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 their market dominance. 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, 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, BU-Harvard-MIT Health Economics Seminar, Bank of Finland Research Seminar, CEPR Health Economics 2023, EARIE 2023, EuHEA 2024, FINHESS, IHEA 2023, IIOC 2024, Midwest Applied Microeconomics Conference 2025, 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: October 13, 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 typically 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 more restricted choice as a control group.

Our DiD design with a binary treatment based on the regional patient choice reform provides a direct policy evaluation of how introducing choice and non-price competition as the market mechanism affects patient allocation 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 heterogeneous effects across different types of producers in the entire reform area. Thus, we differ from previous research on patient choice that has estimated the effect of *market structure* on hospital performance 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). No previous studies have examined patient choice policies introduced in some areas of a country while leaving other areas unaffected, that is, estimated the average treatment effect of a patient choice reform, as such reforms have been implemented simultaneously at the national level (e.g., in the UK and Norway).

We find that the regional patient choice reform affected patients' hospital choices and allocation for planned orthopedic surgery. 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 led to an increase in teaching hospitals' patient volumes and promoted concentration, as measured by the Herfindahl-Hirschman Index, 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 27 days or 22 percent and more patients received care in all the hospitals exposed to the reform on average.[1] 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 approximately $1,600$ patients or 7 percent 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

[1]The volume increase was, however, much larger in teaching hospitals that have both larger capacity and better resources.

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 how hospital performance varies with market structure after the simultaneous adoption of nationwide choice reforms (Cooper et al., 2011; Gaynor, Moreno-Serra and Propper, 2013; Brekke et al., 2021; Moscelli, Gravelle and Siciliani, 2021, 2023).[2] In contrast, we perform a direct policy evaluation and estimate the market-wide average and aggregate effects of enhanced versus more restricted patient choice using a binary DiD approach. We highlight that the DiD approach used in the previous literature estimates a different parameter of interest than the average effect of the reform 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, volumes and market concentration as outcomes.[3] This differs from the previous studies that have used pre-reform market concentration as a right-hand side variable, rather than as an outcome. While previous studies examine the effects of market concentration after choice reforms, they do not address whether greater patient choice itself promotes market dominance by large hospitals, which is what we find.[4]

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 reallo-

[2]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).

[3]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.

[4]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.

cate 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.

Finally, our study links the empirical literature analyzing the impacts of market-based choice reforms with the 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 at a greater rate 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 tax-financed, decentralized health care system, where all residents are entitled to public health care services. Public primary care is organized and financed by the municipalities ($N = 326$ in 2010) for their residents 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.[5]

Prior to the reform, patient choice was highly restricted, and patients were referred to a 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 hospital within and across hospital districts in the reform area,[6] leaving patients and hospitals outside the reform area unaffected.[7]

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

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

[6] Small municipality-owned hospitals were not officially part of the reform. However, those hospitals were indirectly part of the reform as some of the patients were now able to be treated in a university hospital instead of these municipality-owned hospitals. Municipality-owned are included in our estimation samples, but we have confirmed that their exclusion would not alter the results.

[7] 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).

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.[8]

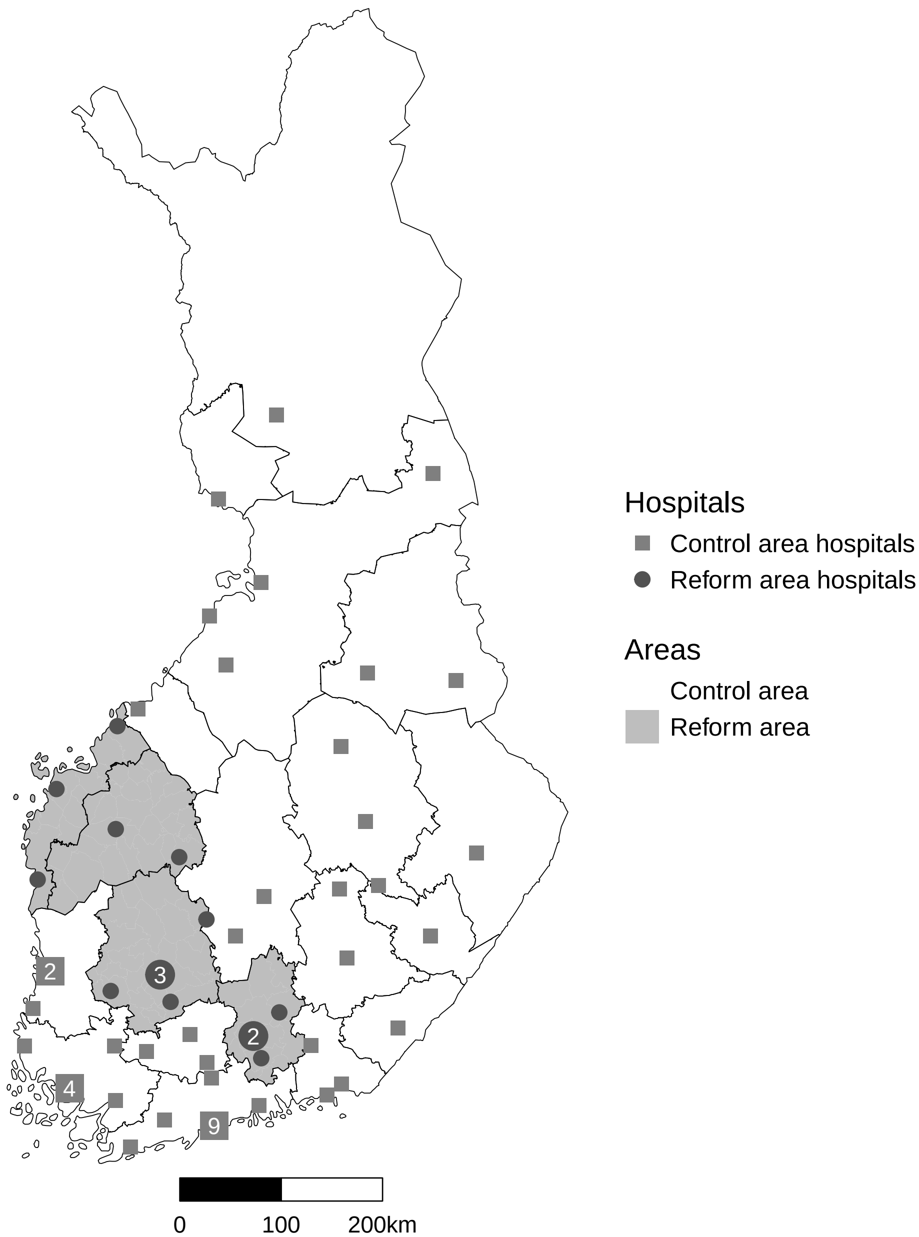
Finnish public hospitals are reimbursed for the services produced from their patients' municipalities of residence and the money follows the patients. Each hospital district sets the reimbursement rates of their own hospitals administratively and bases these on administered prospective payments using diagnosis-related groups (DRGs) (Kautiainen, Häkkinen and Lauharanta, 2011).[9]

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

[8]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.

[9]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.

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.

Decree 464/2008), which at the time corresponded to 4 percent of the monthly disposable cash income of a median household.[10] Hence we do not expect hospital co-payments to affect patients' hospital choices to a great extent. However, more relevant 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).[11] Because money follows patients through municipality reimbursements, hospitals' can increase their revenues by attracting more patients if they have capacity to treat them.[12]

The patient choice reform brought about a substantial shift in hospitals' ability to attract and compete for patients, with potential efficiency gains in health care production (Gaynor, Moreno-Serra and Propper, 2013; Longo et al., 2019). Given that administered co-payments are almost fixed, hospitals can attract more patients mainly by improving the quality of care and/or shortening waiting times. To reduce waiting times, hospitals can increase supply either by expanding resources or by making health care delivery more productive and efficient (OECD, 2020). In the latter case, elective care volume can increase without additional capacity for example, through shorter lengths of stay or more efficient use of existing resources.[13]

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

[10] €83.90/(€28,400/12 months) $\approx$ 0.04 (Statistics Finland, 2010).

[11] 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).

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

[13] One might think of hospital volume as exogenous and fixed, which is a reasonable assumption for emergency patients but not for elective care. In our setting and in theoretical literature (Brekke, Siciliani and Straume, 2008), hospital volume and waiting time are choice variables impacted by the introduction of competition.

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 from the Finnish Institute for Health and Welfare.[14] The data contain public hospital admissions and discharges in Finland in 2004–2010. We focus on patients aged 18–74 years at the time of hospital admission.[15] 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.

We use information on medical procedures produced by hospitals to create our main sample, which includes all planned orthopedic surgeries performed during our study period. These procedures cover a wide range of surgeries that differ in complexity, invasiveness, and recovery time.

[14]The research data was obtained from the Finnish Social and Health Data Permit Authority Findata with data permit Dnro THL/5839/14.06.00/2024.

[15]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).

We focus on planned orthopedic surgeries, as these are the procedures where the choice reform is expected to provide scope for patient choice and producer competition. Orthopedic surgeries are also commonly studied in the previous literature on hospital performance (Moscelli, Gravelle and Siciliani, 2021; Godøy et al., 2024). By including all planned orthopedic surgeries, we capture both minor procedures and more extensive operations, giving a comprehensive view of orthopedic care.[16] In addition, we study whether the effects of the reform are different for more invasive orthopedic surgeries, including hip and knee replacements (Section VI.C).

