipal FEs	✓	✓	✓	✓
Age & sex	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 301 - -317$ depending on the sample).

[a] Equals one if the patient received care in a teaching (university) hospital.

[b] Distance from patient's residence to the hospital in kilometers.

[c] Equals one if the patient received care in the nearest hospital.

[d] Equals one if the patient received care outside their hospital district of residence.

Table A.19: Robustness Check: Excluding Hospitals That Used a Joint Hospital ID (Hospital Performance Outcomes)

	Readmission[a] (1)	Waiting time[b] (2)	Length of stay[c] (3)
Panel A. Hip replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.010	-88.543***	-0.187
	(0.007)	(12.427)	(0.310)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.082	183.094	8.062
N	21,785	15,828	21,785
Panel B. Knee replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.010	-118.658***	-0.377
	(0.009)	(17.846)	(0.324)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.107	226.885	7.709
N	28,093	19,991	28,093
Panel C. All orthopedic surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.003	-32.246***	-0.102*
	(0.003)	(7.254)	(0.052)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.060	140.687	2.111
N	322,286	207,051	322,286
Surgery type / Diagnosis code FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 301 - -317$ depending on the sample).
[a] Equals one if the patient had an emergency readmission (to any hospital) within 30 days after discharge.
[b] Number of days between the specialist's final treatment decision and the surgery. Some of the values are missing, which results in smaller N compared to other columns (see online Appendix Section A1.7 for more details).
[c] Number of days between admission and discharge.

Table A.20: Robustness Check: Tests Regarding Waiting Times

	Baseline	Districts for which <30% missing[a]	Surgeries for which <30% missing[b]	Excluding hospital which did not report waiting times in 2008–2009[c]	2006 vs 2010[d]
	(1)	(2)	(3)	(4)	(5)
Panel A. Hip replacement surgeries					
$\text{Treated}_h \times \text{Post}_t$	-77.914***	-79.024***		-84.936***	-97.505***
	(11.845)	(12.039)		(12.587)	(17.055)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	181.991	186.917		181.744	157.403
N	24,222	20,086		23,364	8,119
Panel B. Knee replacement surgeries					
$\text{Treated}_h \times \text{Post}_t$	-100.740***	-104.392***		-114.063***	-113.779***
	(17.836)	(18.310)		(18.116)	(20.204)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	225.373	229.811		226.159	192.898
N	29,759	24,372		28,704	10,263
Panel c. All orthopedic surgeries					
$\text{Treated}_h \times \text{Post}_t$	-20.723***	-15.076**	-41.037***	-23.994***	-38.758***
	(6.709)	(7.137)	(8.589)	(7.037)	(7.904)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	145.636	144.817	167.604	145.849	129.905
N	316,454	232,753	161,463	308,543	100,109
Surgery type / Diagnosis code FEs	✓	✓	✓	✓	✓
Municipal FEs	✓	✓	✓	✓	✓
Age & sex	✓	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 246$–319 depending on the column). Outcome in all columns is waiting time in days.

[a] Including only hospitals in districts in which less than 30 percent of the waiting time values were missing in 2004–2010.

[b] Including only those surgeries for which less than 30 percent of the waiting time values were missing in 2004–2010.

[c] Excluding one reform area hospital which did not report most of its waiting times in Q1/2008–Q4/2009.

[d] Including only the years 2006 and 2010, when the share of missing waiting time values was generally low across all regions and hospitals. This analysis also mitigates the potential bias which may arise when hospitals' shares of missing waiting time values fluctuate over time.

Table A.21: The Effect of the Reform on Hospital-level Waiting Times After Adjusting to the Introduction of Maximum Waiting Time Target

	Whole time period		After target	
	DiD (1)	Heterogeneity (2)	DiD (3)	Heterogeneity (4)
Panel A. Hip replacement surgeries				
$\text{Treated}_h \times \text{Post}_t$	-83.428**	-68.589	-36.807**	-23.241*
	(36.518)	(46.117)	(15.764)	(12.457)
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		-47.303		-52.728**
		(46.136)		(24.547)
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	172.949	172.949	139.608	139.608
N	791	791	463	463
Panel B. Knee replacement surgeries				
$\text{Treated}_h \times \text{Post}_t$	-113.316**	-101.002*	-43.672***	-37.651*
	(42.295)	(57.356)	(13.599)	(18.469)
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		-38.022		-26.752
		(59.088)		(26.842)
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	208.711	208.711	158.762	158.762
N	812	812	470	470
Panel C. All orthopedic surgeries				
$\text{Treated}_h \times \text{Post}_t$	-14.984*	-10.619	-11.811	-7.303
	(8.308)	(11.679)	(8.534)	(12.440)
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		-33.429**		-31.311**
		(12.385)		(14.233)
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	125.849	125.849	107.167	107.167
N	1,114	1,114	661	661
Hospital and time FEs	✓	✓	✓	✓
Age & sex mix[a]	✓	✓	✓	✓
Surgery types[b]	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Outcome in all columns is mean waiting time in hospital h in quarter t. Columns 1–2 include hospital-quarter-level observations from Q1/2004 to Q4/2010 and Columns 3–4 observations from Q4/2006 to Q4/2010. Standard errors clustered at the hospital district level ($N = 20$).
[a] Shares of females, 18–29, 30–39, 40–49, 50–59, and 70–74-year-old patients of all of the hospital's patients. Baseline = share of 60–69-year-old male patients.
[b] Shares of different procedure codes among the hospital's patients.

Table A.22: The Effect of the Reform on Patient Characteristics at the Hospital Level

	Mean age[a]		Female (%)[b]		Emergency admissions[c]	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Hip replacement surgeries						
$\text{Treated}_h \times \text{Post}_t$	1.237*	1.295	-0.006	-0.012	0.106**	0.088*
	(0.637)	(0.818)	(0.015)	(0.014)	(0.045)	(0.043)
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		-0.562		0.016		0.014
		(0.954)		(0.018)		(0.060)
$\text{mean}(y_{ht} \vert \text{Post}_t = 0)$	63.400	63.400	0.522	0.522	0.469	0.469
N	1,009	1,009	1,009	1,009	1,009	1,009
Panel B. Knee replacement surgeries						
$\text{Treated}_h \times \text{Post}_t$	-0.626*	-0.714*	-0.011	-0.014	-0.005	-0.022
	(0.331)	(0.386)	(0.011)	(0.011)	(0.052)	(0.051)
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		0.490		0.015		0.072*
		(0.488)		(0.009)		(0.041)
$\text{mean}(y_{ht} \vert \text{Post}_t = 0)$	64.631	64.631	0.662	0.662	0.456	0.456
N	1,070	1,070	1,070	1,070	1,070	1,070
Panel C. All orthopedic surgeries						
$\text{Treated}_h \times \text{Post}_t$	-0.522	-0.574	0.004	0.002	0.040	0.034
	(0.393)	(0.439)	(0.008)	(0.008)	(0.049)	(0.049)
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		0.528		0.027***		0.065
		(0.433)		(0.009)		(0.050)
$\text{mean}(y_{ht} \vert \text{Post}_t = 0)$	52.174	52.174	0.524	0.524	0.530	0.530
N	1,295	1,295	1,295	1,295	1,295	1,295
Surgery/diagnosis FEs	✓	✓	✓	✓	✓	✓
Hospital FEs	✓	✓	✓	✓	✓	✓

Notes: Estimated using hospital-quarter-level data in Q1/2004–Q4/2010. t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors clustered at the hospital district level ($N = 20$).

