

Does hospital competition reduce rates of patient harm in the English NHS?

Economics Working Paper

January 2019

Does hospital competition reduce rates of patient harm in the English NHS?

Russell Whitehouse[†]

Pasquale Schiraldi[‡]

[†] Competition and Markets Authority (russell.whitehouse@cma.gsi.gov.uk)

[‡] London School of Economics (pasquale.schiraldi@lse.ac.uk)

Abstract^{*3}

Previous hospital mergers have reduced the number of distinct organisations offering publicly funded and provided care in the English National Health Service, reducing choice and between-hospital competition in some areas. We test the impact of variation in concentration on a new quality indicator: the prevalence of patient harm from falls, pressure ulcers, blood clots and urinary tract infections. We exploit a new source of identifying variation accounting for the multi-product nature of hospitals, using instrumental variables to address endogeneity. We find that hospital mergers in concentrated areas without offsetting clinical benefits could significantly increase rates of patient harm.

JEL codes: C23; H42; I11; I18; L11; L41.

* We are grateful for invaluable advice and support from Kate Collyer and several members of the Economics Group at the Competition and Markets Authority: in particular, we are indebted to Mike Walker, Julie Bon, Tom Kitchen, Paul Reeve, Alex Moore, Cecilia Rossi and Elena Ashtari-Tafti for various helpful comments and individual stata codes. We also thank participants at CMA Economics Seminars and the Competition in Hospital Markets conference in Rotterdam for their insight.. All errors are the authors' alone, and any opinions expressed may not necessarily represent the views of the CMA.

1. Introduction

In the mid-2000s, the UK government introduced reforms to England's taxpayer funded and publicly run healthcare system, the National Health Service (NHS). As part of these reforms, policy makers sought to give patients greater choice, and to strengthen incentives of providers to maintain and improve the quality of care by introducing competition between NHS hospitals to attract patients.

Although choice and competition in the NHS have been in place for over a decade, perceptions vary about their effectiveness. One potentially significant cost of the regime is scrutiny of administrative mergers between the public organisations which run NHS hospitals, called NHS trusts.¹ In particular, plans may be reviewed by the competition and consumer body, the Competition and Markets Authority (CMA).

Since 2010, every hospital merger reviewed bar one² has been permitted either on the basis that it would not substantially reduce competition in the local healthcare economy, or that specified benefits would outweigh any harm arising. This is despite most cases involving reductions in the number of alternative hospital trusts from three to two, or from two to one, in a context of generally low demand elasticity and high barriers to entry: factors which would generally give cause for concern.

There are several reasons why competition may play a reduced role in the NHS relative to other markets, including regulation, staff altruism and other structural factors such as persistent capacity and financial constraints. In spite of these, much (but not all) of the existing economic literature has concluded that competition plays an important role in improving health outcomes for patients.

There remains scope to contribute to the literature however. Our work contributes analysis of a new quality measure which describes whether admitted patients experience new falls, pressure ulcers, blood clots or urinary tract infections. This measure is more comparable across different care settings and is more clearly relatable to the impact of competition than those studied so far. We use a novel identification strategy in which we exploit variation within hospital trusts across departments, rather than within hospital trusts across time, or simply across hospital trusts. This methodology accounts for the multi-product nature of hospital services,

¹ There is a technical distinction between 'NHS trusts' and 'NHS foundation trusts'. However, as we set out later, the distinction has been increasingly irrelevant in substantive if not legal terms for the purposes of this analysis. We therefore refer to both in discussing NHS trusts.

² A previous merger proposal between two hospitals, The Royal Bournemouth and Christchurch Hospitals NHS Foundation Trust and Poole Hospital Foundation Trust, was prohibited by the Competition Commission in a much debated decision in 2013 on the basis of competition concerns and the absence of demonstrated benefits.

and reduces the literature's reliance on one natural experiment. We also study a more recent time period and more treatments relative to previous papers.

Similar to the existing literature, we employ an instrumental variables specification to account for endogeneity of market structure to both choice and entry. We study eight specialties (treatment groupings), which map to the way hospitals generally organise themselves into departments, and which together constitute a majority of acute admissions.