Our main sample of orthopedic surgeries covers 63 hospitals and 473,617 planned orthopedic surgery observations in total. Approximately one fifth of the patients resided in areas affected by the reform. All samples exclude a small portion (1–2 percent) of patients who obtained orthopedic surgery across reform and control area borders. In addition, we exclude all patients who received care in or lived close to a regional hospital that closed during the pre-reform period.

Next, we discuss the construction of the main variables, while leaving the more detailed description of samples and variables to online Appendix Section A1. We conclude this section with descriptive statistics.

III.A. Measures of Hospital Choice and Patient Allocation

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

[16]The most common procedure in our sample is knee meniscus surgery (see Online Appendix Table A.6). Knee meniscus surgery requires only small cuts and is an example of less invasive orthopedic surgeries that often involves shorter hospital stays or even same-day discharge. In contrast, joint replacement surgeries such as hip and knee replacements typically require longer hospital stays and account for approximately 15 percent of the sample.

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.

A choice reform does not necessarily imply that patients will, on average, travel farther to receive care. While some patients may prefer hospitals located further or outside their district, others may continue to use the nearest facility, keeping their travel distances unchanged. In fact, greater choice could even reduce average travel if patients previously had to go to a designated district hospital that was farther away than an equally suitable hospital just across the district boundary. Thus, the reform may change travel distances in both directions, making the overall effect on mean distance theoretically ambiguous.

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 planned surgical patient underwent an emergency readmission within 30 days of discharge. Emergency readmissions are a commonly used measure of hospital quality (Varkevisser, van der Geest and Schut, 2012; Gupta, 2021; Moscelli, Gravelle and Siciliani, 2021). In our analyses for hip and knee replacements, we use more detailed measures of clinical quality such as revision surgeries and mechanical complications and infections in the prosthesis (Section V.B). In addition to the direct quality effects, we also examine the indirect effects of the choice reform on emergency care quality (Section V.B). Emergency care is relevant in this context because, although it was not targeted by the reform, changes in hospital incentives and resource allocation in planned care can also influence emergency care performance.

Waiting time. In addition to clinical quality, waiting time is another commonly studied measure of hospital performance (Propper, Burgess and Gossage, 2008; Moscelli, Gravelle and Siciliani, 2021, 2023). It was also an objective of the choice reform (Section II.B). We measure waiting time by the number of days that the patient spends on the waiting list for surgery.[17] Waiting times

[17]The patient is placed on the waiting list when a specialist has made the final decision on surgery. In most cases,

delay access to medical treatment and can reflect 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, comprehensive information on costs or on the physical and human resources used is not systematically collected in Finland. 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 fewer 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 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

waiting times are recorded comprehensively in the discharge data, though some hospitals record waiting times less comprehensively (online Appendix Section A1). Our results are robust for excluding these hospitals from our samples (Section V.C).

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 and 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 patient allocation outcomes for all planned orthopedic surgeries in the pre-reform period (Q1/2004–Q3/2007). We present the statistics separately for the reform and control areas. We find some differences in pre-reform means between treatment and control areas. For example, the probability of choosing a teaching hospital was lower in the reform than in the control area (23 versus 41 percent), and the distance traveled was on average shorter in the treatment area than in the control area (23 versus 27 kilometers). Similarly, hospital volumes were also smaller in the reform than in the control area. Moreover, the hospital-level means of actual HHI indicate a fairly high degree of market concentration at the level of 0.63–0.71, with variation across hospitals and/or over time in the pre-reform period (SD 0.13–0.14).

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

	Reform area			*Control area*		
	Mean	SD	N	Mean	SD	N
Variable						
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 a continuous variable and the other choice outcomes, including teaching hospital, nearest hospital, and different hospital district, are binary (0/1) variables. 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
Variable						
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 variables and the other variables, including emergency readmission within 30 days and female, are binary (0/1) variables. Some hospitals did not report waiting times in some quarters, which results in a smaller number of observations compared to the other columns (see online Appendix Section A1.4 for more details). We analyze the robustness of the results regarding missing values in waiting times and summarize and show robustness of the results regarding missing values in waiting times in Section V.C.

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 (5–6 percent), and length of stay (2 days). In contrast, waiting times were longer in the reform area than in the control area (153 versus 144 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 hospitals and small regional hospitals (1,281 versus 402 and 168 patients per quarter on average, respectively; online Appendix Table A.7).[18] 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.8 and A.9). This finding is consistent with the evidence for common medical and surgical conditions in other settings (Burke et al., 2019; Silber et al., 2020). The mean waiting times were, however, longer in teaching hospitals than in medium-sized non-teaching hospitals (167 versus 156 days) (online Appendix Table A.8).[19]

[18]Online Appendix Tables A.10 and A.11 further show that teaching hospitals were very similar in terms of their mean performance and patient composition, respectively, between reform and control areas.

[19]Regional hospitals had the shortest mean waiting times. This might reflect lower demand rather than greater supply by regional hospitals.

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.[20] 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. We also include fixed effects for the patient's municipality of residence, μ_m, to control for time-invariant differences between municipalities (and also between the reform and control areas), for example, in patients' location, preferences, or morbidity.[21]

$\mathbf{X}_{it}$ 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 resi-

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

[21] 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).

dence ($N = 319$) to account for the possibility of within-municipality correlation in unobservable variables among patients located in the same geographical area.[22]

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).[23]

Our first key identifying assumption is that the reform did not have 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. We have excluded these observations from our samples (Section III), but our results are robust to inclusion of them.

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 Appendix Figures A.2–A.3 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. 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

[22] It may seem natural to cluster at the treatment level; however, this is not feasible in our setting because we have only one treatment group (cluster).

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

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 all planned orthopedic surgeries in Table 3.[24]

We find that the reform led to the reallocation of patients towards large teaching hospitals (Column 1). The probability of choosing a teaching, instead of a non-teaching, hospital increased by 3 percentage points (7 percent in comparison to the pre-reform mean) for all orthopedic surgeries. The reform, which expanded patients' choice of hospitals both within and across hospital districts, also induced patients to travel longer distances for care (Column 2). On average, travel distance increased by 2.4 kilometers or 9 percent compared to the mean.

We further examine whether the reform affected the probability of receiving care at the nearest hospital. In principle, patients in the treatment area could now choose the nearest hospital located outside their own district, provided it was within the reform area. Although the point estimate is negative, we find no statistically significant effect on the probability of choosing the nearest hospital (Column 3).[25] The 95 percent confidence intervals allow us rule out effects larger than 2 percentage points (3 percent) on this outcome. By contrast, the reform increased the probability of patients seeking care beyond their own hospital district. The DiD point estimate shows 1 percentage point increase in choosing a hospital in a different hospital district (23 percent at the sample mean) (Column 4).

Figure 2 presents the event study estimates for all hospital choice and allocation outcomes by estimating specification (2). We find that the probability of choosing a teaching hospital, travel distance, and the probability of choosing a hospital from a different hospital district began to increase half a year after the implementation of the reform. The lag in the effects may result from waiting

[24]Section VI.C presents heterogeneity by procedure type, based on samples of planned hip and knee replacements.

[25]Both the p-values and sharpened q-values indicate statistically significance at 10 percent or less for all choice outcomes except for the probability of choosing the nearest hospital.

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)
Treated$_i$ × 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}
mean(y_{imht} | 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 from specification (1) using patient-level data. Standard errors clustered at the level of the patient's home municipality ($N = 319$). Sharpened q-values are calculated for the sample.

[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.

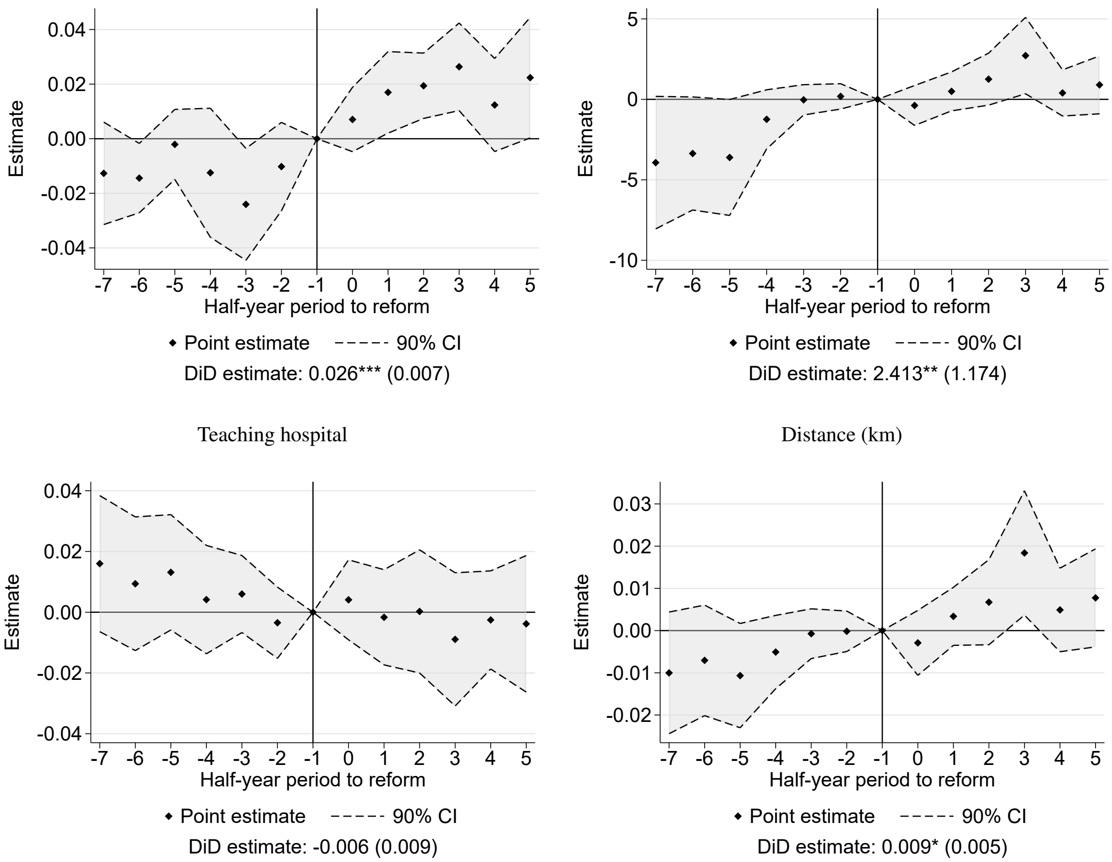
[d] Equals one if the patient received care outside their hospital district of residence.