[a] Mean age of patients at the time of admission (18–74).

[b] Share of females of all patients.

[c] Mean number of emergency admissions hospital's patients had within 1 year before their surgery.

Table A.23: The Effects on Hospitals' Surgical Expenditure

	Total expenditure (millions of €)[a]		Expenditure per patient (€)[b]	
	DiD	Heterogeneity	DiD	Heterogeneity
	(1)	(2)	(3)	(4)
$\text{Treated}_h \times \text{Post}_t$	-0.144	-1.017	4.592	3.795
	(1.142)	(0.668)	(10.029)	(10.921)
$\text{Treated}_h \times \text{Post}_t \times \text{Teaching}_h$		11.059***		3.507
		(3.719)		(32.934)
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	24.419	24.419	411.885	411.885
N	300	300	300	300
Hospital and time FEs	✓	✓	✓	✓
Case mix index	✓	✓	✓	✓

Notes: Estimated using hospital-year-level data in 2004–2010. t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors clustered at the hospital district level ($N = 20$).
[a] Hospital's annual care-related expenditure in the surgical ward (millions of €, deflated using prices in 2000).
[b] Hospital's annual care-related expenditure in the surgical ward (€, deflated using prices in 2000) divided by DRG-weighted number of surgical patients.

Table A.24: Back-of-the-Envelope Calculation

		Hip replacement surgeries		Knee replacement surgeries		All orthopedic surgeries	
Variable	Description	Source	Value	Source	Value	Source	Value
Panel A. Benefits of the reform							
g	Health gain from surgery (2010 euros)[a]	Jansson and Granath (2011); Gyrd-Hansen (2003)	4,386.17	Jansson and Granath (2011); Gyrd-Hansen (2003)	3,112.76	Jansson and Granath (2011); Gyrd-Hansen (2003)	2,546.81
	Average age (years)	Table 2	62	Table 2	65	Table 2	52
T	Discounted life expectancy at ave. age[b]		14.75		13.56		19.76
$w_0 \times 365.25$	Waiting time (days), pre-reform	Table 5	140.99	Table 5	163.07	Table 5	118.06
$w_1 \times 365.25$	Waiting time (days), post-reform	Table 5	85.83	Table 5	104.69	Table 5	91.54
	Average volume per hospital/qrt	Table 6	29	Table 6	34	Table 6	361
y_0	Total volume per year, pre-reform[c]	Table 6	1,752	Table 6	2,015	Table 6	21,675
y_1	Total volume per year, post-reform[d]	Table 6	2,094	Table 6	2,781	Table 6	23,302
$g \times (T - w_0) \times y_0$	Total health benefit per year, pre-reform		110,377,705.72		82,271,856.19		1,073,161,244.91
$g \times (T - w_1) \times y_1$	Total health benefit per year, post-reform		133,343,458.88		114,896,917.92		1,158,060,556.36
	Total benefit change (post − pre)		22,965,753.16		32,625,061.73		84,899,311.44
Panel B. Costs of the reform (2010 euros)							
c_0	Average cost, pre-reform	Remes et al. (2007)	8,376.50	Remes et al. (2007)	8,503.00	Osnes-Ringen et al. (2011)	6,645.23[e]
	Costs related to LOS per patient[f]						
c_0^{LOS}	- Pre-reform	Remes et al. (2007)	4272.02	Remes et al. (2007)	4336.53	Osnes-Ringen et al. (2011), Table 4	3,389.07
c_1^{LOS}	- Post-reform	Remes et al. (2007)	4019.63	Remes et al. (2007)	4051.00	Osnes-Ringen et al. (2011), Table 4	3,141.61
c_1	Average cost, post-reform[g]		8,124.11		8,217.47		6,397.77
$c_1 \times y_1 - c_0 \times y_0$			2,339,808.26		5,713,388.39		5,050,000.91
Panel C. Net effects of the reform							
	Total benefits − total costs		21,802,330.07		25,928,186.40		79,849,310.53

Notes: [a] To measure h, we first use the estimates of the quality of life (EQ-5D index) gains from hip and knee replacements and orthopedic surgeries (index score improvements of 0.31, 0.22, and 0.18, respectively; Jansson and Granath (2011)). Then we measure their monetary value based on the willingness-to-pay estimates by Gyrd-Hansen (2003), which we express in 2010 euros.

[b] The discounted life expectancy is calculated at the average age using a discount factor of 3 percent and expected death at 80 years of age.

[c] y_0 = the average number of patients per hospital and quarter in the pre-reform period × the number of reform area hospitals ($N = 15$) × the number of quarters ($N = 4$).

[d] y_1 = the average number of patients per hospital and quarter in the post-reform period × the number of reform area hospitals ($N = 15$) × the number of quarters ($N = 4$). The post-reform average number of patients is calculated by adding the coefficient estimate of $\text{Treated}_h \times \text{Post}_t$ for the volume (Column 1 of Table 6) to the average pre-reform volume.

[e] c_0 is approximated by taking a weighted average of the costs for different orthopedic surgeries in Osnes-Ringen et al. (2011), using their observation shares as weights and expressed in 2010 euros.

[f] Following Osnes-Ringen et al. (2011), we assume that the costs related to length of stay, c_0^{LOS}, account for 51 percent of the costs in the pre-reform period. In the post-reform period, we take into account shorter length of stays in that period (a decrease of 0.161 days, Table 4). Then the post-reform costs for orthopedic surgeries related to LOS could be $c_1^{LOS} = c_0^{LOS}/\textit{mean LOS} * (\textit{mean LOS} - 0.161)$, where *mean LOS* is the mean length of stay in the pre-reform period (2.205 days).

[g] $c_1 = c_0 + (c_1^{LOS} - c_0^{LOS})$.