We measure market concentration by counting the number of distinct trusts in the neighbourhood of each hospital (called the 'competitor' or 'fascia' count). We also use the Herfindahl-Hirschman Index (HHI). We allow concentration to have a nonlinear effect on quality, and relate our results to merger control, a key area of NHS competition policy.

We find a significant inverse relationship between concentration and quality. Specifically, our estimates imply that a hypothetical merger to monopoly would, on average, be associated with a significant increase in harm rates. This increase becomes smaller when more competitors remain post-merger. The economic and statistical significance of this effect is robust, including when we consider the effect of mergers on in-hospital mortality.

In addition to the direct effect on patients' well-being, even small gains to health from competition would likely give rise to large savings for the taxpayer because of reduced treatment costs. Quantifying this is difficult, however we find some indicative evidence that an average hypothetical merger would result in extra direct costs of £2.5 million annually to the NHS, in the absence of compensating patient benefits.

2. Background to the NHS

Industry landscape

The English NHS is a collection of several interlinked but distinct organisations. Overall responsibility lies with the Secretary of State for Health and Social Care; the Department for Health and Social Care is responsible for strategic leadership and funding; and NHS England is the body which sets priorities and direction of the NHS, acting as the lead commissioner for a wide range of services. Local organisations called Clinical Commissioning Groups are responsible for the health of individual areas, paying for the healthcare received by their patients in providers such as NHS

Trusts, the focus of this paper. NHS Trusts are the organisations which operate hospitals, and make relatively autonomous decisions about budgeting and care.³

Two major reforms relevant to this paper were implemented in the mid-2000s. The first reform gave patients the right to choose the NHS hospital in which they received treatment, at least for their first outpatient appointment.⁴ This was introduced in 2006 and expanded in 2008. Surveys consistently show that around 40-50% of patients recall being offered choice, with their GP exercising choice on behalf of the remaining patients (see for example, Competition Commission (2013); CMA (2015)). In addition, a second reform was to introduce NHS funding for care in privately run facilities named Independent Sector Treatment Centres (ISTCs). Although ISTCs continue to account for a small minority of overall admissions, their significance has grown in recent years.

It is common for NHS trusts to operate several individual hospitals, although most are smaller satellite or community hospitals. Around 30% operate two sites providing mostly secondary acute care (so called “district general hospitals”), and 20% operate more than this.⁵ Since new NHS hospitals are hardly ever built, almost all multisite trusts were formed through mergers. Consolidation in the late 1990s led to fewer distinct NHS trusts, although the number has since has remained fairly static. Recent financial pressure in the NHS has given rise to an uptick in the numbers of trusts considering merging however.

There are many reasons trusts decide to merge, including financial or clinical benefits to improve care for patients. Reasons for merging have included increasing the ability of trusts to recruit and retain staff; consolidating and introducing new services; and integrating back office functions (NHS Improvement, 2016a). However, there can be challenges. Some NHS merger plans have been found to extend little beyond the initial

³ More precisely, hospitals may be operated either by NHS Trusts or by NHS Foundation Trusts. Although only the latter are technically self-governing legal entities, in practice both make independent decisions about the care given to patients, and the organisational strategy pursued. Although the approvals process to convert from an NHS Trust to an NHS Foundation Trust meant that Foundation Trusts used to be in stronger clinical, financial and governance positions, by 2015 most of the differences between Trusts and Foundation Trusts had been eroded. For simplicity, throughout the rest of this paper we use the term ‘NHS Trust’ to refer to Foundation Trusts as well. We have focussed on acute trusts. There are a number of additional trust types which focus on types of care outside of our focus, secondary acute care. Since these would not generally be alternatives for patients requiring treatment of this type, we exclude these from our analysis.

⁴ Patients have the right to choose any provider in England that has been commissioned by a CCG or NHS England for their first outpatient appointment for NHS elective services. This is enshrined in the NHS Constitution (Department of Health, 2015). First outpatient appointments are the start of patients’ secondary care pathway. Few patients who require an admission to hospital for further treatment later change organisations. Further, where choice is exercised, it might also be done so taking into account factors relating to admitted care. In most cases, we consider that the right to choose affects the whole of patients’ secondary care pathways.