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 most choice and allocation outcomes, providing supporting evidence for the credibility of our research design. There is, however, evidence of a pre-trend in distance traveled during earlier periods (5—7 half-years) before the reform. In contrast, the pre-trend coefficients in the periods closer to the reform are near zero and estimated with relatively high precision.[26]

In sum, our results suggest that planned surgery patients responded to the reform by traveling longer distances and across hospital districts to receive surgery at large teaching hospitals. This pattern highlights how patients balance the trade-off between travel distance and the superior resources and reputation associated with teaching hospitals. We provide further evidence of this mechanism in Section VI.C.

[26]Descriptive evidence shows that the pre-trend is driven by a declining trend in the control group's travel distance (online Appendix Figure A.2). The trend in the treatment group's travel distance was stable before the reform and increased by up to 3 kilometers post-reform. The DiD estimate (a 2.4 kilometer increase in distance) indicates a fairly similar increase, though it is likely to be slightly downward biased.

Figure 2: The Effect of the Reform on Choice and Allocation Outcomes



Teaching hospital

Distance (km)

Nearest hospital

Different hospital district

Notes: Estimated using specification (2) using patient-level data. Includes the DiD estimates (in percentage points) shown in Columns 1–4 of Table 3.

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)
$\text{Treated}_i \times \text{Post}_t$	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}
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.060	145.636	2.205
N	473,617	316,454	473,617
Surgery type FEs	✓	✓	✓
Municipal FEs	✓	✓	✓
Age & sex	✓	✓	✓

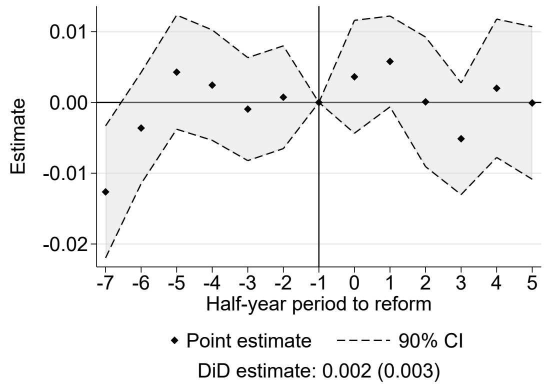
Notes: t-test level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Estimated from specification (1) using patient-level data. Standard errors clustered at the level of patient's home municipality ($N = 319$). Sharpened q-values are calculated for the sample.
[a] Equals one if the patient 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 a smaller N compared to the other columns (see online Appendix Section A1.4 for more details).
[c] Number of days between admission and discharge.

V.B. Effects on Hospital Performance

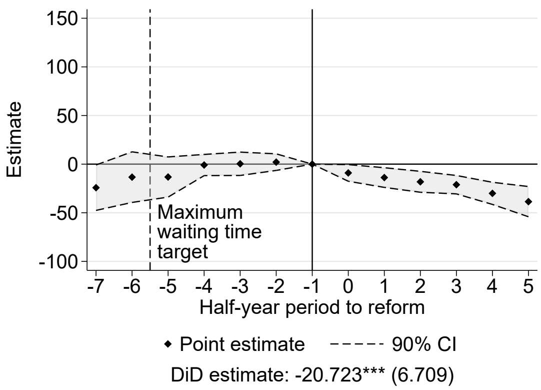
Table 4 displays the results on the effects of the choice reform on hospital performance, as measured by the quality of care, waiting time, and length of stay. Column 1 shows that the reform had little effect on the quality of care, as measured by the 30-day emergency readmission probability The point estimates are small in magnitude and not statistically significant. The 95 percent confidence intervals allow us to rule out effects larger than 1 percentage points (17 percent compared to the baseline). Figure 3 also shows that none of the post-reform point estimates are statistically different from zero.

We find that waiting times decreased substantially (a decrease of 21 days or 14 percent) post-reform (Column 2 in Table 4 and Figure 3). The waiting time estimates are precise and there is only little evidence of pre-trends. The decrease in waiting time is fairly similar or even larger (27 days or 22 percent) when we use data from periods after the markets had already adjusted to the introduction of the nationwide maximum waiting time target, Q4/2006–Q4/2010 (online Appendix Table A.12).

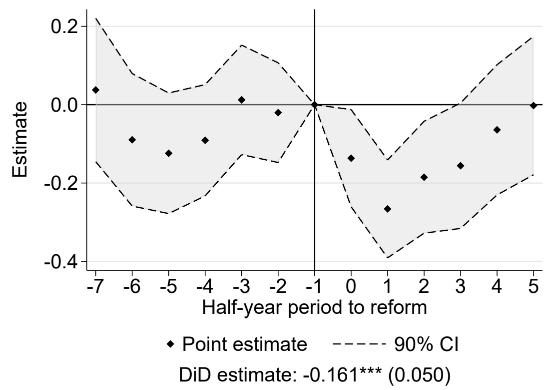
Figure 3: The Effects of the Reform on Quality of Care, Waiting Time, and Length of Stay



Readmission



Waiting Time



Length of Stay

Notes: Estimated from specification (2) using patient-level data. Includes the DiD estimates (in percentage points or number of days) shown in Columns 1–3 of Table 4. The dashed vertical line at $l = -5$ indicates the introduction of a maximum waiting time target. This national law change, implemented in early 2005 before the patient choice reform, guaranteed access to medical treatment within six months of referral.

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 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 (Column 3 in Table 4 and Figure 3). This suggests that the reform may have incentivized hospitals to discharge orthopedic patients earlier. 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.

Spillover effects to emergency care. 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 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.[27] The DiD estimates in online Appendix Table A.13 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 A3.1). 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 (an 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 are similar to our baseline results (Tables A.14 and A.15).

We also study the robustness of our results regarding our baseline sample. 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. The baseline results are, however, robust to the inclusion of these patients to the estimation sample (Tables A.16 and A.17). 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

[27] 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).

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.18–A.21). Third, some hospitals had a joint hospital identifier in the data and cannot be distinguished from each other. 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.22 and A.23).

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 effects are similar or even larger than our baseline results (Table A.24).

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 different types of 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 identifies the average treatment effect of the patient choice reform for non-teaching hospitals, and β_3 identifies the differential or heterogeneous effect of the patient choice reform in teaching hospitals compared to non-teaching hospitals. For the specification estimating the average hospital-level effects across all hospitals (without differential effect), we set $\beta_2, \beta_3 = 0$.[28] We cluster the standard errors at the hospital district level ($N = 20$).

Table 5 presents the estimated effects of the patient choice reform on hospital volume and market concentration, based on hospital-level data with outcomes defined as the number of patients and the actual hospital-level HHI. The results for the average effects across all hospitals (without differential effects) are shown in Columns 1 and 3, while the differential effects by hospital type are shown in Columns 2 and 4. We find that hospital volume increased by 27 patients (8 percent) on average after the reform (Column 1). Moreover, the results suggest a volume increase in non-teaching hospitals but the estimate of β_2 is not statistically significant (Column 2). Hospital volumes increased disproportionally in teaching hospitals, with an increase of 132 patients per quarter (Column 2).[29] In relative terms, it corresponds to an increase of 36 percent compared with the mean volume of all hospitals (361 patients per quarter) and 10 percent compared with the mean volume of teaching hospitals (1,281 patients per quarter, see online Appendix Table A.7).

The DiD point estimates show a 4 percent decrease in market concentration (Column 3). The average treatment effects, however, mask considerable heterogeneity in the effects across different types of hospitals and their markets (Column 4). The point estimates for non-teaching hospitals

[28] More specifically, we estimate the following specification: $y_{ht} = \beta_1 \text{Treated}_h \times \text{Post}_t + \mathbf{X}'_{ht}\gamma + \lambda_t + \eta_h + \varepsilon_{ht}$.
[29] The sum of the coefficients on Treated$_h \times$ Post$_t$ and Treated$_h \times$ Post$_t \times$ Teaching$_h$: $\beta_1 + \beta_3$.