Table A.25: Differential Change in Quality of Care, Waiting Time, and Length of Stay by Market Structure After the Choice Reform

	Readmission[a]		Waiting time[b]		Length of stay[c]	
	Reform area (1)	Full sample (2)	Reform area (3)	Full sample (4)	Reform area (5)	Full sample (6)
Panel A. Hip replacement surgeries						
$\text{Treated}_i \times \text{Post}_t$		0.005		-139.018**		3.675***
		[-0.049, 0.067]		[-218.266, -43.037]		[2.027, 5.216]
$\widehat{\text{HHI}}_h \times \text{Post}_t$	-0.005	0.025	51.364	-32.472	-6.462***	-0.598
	[-0.072, 0.054]	[-0.032, 0.080]	[-100.053, 190.093]	[-70.632, 5.504]	[-8.750, -3.905]	[-1.922, 0.733]
$\widehat{\text{HHI}}_h \times \text{Post}_t \times \text{Treated}_i$		-0.029		83.835		-5.864***
		[-0.114, 0.051]		[-73.626, 229.434]		[-8.383, -3.224]
$\text{mean}(y_{imht} \| \text{Post}_t = 0)$	0.070	0.080	213.560	181.959	7.658	8.091
N	6,759	31,404	6,149	24,224	6,759	31,404
Panel B. Knee replacement surgeries						
$\text{Treated}_i \times \text{Post}_t$		0.003		-215.668**		4.196***
		[-0.060, 0.073]		[-323.042, -77.467]		[2.558, 5.682]
$\widehat{\text{HHI}}_h \times \text{Post}_t$	0.034	0.039	114.526	-102.672**	-5.846***	1.289*
	[-0.056, 0.127]	[-0.018, 0.095]	[-66.117, 278.178]	[-168.059, -38.301]	[-8.256, -3.175]	[0.089, 2.521]
$\widehat{\text{HHI}}_h \times \text{Post}_t \times \text{Treated}_i$		-0.005		217.198*		-7.135***
		[-0.110, 0.102]		[34.143, 387.838]		[-9.815, -4.281]
$\text{mean}(y_{imht} \| \text{Post}_t = 0)$	0.103	0.102	264.883	225.345	7.364	7.609
N	8,202	39,427	7,290	29,760	8,202	39,427
Panel C. All orthopedic surgeries						
$\text{Treated}_i \times \text{Post}_t$		0.041***		-58.823*		0.865***
		[0.017, 0.062]		[-113.712, -1.373]		[0.523, 1.228]
$\widehat{\text{HHI}}_h \times \text{Post}_t$	-0.020	0.035***	-11.237	-82.182***	-1.197***	0.110
	[-0.047, 0.009]	[0.017, 0.055]	[-75.074, 49.817]	[-117.893, -46.198]	[-1.708, -0.751]	[-0.161, 0.391]
$\widehat{\text{HHI}}_h \times \text{Post}_t \times \text{Treated}_i$		-0.055**		70.946*		-1.308***
		[-0.088, -0.019]		[0.971, 140.347]		[-1.899, -0.776]
$\text{mean}(y_{imht} \| \text{Post}_t = 0)$	0.049	0.060	152.738	145.636	2.268	2.205
N	87,417	473,617	60,460	316,454	87,417	473,617
Surgery type FEs	✓	✓	✓	✓	✓	✓
Municipal FEs	✓	✓	✓	✓	✓	✓
Age & sex	✓	✓	✓	✓	✓	✓
Control function residuals	✓	✓	✓	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Columns (1), (3) and (5) estimated using equation (4) and data from reform area only ($\text{Treated}_i = 1$). Columns (2), (4) and (6) estimated using equation (5) and data from the whole country. Standard errors clustered at the level of the patient's home municipality ($N = 67$). HHI measured on a 0–1 scale, where higher value indicates more market concentration. See online Appendix A2.2 for further details on the control function approach.

[a] Equals one if the patient had an emergency readmission (to any hospital) within 30 days after discharge.

[b] Number of days between the specialist's final treatment decision and the surgery. Some of the values are missing, which results in smaller N compared to other columns (see online Appendix Section A1.7 for more details).

[c] Number of days between admission and discharge.

Table A.26: Robustness Check: Effect of the Reform when Controlling for First-Stage Residuals from a Hospital Choice Model (Hospital Performance Outcomes)

	Readmission[a] (1)	Waiting time[b] (2)	Length of stay[c] (3)
Panel A. Hip replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	-0.001	-75.443***	-0.512*
	(0.007)	(12.160)	(0.301)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.080	181.991	8.091
N	31,404	24,222	31,404
Panel B. Knee replacement surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.002	-98.435***	-0.556*
	(0.008)	(17.002)	(0.291)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.102	225.373	7.609
N	39,427	29,759	39,427
Panel C. All orthopedic surgeries			
$\text{Treated}_i \times \text{Post}_t$	0.003	-18.680***	-0.145***
	(0.003)	(6.856)	(0.052)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.060	145.636	2.205
N	473,617	316,454	473,617
Surgery type FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓
Control function residuals	✓	✓	✓

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (1). Standard errors clustered at the level of the patient's home municipality ($N = 319$). See online Appendix A2.2 for further details on the control function approach.

[a] Equals one if the patient had an emergency readmission (to any hospital) within 30 days after discharge.

[b] Number of days between the specialist's final treatment decision and the surgery. Some of the values are missing, which results in smaller N compared to other columns (see online Appendix Section A1.7 for more details).

[c] Number of days between admission and discharge.

Table A.27: Differential Change by Market Structure Without Controlling for First-Stage Residuals from a Hospital Choice Model

	Readmission[a] (1)	Waiting time[b] (2)	Length of stay[c] (3)
Panel A. Hip replacement surgeries			
$\widehat{\text{HHI}}_h \times \text{Post}_t$	-0.005	111.933	-5.993***
	[-0.060, 0.042]	[-49.076, 258.576]	[-8.112, -3.675]
mean(y_{imht}|$\text{Post}_t = 0$)	0.070	213.560	7.658
N	6,759	6,149	6,759
Panel B. Knee replacement surgeries			
$\widehat{\text{HHI}}_h \times \text{Post}_t$	-0.015	215.717**	-5.878***
	[-0.100, 0.069]	[45.290, 378.912]	[-8.199, -3.193]
mean(y_{imht}|$\text{Post}_t = 0$)	0.103	264.883	7.364
N	8,202	7,290	8,202
Panel C. All orthopedic surgeries			
$\widehat{\text{HHI}}_h \times \text{Post}_t$	-0.029*	10.030	-1.204***
	[-0.052, -0.001]	[-50.913, 69.490]	[-1.726, -0.765]
mean(y_{imht}|$\text{Post}_t = 0$)	0.049	152.738	2.268
N	87,417	60,460	87,417
Surgery type FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓
Control function residuals			

Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated using equation (4). Standard errors clustered at the level of the patient's home municipality ($N = 67$). HHI measured on a 0–1 scale, where higher value indicates more market concentration. See online Appendix A2.2 for further details on the control function approach.

