Does competition from private surgical centres improve public hospitals’ performance? Evidence from the English National Health Service

Zack Cooper a, Stephen Gibbons b, Matthew Skellern b,⁎

a Yale School of Public Health and Yale University, United States of America
b London School of Economics, United Kingdom of Great Britain and Northern Ireland

Abstract

Article history:
Received 4 November 2016
Received in revised form 31 July 2018
Accepted 7 August 2018
Available online 11 September 2018

JEL classification:
C23
H57
I11
L1
L2
L3
R12

Keywords:
Hospital competition
Public-private competition
Market entry
Market structure
Outsourcing
Hospital efficiency
Risk selection
Cream skimming
Public services

1. Introduction

In the 2000s, there was a widespread push in Europe and the United States to increase the role of user choice and provider competition in public services. In general, these pro-market reforms were designed to increase the quality and efficiency of public services like health care and education, which had previously been run through non-market means like performance management (Gaynor and Town, 2011; Propper et al., 2007, 2010). Often, as part of these market-based reforms, policymakers encouraged the entry of private, for-profit firms to compete against public sector providers. These efforts are exemplified by the growing use of charter schools in the United States and private health care providers in publicly funded health systems in Western Europe (Jost et al., 2006; Fryer Jr, 2012). This paper explores how competition generated by the government-facilitated entry of private, for-profit firms affects the performance of incumbent public providers. In particular, we estimate the impact of the entry of a series of private, for-profit surgical centres in the English National Health Service (NHS). Policymakers steered the entry of these surgical centres to areas with high patient waiting times, with the aims of increasing surgical capacity and stimulating competition. We estimate the impact of this private provider entry on the efficiency of incumbent public hospitals, and examine whether it left incumbents with a riskier and more costly mix of patients.

Advocates of diversifying the supply of public services providers argue that private, for-profit entrants will innovate and offer higher quality than incumbents, and that entry of private providers will create competitive pressure on public providers to raise their own...
performance (Le Grand, 2009; Seddon, 2007). We are particularly focused on testing this latter claim: can the entry of private, for-profit surgical centres improve the performance of incumbent public hospitals?

Critics of market-based reforms generally cite the many ways that public services, and health care in particular, differ from highly stylised, perfectly competitive markets, and argue that competition will not improve performance (Jones and Mays, 2009; Fotaki et al., 2008). Moreover, it is sometimes argued that, because new entrants are often much smaller than incumbents (in our case, we analyse surgical centres competing against hospitals), they may not have sufficient scale to affect the behaviour of existing providers (Goddard, 2015). A third criticism is that private, for-profit firms may select customers with desirable characteristics (e.g. better students or less risky patients), leaving public providers treating a riskier or costlier group of users (Los Angeles Times Editorial Board, 2016; Bardsley and Dixon, 2011). More generally, it is not clear that governments are well equipped to determine where to locate entrants in such a way as to engineer effective competition.

The English NHS provides a unique environment in which to test the effect of private, for-profit provider entry on public service providers’ performance, and in so doing to analyse the extent to which governments can ‘create’ competition. In the 2000s, the British government facilitated the entry of Independent Sector Treatment Centres (ISTCs). ISTCs are private, for-profit centres focused on provision of routine, high volume elective (i.e. medically necessary, non-emergency, scheduled in advance) surgical procedures to public (NHS) patients. This policy was part of a wider policy package designed to tackle waiting times within the English NHS, the centrepiece of which was an ambitious set of targets to reduce waiting times for surgery. ISTCs were established to rapidly expand capacity in regions deemed at risk of not meeting these targets (Naylor and Gregory, 2009). As we demonstrate, while the placement of these specialty surgical centres was correlated with local public hospital waiting times during the pre-policy period, their placement was uncorrelated with measures of the efficiency and clinical quality of these incumbents over the same period. This implies that treatment assignment was unrelated to the pre-policy levels of the outcome variables we study. In addition, we demonstrate that public hospitals close to ISTC entrants had nearly identical pre-entry trends to public hospitals unexposed to ISTC entry across a range of performance measures (other than waiting times). We use this observation to motivate a difference-in-difference (DID) strategy to estimate the causal effect of ISTC entry on outcomes at nearby public hospitals and highlight that our control group serves a good counterfactual for what would have occurred to the treatment group after 2004/5 in the absence of the entry of ISTCs.

Measuring efficiency of health care provision is a long-standing challenge because of the absence or poor standard of data on costs and quality. Faced with these problems, researchers have frequently used patient length of stay (LOS) as a proxy for efficiency (Fenn and Davies, 1990; Martin and Smith, 1996; Gaynor et al., 2013) on the grounds that, provided clinical quality can be maintained, shorter LOS implies lower costs for the same outcomes. However, a key difficulty with using LOS to capture efficiency is that it is heavily influenced by patient characteristics – patients in poorer health before surgery will tend to have longer lengths of stay for reasons unrelated to hospital performance. In this study, we use an innovative approach to address the influence of patient characteristics on LOS-based efficiency measures by disaggregating LOS into two components: time from admission until surgery (‘pre-surgery LOS’), and time from surgery until discharge (‘post-surgery LOS’). We show that pre-surgery LOS is less affected by patient characteristics than other components of LOS, and use it – or alternatively, the percentage of patients treated on the day of admission – as a proxy for hospital efficiency.

In what follows, we show that the entry of private, for-profit specialty surgical centres led to a 16% reduction in pre-surgery LOS at nearby public hospitals – which translates to a 24 percentage point increase in the proportion of patients treated on the day of admission. However, we also find evidence that these entrants engaged in risk selection, leaving nearby public hospitals with a sicker (and therefore costlier) mix of patients. In particular, public hospitals exposed to the entry of private specialty surgical centres experienced an 11.6% deterioration in average patient health status as captured by the Charlson score (defined in Section 4). This increase in patient severity likely led to an increase in post-surgery LOS at incumbent NHS hospitals. Finally, while ISTC entry may have led to reduced case loads at some public hospitals with which they shared a market, we show that our estimated treatment effects are not driven by changes in volume caused by ISTC entry.

This paper adds to several literatures. First, it builds on previous work assessing how the entry of private, for-profit firms impacts the performance of incumbent public service providers (Hoxby, 1994; Barro et al., 2006; Cutler et al., 2010; Sass, 2006). In general, researchers have struggled to assess the causal impact of competition from new market entrants (e.g. surgical centres and charter schools) into markets for public services because the entry location of private firms is usually endogenous. We exploit the fact that siting of surgical centres in England was driven by government policy tied to waiting times, not our efficiency measure, and show that the entry of ISTCs raised incumbent hospitals’ productivity. Second, it adds to the literature assessing the impact of hospital competition (Kessler and McClellan, 2000; Gaynor et al., 2015; Cooper et al., 2011). We illustrate that, in markets where payments are regulated, competition can raise hospitals’ efficiency. Moreover, we find that smaller entrants can affect the behaviour of larger incumbents. Third, it adds to the literature analysing whether private, for-profit surgical centres offering public services risk-select against public incumbents (Barro et al., 2006; Winter, 2003; Gram et al., 2005; Street et al., 2010; Zimmer and Guarino, 2013; Bifulco and Reback, 2014). We find that the entry of ISTCs left public hospitals with a riskier mix of patients. To some extent, this was by design: ISTCs in England were focused on treating uncomplicated cases. While the entry of specialist surgical centres focused on routine procedures could in theory represent efficient patient sorting, such an arrangement is likely to leave existing providers treating a sicker patient mix and worse off financially, unless it is accompanied by a reimbursement system that adequately adjusts payments to reflect patient severity. The consensus is that NHS payments were not adequately risk adjusted during the period we investigate (Mason et al., 2008), meaning that NHS hospitals that had an ISTC enter nearby were likely left worse off as a result of being left with a sicker mix of patients.

More generally, this paper highlights the trade-offs that policymakers face when considering policies to encourage the entry of for-profit firms to compete with public service providers. Facilitating entry can lead to competition, which can prompt incumbent providers to raise their performance. However, these for-profit entrants may have very different objectives than incumbent providers, and may have a higher propensity to risk-select in order to draw a more advantageous mix of patients. Our work highlights the need for policy-makers to take risk-adjustment of payments seriously when considering policies to promote competition between firms with different objectives and differing abilities to treat complicated cases.

The remainder of this paper is structured as follows. Section 2 presents background information on recent NHS reforms, with particular focus on the ISTC programme. Section 3 explores the potential impact of ISTC entry on incumbents’ performance. Section 4 presents the data and empirical strategy. Section 5 reports the results, while Section 6 discusses and concludes.

2. Recent NHS reforms and the ISTC programme

The English NHS, founded in 1948, is funded through general taxation and, with few exceptions, offers health care that is free at the point of use. Patients must register with a single general practice (GP) clinic for the provision of primary care, and GPs act as ‘gatekeepers’ to
the secondary care system. For the most part, secondary care in England is organised around large public hospitals.

The NHS long struggled with waiting times for elective surgery, which, in some cases, could exceed a year. In 1997, a new Labour government was elected promising quick action to reduce waiting times. However, 1 year later, waiting times had increased (Klein, 2013, p.200). Concerns over waiting times became the catalyst for a series of reforms from 2000 onwards, which included rigorous performance management of public hospitals; introduction of patient choice and hospital competition underpinned by prospective reimbursement; and facilitated entry of specialist private surgical centres to compete with larger public hospital incumbents. The new prospective reimbursement system (known as Payment by Results or PbR) was modelled on the Diagnosis-Related Group (DRG) system used in Medicare in the United States (US). Under PbR, hospital reimbursement is tied to activity rather than to annual budgets or block contracts as was previously the case (DH, 2011).