Table 5: 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)
$\text{Treated}_h \times \text{Post}_t$	27.128***	16.467	-0.036**	-0.042**
	(7.826)	(11.037)	(0.015)	(0.015)
p-value	[0.003]	[0.152]	[0.023]	[0.014]
Sharpened q-value	{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)
p-value		[0.005]		[0.030]
Sharpened q-value		{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 are in parentheses, standard p-values are in square bracket, and sharpened q-values are 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 A1.5 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.

show that concentration decreased by 4 percent in their markets. In contrast, reflecting the disproportionate increase in teaching hospitals' volumes, concentration in their markets increased by 1 percentage point (1.3 percent compared with the mean HHI) after the choice reform.

Table 6 presents the effects of the patient choice reform estimated from a specification identifying the average effects across all hospitals and from the specification allowing differential effects by hospital type. 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) and remain fairly similar in magnitude when we use data from periods after the markets had already adjusted to the introduction of the nationwide maximum waiting time target, Q4/2006–Q4/2010 (online Appendix Table A.25).

Table 6: 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)
$\text{Treated}_h \times \text{Post}_t$	0.001	0.000	-14.984*	-10.619	-0.036	-0.017
	(0.004)	(0.005)	(8.308)	(11.679)	(0.095)	(0.111)
p-value	[0.841]	[0.927]	[0.087]	[0.375]	[0.712]	[0.879]
Sharpened q-value	{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)
p-value		[0.284]		[0.014]		[0.157]
Sharpened q-value		{0.890}		{0.147}		{0.646}
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	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 are in parentheses, standard p-values are in square bracket, and sharpened q-values are 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 A1.4 for more details). We analyze the robustness of the results regarding missing values in waiting times and summarize and show robustness of the results regarding missing values in waiting times in Section V.C. See also online Appendix Table A.25 for the waiting time estimates unaffected by the introduction of the maximum waiting time target policy (during pre-period).

[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.

Our results also suggest that all hospitals, regardless of type, improved their performance in response to increased choice and competition. The statistically insignificant point estimates further suggest that the reform reduced the mean length of stays in both teaching and non-teaching hospitals (Columns 5 and 6)—without major changes in clinical quality (Columns 1 and 2).

VI.B. Patient Composition, Resource Use, and Mechanisms

A potential concern is that our results reflect the sorting of patients by severity across hospitals after the choice reform, rather than improvements in hospital performance. The reform may have changed both hospital volumes and the composition of patients treated, as some hospitals attracted healthier patients while others treated relatively sicker patients. For example, hospitals could have increased volumes without expanding capacity by admitting and treating healthier patients with shorter length of stays after the reform.[30] To examine whether patient composition changed after the reform, 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.26 show little change in patient composition on average and across hospital types post-reform. The only exception is teaching hospitals, which treated a higher share of female patients post-reform.

Hospitals can also increase volumes by expanding resources or capacity, which would typically involve higher costs. We therefore 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.27).[31] 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.The reform might have incentivized hospitals to reduce potential spare capacity that persisted despite long waiting times, for example, due to inefficiencies in resource use, patient management, discharge practices, or the reservation of capacity for more severe cases. Using information in hospital districts' annual reports, we confirm that teaching hospitals neither operated at full capacity or expanded their capacity in terms of the number of beds post-reform. Surgical expenditure increased over 40 percent at teaching hospitals (online Appendix Table A.27), likely reflecting patient reallocation and a disproportionate increase in volumes. However, expenditure per patient did not change in an economically or statistically significant way, suggesting that improved outcomes were achieved without using additional resources. Shorter hospital stays and more efficient resource use may

[30]Patient composition could also have changed for reasons unrelated to the reform, such as differences in aging and demographic trends between the reform and control areas. However, at the baseline, we find that the mean age and sex compositions were similar between the two areas (Section III.E).

[31]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.

have allowed hospitals to treat more patients without raising the cost per per patient.

VI.C. Heterogeneity by Procedure Type: Hip and Knee Replacements

We examine heterogeneity across procedure types, as differences in clinical complexity may influence both patient choice and provider responses. We conduct all analyses separately for hip and knee replacement surgeries, which are more invasive surgeries with longer length of stays compared with many minor orthopedic surgeries.

Because more invasive surgeries demand specialized expertize and resources, patients may place greater value on the reputation of teaching hospitals. Consistent with this, we find a larger increase in the probability of choosing a teaching hospital for hip and knee replacements (an increase of 4—5 percentage points or 10-–12 percent in online Appendix Table A.28 and Figures A.4-A.5) compared with all orthopedic surgeries (an increase of 3 percentage point or 7 percent in Table 3). Hip and knee replacement patients were also more willing to travel longer distances and across hospital district borders after the reform, although the coefficient estimates are not statistically significant for knee replacement surgeries.

We then study the effects on hospital performance outcomes. In terms of clinical quality outcomes, we find little effect on 30-day emergency readmissions for hip and knee replacements using patient-level data (Online Appendix Table A.29 and Figure A.6). We further examine other clinical quality outcomes such as revision surgeries, mechanical complications, and prosthesis infections or inflammation. Since hip and knee replacement surgeries are relatively invasive and resource-intensive, these outcomes are particularly well-related to the quality of surgical and post-operative care (Urquhart et al., 2010; Mäkelä et al., 2011; Bayliss et al., 2017; Fleischman et al., 2019). Most estimates are statistically insignificant, except for knee replacements, where mechanical complications decreased by 1 percentage points (22 percent) (online Appendix Table A.30). In terms of other hospital performance measures, we find that the reform also shortened waiting times for hip and knee replacements (Online Appendix Table A.29), coinciding with an increase in hospital volumes (Online Appendix Table A.31).[32] The volume increased disproportionally in teaching hospitals (an increase of up to 29 patients per quarter), increasing concentration in their markets by up to 7 percent. Increasing volumes and improving timely access to care can be particularly critical for hip and knee replacement patients, as prolonged waits can exacerbate pain and reduce quality of life.[33]

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.33.

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.[34] 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 A.12).[35] 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

[32]Table A.32 shows that all hospital types reduced waiting times without changing quality, although many hospital-type–specific estimates are imprecise.

[33]The results also indicate a decline in length of stays for hip and knee replacements, though it appears to be imprecisely estimated based on the event study Figure A.6.

[34]For comparison, g would be $4,386$ and $3,113$ euros for hip and knee replacements, respectively (online Appendix Table A.33).

[35]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 orthopedic surgery. Then, g would be reduced by $w_t/365.25 \times 1.27$ euros, depending on the waiting time in *years*, w_t ($145.636 \times 1.27 \approx 185$ euros in the pre-reform period and $(145.636 - 20.723) \times 1.27 \approx 159$ euros in the post-reform period for orthopedic surgery).

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 5), and there were 15 hospitals in the reform area. The approximated total volume per year would then be $y_0 = 361 \times 15 \times 4 \approx 21,700$ 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 $y_1 - y_0 \approx 1,600$ patients or 7 percent annually). Shorter waiting times also enable patients to benefit from health gains for a longer period due to earlier treatment.

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.[36]

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).[37] 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 mechanism 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

[36] $[g \times (T - w_1) \times y_1] - [g \times (T - w_0) \times y_0] = [2,547 \times (19.8 - 0.25) \times 23,300] - [2,547 \times (19.8 - 0.32) \times 21,700] \approx$ 85 million euros.

[37] 85 million euros $- (y_1 - y_0) \times c =$ 85 million euros $- (23,300 - 21,700) \times$ €6,645 $\approx$ 74 million euros.

euros ($3,400$ euros per patient).

VIII. Effect of Market Structure After Choice Reform

Next, we highlight the differences between our binary DiD design and the DiD approach used by the previous studies estimating *the effect of market structure* on hospital performance after nationwide choice reforms (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 patient choice treated and untreated areas. In contrast, previous studies have exploited variation in market structure after nationwide reforms, where all areas were treated simultaneously (e.g., in the English National Health Service, the Netherlands, and Norway), leaving no untreated control group. Thus, our design identifies the average treatment effect of introducing patient choice, rather than post-reform heterogeneity in hospital performance across different market structures.

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).[38] We bootstrap the clustered standard errors ($N = 67$ municipalities in the reform area) because $\widehat{\mathrm{HHI}}_h$ 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, estimates how hospital performance varies 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 effects of market structure post-reform in Columns 1, 3 and 5 of online Appendix Table A.34. The coefficient estimates for the measure of clinical quality, the readmission probability, are not statistically significant. The coefficient estimates for waiting time are not statistically significant but strongly positive, suggesting that waiting times decreased more in *less* concentrated markets after the reform; the DiD estimate of 71 indicates that a one standard deviation (0.15) decrease in the predicted HHI is associated with a $71 \times 0.15 \approx 11$ 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

[38]The results obtained using a binary DiD approach and a DiD approach using variation in the effects of choice reform by pre-reform market structure are not sensitive to the inclusion or exclusion of the control function residuals (online Appendix Tables A.35–A.36). Thus, changes in patient allocation to hospitals based on morbidity do not seem to drive our results.

choice literature relates to our binary DiD approach. We use the following specification:

$$
\begin{aligned}
y_{ht} = & \kappa_0 \widehat{\mathrm{HHI}}_h + \kappa_1 \mathrm{Post}_t + \kappa_2 \mathrm{Treated}_i \times \mathrm{Post}_t + \kappa_3 \mathrm{Treated}_i \times \widehat{\mathrm{HHI}}_h \\
& + \kappa_4 \widehat{\mathrm{HHI}}_h \times \mathrm{Post}_t + \kappa_5 \mathrm{Treated}_i \times \widehat{\mathrm{HHI}}_h \times \mathrm{Post}_t \\
& + \mathbf{X}'_{it} \nu_{\mathbf{1}} + \hat{r}'_{it} \theta_1 + \mathrm{Treated}_i \times (\mathbf{X}'_{it} \nu_{\mathbf{2}} + \hat{r}'_{it} \theta_2) + \mu_m + \varepsilon_{imht}.
\end{aligned} \tag{5}
$$