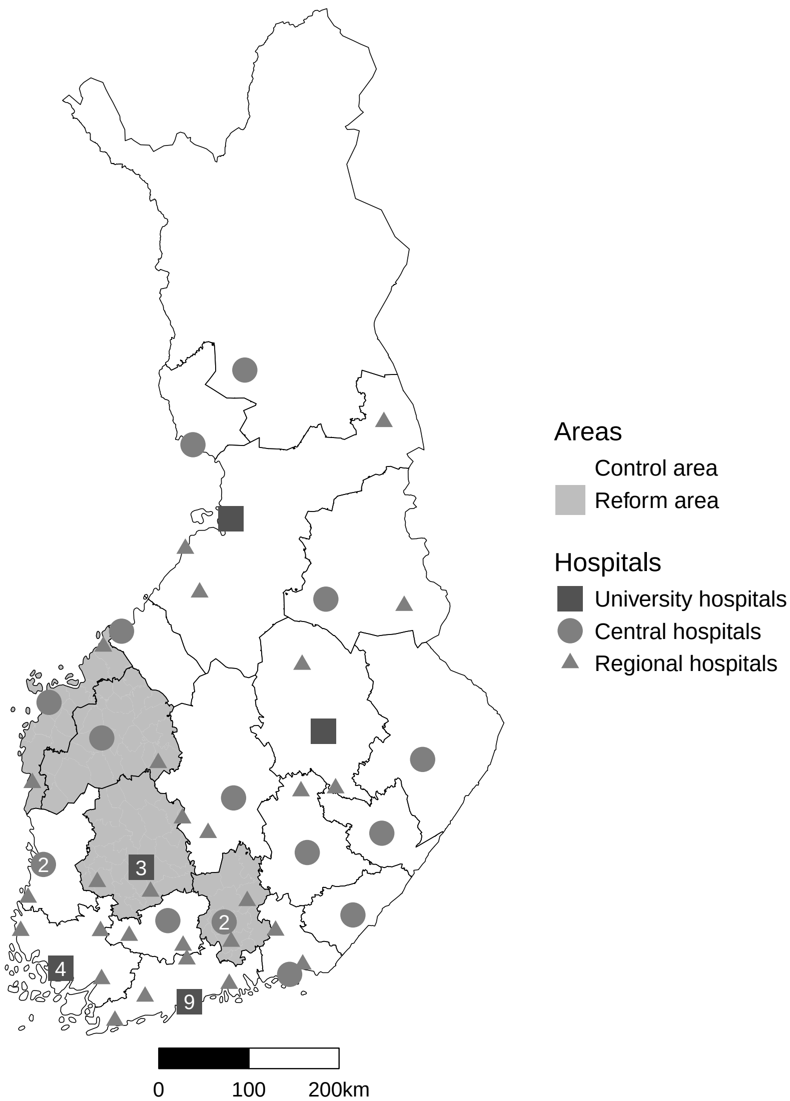
[a] Equals one if the patient had an emergency readmission (to any hospital) within 30 days after discharge.

[b] Number of days between the specialist's final treatment decision and the surgery. Some of the values are missing, which results in smaller N compared to other columns (see online Appendix Section A1.7 for more details).

[c] Number of days between admission and discharge.

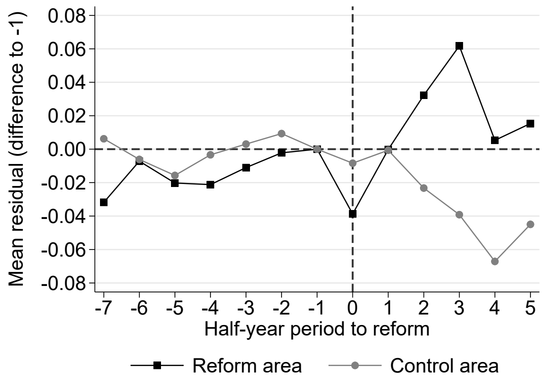
A4. Figures

Figure A.1: Teaching, Central, and Regional Hospital Locations

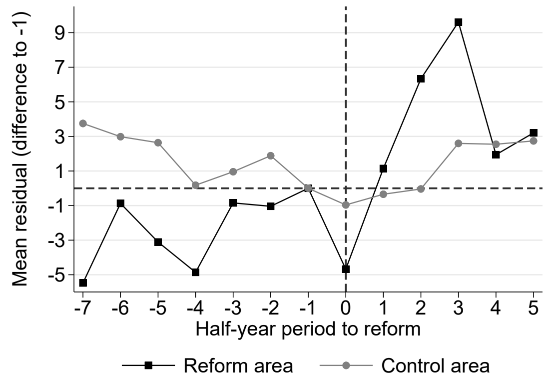


Notes: Borders indicate hospital districts in 2007 and the shaded area constitutes the 2007 reform area. Squares are teaching hospitals, dots central hospitals, and triangles regional hospitals. The squares and circles with a number mark the regions which had multiple (2–9) hospitals, all of which had one or more regional hospitals in addition to one central or university hospital. In total, there were $N = 15$ hospitals in the reform area and $N = 48$ hospitals in the control area. The figure includes all public hospitals which performed planned orthopedic surgeries, although some of them did not perform hip and/or knee replacement surgeries.

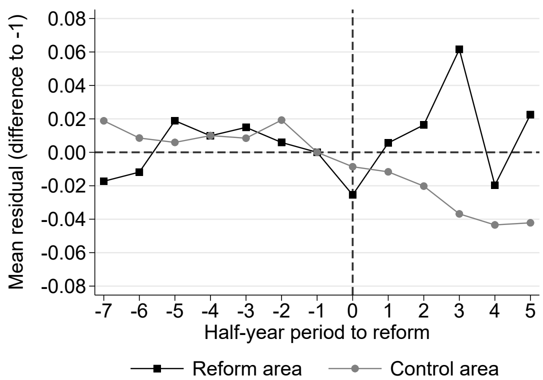
Figure A.2: Mean Outcome Trends: Probability of Choosing a Teaching Hospital and Distance to the Hospital



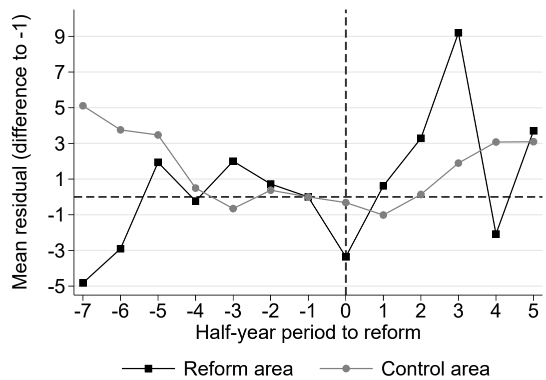
Panel A. Hip replacements: teaching hospital