In 2000, The Secretary of State released The NHS Plan (Secretary of State for Health, 2000), in which the government committed to cutting maximum waiting times for elective surgery from 18 months to 6 months by the end of 2005 (later reduced to 18 weeks, by 2008) using a series of targets tied to rewards and punishments. There is substantial evidence that the targets and performance management regime was extremely effective at reducing waiting times (Propper et al., 2008, 2010; Besley et al., 2009).

As part of its reform programme, in April 2002 the government announced that it was facilitating the entry of a series of privately run surgical centres (ISTCs) to deliver routine, high-volume diagnostic and elective surgical procedures to English NHS patients. Like other NHS services, NHS-funded patients could use ISTCs free of charge.

Although the NHS had long made use of private providers in England, ISTCs were distinctive in three ways. First, they were created as a deliberate policy of government, as opposed to being a result of decisions by local commissioners of care. Second, they provided services exclusively to NHS patients, as opposed to earlier arrangements in which NHS patients were treated in settings mainly focused on treatment of private patients (Naylor and Gregory, 2009). Third, whereas NHS physicians are in general permitted to also work in private settings, the first wave of ISTCs (which are the focus of this paper) were not allowed to use NHS doctors. This restriction ensured that ISTCs represented genuine new additions to capacity, rather than drawing away physician labour from nearby public hospitals.

More than any other factor, it was local waiting times that influenced where the government sought to locate the new private surgical centres (HCHC, 2006). According to government officials, “In October 2002, the Department [of Health] conducted an extensive forward planning exercise, during which all Strategic Health Authorities were asked to identify, in conjunction with their respective Primary Care Trusts, any anticipated gaps in their capacity need to meet the 2005 waiting times targets. The results of this exercise led to the identification of capacity gaps across the country, particularly in specialties such as cataract removal and orthopaedic procedures, where additional capacity was needed” (Anderson, 2006). Following this planning exercise, in December 2002 the Department of Health invited expressions of interest to run the first Wave of ISTCs. These invitations indicated

---

1 During this period, newspaper stories about excessive waiting times appeared regularly in the popular press. As Klein (2013, p.202) writes, “No matter that the lengths of the [waiting] lists were an ambiguous indicator of performance. No matter that they were, if anything, a misleading measure of the NHS’s ability to meet demands. Waiting lists were confirmed as the symbol of the NHS’s inability to meet public expectations of quick and ready access to treatment.”

2 DH, 2002. This section also draws on Naylor and Gregory, 2009; Allen and Jones, 2011; Anderson, 2006; BSG, 2005; and HCHC, 2006. ISTCs were also established in Wales and Scotland, but are outside the scope of this paper, given the devolution of the NHS to the constituent countries of the United Kingdom during this period.

3 Two of the 27 Wave 1 ISTCs had parent companies with more mixed ownership structures. Circle Health, which co-ran Nottingham ISTC with the for-profit Nations Healthcare, was to a limited extent a doctors’ mutual. The New York Presbyterian Healthcare System, a charity, owned Specialist Hospitals Ltd., a for-profit company that ran Shetton Mallet ISTC. All other Wave 1 ISTCs were run on an exclusively for-profit basis, and even these two exceptional cases were run with a significant for-profit component.

4 To prevent our estimates being contaminated by effects arising from the Phase 2 ISTC programme, we drop from our estimation sample any NHS hospital whose closest ISTC was a Phase 2 ISTC, provided that this ISTC was close enough to act as a competitor (using a criterion set out in Section 4).

5 ASA 1: Healthy patient with localized surgical pathology and no systemic disturbance; ASA 2: Patient with mild to moderate systemic disturbance (i.e. surgical pathology or other disease process); ASA 3: Patient with severe systemic disturbance from any cause; ASA 4: Patient with life threatening systemic disorder which severely limits activity; ASA 5: Gravely ill patient with little chance of survival.
costlier, more complex patients onto the public hospital system (Wallace, 2006; Kmitowicz, 2006).

3. Hypotheses on the impact of ISTC entry on incumbent public hospitals

This section examines the likely response of public (NHS) hospitals to the entry of private, for-profit surgical centres (ISTCs). In understanding the impact of the ISTC programme, it is important to note that, although public NHS hospitals are run on a not-for-profit basis, they are financially and managerially independent of central government, and during this period had strong incentives to generate a financial surplus, or at least not to make substantial losses.

In the early 2000s, the government introduced a system of ‘star rating’ of NHS hospitals, in which financial performance was a major factor (Bevan and Hood, 2006a, 2006b; DH, 2002). Hospitals given a zero-star rating were ‘named and shamed’, and their chief executives were at risk of losing their jobs. Later, high-performing hospitals (those with Foundation Trust status) were given additional freedoms to retain financial surpluses across financial years. Other hospitals were eventually able to achieve Foundation Trust status in part through good financial performance. These factors meant that, during this period, public hospitals had a strong incentive to generate operating surpluses. It has therefore been argued that it is reasonable to think of public hospitals during this period as maximising profits plus some additional term reflecting altruistic valuation of quality and/or quantity (Gaynor et al., 2013).

Ultimately, Wave 1 ISTCs differed from public hospitals in three key dimensions. First, they were explicitly for-profit ventures. Second, they were narrowly focused on offering a small range of elective surgical procedures. Third, given their inability to hire NHS staff, their institutional cultures may have differed sharply from those at NHS hospitals. In what follows, keeping in mind these three differences, we present hypotheses about the response of NHS hospitals to the ISTC programme.

3.1. Efficiency

We expect ISTC entry to lead to efficiency improvements at nearby incumbents. As mentioned in Section 1, we measure hospital efficiency using pre-surgery LOS. Prospective reimbursement systems (like PbR in England) pay hospitals on the basis of outputs rather than inputs. This creates incentives for hospitals to reduce marginal costs by shortening patient LOS (Cutler, 1995). Empirical studies of England (Farrar et al., 2009), the United States (Feder et al., 1987; Guterman and Dobson, 1986; Feinglass and Holloway, 1991; Kahn et al., 1990), Israel (Shmuelli et al., 2002) and Italy (Louis et al., 1999) provide evidence that prospective reimbursement leads to shorter LOS.

While prospective reimbursement systems provide incentives for all hospitals to reduce patient LOS, these incentives will likely be particularly sharp in more competitive markets. Hospitals located in less competitive markets likely have limited scope to expand their activity because they are constrained by the relative inelasticity of clinical demand within their catchment areas. By contrast, hospitals in more competitive markets have greater opportunity to expand activity by capturing market share from other hospitals. To create capacity for such expansion, in health care systems with prospective reimbursement, hospitals in more competitive markets are likely to take stronger action to reduce patient LOS, so that they can treat additional patients. Consistent with this hypothesis, studies of the 2006 patient choice reforms in the English NHS found that hospitals located in more competitive markets decreased their LOS by larger amounts than hospitals in less competitive markets (Cooper et al., 2012; Gaynor et al., 2013). In light of this theoretical prediction and empirical evidence, we hypothesise that incumbent hospitals exposed to entry by an ISTC will have reduced patient LOS over and above any secular decreases in LOS resulting from the introduction of prospective reimbursement. We therefore identify the effect of ISTC entry on the efficiency of nearby public hospitals using a DID estimator in which the treatment effect equals the change in efficiency at ISTC-exposed public hospitals minus the change in efficiency at unexposed public hospitals.

During the 2000s, the government announced that performing elective surgery on the day of a patient’s admission was a key measure of hospital performance, and highlighted that ISTCs would be particularly effective at this. The NHS Institute for Innovation and Improvement (2006, 2008a, 2008b) identified surgery on day of admission as one of the six characteristics of high-performing orthopaedic surgical facilities and argued (2006, p.20) that public hospitals would have to respond to competition from private entrants by streamlining their production: “Same-day admissions [i.e. admission on day of surgery] are seen as imperative by independent [private] providers. Acute [public] trusts will need to reflect this as an integral element of any market strategy when seeking to demonstrate competitive advantage.” This explicit focus on admission on day of surgery means that, in addition to the more general incentives to increase efficiency brought about by surgical centre entry, we expect public hospitals facing increased pressure from private surgical centres to have improved their performance in this dimension in particular.

3.2. Case mix

The entry of private, for-profit surgical centres could also change the case mix at nearby incumbents due to risk selection by entrants. Whereas in classical private goods markets the profitability of selling to a particular customer is determined solely by their willingness to pay, in health care markets – as in many other markets for the provision
of public services, such as social care and education – the profitability of treating a given customer will be influenced by characteristics of the customer that are often imperfectly observed.

The influence of patient characteristics on profitability provides all hospitals with an incentive to refuse to treat the sickest patients. However, private, for-profit entrants like ISTCs are likely to be more willing than public hospital incumbents to actively select against costly patients, as for-profit firms are able to redistribute profits to shareholders, whereas public hospitals are, at most, only allowed to reinvest profits into the organisation. The literature on specialty hospitals in the US, for example, has found evidence that these providers select low-risk patients, leaving the sickest patients to nearby general hospitals (Barro et al., 2006; Winter, 2003; Cram et al., 2005).