⁵ We have considered only acute trusts, and counted only sites classified as “General Acute”, “Mixed Service” or “Multi Service”. The percentage of ISTCs operating more than one site is much higher due to the formation of hospital chains. Most are relatively geographically dispersed however.

amalgamation of organisations and their back-office functions, and some providers have found there to be significant challenges in operating across multiple sites (NHS Improvement, 2016b).

Partly as a result of these challenges, mergers have several layers of internal and external review. This may include a competition assessment: for mergers within its jurisdiction, the UK's competition and consumer body, the Competition and Markets Authority (CMA) is charged with assessing whether a proposed merger is likely to substantially reduce competition between hospitals.⁶

Role of competition

In an industry with fixed prices, competition may still occur on quality, provided some necessary conditions are met on both the demand and supply side of the market.

On the demand side, quality of hospital care must influence the choices patients make. There are mechanisms through which information can reach patients, including word of mouth, an NHS comparison website, and GP advice.⁷ Empirical studies have generally confirmed this by finding significant elasticities of demand with respect to a wide range of quality measures, including for example waiting times, mortality rates and staffing rates, as well as on other drivers such as patient-hospital distance.⁸ As such, a hospital which wanted to attract more patients could do so by improving its quality.

On the supply side, hospitals must have incentives to attract more patients. Over the period considered, most revenue received by hospitals in our specialties was determined by the volume of patients which they treated.⁹ Each patient stay attracts a fixed payment valued according to an annually-revised national tariff. Hospitals therefore receive additional revenue for attracting new patients. Where the adjusted

⁶ NHS Improvement, the sectoral regulator for the NHS, also plays a role preventing anti-competitive behaviour against the interests of people who use NHS services, and advises the CMA on the potential benefits of any given merger.

⁷ Since 2007, an array of quality indicators has been published online through NHS Choices, a web portal which allows patients to research and compare different NHS trusts. Other potential sources of quality information include patient reviews online, local newspapers (where significant quality events are often reported) and information given out in hospitals and through other healthcare providers.

⁸ See for example Beckert, Christensen and Collyer (2011) and Gaynor, Propper and Seiler (2016).

⁹ A portion of many hospitals' revenue, is determined according to block payments made by commissioners. These payments are generally put in place where demand is projected to exceed the budget to pay for that demand. Alternative structures exist such as cap and collar schemes, whereby revenues of the hospital are constrained to lie between certain bounds even if demand exceeds or falls below these limits. Lastly, there exist also volume-based payments for some services not covered by the national tariff, however these may vary from area to area. Nevertheless, a large portion of most hospitals' revenues are made according to the national tariff.

tariff price exceeds the marginal cost of treating an additional patient, hospitals are also incentivised to attract more patients.

Even if these basic conditions are met, there are still sector-specific reasons why competition might not drive improvement in health outcomes, or may be less important than in other industries. In particular, regulation, altruism and capacity constraints could be important. We discuss each in turn.

First, regulation might constrain the ability for providers to alter quality. Indeed, analysis of hospitals' internal documents has previously shown that hospitals boards consider regulatory matters significantly more frequently than competition matters (CMA, 2017). However, whilst regulation plays a key role in protecting patients from low quality care, it does not necessarily incentivise high quality care by itself. Even where it does so, there is potentially scope for competition to work alongside regulation in driving patient outcomes.

Second, hospital staff are generally seen as altruistic and the NHS is a not-for-profit system. But models of competition tend to assume profit maximisation. The two facts can be reconciled: Gaynor and Town (2011) introduce a simple model whereby hospitals maximise an objective function which accounts for revenue net of costs (since hospitals are required not to run large deficits, and managers are penalised reputationally should they do so), but also altruistic behaviour (treated as a reduction in the marginal cost of quality). They show that altruism can indeed raise quality levels in this scenario. Independently of this, greater competition (assumed to increase the elasticity of demand) still raises quality levels, provided the regulated price exceeds marginal cost.¹⁰

Third, capacity constraints might prevent hospitals from accepting additional patients attracted by higher quality, diluting incentives to improve their offering in the first place. High occupancy rates and waiting times are often cited as evidence in favour of this proposition. Analysis of capacity in hospitals has however shown that in general, hospitals can sometimes unlock capacity by reducing length of stay and managing beds more effectively, by investing, or by innovating (CMA, 2017).¹¹ Nevertheless, there are clearly limits to hospitals' abilities to do this.