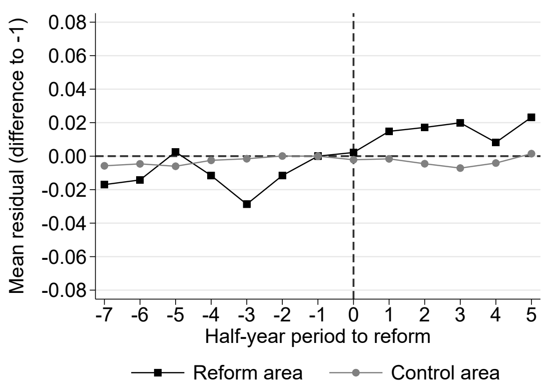
Panel B. Hip replacements: distance



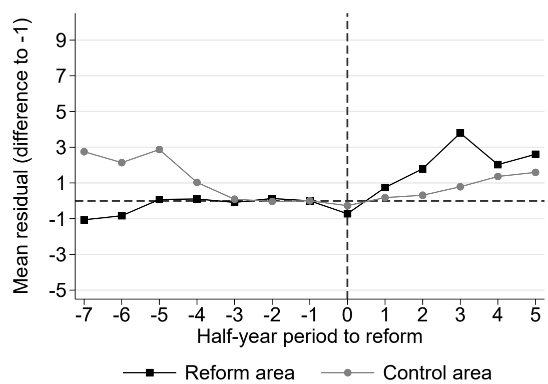
Panel C. Knee replacements: teaching hospital



Panel D. Knee replacements: distance



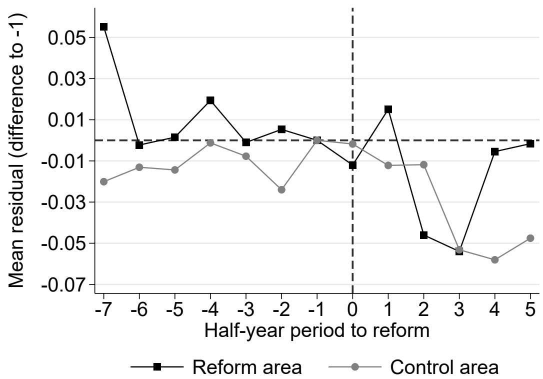
Panel E. All orthopedic surgeries: teaching hospital



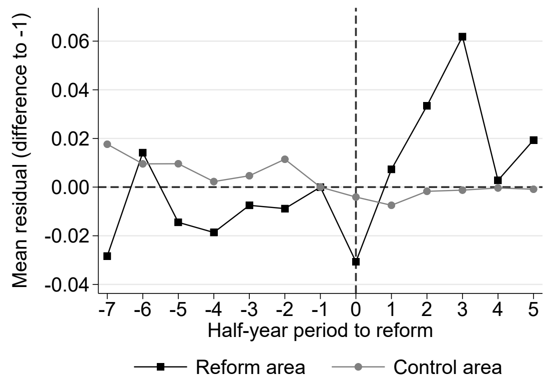
Panel F. All orthopedic surgeries: distance

Notes: The figure plots the evolution of residualized outcome means over time relative to the period of reform adoption in the reform and control areas. The area-level means of the outcomes' residuals result from estimations in which the outcome is regressed by only fixed effects for the patient's municipality of residence. We normalized the outcome in "one period before policy adoption" ($l = -1$) to zero.

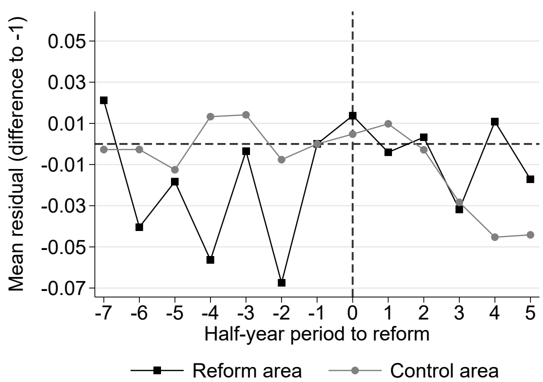
Figure A.3: Mean Outcome Trends: Probability of Choosing the Nearest Hospital or Hospital From a Different Hospital District



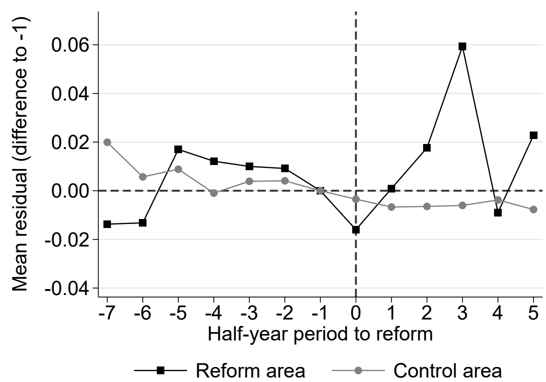
Panel A. Hip replacements: nearest hospital



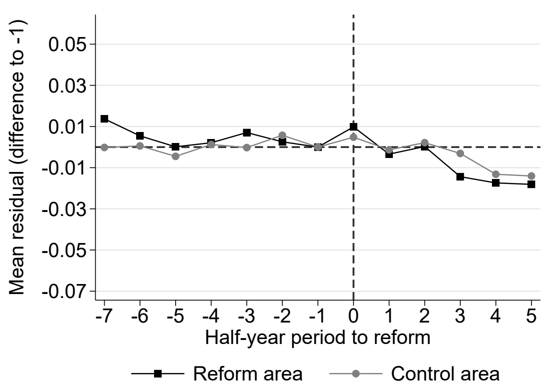
Panel B. Hip replacements: different hospital district



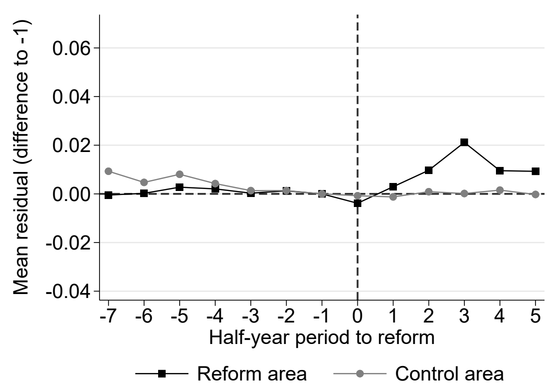
Panel C. Knee replacements: nearest hospital