The results are presented in Columns 2, 4 and 6 of online Appendix Table A.34. 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.34. However, α_2 does not capture the average treatment effect of the choice reform, which is instead captured by $\kappa_2 + (\kappa_5 \times \widehat{\mathrm{HHI}}_h)$, where $\widehat{\mathrm{HHI}}_h$ is evaluated at the sample mean. Taking waiting times as an example, the choice reform reduced on average waiting times for orthopedic surgery by 11 days [39] (in line with Table 4 and Figure 3). The reduction was more pronounced in less concentrated markets (an additional 1.6 days[40] reduction for one standard deviation decrease 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 is informative about the effects of market structure and can be interpreted as a heterogeneity analysis across more and less concentrated markets, whereas we identify the average treatment effect of the reform using 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 surgery,

[39] $\kappa_3 + \kappa_5 \times \mathrm{mean}(\widehat{\mathrm{HHI}}_h) = -58.823 + 70.946 \times 0.674 \approx -11.$

[40] $(\kappa_4 + \kappa_5) \times \mathrm{SD}(\widehat{\mathrm{HHI}}_h) = (-82.182 + 70.946) \times 0.146 \approx -1.6.$

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 large teaching hospitals' market dominance by reallocating patients towards them, potentially due to their better resources and higher perceived quality. However, this does not necessarily imply that the choice reform incentivized hospitals to improve clinical quality.[41] 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 approximately $1,600$ patients (7 percent) annually and waiting times (i.e., non-monetary costs) became much shorter.[42]

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 to other systems with administratively set reimbursements to producers (such as the Medicare program 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

[41] 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.

[42] 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.

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.

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.

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.[43] 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.

We analyze treatment spells that began between January 1st 2004 and December 31st 2010 (our observation period).

A1.2. Sample Construction

Table A.1 summarizes the inclusion criteria for different samples. For our baseline estimation we construct an estimation sample from the discharge data containing all planned orthopedic surgeries. We identified the surgeries based on the main procedure code and the type of hospital admission: non-emergency/planned.[44] In total, the full orthopedic sample includes 658 surgery types.

We construct two other subsamples for our heterogeneity analyses: a sample of planned (primary) hip replacement surgeries and a sample of planned (primary) knee replacement surgeries.

[43] 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.

[44] 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.

In total, the hip replacement sample includes 8 surgery types and the knee replacement sample includes 7 surgery types.

For our analyses on the quality spillovers on emergency care, we also constructed three emergency care samples based on the main diagnosis code (ICD-10): AMI, stroke, and hip fracture admissions. We also identified all planned surgeries for the purpose of calculating the predicted HHI that we use in the DiD approach using variation in the effects of choice reform by pre-reform market structure.

Table A.1: 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.

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. Variable Construction

The details on our variable construction are provided in Tables A.2–A.5. In terms of our hospital performance measures, 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.[45] We implement comprehensive robustness checks in Section VIII, and confirm that the missing values do not bias the estimates.

Table A.2: 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 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 at the 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.

Hospital Volume and Market Structure Measures

	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 A1.5.

[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.

Table A.3: Hospital Performance Measures

	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.

Table A.4: Other Variables

	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.

Table A.5: Variables Measuring Resource Use

	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.

A1.5. 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:

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

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. Effect of Market Structure After Choice Reform: Further Details

A2.1. Predicted HHI

Following the standard practice in the literature (Kessler and McClellan, 2000; Gaynor, Moreno-Serra and Propper, 2013; Moscelli, Gravelle and Siciliani, 2021), we construct a predicted HHI which is used in specification (4). 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:

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

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.[46] Patient i 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.[47] 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.

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

[47] 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.

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:[48]

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

We refer to $\widehat{HHI}_h^{HOSP}$ as the *predicted* HHI.

A2.2. Control Function 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:

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

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.[49] The chosen covariates reflect geographical access (km_{ih},

[48] 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.

[49] 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.

nearest$_{ih}$) and factors potentially relevant to hospital choice post-reform (sameHD$_{ih}$, teaching$_h$, regional$_h$).[50] 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]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. Main Specification: Additional Tables and Figures

A3.1. Tables

Table A.6: 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.7: 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
Mean	361	1281	402	168	0.69	0.80	0.81	0.61
SD	(432)	(741)	(188)	(100)	(0.13)	(0.14)	(0.07)	(0.09)
N	705	75	225	405	705	75	225	405

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.8: 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
Mean	0.06	0.07	0.06	0.05	126	167	156	94	1.91	2.89	2.29	1.52
SD	(0.01)	(0.01)	(0.01)	(0.01)	(65)	(44)	(74)	(44)	(0.96)	(1.00)	(0.60)	(0.91)
N	705	75	225	405	610	75	223	312	705	75	225	405

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 waiting times in some quarters, which results in a smaller N compared to the other columns (see online Appendix Section A1.4 for more details). We analyze the robustness of the results regarding missing values in waiting times and summarize and show robustness of the results regarding missing values in waiting times in Section V.C.

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

Table A.9: 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.10: 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
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 waiting times in some quarters, which results in a smaller *N* compared to the other columns (see online Appendix Section A1.4 for more details). We analyze the robustness of the results regarding missing values in waiting times and summarize and show robustness of the results regarding missing values in waiting times in Section V.C.
[c] Mean length of stay in days (no risk-adjustment).

Table A.11: 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
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.12: The Effect of the Choice Reform on Waiting Time Estimates After Adjusting to the Introduction of the Maximum Waiting Time Target

	Whole time period (1)	After target (2)
$\text{Treated}_i \times \text{Post}_t$	-20.723***	-26.522***
	(6.709)	(5.212)
p-value	[0.002]	[0.000]
Sharpened q-value	{0.002}	{0.001}
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	145.636	118.062
N	316,454	199,492

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 from specification (1) using patient-level data. Standard errors clustered at the level of the patient's home municipality ($N = 319$). Column 1 re-displays baseline estimates from Column 2 of Table 4. Column 2 displays estimates using data from Q4/2006–Q4/2010 only. Sharpened q-values calculated sample-wise (i.e., jointly for (1)–(2)).

Table A.13: 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 from specification (1) using patient-level data of emergency care patients. 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.

A4. Additional Specifications, Robustness and Samples

Table A.14: Robustness Check: Additional Controls (Choice Outcomes)

	Teaching hospital[a] (1)	Distance (km)[b] (2)	Nearest hospital[c] (3)	Different hospital district[d] (4)
$\text{Treated}_i \times \text{Post}_t$	0.026^{***}	2.413^{**}	-0.006	0.009^{*}
	(0.007)	(1.179)	(0.009)	(0.005)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.377	26.481	0.794	0.039
N	473,617	473,617	473,617	473,617
Surgery type / 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 from specification (1) using patient-lelve data. 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.15: Robustness Check: Additional Controls (Hospital Performance Outcomes)

	Readmission[a] (1)	Waiting time[b] (2)	Length of stay[c] (3)
$\text{Treated}_i \times \text{Post}_t$	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 from specification (1) using patient-level data. 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 A1.4 for more details).
[c] Number of days between admission and discharge.

Table A.16: Robustness Check: Including Patients Affected by the Closed-Down Hospital (Choice Outcomes)

	Teaching hospital[a] (1)	Distance (km)[b] (2)	Nearest hospital[c] (3)	Different hospital district[d] (4)
$\text{Treated}_i \times \text{Post}_t$	0.026***	2.133*	0.015	0.007
	(0.007)	(1.201)	(0.014)	(0.006)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	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 from specification (1) using patient-level data. 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.
[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.17: 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)
$\text{Treated}_i \times \text{Post}_t$	0.002	-17.408**	-0.164***
	(0.003)	(6.945)	(0.049)
$\text{mean}(y_{imht} \| \text{Post}_t = 0)$	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 from specification (1) using patient-level data. 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.

[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 A1.4 for more details).

[c] Number of days between admission and discharge.

Table A.18: Robustness Check: Excluding Three Western Hospital Districts With Possibilities to Obtain Referral Outside Own Hospital District (Choice Outcomes)

	Teaching hospital[a] (1)	Distance (km)[b] (2)	Nearest hospital[c] (3)	Different hospital district[d] (4)
$\text{Treated}_i \times \text{Post}_t$	0.028***	1.594	-0.004	0.011*
	(0.009)	(1.018)	(0.010)	(0.006)
$\text{mean}(y_{imht} \| \text{Post}_t = 0)$	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 from specification (1) using patient-level data. 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.19: 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)
$\text{Treated}_i \times \text{Post}_t$	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 from specification (1) using patient-level data. 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 A1.4 for more details).
[c] Number of days between admission and discharge.