Two further factors add weight to the hypothesis that ISTCs had stronger incentives to risk-select than public hospital incumbents. First, ISTCs could legally decline to treat complicated cases, whereas public hospitals were formally prohibited from doing so. Second, as mentioned previously, ISTCs were prohibited from using NHS doctors, so their workplace culture likely differed sharply from that at incumbents. As Rose-Ackerman (1996) notes, the culture of staff plays a key role in dictating firm behaviour – thus these cultural differences may have led ISTCs to be more willing than NHS providers to engage in profit-driven risk-selection.

Prospective reimbursement encourages cream-skimming, since it provides incentives for hospitals to avoid admitting patients whose cost of treatment is likely to exceed the regulated payment (Allen and Gertler, 1991; Ellis and McGuire, 1986; Ellis, 1998; Newhouse, 1989). We use DiD methods to estimate the extent to which ISTCs left incumbent NHS hospitals with a sicker, costlier mix of patients, over and above any secular changes in case mix over this period (either as a result of the introduction of prospective reimbursement, or for other reasons).

Previous studies have confirmed that ISTCs treated healthier and less complex patients than nearby public hospitals (Street et al., 2010; Mason et al., 2008, 2010; Browne et al., 2008; Chard et al., 2011; Fagg et al., 2012). However, no one has yet compared the evolution of average patient severity at ISTC-exposed public hospitals with that at public hospitals unaffected by the ISTC programme, and shown that ISTC-exposed public hospitals experienced a larger reduction in average patient health status (measured using a Charlson Index) than public hospitals not exposed to the entry of an ISTC. Providing evidence of such an effect of ISTC entry is important because the case mix differences between ISTCs and nearby public hospitals documented by the existing literature may simply reflect the fact that ISTCs attracted patients who would not otherwise have undergone surgery.7

4. Data, definition of treatment group, and estimation strategy

Our aim is to estimate the causal effect of the entry of private surgical centres on the efficiency, case mix, and case load of nearby incumbent public hospitals. We use difference-in-difference (DiD) regressions in which the impact of ISTC exposure is estimated from the mean change in outcomes for public hospitals in a treatment group (those that had a private surgical centre placed nearby) minus the mean change in outcomes for public hospitals in a control group (those that did not have a private surgical centre placed nearby) before and after entry occurred. This section describes our outcome measures, construction of treatment groups, and identification strategy.

4.1. Data and outcome variables

Our dataset is derived from the NHS Hospital Episode Statistics (HES) (HSCIC, 2016), which contains the universe of government-funded hospital admissions in England.8 Our data extract consists of all elective hip and knee replacements on patients aged 55–100 performed between financial years 2002/3 and 2008/9 (see Table 1). We focus on hip and knee replacements for two reasons. First, orthopaedic surgery was a major focus of the ISTC programme, as it was recognised in the early 2000s that achieving the government’s waiting time targets was going to be more challenging in this surgical specialty than in any other area (Harrison and Appleby, 2005). Second, clinical practice in relation to hip and knee replacements did not change significantly during this period in ways that could affect LOS. As a result, any observed changes in LOS will likely be driven by NHS reforms, not by differential uptake of new medical technologies.

We focus on hip and knee replacements performed in NHS hospitals. NHS hospital trusts (firms) often consist of multiple hospitals (individual sites) that can be located up to 100 km away from each other. We therefore analyse the data at site (hospital) level rather than trust (firm) level, and assign hospitals (sites) to treatment and control groups based on the site’s proximity to the nearest ISTC. All references to ‘hospitals’ in this paper are therefore to hospital sites, not to trusts (firms). After cleaning and imputing missing values for the site code field, and applying exclusion criteria detailed below, there are 166 public hospitals treating hip and knee replacement patients from 2002/3 to 2008/9.

Researchers have generally struggled to quantify hospital efficiency. In the absence of hospital cost data, many studies use proxy measures of efficiency such as LOS (Fenn and Davies, 1990; Martin and Smith, 1996; Gaynor et al., 2013). The logic underlying this measure is that, if a hospital can treat patients more quickly without any deterioration in clinical quality, then it must have become more efficient. However, a critical shortcoming of overall LOS as an efficiency measure is that recovery time after surgery is also heavily dependent on patient characteristics and health status. Moreover, a hospital’s average LOS may reflect undesirable hospital behaviour such as cream skimming (prioritising treatment of less costly patients); dumping (avoiding treatment of costlier patients); and quality skimming (discharging patients ‘sicker and quicker’) (Epstein et al., 1990; Martin and Smith, 1996; Sudell et al., 1991).

In this study, we use an innovative method to obtain a cleaner proxy for hospital efficiency. We decompose LOS for hip and knee replacements into two parts: the time from admission to surgery (pre-surgery LOS), and the time from surgery until discharge (post-surgery LOS). We hypothesise that, for elective orthopaedic surgery, pre-surgery LOS is not significantly influenced by patient characteristics, as there is rarely a clinical rationale for admitting an elective orthopaedic surgery patient before the scheduled day of their operation. In the early 2000s, fewer than 20% of hip and knee replacement patients had surgery on the day they were admitted to the hospital. Patients were often kept overnight before elective surgery not for clinical reasons, but because operating...
denoting the absence of any predictors of mortality (HSCIC, 2013). As proxies for health status and clinical risk, we also use the patient's age, as well as the IMD income domain (Noble et al., 2004), which reports the percentage of households in the patient's residential Lower Super Output Area (LSOA, a statistical geographical areas containing around 1500 residents) that are income deprived (in our dataset this variable ranges from 0 to 83).

The data includes 478,226 hip and knee replacements performed from 2002/3 to 2008/9 that met the sample restrictions. As Table 1 illustrates, during the analysis period pre-surgery LOS, post-surgery LOS, and total LOS fell considerably.  

4.2. Treatment assignment

We assign public hospitals to treatment or control groups based on their geographical proximity to the new market entrants, on the assumption that exposure to competition from these entrants is a product of proximity. In particular, we assign treatments by comparing the distance from an NHS hospital to its nearest ISTC with the percentiles of distance travelled by that hospital’s hip and knee replacement patients.

We measure the distance travelled by each hip and knee replacement patient to hospital using the centroids of a patient’s residential LSOA to define home location. We then calculate, for each NHS hospital, percentiles of patient distance travelled (e.g. the distance that captures 25% of a hospital’s hip and knee replacement patients). Percentiles of patient distance travelled can be endogenous to hospital quality – for example, a high-quality hospital may attract patients from further afield. To ameliorate this concern, we use percentiles of patient distance travelled based on patient flows from 2002/3 to 2004/5 (i.e. before implementation of either the ISTC programme or patient choice of hospital for elective surgery).

Table 3 presents descriptive statistics for quantiles of patient distance travelled and the exposure of NHS hospitals to Wave 1 ISTCs within each of these quantile bands. Panel A presents the mean kilometre distances (averaging over the NHS hospitals in our estimation sample) corresponding to these patient percentile travel distances. The mean value of the 25th percentile of travel distance for hip and knee replacement patients is 4.25 km, while for the 95th percentile it is 9.6 km.

9 Personal communication with NHS physicians supported this claim. We recognise too that not all pre-surgery LOS is wasted time and there may, in rare cases, be clinical reasons to admit patients before the day of surgery. However, unless this clinical need is correlated with ISTC treatment status, this issue will not threaten our strategy for identification of the effects of ISTC treatment on efficiency.

10 The Charlson score ranges from 0 to 130. However, only around 8% of patients have a score above 6. We therefore cap the score at 6, to ensure that our results for this outcome variable are not driven by outliers.

11 Appendix A provides further information about how hip and knee replacement patients are identified, how the LOS-related variables are constructed, and how missing site codes are imputed. Separate versions of Table 1 for hip and knee replacements are reported in Table B.1. Descriptive statistics for all variables used in this paper are reported in Table B.2. Because our results for hip replacements and knee replacements were very similar, we pooled together these surgical procedures in the estimates reported in the main body of this paper and included a dummy variable indicating hip replacement, to capture level differences in the outcome variable across the two surgical procedures. Tables F.1 to F.4 show that our results are qualitatively unchanged when separate treatment effects are estimated for hip replacements and knee replacements.
near inef
enous to our primary outcome (LOS) because ISTCs may have opened
estimates.13

high LOS. Indeed, as we observe, hospitals’ LOS is uncorrelated with
reason to expect hospitals with high waiting times to necessarily have
are dependent on a wide range of demand and supply side capacity-
waiting times. In the robustness section, we show the effect of
NHS hospitals in the High Treatment group, 51 in the Low Treatment
and control groups will confound our DiD estimates.

We allocate NHS hospitals to treatment categories by comparing dis-
tance to ISTC with percentiles of patient distance travelled, not with
kilometre distances, to control for rural-urban differences – treatment
assignment based on fixed distances will over-estimate the size of
markets in urban areas relative to rural areas, given the impact of
urban congestion on travel speeds. In the robustness tests, we examine
whether our results change if we use a treatment assignment strategy
based on fixed distances from public hospital to ISTC.

4.3. Treatment start and end dates

There is some ambiguity as to the appropriate way to define the
policy-on and policy-off dates for a given public hospital exposed to
ISTC entry. As the first ISTCs in our analysis opened in April 2005
(financial year 2005/6), we define 2004/5 as the last pre-treatment
(pre-ISTC-programme) financial year.14 However, some ISTCs did not
begin operations until 6 months to a year after their contracted start
date. Moreover, when the initial ISTC contracts (generally around 4 or
5 years in length) were completed, some managed to secure further
contracts, but others were shut down or absorbed into neighbouring
NHS trusts. The fate of an ISTC was generally announced in the last
year of the contract. Thus, if contract end date were used as treatment
end date, estimates of treatment effects could be confounded by
changes in behaviour due to anticipated contract completion.