Panel D. Knee replacements: different hospital district



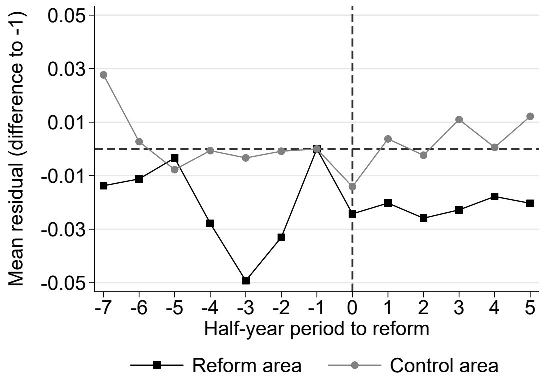
Panel E. All orthopedic surgeries: nearest hospital



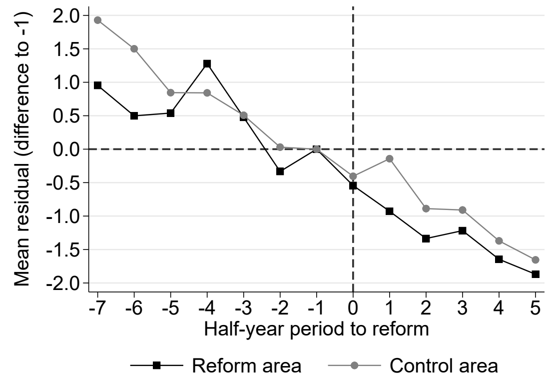
Panel F. All orthopedic surgeries: different hospital district

Notes: The figure plots the evolution of residualized outcome means over time relative to the period of reform adoption in the reform and control areas. The area-level means of the outcomes' residuals result from estimations in which the outcome is regressed by only fixed effects for the patient's municipality of residence. We normalized the outcome in "one period before policy adoption" ($l = -1$) to zero.

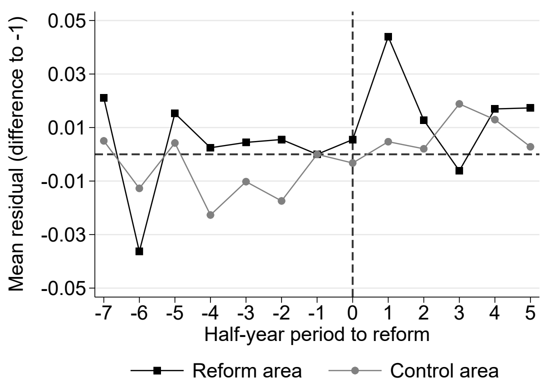
Figure A.4: Mean Outcome Trends: Probability of Emergency Readmission and Length of Stay



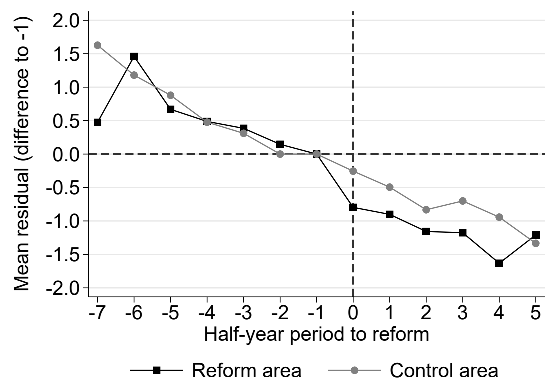
Panel A. Hip replacements: emergency readmissions



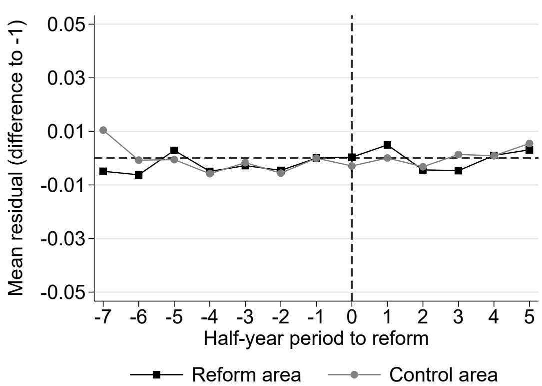
Panel B. Hip replacements: length of stay



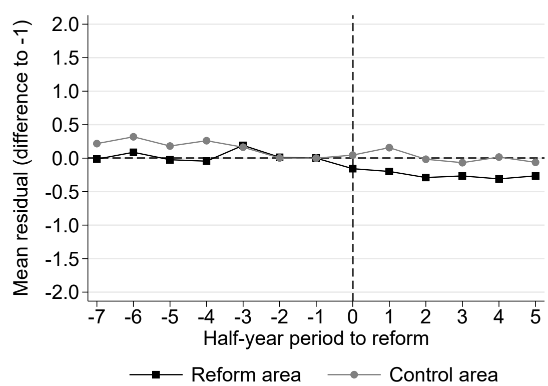
Panel C. Knee replacements: emergency readmissions



Panel D. Knee replacements: length of stay



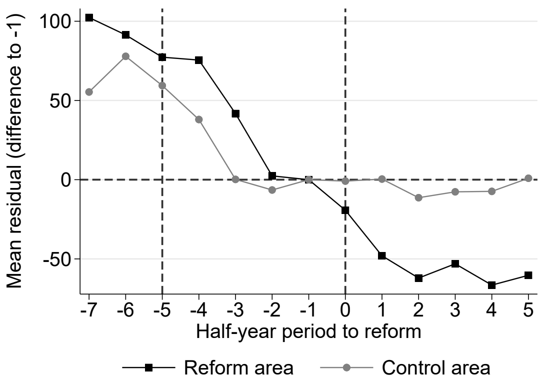
Panel E. All orthopedic surgeries: emergency readmissions



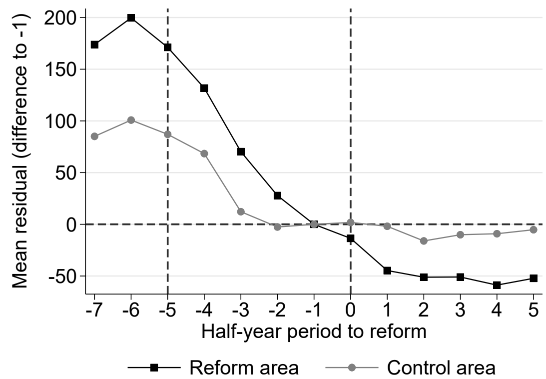
Panel F. All orthopedic surgeries: length of stay

Notes: The figure plots the evolution of residualized outcome means over time relative to the period of reform adoption in the reform and control areas. The area-level means of the outcomes' residuals result from estimations in which the outcome is regressed by only fixed effects for the patient's municipality of residence. We normalized the outcome in "one period before reform adoption" ($l = -1$) to zero.