Table A.20: Robustness Check: Excluding Four Northern Hospital Districts Where Possible to Obtain Referral Outside Own Hospital District (Choice Outcomes)

	Teaching hospital[a] (1)	Distance (km)[b] (2)	Nearest hospital[c] (3)	Different hospital district[d] (4)
$\text{Treated}_i \times \text{Post}_t$	0.026***	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 from specification (1) using patient-level data. 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.21: 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)
$\text{Treated}_i \times \text{Post}_t$	0.002	-17.366***	-0.173***
	(0.003)	(6.664)	(0.052)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	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 from specification (1) using patient-level data. 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 A1.4 for more details).
[c] Number of days between admission and discharge.

Table A.22: Robustness Check: Excluding Hospitals That Used a Joint Hospital ID (Choice Outcomes)

	Teaching hospital[a] (1)	Distance (km)[b] (2)	Nearest hospital[c] (3)	Different hospital district[d] (4)
$\text{Treated}_i \times \text{Post}_t$	0.032***	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 from specification (1) using patient-level data. Standard errors clustered at the level of the patient's home municipality ($N = 317$).
[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.23: 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)
$\text{Treated}_i \times \text{Post}_t$	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 from specification (1) using patient-level data. Standard errors clustered at the level of the patient's home municipality ($N = 317$).
[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 A1.4 for more details).
[c] Number of days between admission and discharge.

Table A.24: 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)
$\text{Treated}_h \times \text{Post}_t$	-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 from specification (1) using patient-level data. Standard errors clustered at the level of the patient's home municipality ($N = 319$). 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.25: 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)
$\text{Treated}_h \times \text{Post}_t$	-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.26: 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)
$\text{Treated}_h \times \text{Post}_t$	-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)
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	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.27: The Effects on Hospitals' Surgical Expenditure

	Total expenditure (millions of €)[a]		Expenditure per patient (€)[b]	
	DiD (1)	Heterogeneity (2)	DiD (3)	Heterogeneity (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)
mean(y_{ht}|$\text{Post}_t = 0$)	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.28: The Effects of the Reform on Hospital Choice and Allocation Outcomes for Hip and Knee Replacements

	Teaching hospital[a] (1)	Distance (km)[b] (2)	Nearest hospital[c] (3)	Different hospital district[d] (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}
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.429	29.091	0.818	0.035
N	31,404	31,404	31,404	31,404
Panel B. Knee replacement surgeries				
$\text{Treated}_i \times \text{Post}_t$	0.039***	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}
$\text{mean}(y_{imht} \vert \text{Post}_t = 0)$	0.393	28.399	0.819	0.034
N	39,427	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 from specification (1) using patient-level data. Standard errors clustered at the level of the patient's home municipality ($N = 319$). Sharpened q-values are calculated for the sample.

[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.29: The Effects of the Reform on Quality of Care, Waiting Time, and Length of Stay for Hip and Knee Replacements

	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}
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	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	-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}
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	0.102	225.373	7.609
N	39,427	29,759	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 from specification (1) using patient-level data. Standard errors clustered at the level of patient's home municipality ($N = 319$). Sharpened q-values are calculated for 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 a smaller N compared to the other columns (see online Appendix Section A1.4 for more details).

[c] Number of days between admission and discharge.

Table A.30: The Effects of the Reform on Additional Quality Measures for Hip and Knee Replacements

	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)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	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)
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	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 from specification (1) using patient-level data. 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.31: The Effects of the Reform on Hospital Volumes and Market Structure for Hip and Knee Replacements

	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)
p-value	[0.027]	[0.078]	[0.174]	[0.041]
Sharpened q-value	{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)
p-value		[0.000]		[0.006]
Sharpened q-value		{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)
p-value	[0.004]	[0.000]	[0.614]	[0.453]
Sharpened q-value	{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)
p-value		[0.000]		[0.092]
Sharpened q-value		{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
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 are in parentheses, standard p-values are in square bracket, and sharpened q-values are 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 A1.5 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.

Table A.32: The Effects of the Reform on Hospital-level Means of Quality of Care, Waiting Times, and Length of Stay for Hip and Knee Replacements

	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)
p-value	[0.797]	[0.773]	[0.034]	[0.153]	[0.062]	[0.000]
Sharpened q-value	{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)
p-value		[0.690]		[0.318]		[0.003]
Sharpened q-value		{0.549}		{0.361}		{0.012}
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	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)
p-value	[0.616]	[0.666]	[0.015]	[0.094]	[0.081]	[0.036]
Sharpened q-value	{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)
p-value		[0.765]		[0.528]		[0.124]
Sharpened q-value		{0.516}		{0.500}		{0.211}
$\text{mean}(y_{ht} \mid \text{Post}_t = 0)$	0.109	0.109	208.711	208.711	8.313	8.313
N	1,070	1,070	812	812	1,070	1,070
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 are in parentheses, standard p-values are in square bracket, and sharpened q-values are 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 A1.4 for more details). We analyze the robustness of the results regarding missing values in waiting times and summarize and show robustness of the results regarding missing values in waiting times in Section V.C. See also online Appendix Table A.25 for the waiting time estimates unaffected by the introduction of the maximum waiting time target policy (during pre-period).

[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.

Table A.33: Back-of-the-Envelope Calculation

Variable	Description	Source	Value
Panel A. Benefits of the reform			
g	Health gain from surgery (2010 euros)[a]	Jansson and Granath (2011); Gyrd-Hansen (2003)	2,546.81
	Average age (years)	Table 2	52
T	Discounted life expectancy at ave. age[b]		19.76
$w_0 \times 365.25$	Waiting time (days), pre-reform	Table A.12	118.06
$w_1 \times 365.25$	Waiting time (days), post-reform	Table A.12	91.54
	Average volume per hospital/qrt	Table 5	361
y_0	Total volume per year, pre-reform[c]	Table 5	21,675
y_1	Total volume per year, post-reform[d]	Table 5	23,302
$g \times (T - w_0) \times y_0$	Total health benefit per year, pre-reform		1,073,161,244.91
$g \times (T - w_1) \times y_1$	Total health benefit per year, post-reform		1,158,060,556.36
	Total benefit change (post − pre)		84,899,311.44
Panel B. Costs of the reform (2010 euros)			
c_0	Average cost, pre-reform	Osnes-Ringen et al. (2011)	6,645.23[e]
	Costs related to LOS per patient[f]		
c_0^{LOS}	- Pre-reform	Osnes-Ringen et al. (2011), Table 4	3,389.07
c_1^{LOS}	- Post-reform	Osnes-Ringen et al. (2011), Table 4	3,141.61
c_1	Average cost, post-reform[g]		6,397.77
$c_1 \times y_1 - c_0 \times y_0$			5,050,000.91
Panel C. Net effects of the reform			
	Total benefits − total costs		79,849,310.53

Notes: [a] To measure h, we first use the estimates of the quality of life (EQ-5D index) gains from orthopedic surgeries (index score improvement of 0.18, Jansson and Granath (2011)). Then we measure their monetary value based on the willingness-to-pay estimates by Gyrd-Hansen (2003), expressed 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 $\times$ the number of reform area hospitals ($N = 15$) $\times$ the number of quarters ($N = 4$).

[d] y_1 = the average number of patients per hospital and quarter in the post-reform period $\times$ the number of reform area hospitals ($N = 15$) $\times$ the number of quarters ($N = 4$).

[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 $c_1^{LOS} = c_0^{LOS} / mean\ LOS * (mean\ LOS - 0.161)$, 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.34: 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)
$\text{Treated}_i \times \text{Post}_t$		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]
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	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 from specification (4) and data from reform area only ($\text{Treated}_i = 1$). Columns (2), (4) and (6) estimated from specification (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 A1.4 for more details).

[c] Number of days between admission and discharge.

Table A.35: Robustness Check: Effects 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)
$\text{Treated}_i \times \text{Post}_t$	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 from specification (1) using patient-level data. 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 A1.4 for more details).
[c] Number of days between admission and discharge.

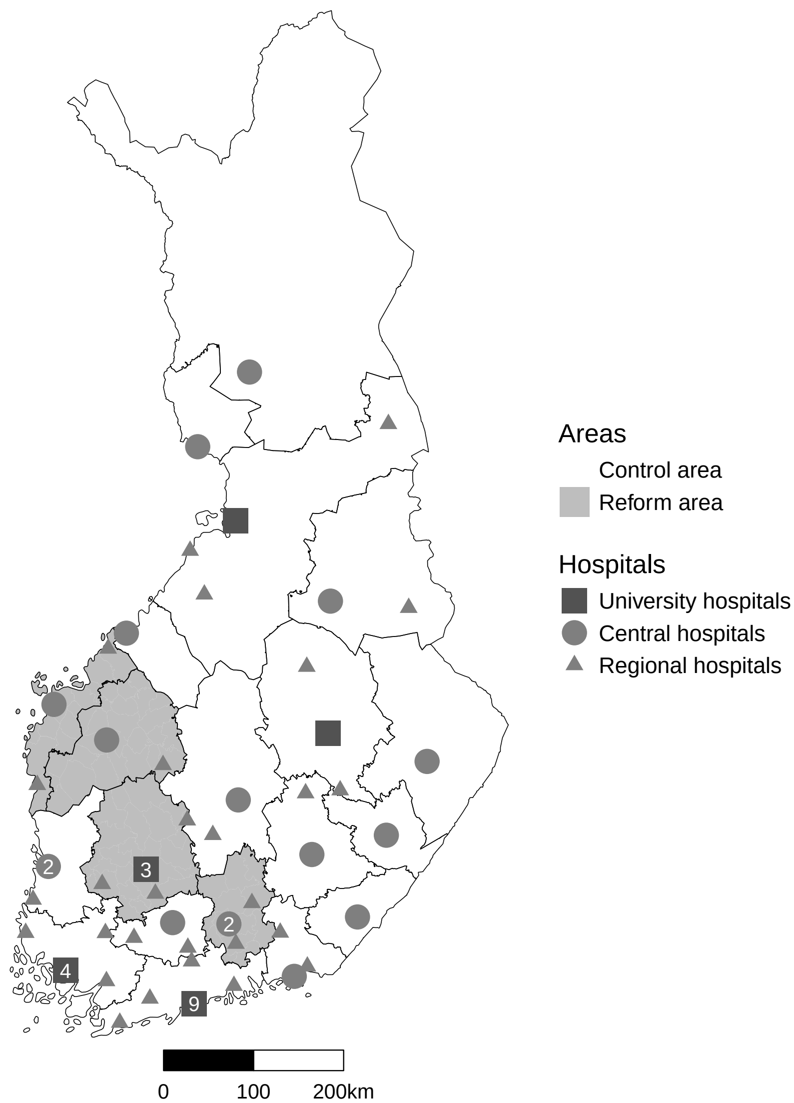
Table A.36: The Effects of Market Structure Without Controlling for First-Stage Residuals from a Hospital Choice Model