In response to these ambiguities, we employ a long differences
specification using data from the 2004/5 and 2008/9 financial years.
We choose 2004/5 as the pre-treatment period in our main specification
because it is the last year before the first contract start date amongst
the ISTCs we use for treatment assignment, and thus most likely to capture
the effect of ISTC exposure as distinct from the effect of other elements
of the government’s reform programme. We choose 2008/9 as the post-
treatment period to allow time for treatment effects to be realised,
while avoiding contamination from responses to announcements
concerning extension or non-extension of ISTC contracts. For

---

13 The correlation between hospitals’ average total LOS and average waiting time for hip
and knee replacement surgery in 2002/3 is 0.06, p-value 0.46. Simple bivariate regres-
sions of the log of average total LOS on log of average waiting time during this period yield
a tiny and statistically insignificant coefficient (0.03, p-value 0.275). We take this as evi-
dence that selection for ISTC placement on the basis of the average waiting times of nearby
NHS hospitals does not imply selection, via correlation, on the basis of nearby hospitals’
average LOS.

14 All references to years in this paper are to financial years. Three Wave 1 ISTCs had con-
tact start dates before April 2005 but are not included in our analysis, either because they
did not have (sufficient) orthopaedic capacity or because of ambiguity over location and
differentiation from a nearby NHS hospital. See Appendix A for further details.
robustness, we show how treatment effects change annually in the post-reform years from 2005/6 to 2008/9, and report estimates that define a public hospital’s treatment start date as the contract start date of its nearest ISTC.

4.4. Regression specification

We identify the impact of hospital market entry using a DiD regression framework where dummy variables indicating treatment group membership are interacted with a post-policy dummy, which is switched on for the 2008/9 financial year. Regressions are run at the patient level and any non-binary dependent variables are log-transformed (after adding 1 to any variables that have a minimum value of zero) such that the treatment effects are interpretable as percentage changes.

Our basic DiD specification is:

\[ y_{ijt} = \beta_0 + \beta_1 \text{post}_t + \beta_2 \text{high}_j + \beta_3 \text{low}_j + \beta_4 \left( \text{high}_j \times \text{post}_t \right) \\
+ \beta_5 \left( \text{low}_j \times \text{post}_t \right) + \epsilon_{ijt} \]  

(1)

In this specification, \( t \) denotes the time period (financial year), \( \text{post}_t \in \{0,1\} \) denotes whether an observation occurs in the post-reform period, \( y_{ijt} \) denotes the outcome variable under consideration for patient \( i \) attending hospital \( j \) at time \( t \), and \( \text{high} \) and \( \text{low} \) denote dummies for the High and Low Treatment groups respectively. Treatment effects are given by the coefficients on the interaction terms, \( \beta_4 \) and \( \beta_5 \). We also include a dummy in our regressions to control for the type of procedure a patient undergoes (hip or knee surgery) but suppress this in the notation for simplicity.

Our second specification includes hospital fixed effects (\( \mu_j \)) in place of the treatment group indicators to capture all time-invariant hospital and spatial characteristics, and time-period-specific (month-year) fixed effects (\( \eta_h \)) in place of the post-policy period dummy:

\[ y_{ijt} = \beta_0 + \beta_1 \left( \text{high}_j \times \text{post}_t \right) + \beta_2 \left( \text{low}_j \times \text{post}_t \right) + \eta_h + \mu_j + \epsilon_{ijt} \]  

(2)

Our third specification is identical to (2), but includes an extensive set of controls for patient and hospital characteristics.\(^\text{15}\) All specifications are estimated using ordinary least squares (OLS), with standard errors clustered at the hospital level to account for correlation in unobservables within hospitals (between patients and over time).

There are two core threats to our identification strategy. The first is that trends in outcome variables may not have been parallel between treatment and control groups prior to ISTC entry. The second is that there may have been other time-varying policy changes that also affected outcomes concurrently with the ISTC programme. To address the first possibility, we demonstrate that treated and control hospitals had parallel trends for our key dependent variables before the ISTC programme was launched by showing trends in the data graphically, and formally testing for statistically significant differences in trends. To address the risk that concurrent and correlated policy shocks drive our results, Section 5.4 discusses the two most prominent policy

\(^{15}\) Thus, while our second specification estimates the effect of ISTC exposure on public hospital performance inclusive of any effects via changing patient characteristics (e.g. due to risk selection of patients by ISTCs), our third specification estimates the effect of ISTC exposure on public hospital performance conditional on observable patient characteristics.
changes that could have affected outcomes contemporaneously with ISTC entry – the introduction of hospital competition via patient choice of hospital for elective surgery, and the enactment of differential health policies by Strategic Health Authorities (SHAs) – and illustrates that controlling for these policy changes does not materially influence our main estimates.

5. Results

5.1. Descriptive evidence

Fig. 4 presents the evolution of key outcome variables – pre-surgery LOS, percentage treated on day of admission, post-surgery LOS, total LOS and Charlson score – between 2002/3 and 2008/9, for treatment and control groups. The shaded area represents the range of treatment start dates for the Wave 1 ISTCs. We expect that any treatment effects will arise either within the time period captured by shaded region or, if behavioural responses took place with a lag, some time thereafter. Each data point represents a month, but the plots are smoothed using a moving average of the month and the previous quarter. These graphs allow visual examination of whether treated and control groups followed similar trends prior to the entry of ISTCs.

Panel A shows changes in pre-surgery LOS and illustrates that High Treatment group, Low Treatment group, and Control group hospitals follow similar trends before ISTC entry. Over and above a secular downward trend, reflecting general improvements in turnaround time, there is evidence of a treatment effect from ISTC entry. After ISTC entry, trends diverge, and by the end of the treatment period the reduction in pre-surgery LOS is notably larger for the High Treatment group than for the Control group. There also appears to be a smaller effect for the Low Treatment group. Panel B shows trends in the percentage of patients treated on the day of admission. All three groups have similar pre-entry trends, but after ISTC entry the percentage of patients treated on the day of admission increases more quickly for High Treatment group hospitals. Overall, Panels A and B provide visual evidence that the entry of private specialty surgical centres in the English NHS made nearby public hospitals more efficient, by reducing pre-surgery delays.

Panel C shows trends in post-surgery LOS, which have a markedly different pattern. The High Treatment group, Low Treatment group, and Control group hospitals have similar pre-entry trends. However, there is a sharp increase in post-surgery LOS for the High Treatment group from the middle of the ISTC entry period onwards. Overall, after ISTC entry post-surgery LOS decreases in the High Treatment group less than in the Control group. Panel D presents trends in total LOS, which follow a similar pattern to post-surgery LOS. As discussed in Section 4.1, post-surgery LOS (and therefore total LOS) will be influenced both by changes in hospital efficiency due to increased competitive pressure from the entry of private surgical centres, and by changes in patient characteristics due to cream skimming by entrants. Panels C and D therefore provide suggestive evidence that the negative impact of ISTC cream skimming on nearby public hospitals’ LOS may have outweighed any efficiency improvements with respect to LOS arising from competitive pressure from these new market entrants.

Panel E looks more directly at the impacts of ISTC entry on public hospitals’ case mix by plotting the evolution of average Charlson scores. The pre-policy levels and trends of the Charlson score are similar across treatment and control groups. However, the High Treatment group starts receiving sicker patients from early in the ISTC entry period. This evidence is consistent with our hypothesis that selection of less risky patients by ISTCs left a residual pool of higher-risk patients to be treated by public hospitals. Graphical evidence for other case mix variables is presented Appendix D.

Overall, the similar pre-policy trends in treatment and control groups for all outcome variables provides strong support for our argument that DID estimates are likely to provide an unbiased estimate of treatment effects from ISTC entry. The fact that pre-policy trends (and in many cases levels) of our outcome variables are similar across treated and untreated groups is consistent with our argument that the principal target of ISTC placement was to reduce waiting times for admission to hospital, not to reduce time spent in hospital or to improve clinical quality.16

5.2. Regression-based difference-in-difference estimates

Table 5 presents our main difference-in-difference estimates of the effect of ISTC entry on log of pre-surgery LOS, percentage of patients treated on day of admission, and log of post-surgery LOS. The sample includes hip and knee replacements, with a hip replacement dummy included to account for level differences in outcomes between the two procedures.

In Columns (1) to (3), the dependent variable is log of pre-surgery LOS. Column (1) presents estimates of Eq. (1), without hospital fixed effects or patient controls. The estimate implies that ISTC entry led to a 14.4% (=100(e^{0.156} – 1)) reduction in pre-surgery LOS. In Column (2), we estimate Eq. (2), adding hospital and month × year fixed effects. The results are qualitatively unchanged and imply that ISTC entry led to a 16.1% reduction in pre-surgery LOS. In Column (3), we add patient controls and the results again remain qualitatively unchanged (we find a 16.6% reduction). That controlling for patient characteristics barely shifts the estimated treatment effects for the High Treatment group suggests that there is little selection into treatment on the basis of these observable demographic characteristics. This, in turn, implies that there is likely to be little selection into treatment on the basis of unobservable patient characteristics (Altonji et al., 2005). The impact on the Low Treatment group is of the same sign and around one third the magnitude of the High Treatment group effect, but is imprecisely measured and never significant at conventional levels. The most likely interpretation is that there were moderate

<table>
<thead>
<tr>
<th></th>
<th>(1) High treatment group</th>
<th>(2) Low treatment group</th>
<th>(3) Control group</th>
<th>(4) Low treatment group + control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waiting times</td>
<td>271.4</td>
<td>248.5</td>
<td>259.8</td>
<td>256</td>
</tr>
<tr>
<td>Pre-surgery length of stay</td>
<td>0.889</td>
<td>0.923</td>
<td>0.925</td>
<td>0.925</td>
</tr>
</tbody>
</table>

Notes: Table reports average waiting times, total length of stay (LOS), pre-surgery LOS, and post-surgery LOS for hip and knee replacement patients at treatment and control group hospitals in 2002/3 (i.e. before commencement of the ISTC programme).