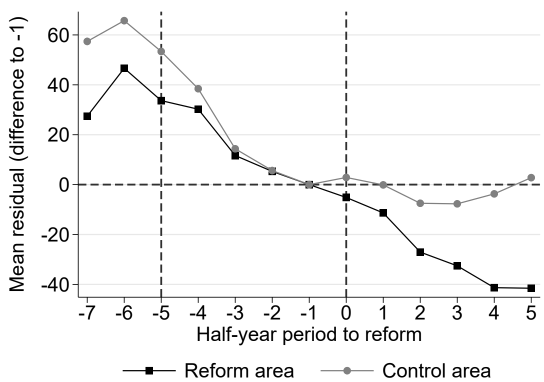
Figure A.5: Mean Outcome Trends: Waiting Time



Panel A. Hip replacements



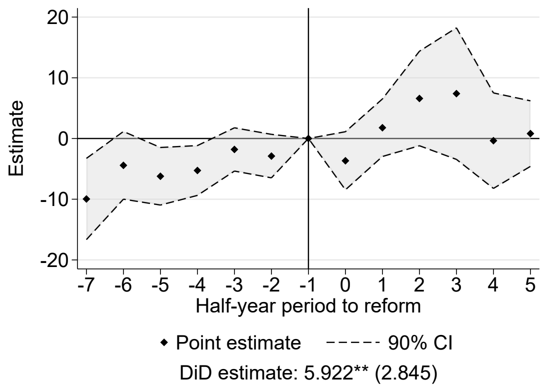
Panel B. Knee replacements



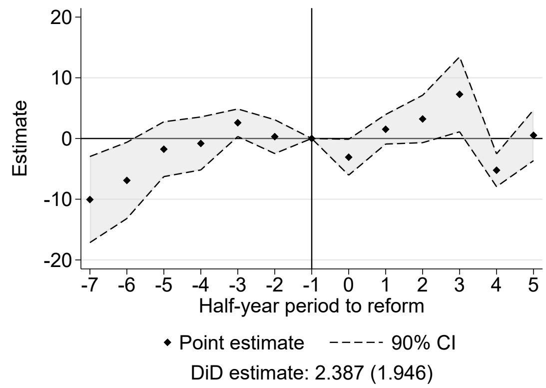
Panel C. All orthopedic surgeries

Notes: The figure plots the evolution of residualized outcome means over time relative to the period of reform adoption in the reform and control areas. The area-level means of the outcomes' residuals result from estimations in which the outcome is regressed by only fixed effects for the patient's municipality of residence. We normalized the outcome in "one period before reform adoption" ($l = -1$) to zero. The dashed vertical line at $l = -5$ indicates the introduction of a maximum waiting time target that guarantees medical treatment within 6 months.

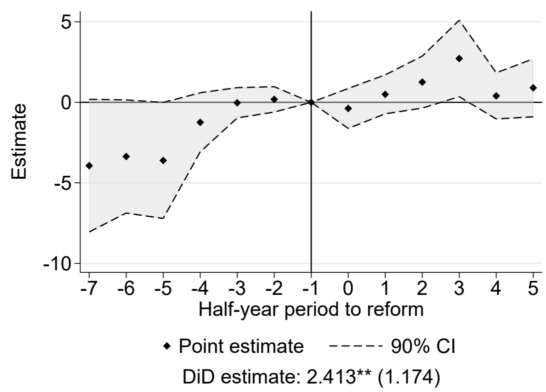
Figure A.6: The Effect of the Reform on Travel Distance of Surgical Patients



Panel A. Hip replacements



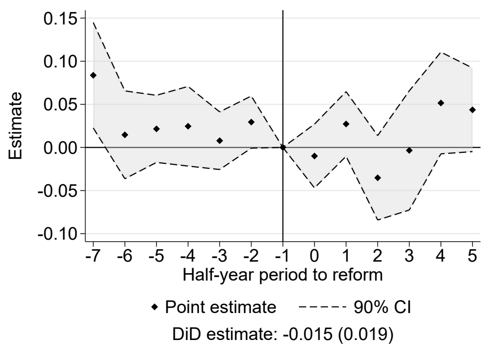
Panel B. Knee replacements



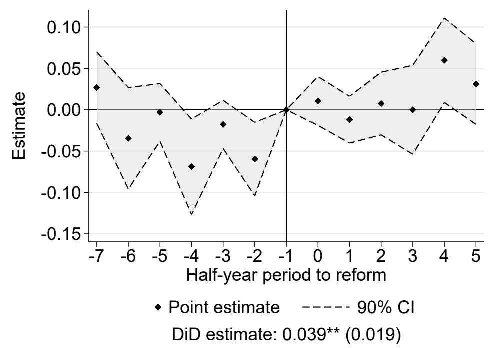
Panel C. All orthopedic surgeries

Notes: Estimated using equation (2). Includes the DiD estimates (in kilometers) shown in Column 2 of Table 3.

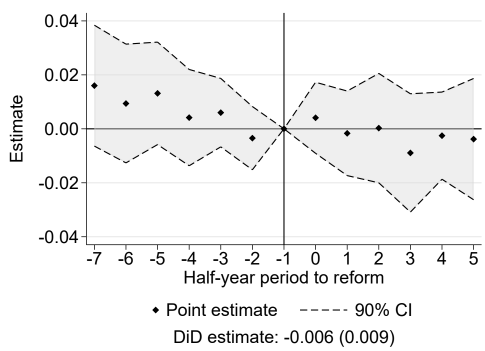
Figure A.7: The Effect of the Reform on Probability of Surgical Patients Receiving Care at Their Nearest Hospital



Panel A. Hip replacements



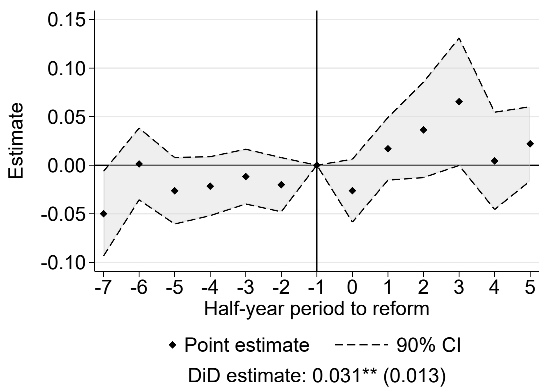
Panel B. Knee replacements



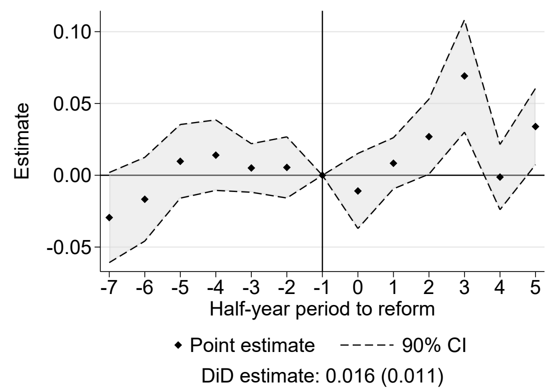
Panel C. All orthopedic surgeries

Notes: Estimated using equation (2). Includes the DiD estimates (in percentage points) shown in Column 3 of Table 3.