¹⁰ We expect this to be true for most if not all hospitals over this period because prices are set on the basis of average costs, including fixed costs. Because fixed costs (such as building maintenance and equipment) are likely to be appreciable in healthcare, the marginal cost of treating a patient (which does not include this fixed element) should be lower than the average cost, and therefore the price received.

¹¹ See for example, CMA: Manchester Hospitals paragraph 10.25, page 73

As a consequence, regulation and altruism are very important in setting quality, but still allow for competition to play a role. Capacity constraints have the potential to dampen the role of competition. However NHS providers generally achieved important benchmarks relevant to capacity over the period studied: over 92% of patients on incomplete pathways were seen within the 18 week referral to treatment target between March 2012 and March 2015, a key target (NHS England, 2017). Therefore, there is scope for competition to have played a role.

3. Existing literature and our contribution

Several papers have sought to quantify the benefits of competition in the English NHS under the above institutional framework. The literature generally (but not exclusively) suggests that competition can improve market outcomes.¹² Comprehensive summaries of the existing literature already exist: see Gaynor et al (2013) and Skellern (2018), for example. For brevity, we provide only a short general summary of the literature.

The results presented in Cooper et al (2011) and Gaynor et al (2013) have been widely cited. Both papers found that for the types of care studied, the introduction of choice in 2006-8 led to greater improvements in quality in less concentrated areas than more concentrated areas, implying competition produced real benefits for patients. Gaynor et al (2016) extended these results to show that patients' elasticity of demand increased following the choice reform. This this led to a substantial increase in patient welfare, both due to patients selecting higher quality hospitals, and due to hospitals responding to the enhanced incentives by improving quality. Additionally, using cross-sectional variation and an instrumental variables strategy, Bloom et al (2015) found more competition was associated with higher management quality.

However, some recent work has called these conclusions into question.¹³ Skellern (2018) uses cross sectional variation to assess whether concentration had an effect on health gains using patient reported outcome measures (PROMS). He finds that competition may have had a negative effect on elective quality, arguing that hospitals have been incentivised to improve non-elective quality *at the expense of elective quality*. Moscelli et al (2017) assess in-hospital mortality for hip fracture and stroke

¹² We do not discuss the US-focussed literature due to the differing institutional relationships.

¹³ Propper et al (2004) and Propper et al (2008) both also find results suggesting that some types of competition policy are ineffective in driving NHS quality. Both papers focus on the "internal market" introduced in 1990, where hospitals were encouraged to compete on price charged to local authority purchasers. However, the significant differences in the institutional structure limit the applicability of these papers to the policy in place after 2006, where competition was to be focussed on quality.

patients using a similar difference-in-differences approach to that of Cooper et al (2011) and Gaynor et al (2016). The authors find a *detrimental* effect of the number of rivals which declined with time, although the introduction of choice did not appear to play a role in this decline.

Our paper contributes a new approach to the literature. We make two particular contributions:

- (i) First, where the literature has so far focussed on a small set of treatment-specific health outcome measures, we analyse a new and broader dimension of quality, namely four key ‘harms’ which are comparable across treatments comprising around 29% of departments, but 60% of all hospital admissions.¹⁴ These harms are important for patients’ health, costly for the NHS, and likely to be closely related to the mechanisms by which competition could influence quality.
- (ii) Second, we use a novel identification strategy in which we exploit variation within hospital trusts across departments, rather than within hospital trusts across time, or simply across hospital trusts. This allows us to test the hypothesis that hospitals optimise quality with respect to competition across their various departments, and to introduce a battery of fixed effects to control for unobservable confounding factors.

In addition, whereas most studies have focussed on the period in which choice was introduced, we study the 2013/14 and 2014/15 financial years (using quarterly data). By this time the policy should have been well established. Furthermore, over this period the NHS was not yet subject to the degree of supply side pressures faced by the service in years immediately following.¹⁵ This represents something of a “sweet spot” in the pre-conditions for effective competition.