16 We deliberately choose not to study the effect of ISTC entry on nearby public hospital waiting times in the main body of this paper because ISTCs were strategically located near hospitals with high waiting times, implying that treatment is endogenous with respect to this outcome variable. Nevertheless, for interested readers, the Appendix presents graphical evidence concerning the evolution of waiting times at treatment and control group hospitals, and DID estimates corresponding to this graphical evidence. The accompanying text explains in further detail why our DID estimates cannot be interpreted as causal effects of ISTC entry on waiting times.
impacts of ISTC entry on the Low Treatment group, but that our research design does not have sufficient power to detect them.

Columns (4) to (6) examine the effect of ISTC entry on the proportion of patients who were treated on the day of admission. The results are similar with and without hospital fixed effects, and with and without patient controls. The results in Column (5), which estimates Eq. (2), shows that ISTC entry led to a 24.3 percentage point reduction in the proportion of patients treated on the day of admission at High

Fig. 4. Trends in key outcomes variables. Notes: Sample includes hip and knee replacement patients. Treatment groups are as defined in Table 5 notes. Graphs show moving averages of hospital means calculated over a month and the previous quarter. Shaded area marks main period of Wave 1 ISTC entry. Figs. F.1 and F.2 provide versions of figures for hip replacement patients and knee replacement patients separately.
Treatment group hospitals, from a baseline of 22.8% in 2004/5. As with pre-surgery LOS, the impact on the Low Treatment group is same-signed, smaller, and never significant at conventional levels.

Columns (7) to (9) present estimates of the effect of ISTC entry on post-surgery LOS. In Column (7), we estimate that the entry of ISTCs led to an 8.47% increase in post-surgery LOS at High Treatment group hospitals. However, this effect is only significant at the 10% level. The precision and magnitude of our estimates is reduced when hospital and month × year fixed effects are included, in Column (8). Adding patient controls further reduces the size of the point estimate and decreases precision, as we would expect if patient characteristics affect post-surgery LOS and it is selection of less riskier patients into ISTCs and out of NHS hospitals which drives the treatment effects on post-surgery LOS. That patient controls reduce the magnitude of the post-surgery LOS estimates, but have little impact on the pre-surgery LOS estimates, provides further evidence that patient characteristics are a major driver of post-surgery LOS, but have little influence on pre-surgery LOS.

We interpret changes in post-surgery LOS resulting from ISTC entry as a joint product of (i) changes in the mix of patients being treated by public hospitals, due to cream skimming by neighbouring ISTCs and (ii) behavioural responses by public hospital managers and clinicians to competition from new private entrants. Although only significant at the 10% level, the estimates in Column (7) suggest that the increases in post-surgery LOS generated by cream-skimming were larger than the reduction in LOS generated from any efficiency improvements reported in Columns (1) to (6) of Table 5. We estimate Eq. (1) and Eq. (2) with hospital and month × year fixed effects; no specifications include patient controls. Columns (1) to (4) indicate that ISTC entry led to an 11.6% increase in the average Charlson score at High Treatment group hospitals — or a 6.2 percentage point increase in the proportion of patients with a Charlson score of three or more — significant at the 5% level. Column (5) indicates that ISTC entry led to a 5.54% increase in the IMD income deprivation score at High Treatment group hospitals, although this point estimate reduces in magnitude and becomes imprecise when hospital and month × year fixed effects are added. We do not find that ISTC entry led to a precisely estimated increase in patient age at nearby NHS hospitals.

Table 7 presents event study estimates of year-by-year effects of ISTC entry on log of pre-surgery LOS, percentage of patients treated on day of admission, and log of Charlson score at exposed NHS hospitals, from 2004/5 to 2008/9. The estimates mirror our graphical evidence in Fig. 4 and show that ISTC entry had a statistically significant effect on these outcomes at exposed NHS hospitals from 2007/8 onwards.

Appendix J explores the impact of ISTC entry on clinical quality at NHS hospitals by analysing changes in 30-day in-hospital mortality from acute myocardial infarction (AMI) at nearby public hospitals. We find that, after controlling for patient characteristics, ISTC entry did not have a statistically significant effect on AMI mortality at nearby public hospitals. The results suggest that the efficiency improvements reported in Columns (1) to (6) of Table 5 were achieved without any evidence of concomitant deterioration in clinical quality.

5.3. Treatment assignment using fixed distances

Table 8 presents estimates of the effect of ISTC entry when treatment assignment is based not on patient flows but on fixed kilometre distances between ISTCs and NHS hospitals. Specifically, the High Treatment group comprises NHS hospitals that had an ISTC enter within 5 km, the Low Treatment group comprises NHS hospitals that had an ISTC enter within 30 km (but not within 5 km), and the Control group comprises NHS hospitals that did not have an ISTC enter within 30 km. Using this definition, the High Treatment group contains 14 hospitals, the Low Treatment group 78 hospitals, and the Control group 77 hospitals. Table 8 indicates that ISTC entry within 5 km of an NHS

---

Table 5: Impact of ISTC entry on length of stay at nearby public hospitals (treatment defined using patient flows).

<table>
<thead>
<tr>
<th></th>
<th>Log of pre-surgery length of stay</th>
<th>% treated on day of admission</th>
<th>Log of post-surgery length of stay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>High treatment × post</td>
<td>−0.156**</td>
<td>−0.176***</td>
<td>−0.181***</td>
</tr>
<tr>
<td></td>
<td>(0.0636)</td>
<td>(0.0659)</td>
<td>(0.0663)</td>
</tr>
<tr>
<td>Low treatment × post</td>
<td>−0.0455</td>
<td>−0.0642</td>
<td>−0.0649</td>
</tr>
<tr>
<td></td>
<td>(0.0480)</td>
<td>(0.0459)</td>
<td>(0.0465)</td>
</tr>
<tr>
<td>Hospital fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>141,661</td>
<td>141,661</td>
<td>141,443</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.204</td>
<td>0.394</td>
<td>0.401</td>
</tr>
<tr>
<td>Mean dependent var</td>
<td>0.846</td>
<td>0.846</td>
<td>0.846</td>
</tr>
</tbody>
</table>

Notes: Sample includes hip and knee replacement patients. High and Low Treatment groups are defined using percentiles of patient distance travelled in 2002/3 to 2004/5. Treatment groups are: High — hospitals with Wave 1 ISTC within their 25th percentile patient travel distance; Low — hospitals with a Wave 1 ISTC between their 25th and 95th percentile patient travel distance; Untreated — hospitals without an ISTC in their 95th percentile patient travel distance. Our pre-reform period is 2004/5; our post-reform period is 2008/9. Mean Dependent Variable reports the average value of the outcome variable (not logged) in our pre-reform period, 2004/5. Standard errors are clustered at the hospital level. **p < 0.01, ***p < 0.05, *p < 0.1. All regressions include a dummy variable indicating hip replacement. The regressions in Columns (1), (4) and (7) include dummy variables indicating the High Treatment group, Low Treatment group and post-reform period. Coefficients on these variables are always near-zero and statistically insignificant. Table E.1 reports coefficient estimates for these variables. The specifications reported in Columns (3), (6) and (9) include Charlson score; number of diagnoses; IMD income deprivation score; IMD health and disability deprivation score; dummy variables indicating procedure type (hip replacement cemented, hip replacement uncemented, revision to hip replacement, knee procedure revisions); indicators for self-discharge, urban residence, mixed ethnicity, Asian ethnicity, black ethnicity, other ethnicity, and unknown ethnicity; and a full set of case mix dummies with gender interacted with five-year age bins. Table E.3 reports coefficient estimates for these variables. Tables F.1 and F.2 report separate estimates for hip replacements and knee replacements respectively.

---

17 The coefficients on the High Treatment and Low Treatment indicators are reported in Table E.1 of Appendix E.
hospital led to a 14.7% reduction in pre-surgery LOS, a 21.9 percentage point increase in the share of patients treated on the day of admission, and an 11.2% increase in the average Charlson score.

5.4. Controlling for contemporaneous NHS policy changes

One concern with our DiD identification strategy is that the resulting estimates may be biased by other policies, implemented concurrently with the ISTC programme, that had a differential effect on our outcome variables across treated and control groups. The most prominent such policy is the 2006 introduction of hospital competition within the NHS via patient choice of hospital for elective surgery. In addition, during this era much NHS policy was dictated by ten regional Strategic Health Authorities (SHAs) – differences in SHA policies implemented during the ISTC period may bias our results, to the extent that ISTC entry was differentiated across regions of England. We investigate these possibilities in Table 9; all reported estimates use as their baseline Eq. (2).

Columns (1) to (3) test whether the estimates reported in Tables 5 and 6 are robust to controlling for the 2006 patient choice reforms. We control for overall competition intensity by including a measure of market concentration (a time-invariant negative log of hospital HHI) interacted with a post-ISTC-entry dummy variable.18 If the patient choice reforms were driving the results, inclusion of this interaction term would severely attenuate our estimates. However, including this interaction term does not materially change the estimates – they remain precisely estimated and similar in magnitude.