	Readmission[a] (1)	Waiting time[b] (2)	Length of stay[c] (3)
$\widehat{\text{HHI}}_h \times \text{Post}_t$	-0.029*	10.030	-1.204***
	[-0.052, -0.001]	[-50.913, 69.490]	[-1.726, -0.765]
$\text{mean}(y_{imht} \mid \text{Post}_t = 0)$	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 from specification (4) using patient-level data. 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 A1.4 for more details).
[c] Number of days between admission and discharge.

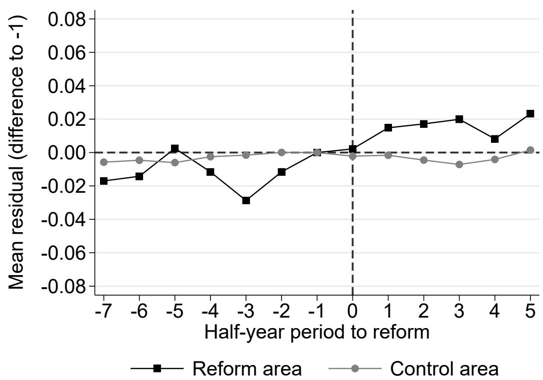
A5. Figures

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

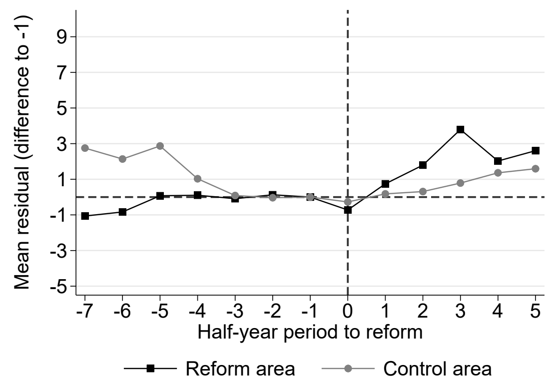


Notes: Borders indicate hospital districts in 2007 and the shaded area constitutes the 2007 reform area. Squares are teaching hospitals, dots central hospitals, and triangles regional hospitals. The squares and circles with a number mark the regions which had multiple (2–9) hospitals, all of which had 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.

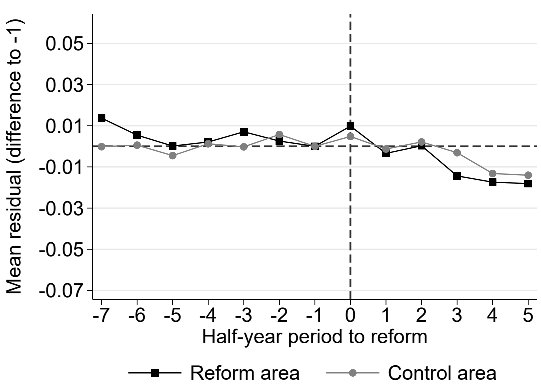
Figure A.2: Mean Outcome Trends: Choice and Allocation Outcomes



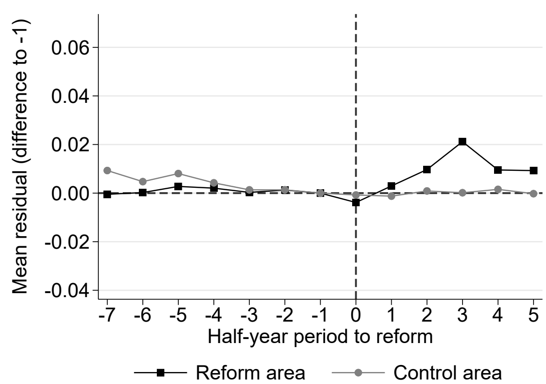
Teaching hospital



Distance



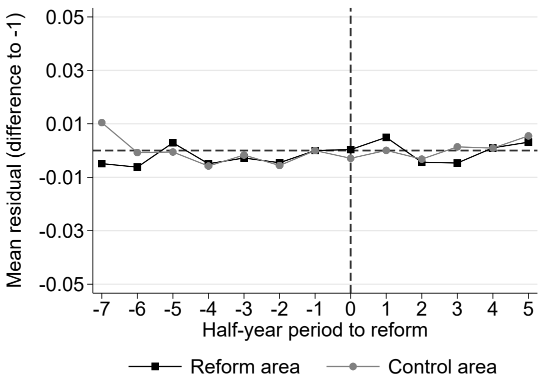
Nearest hospital



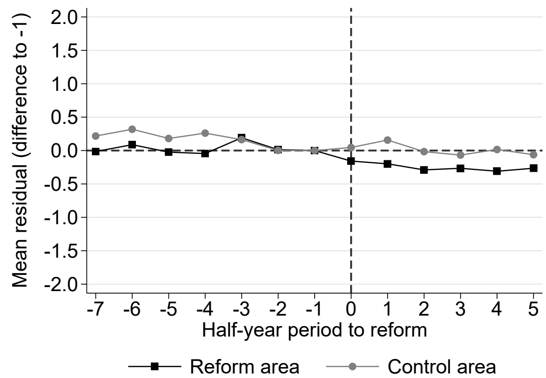
Different hospital district

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 for all orthopedic surgery patients. 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.

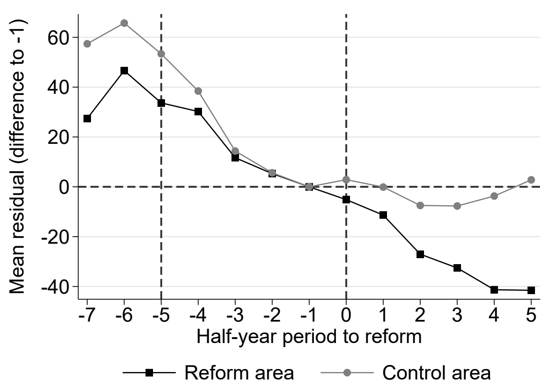
Figure A.3: Mean Outcome Trends: Probability of Emergency Readmission and Length of Stay



Emergency readmissions



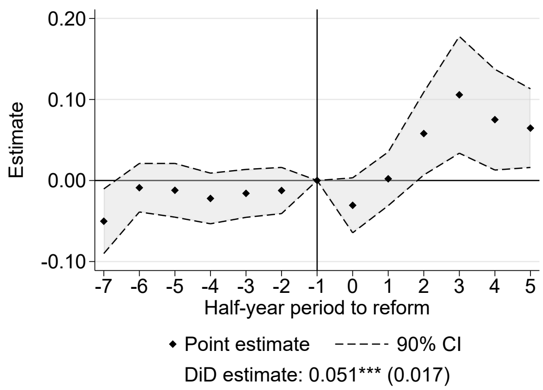
Length of stay



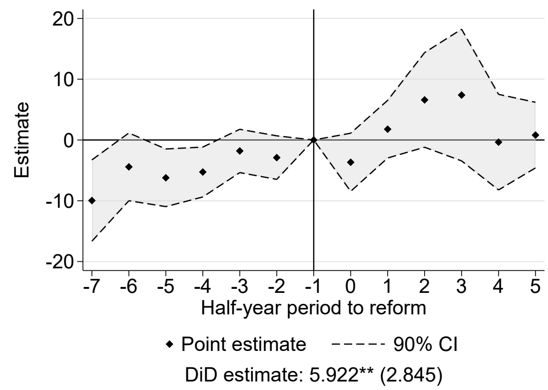
Waiting time

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 for all orthopedic surgery patients. 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.