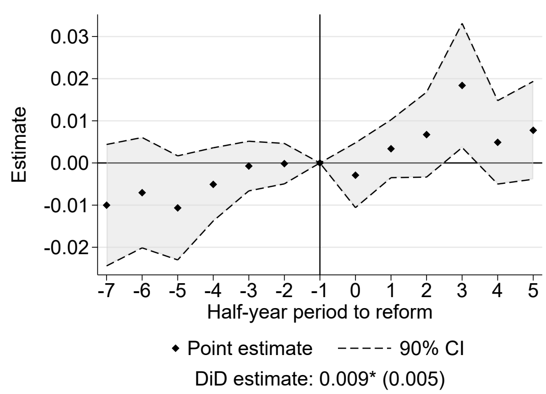
Figure A.8: The Effect of the Reform on Probability of Surgical Patients Receiving Care Outside Their Own Hospital District



Panel A. Hip replacements



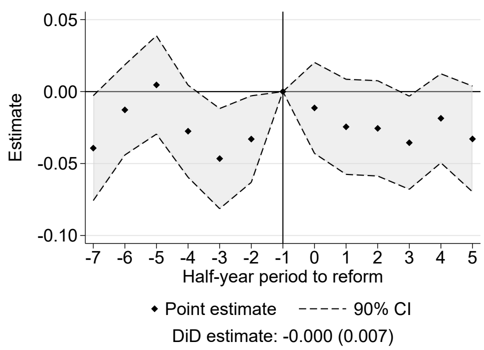
Panel B. Knee replacements



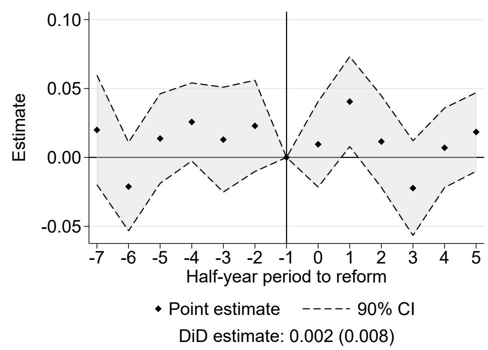
Panel C. All orthopedic surgeries

Notes: Estimated using equation (2). Includes the DiD estimates (in percentage points) shown in Column 4 of Table 3.

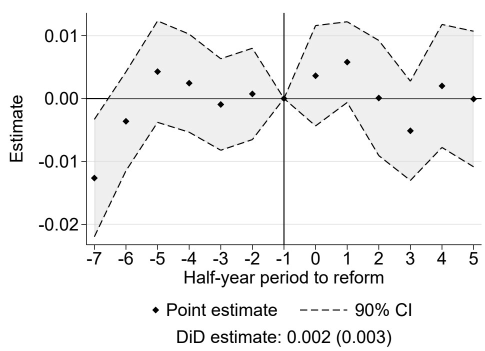
Figure A.9: The Effect of the Reform on Emergency Readmissions of Surgical Patients



Panel A. Hip replacements



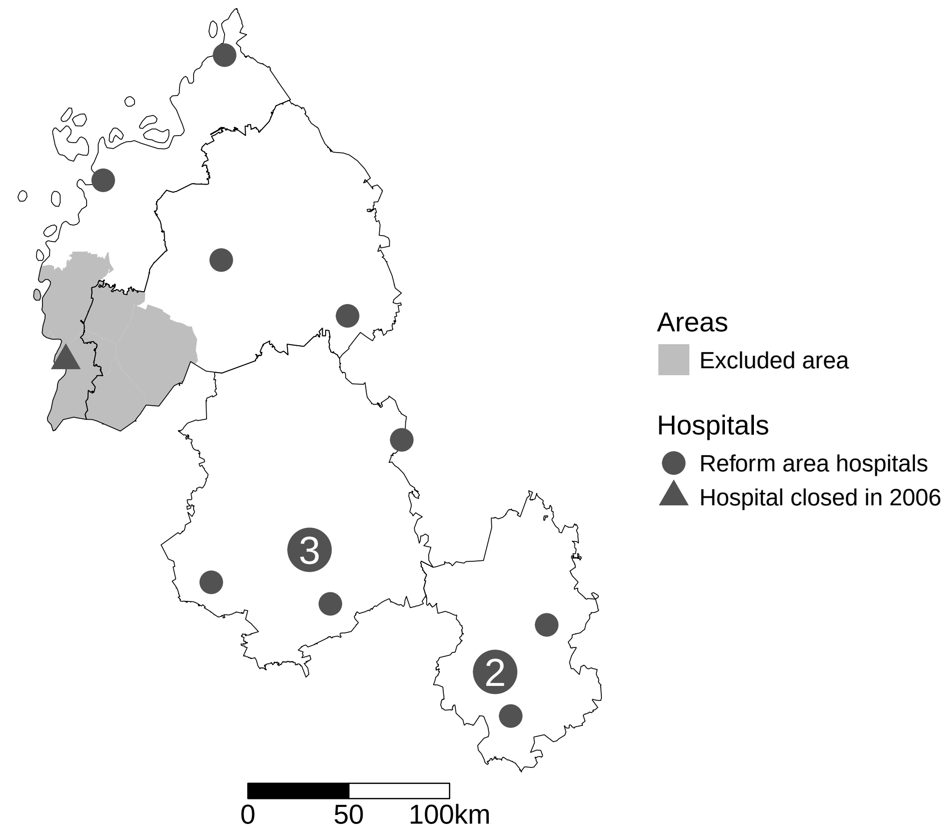
Panel B. Knee replacements



Panel C. All orthopedic surgeries

Notes: Estimated using equation (2). Includes the DiD estimates (in percentage points) shown in Column 1 of Table 4.

Figure A.10: Areas Excluded for Being Affected by the Closure of a Reform Area Hospital



Notes: The figure plots the map of the reform area in 2007 with hospital district borders. Dots mark reform area hospitals and triangle the hospital which closed down in the pre-reform period. Large circles mark regions with multiple (2–3) hospitals. Patients living in the shaded area around the closed hospital were excluded from the estimation samples.