Columns (4) to (6) test for other region-specific policy changes that could be driving our results. To do so, we interact dummies for each of the ten English SHAs with a post-ISTC-entry dummy. These additional controls do not materially impact our results. The results are also robust to controlling more flexibly for differential SHA policies and regional trends via separate SHA × year or SHA × year × month interaction terms (see Appendix G). Overall, the estimates reported in Table 9 provide assurance that our main estimates are not driven by the most worrisome potential sources of bias from contemporaneous policy changes.

5.5. Altering the threshold used to define treatment exposure

So far, we have defined the High Treatment group to include any public hospital that had an ISTC enter within the 25th percentile of patient distance travelled between 2002/3 and 2004/5, or, in the alternative specification reported in Table 8, within 5 km. Figs. 5 and 6 show how the estimates of Eq. (2) change as we vary the treatment group definition thresholds. Fig. 5 (Panel A for log of pre-surgery LOS and Panel B for log of Charlson score) shows how the estimates change when the threshold used to define the treatment group changes from one that captures 15% of a hospital’s hip and knee replacement patients, to one that captures 95%. Panels A and B illustrate that the estimated treatment effects decrease as the treated group is defined more widely.

Notes: Sample includes hip and knee replacement patients. Table reports treatment effects in the 4 years after commencement of the ISTC programme, with 2004/5 used as the base year. Treatment groups, mean dependent variable, standard errors and statistical significance are defined as in the Table 5 notes. The reported treatment effects are not cumulative (e.g. the coefficient on High Treatment × 2005/6 is not turned on for observations in years after 2005/6). The coefficients on High Treatment × 2008/9 and Low Treatment × 2008/9 estimate the same effect as our main results reported in Tables 5 and 6, but differ slightly because the estimated hospital fixed effects are different due to the inclusion of additional years of data. All regressions include a dummy variable indicating hip replacement.

Notes: Sample includes hip and knee replacement patients. Table reports treatment effects in the 4 years after commencement of the ISTC programme, with 2004/5 used as the base year. Treatment groups, mean dependent variable, standard errors and statistical significance are defined as in the Table 5 notes. All regressions include a dummy variable indicating hip replacement. In addition, the regressions in Columns (1), (3) and (5) and (7) include dummy variables indicating the High Treatment group, Low Treatment group and post-reform period. Coefficients for these variables are reported in Table E2. Tables F3 and F4 report separate estimates for hip replacements and knee replacements respectively.

Table 6

<table>
<thead>
<tr>
<th>Log of Charlson</th>
<th>Charlson score 3 or more</th>
<th>Log of IMD income deprivation score</th>
<th>Log of age (coefficients × 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>High treatment × post</td>
<td>0.111** (0.0522)</td>
<td>0.0628** (0.0297)</td>
<td>0.0539** (0.0254)</td>
</tr>
<tr>
<td>Low treatment × post</td>
<td>−0.00387 (0.0244)</td>
<td>0.000865 (0.0144)</td>
<td>−0.0266 (0.0262)</td>
</tr>
<tr>
<td>Hospital fixed effects</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Month × year fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>141,661</td>
<td>141,661</td>
<td>141,443</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.008</td>
<td>0.035</td>
<td>0.019</td>
</tr>
<tr>
<td>Mean dependent variable</td>
<td>0.005</td>
<td>0.035</td>
<td>0.035</td>
</tr>
</tbody>
</table>

Notes: Sample includes hip and knee replacement patients. Treatment groups, pre-reform and post-reform period, mean dependent variable, standard errors and statistical significance are defined as in the Table 5 notes. All regressions include a dummy variable indicating hip replacement. In addition, the regressions in Columns (1), (3) and (5) and (7) include dummy variables indicating the High Treatment group, Low Treatment group and post-reform period. Coefficients for these variables are reported in Table E2. Tables F3 and F4 report separate estimates for hip replacements and knee replacements respectively.

Table 7

<table>
<thead>
<tr>
<th>Log of pre-surgery LOS</th>
<th>% treated on day of admission</th>
<th>Log of Charlson score</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>High treatment × 2005/6</td>
<td>−0.00843 (0.0405)</td>
<td>0.0114 (0.0566)</td>
</tr>
<tr>
<td>High treatment × 2006/7</td>
<td>−0.0904 (0.0778)</td>
<td>0.127 (0.107)</td>
</tr>
<tr>
<td>High treatment × 2007/8</td>
<td>−0.149 (0.0798)</td>
<td>0.205 (0.111)</td>
</tr>
<tr>
<td>High treatment × 2008/9</td>
<td>−0.171* (0.0661)</td>
<td>0.236** (0.0921)</td>
</tr>
<tr>
<td>Low treatment × 2005/6</td>
<td>−0.0129 (0.0358)</td>
<td>0.0214 (0.0510)</td>
</tr>
<tr>
<td>Low treatment × 2006/7</td>
<td>−0.0369 (0.0444)</td>
<td>0.0552 (0.0633)</td>
</tr>
<tr>
<td>Low treatment × 2007/8</td>
<td>−0.0658 (0.0450)</td>
<td>0.0961 (0.0642)</td>
</tr>
<tr>
<td>Low treatment × 2008/9</td>
<td>−0.0554 (0.0450)</td>
<td>0.079 (0.0642)</td>
</tr>
<tr>
<td>Hospital fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Month × year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Patient controls</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>355,551</td>
<td>355,551</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.353</td>
<td>0.398</td>
</tr>
<tr>
<td>Mean dependent variable</td>
<td>0.086</td>
<td>0.228</td>
</tr>
</tbody>
</table>

Notes: Sample includes hip and knee replacement patients. Table reports treatment effects in the 4 years after commencement of the ISTC programme, with 2004/5 used as the base year. Treatment groups, mean dependent variable, standard errors and statistical significance are defined as in the Table 5 notes. All regressions include a dummy variable indicating hip replacement.
Table 8
Impact of ISTC entry on length of stay and case mix (treatment defined using straight-line distances between public hospitals and ISTCs).

<table>
<thead>
<tr>
<th></th>
<th>(1) Log of pre-surgery LOS</th>
<th>(2) % treated on day of admission</th>
<th>(3) Log of Charlson score</th>
</tr>
</thead>
<tbody>
<tr>
<td>High treatment</td>
<td>−0.159**</td>
<td>0.219**</td>
<td>0.106**</td>
</tr>
<tr>
<td>Low treatment</td>
<td>−0.0180</td>
<td>0.0263</td>
<td>0.00169</td>
</tr>
<tr>
<td>Hospital fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Month × year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Patient controls</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>141,453</td>
<td>141,453</td>
<td>141,453</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.396</td>
<td>0.446</td>
<td>0.036</td>
</tr>
<tr>
<td>Mean dependent variable</td>
<td>0.846</td>
<td>0.228</td>
<td>0.885</td>
</tr>
</tbody>
</table>

Notes: Sample includes hip and knee replacement patients. The High and Low Treatment groups are defined using straight-line distances from public hospital to ISTC. The High Treatment group consists of NHS hospitals that had a Wave 1 ISTC enter within the circle defined by a 5 km radius. The Low Treatment group consists of NHS hospitals that had a Wave 1 ISTC enter within the circle defined by a 30 km radius, but not within the circle defined by a 5 km radius. The Control group consists of NHS hospitals that did not have an ISTC enter within the circle defined by a 30 km radius. Pre-reform and post-reform period, mean dependent variable, standard errors and statistical significance are defined as in the Table 5 notes. All regressions include a dummy variable indicating hip replacement.

Fig. 6 performs the same exercise when defining treatment exposure using fixed distances from NHS hospital to ISTC, as in Table 8. Panel A shows that ISTC entry within 5 km of an NHS hospital leads to a precisely estimated reduction in log of pre-surgery LOS at incumbents. As the threshold used to define the treatment group increases, the estimated treatment effects lose precision and asymptote to zero. Panel B, for log of Charlson score, shows a similar sensitivity to the threshold used to define treatment exposure. ISTC entry within 5 km of an NHS hospital leads to a large and statistically significant increase in the average Charlson score at incumbents, with estimated treatment effects decreasing as the distance to ISTC used to define the treatment group increases.

Overall, Figs. 5 and 6 demonstrate that our estimated treatment effects are robust to the exact threshold used to define treatment exposure. Furthermore, as the treatment group is expanded, there is a negative gradient to our estimates, which is broadly supportive of the argument that our estimated treatment effects are driven by ISTC exposure.

5.6. Additional tests of robustness

Table 10 reports additional robustness tests for our main estimates of the impact of ISTC entry on log of pre-surgery LOS, percentage of patients treated on day of admission, and log of Charlson score at High Treatment group hospitals. We estimate Eq. (2) unless otherwise noted.

Panel A formally tests for parallel pre-reform trends for the High Treatment group relative to the Control group using a ‘placebo’ DiD regression where 2002/3 is the base year and 2004/5 (the last year before ISTC entry) is the treated year. This regression allows us to explore whether there was a statistically significant difference in the change in key outcomes in our High Treatment group relative to the change in key outcomes for the Control group during our pre-period from 2002/3 to 2004/5. If there were statistically significant differences (i.e. a ‘placebo’ treatment effect), it would illustrate that there was a difference in trends over the pre-reform period that would potentially invalidate our identification strategy. Consistent with the graphical evidence in Fig. 4, none of the 2002/3 to 2004/5 point estimates are statistically significant at conventional levels, confirming that there were no statistically significant differences in trends for key outcomes between our control group and high treatment group from 2002/3 to 2004/5 (our pre-period). Though not reported, we also find no statistically significant treatment effects for the Low Treatment group relative to the Control group over the same period.