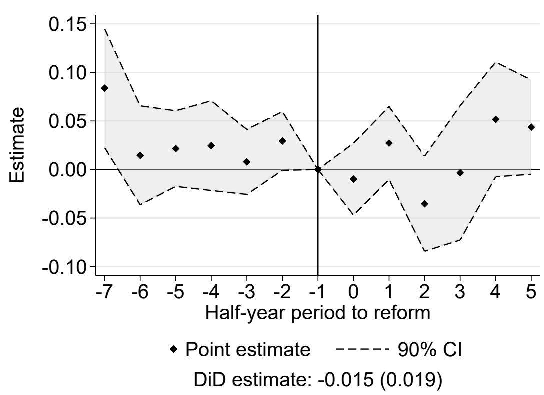
Figure A.4: The Effect of the Reform on Choice and Allocation Outcomes



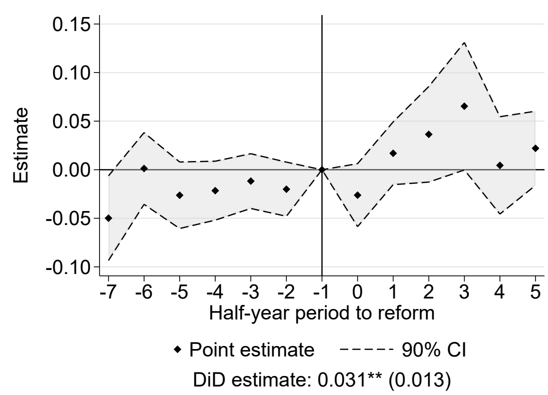
Teaching hospital



Distance (km)



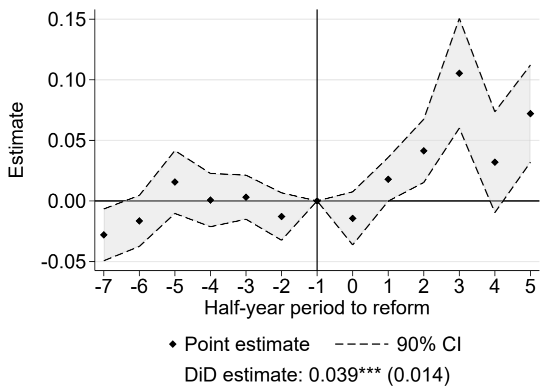
Nearest hospital



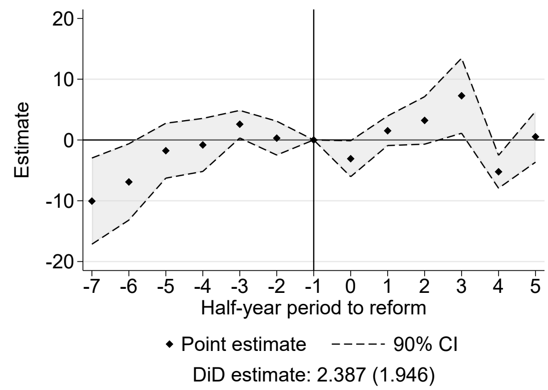
Different hospital district

Notes: Estimated from specification (2) for planned hip replacement surgery patients. Includes the DiD estimates (in percentage points for the binary outcomes and in kilometers for the distance traveled) shown in Columns 1–4 of Table A.28.

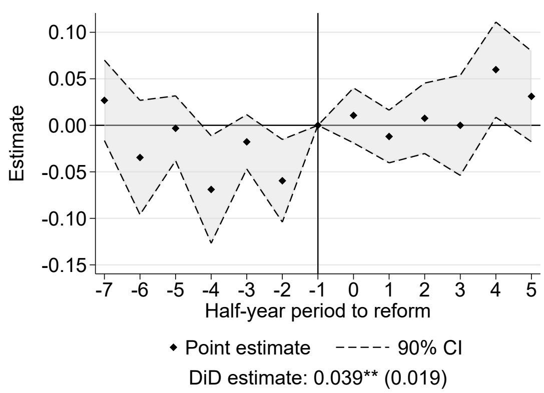
Figure A.5: The Effect of the Reform on Choice and Allocation Outcomes



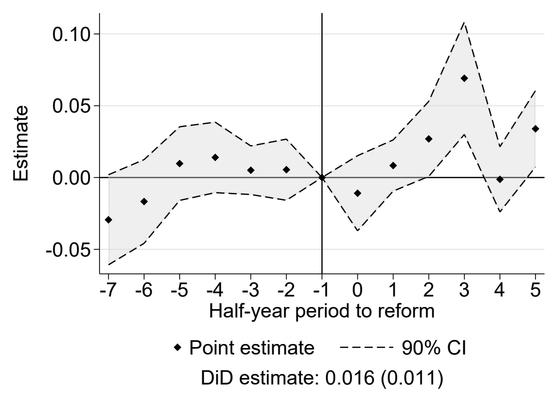
Teaching hospital



Distance (km)



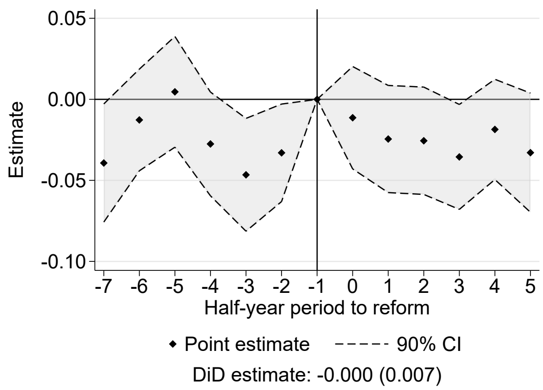
Nearest hospital



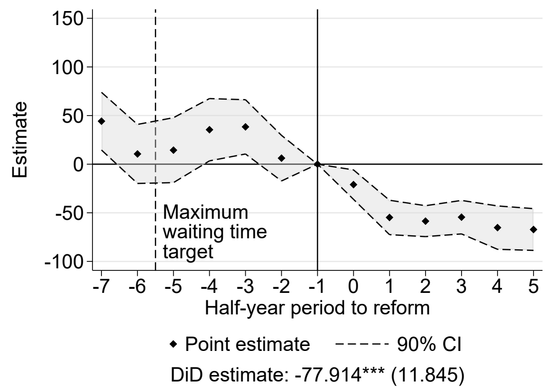
Different hospital district

Notes: Estimated from specification (2) for planned knee replacement surgery patients. Includes the DiD estimates (in percentage points for the binary outcomes and in kilometers for the distance traveled) shown in Columns 1–4 of Table A.28.

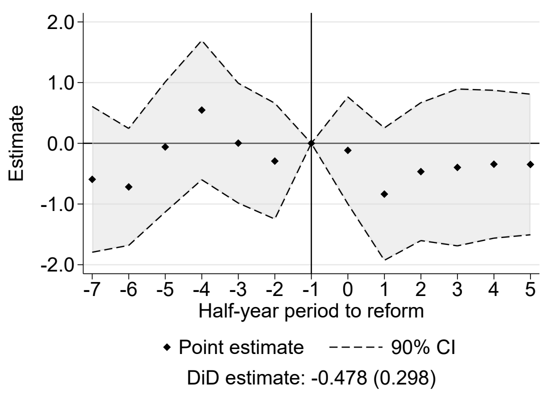
Figure A.6: The Effect of the Reform on Hospital Performance Outcomes



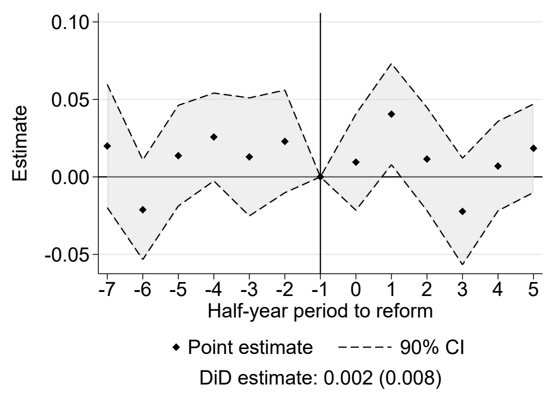
Panel A: Readmission, Hip replacement



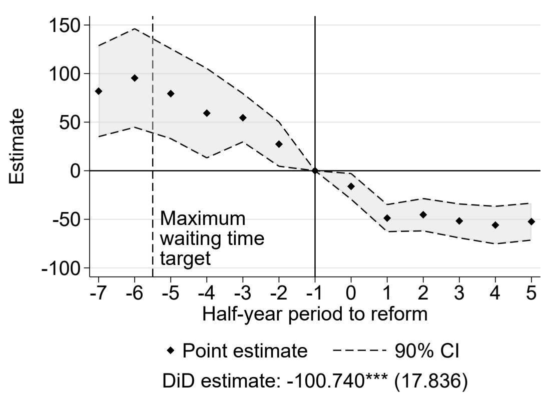
Panel B: Waiting Time, Hip replacement



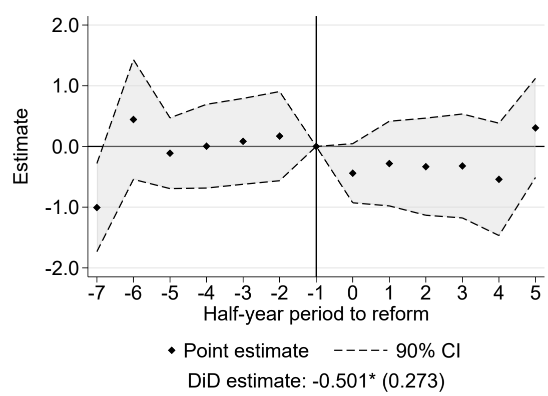
Panel C: Length of Stay, Hip replacement



Panel D: Readmission, Knee replacement



Panel E: Waiting Time, Knee replacement



Panel F: Length of Stay, Knee replacement

Notes: Estimated from specification (2) hip and knee replacement surgery patients. Includes the DiD estimates (in number of days for the waiting time and length of stay) shown in Table A.29.