To address any residual concerns about comparability of treated and untreated hospitals, Panel B reports inverse propensity score weighted estimates where we weight our treatment and control groups by the inverse of the probability that an observation belongs in its group. To do this, we first calculate the probability of a hospital being assigned to the High and Low Treatment groups based on hospital and average patient characteristics. We then run separate regressions to estimate treatment effects for the High and Low Treatment groups, using as probability weights the inverse of the probability of assignment to the group that the observation was actually assigned to. While it is impossible to reject a null hypothesis of no significant differences between treated and control groups with respect to the above determinants of treatment assignment even before these weights are applied, weighting by the inverse propensity score increases the similarity of the treated and control groups further still, by giving higher weight to control observations that look more like treated observations, and vice versa. The resulting estimates are similar to the headline findings, but more precise. The corresponding (unreported) estimates for the Low Treatment group are not statistically significant.

Our main specification accounts for the fact that different ISTCs commenced operations at different times by using a long differences estimation strategy with 2004/5 as the pre-reform year and 2008/9 as the post-reform year. An alternative approach is to define $t = 0$ (the treatment start date) for each public hospital as the contract start month (or month of ‘full service commencement’) of the nearest ISTC, and to use pre- and post-reform periods defined relative to $t = 0$ rather than using calendar time. We do not use this as our main specification because the contract start date is not always an accurate indicator of when an ISTC started treating patients. Nonetheless, Panel C presents estimates from such a specification, with months $-12$ to $-1$ before ISTC entry designated as the pre-reform period, and months 24 to 35 after entry designated as the post-reform period. Hospitals in the Control group are allocated a placebo ‘treatment’ start date equal to the contract start date of the nearest ISTC, even though this ISTC lies outside the 95th percentile of patient distance travelled. The resulting estimates are very similar to our main results.

Panel D reports estimates using a treatment assignment strategy that centres hospital markets on GP surgeries rather than hospitals. Hospital-centred measures of market size based on percentiles of patient distance travelled are potentially endogenous to hospital performance. While we address this concern by basing treatment assignment on percentiles of patient distance travelled between 2002/3 and 2004/5 – before the introduction of patient choice of hospital or the ISTC programme – concerns may remain. To address these concerns, this check assigns treatments by constructing a list of all the NHS hospitals and ISTCs that fall within each GP surgery’s market (95th percentile of distance from GP surgery to NHS hospital for that GP surgery’s hip and knee replacement patients). If an ISTC is in 95% of the GP surgery markets that an NHS hospital falls within, that NHS hospital is assigned to the High Treatment group. If an ISTC is in 75% of the GP surgery markets that an NHS hospital falls within, but not 95%, that NHS hospital is assigned to the Low Treatment group. All other NHS hospitals are assigned to the Control group. The estimates reported in Panel D are consistent with our main results, providing assurance that they are not driven by assignment of treatments based on hospital-centred market definitions.

19 Hospital-level characteristics used to predict treatment assignment include local hospital waiting times, hip and knee replacement case load, and dummy variables for – located in London, teaching hospital, standard acute hospital, university hospital. Average hip and knee replacement patient characteristics used include mortality rates, total LOS, pre-surgery LOS, post-surgery LOS, Charlson score, IMD income deprivation score, IMD health and disability deprivation score, and IMD overall deprivation ranking.
Panel E reports estimates when we do not take logs of the outcome variables (pre-surgery LOS and Charlson score). We continue to find that ISTC entry made nearby public hospitals more efficient, but also left them with sicker patients; the estimates are nearly equal to the exponent of our main results. A number of other checks are reported in the online Supplementary Material (see Appendices E through G); they provide further confirmation that our results are robust to a wide range of specifications.

5.7. Ruling out the possible confounding effect of changes in patient volumes

Increases in local capacity could potentially affect incumbents’ efficiency by reducing congestion and overcrowding, independent of any competitive pressure exerted by entrants. Moreover, public hospitals located near new private entrants may have experienced a reduction in demand. Any resulting reduction in case loads at nearby public hospitals could affect average pre-surgery LOS at these incumbents, given the important influence of volume on efficiency.

We therefore investigated the impact of ISTC exposure on case load using similar regressions to those used for our main estimates, but found no statistically significant evidence of case load reductions in our High Treatment group, which had the biggest reductions in pre-surgery LOS (results available in Table H.1 in the online Supplementary Material). This suggests that the increases in local capacity brought about by private surgical centre entry did not lead to any reduction in the volume of patients treated at High Treatment group hospitals, but, instead, served to take people off waiting lists and reduce waiting times. That is, ISTC exposure led to shorter pre-surgery LOS in close-neighbouring hospitals without any concomitant reduction in the volume of patients being treated.

We did find, however, significant reductions in the volume of patients being treated at Low Treatment group hospitals. This finding seems to suggest that ISTC entry did not simply add to overall clinical capacity, but, at least to some extent, may have reduced patient volume at public hospitals with which they shared a market – although, crucially, these patients seem not to have been drawn from the closest hospitals (i.e. those in the High Treatment group) in which we observe statistically significant reductions in pre-surgery LOS. These findings are broadly supportive of our conjecture that the reductions in pre-surgery LOS reported in Table 5 arose primarily through competitive incentives.20

6. Discussion and conclusions

This paper examines the effect of a UK government programme designed to increase capacity and competition by facilitating the entry of private, for-profit specialty surgical centres into the English NHS. We test the impact that the entry of these facilities – ISTCs – had on incumbent public hospitals’ efficiency, case mix, and case load. We exploit the fact that ISTC location decisions were driven by local waiting times, not by other hospital characteristics such as LOS or clinical performance, to construct treatment and control groups that are comparable with respect to the outcome variables examined. Indeed, we demonstrate that trends of key outcome variables – including pre-surgery LOS, post-surgery LOS, and patient case mix – were the same for public hospitals that had an ISTC enter nearby as for those that did not.

We find that public hospitals that had a private, for-profit surgical centre enter close by experienced substantial reductions in pre-surgery length of stay for hip and knee replacement surgery. The addition of an ISTC to a public hospital’s immediate neighbourhood led to a decrease in pre-surgery LOS of around 16% – or a 24 percentage point increase in the proportion of patients treated on the day of surgery.

---

Notes: Sample includes hip and knee replacement patients. Columns (1) to (3) control for the 2006 introduction of hospital competition by including a control for overall competition intensity (Negative Log of HHI × Dummy variable indicating period after introduction of competition). Columns (4) to (6) control for changes in policy at the Strategic Health Authority (SHA) level, and for differential regional trends generally, by including controls (the coefficients on which are not included) for SHA (region of England) × Dummy variable indicating period after introduction of ISTC programme. Treatment groups, pre-reform and post-reform period, mean dependent variable, standard errors and statistical significance are defined as in the Table 5 notes. All regressions include a dummy variable indicating hip replacement. Table G.1 provides alternative specifications of these robustness tests.

---

Table 9

<table>
<thead>
<tr>
<th></th>
<th>Controlling for introduction of choice and competition</th>
<th>Controlling for changes in strategic health authority (SHA) policy &amp; differential regional trends</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Log of pre-surgery LOS</td>
<td>High treatment × post</td>
<td>−0.175**</td>
</tr>
<tr>
<td></td>
<td>Low treatment × post</td>
<td>−0.0589</td>
</tr>
<tr>
<td></td>
<td>NegLogHHI × post</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>SHA × post-dummies</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Hospital fixed effects</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Month × year fixed effects</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Patient controls</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Observations</td>
<td>141.661</td>
</tr>
<tr>
<td></td>
<td>R-squared</td>
<td>0.394</td>
</tr>
<tr>
<td></td>
<td>Mean dependent variable</td>
<td>0.846</td>
</tr>
</tbody>
</table>

Notes: Sample includes hip and knee replacement patients. Columns (1) to (3) control for the 2006 introduction of hospital competition by including a control for overall competition intensity (Negative Log of HHI × Dummy variable indicating period after introduction of competition). Columns (4) to (6) control for changes in policy at the Strategic Health Authority (SHA) level, and for differential regional trends generally, by including controls (the coefficients on which are not included) for SHA (region of England) × Dummy variable indicating period after introduction of ISTC programme. Treatment groups, pre-reform and post-reform period, mean dependent variable, standard errors and statistical significance are defined as in the Table 5 notes. All regressions include a dummy variable indicating hip replacement. Table G.1 provides alternative specifications of these robustness tests.

---

20 Increases in local capacity could potentially affect the efficiency of incumbents even in the absence of case load reductions at these incumbents. For example, ISTC entry may have reduced pressure on waiting lists at nearby public hospitals, thus making it easier for incumbents to meet the national waiting time targets. If these incumbents had previously been avoiding target breaches by admitting patients inappropriately early and then making them wait in hospital until a surgery slot became available, then ISTC entry could have ameliorated the need for such premature admissions. We think it is unlikely that our estimated reductions in pre-surgery LOS at ISTC-exposed hospitals could be explained by this channel, for two reasons. First, blocking a bed until a surgery slot becomes available would be a very costly way of avoiding waiting time target breaches. Secondly, there is a substantial literature on NHS waiting time targets, which describes a wide range of gaming activities undertaken by hospitals in order to meet waiting time targets, but never (to the best of our knowledge) refers to hospitals admitting patients prematurely, and making them wait for a surgery slot to become available, in order to ensure that they did not breach the maximum permitted waiting time (National Audit Office, 2001; Audit Commission, 2003; Bevan and Hood, 2006a; Besley et al., 2009; Propper et al., 2010; EWHC 2787 (QB), 2005).
admission. Given that these faster turnaround times were achieved without additional expenditure (as they occurred in an environment with fixed payments per procedure), they suggest that hospitals exposed to competition from new private entrants became more efficient.

As well as investigating possible positive effects of ISTC entry on the efficiency of incumbent public hospitals, we looked for evidence of possible negative effects in the form of worsened case mix. We find that ISTC entry led nearby public hospitals to experience an 11.6% increase in patients’ average illness severity as captured by the Charlson score – or a 6.2 percentage point increase in the proportion of patients with a Charlson score of three or more. We also find suggestive evidence that this increase in the sickness of incumbent hospitals’ patients led to an increase in post-surgery LOS. While our identification strategy is unable to pinpoint how much of this increase in patient complexity at incumbent NHS hospitals is due to ISTCs actively applying their exclusion criteria – as opposed, for example, to differential choice behaviour by sick and healthy patients – the fact remains that surgical centre entry appears to have led to worsened case mix at nearby incumbents, irrespective of the exact channel by which this effect arose.

In principle, this sorting of patients between private surgical centres and public hospitals could represent an efficiency-improving division of responsibility between routine and more complex cases. Indeed, this appears to have been the government’s rationale for including wide-ranging exclusion criteria in ISTC contracts, which would allow these private entrants to focus on routine cases. Thus, the risk-selection we document was to some extent an intended policy outcome. However, the fact that policymakers explicitly intended such a division of responsibility between private entrants and public incumbents does not automatically imply that the division was devoid of negative consequences. To ensure that a division of responsibility between complex and straightforward cases does not have a negative effect on the financial position of providers that receive the most

Fig. 5. Treatment effects with alternative treatment group definitions (treatment defined using patient flows). Notes: Figures plot treatment effects (coefficients and 95% confidence intervals) as the cutoff distance from home to hospital used to define the treatment group (measured in percentiles of distance travelled) changes. For example, a percentile cutoff of 25 implies that the treatment group consists of all public hospitals that had an ISTC enter within the circular area around the hospital defined by a radius equal to the 25th percentile of patient distance travelled, defined using patient flows from 2002/3 to 2004/5. In all cases, the comparator (control) group consists of all public hospitals that did not have an ISTC enter within the 95th percentile of patient distance travelled. Patient’s home address is proxied by centroid of residential LSOA. Standard errors are clustered around hospitals.

Fig. 6. Treatment effects with alternative treatment group definitions (treatment defined using straight-line distances). Notes: Figures plot treatment effects (coefficients and 95% confidence intervals) as the cutoff distance from home to hospital used to define the treatment group (measured in kilometre distances) changes. For example, a distance of 10 implies that the treatment group consists of all public hospitals that had an ISTC enter within the circular area around the hospital defined by a 10 km radius. In all cases, the comparator (control) group consists of all public hospitals that did not have an ISTC enter within a 30 km radius. Patient’s home address is proxied by centroid of residential LSOA. Standard errors are clustered around hospitals.
than logs. All regressions include a dummy variable indicating hip replacement.

ignated as the pre-reform period, while months 24 to 35 are designated as the post-reform period. With the exception of Panel D, treatment groups are defined as in the Table 5 notes. Standard errors and statistical significance are defined as in the Table 5 notes. Panel A is a placebo regression with 2002/3 as the pre-reform year and 2004/5 as the post-reform period. With the exception of Panel D, treatment groups are defined as in the Table 5 notes. Standard errors and statistical significance are defined as in the Table 5 notes. Panel A is a placebo regression with 2002/3 as the pre-reform year and 2004/5 as the post-reform year. Panel B weights observations by the inverse of the probability of assignment to the treatment or control group that they were actually assigned to. Panel C defines the post-reform period as the period after the contract start date of the nearest ISTC and defines time relative to this date instead of using calendar time; months −12 to −1 are designated as the pre-reform period, while months 24 to 35 are designated as the post-reform period. Panel D assigns treatments by centring hospital markets on GP surgeries rather than hospitals themselves. Panel E runs the regression on levels of the outcome variable rather than logs. All regressions include a dummy variable indicating hip replacement.

complex patients, hospitals that treat sicker patients must be appropriately compensated.21

Unfortunately, NHS reimbursement rates during the period we study did not adequately adjust for patient severity (Mason et al., 2008). This situation not only provided private surgical centre entrants with an added impecunious to risk select, but it also meant that nearby public hospitals were left treating a costlier mix of patients without adequate financial compensation. While the prospective reimbursement regime (Payment by Results) was updated in April 2009 to include a more dramatic adjustment for patient severity, Mason et al. (2008) state that providers were still likely underpaid for treating sicker patients, and note that is unlikely that a prospective reimbursement system can ever be designed to fully compensate hospitals for a more costly case mix.

Our work highlights one trade-off that arises from the entry of for-profit surgical centres. We show that entry can stimulate improvements in efficiency and clinical quality, not by selecting against certain complex patients, hospitals that treat sicker patients must be appropriately compensated.21

We gratefully acknowledge excellent technical and coding contributions from Simon Jones and Stuart Craig. We appreciate helpful advice and suggestions from Martin Gaynor, Maitreesh Ghatak, Henrik Kleven, Amanda Kowalski, Camille Landais, Alistair McGuire, Tom O’Keefe, Sarah Sandford, Fiona Scott Morton, George Stove, Matt Sutton and Mohammad Vesan. We also thank Chris Buckingham of Ramsey Health Care UK for invaluable first-hand information about the ISTC programme. Charles Gray provided excellent research assistance. All errors of fact and interpretation are the sole responsibility of the authors.

Matthew Skellern conducted this research as part of a PhD funded by a Commonwealth Scholarship, a Newman College Archbishop Mannix Travelling Scholarship, a Goodenough College Fox Bursary, and an LSE Economics Departmental Award.

This paper was produced using Hospital Episode Statistics provided by NHS Digital under Data Sharing Agreement NIC-354497-V29JP. This paper has been screened to ensure no confidential information is revealed. Funding for this research was provided under the UK Economic and Social Research Council grant ES/M010341/1.

Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpubeco.2018.08.002.

References


Table 10

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Log of pre-surgery LOS</td>
<td>% treated on day of admission</td>
<td>Log of Charlson score</td>
</tr>
<tr>
<td>A. Test for parallel trends during pre-reform period (2002/3 to 2004/5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High treatment × post 2004/5 dummy</td>
<td>0.032</td>
<td>−0.0436</td>
<td>0.0659</td>
</tr>
<tr>
<td>2004/5 dummy</td>
<td>(0.0265)</td>
<td>(0.0387)</td>
<td>(0.0261)</td>
</tr>
<tr>
<td>B. Estimation using inverse propensity score weights</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High treatment × post</td>
<td>−0.167***</td>
<td>0.232***</td>
<td>0.109***</td>
</tr>
<tr>
<td></td>
<td>(0.0509)</td>
<td>(0.0740)</td>
<td>(0.0363)</td>
</tr>
<tr>
<td>C. Contract start date as t = 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High treatment × post</td>
<td>−0.175*</td>
<td>0.245**</td>
<td>0.123**</td>
</tr>
<tr>
<td></td>
<td>(0.0741)</td>
<td>(0.103)</td>
<td>(0.0543)</td>
</tr>
<tr>
<td>D. GP-centred treatment assignment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High treatment × post</td>
<td>−0.186***</td>
<td>0.259***</td>
<td>0.125**</td>
</tr>
<tr>
<td></td>
<td>(0.0657)</td>
<td>(0.0903)</td>
<td>(0.0526)</td>
</tr>
<tr>
<td>E. Outcomes in levels, not logs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High treatment × post</td>
<td>−0.264***</td>
<td>−0.309***</td>
<td>−0.147</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.147)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Sample includes hip and knee replacement patients. The table reports robustness tests of the main estimates for the High Treatment group based on the ‘headline’ regression specification (Eq. (2)), with hospital fixed effects and a full set of month-year dummies. With the exception of Panels A and C, this table uses 2004/5 as the pre-reform period and 2008/9 as the post-reform period. With the exception of Panel D, treatment groups are defined as in the Table 5 notes. Standard errors and statistical significance are defined as in the Table 5 notes. Panel A is a placebo regression with 2002/3 as the pre-reform year and 2004/5 as the post-reform year. Panel B weights observations by the inverse of the probability of assignment to the treatment or control group that they were actually assigned to. Panel B defines the post-reform period as the period after the contract start date of the nearest ISTC and defines time relative to this date instead of using calendar time; months −12 to −1 are designated as the pre-reform period, while months 24 to 35 are designated as the post-reform period. Panel D assigns treatments by centring hospital markets on GP surgeries rather than hospitals themselves. Panel E runs the regression on levels of the outcome variable rather than logs. All regressions include a dummy variable indicating hip replacement.

Acknowledgements

21 When certain types of care are removed from public facilities and shifted to the private sector, it can also have negative dynamic effects through missed learning opportunities for public sector clinicians and care professionals. Although not part of our formal analysis, personal conversations with clinicians indicate that an additional negative effect of ISTC entry was that it reduced the number of straightforward cases at nearby NHS teaching hospitals, which are essential for the training of surgical registrars.