We also extend somewhat the approaches taken in the literature to date to address ‘reverse causality’ issues. In particular, whilst the number of competitors may influence quality, quality may influence both the distance over which hospitals compete for patients and the number of hospitals within any given distance. For the first time, we

¹⁴ That is, of all admissions in hospitals we have identified as ‘acute’ hospitals. These figures relate to the proportion of all of England’s acute admissions in these specialties in departments which treat at least 100 patients a quarter, we discuss matching rates into our final dataset later.

¹⁵ In particular, levels of capacity utilisation and financial pressure appear to be higher in the period post 2015. The net aggregate surplus (before impairments and transfers) of all NHS foundation trusts was £134m in the 2013/14 financial year (Monitor, 2014), the first year of our data. However, by the 2017/18 financial year the equivalent surplus was negative £345 million (NHS Improvement, 2018). Financial positions have also deteriorated markedly after impairments and transfers for this group of trusts.

address simultaneity in ISTC entry. We present results using both fascia count and the HHI, and allow for nonlinearities in the effect of concentration. This approach allows us to test whether changes in the degree of competition have a greater effect on quality where there are fewer alternatives for patients.

Our approach allows us to draw particularly clear conclusions as to the effect of mergers between NHS trusts: a topic which has received only passing reference in the existing literature, despite the fact that merger assessment is of central relevance to the operation of competition policy in the NHS, involving usually permanent structural change and with the potential to affect both patient well-being and millions of pounds of public money.

4. Quality measure

Even relatively simple and relatively homogenous goods can have multiple dimensions of quality. Healthcare is both complex and heterogeneous, and also suffers from the challenge of observing quality itself. As a consequence, almost all studies focus on the quality of health outcomes rather than healthcare, standardising insofar as possible pre-treatment health to allow inference to the quality of the interventions patients receive.

The existing literature has generally focussed on mortality and readmission rates. Each of these relates relatively closely to the effectiveness of the hospital's front line medical staff in performing the clinical interventions for which the patient was admitted. Studies considering mortality generally focus on Acute Myocardial Infarctions (AMI), whilst readmission rates and PROMs are usually confined to orthopaedic procedures.

However, health outcomes not relating to the main clinical intervention are also important. Having undergone their primary treatment, patients often spend time on a ward recovering or receiving follow-up care. The training and resourcing of nursing staff, infection control, clinical leadership and other process management activities can therefore also have material implications for patients' health. But these have been little studied so far.

We contribute a new quality measure of this type, namely the probability that a given patient will experience one of four particular types of harm whilst in hospital: falls, pressure ulcers, venous thromboembolism and urinary tract infections.¹⁶ These four

¹⁶ Although some level of harm may be unavoidable, NHS documentation reveals that commissioners considered there was "no evidence to suggest the harms recorded in the NHS Safety Thermometer [were] anywhere near a clinically irreducible minimum" over the time period studied (Harm Free Care, 2014)

types of harm are combined into one indicator which is tracked across all acute hospitals. Whilst these harms are less related to primary health interventions, it has three specific advantages over conventionally studied quality measures, such as AMI mortality and hip readmissions.

- (i) First, competition policy applies across a broad spectrum of elective care. However, the specificity of the measures so far studied means it is difficult to extrapolate evidence beyond a small fraction of patients. For example, there is no consistent evidence on the degree of correlation in quality across departments. By contrast, harm rates are relevant and comparable across a wide range of departments.
- (ii) Second, patients are often admitted precisely because they have health complications in the dimensions along which these outcomes are measured. A hospital's job is to reduce the risk of readmission and mortality. But since the existing risk is likely to vary across patients and is unlikely to be orthogonal to concentration, there exists a strong possibility of bias. This is much reduced in our case because the harm measure focusses on health outcomes in dimensions where patients are already healthy.
- (iii) Third, the mechanisms by which competitive incentives filter down to front line medical staff are not fully clear. As a result, competition could only have an indirect impact on the quality of primary interventions. By contrast, managers often monitor competitors and have clear incentives to respond where there is available capacity by improving processes to attract additional patients and therefore revenue. The effect of these improvements are likely to be manifested in our four harm indicators.¹⁷

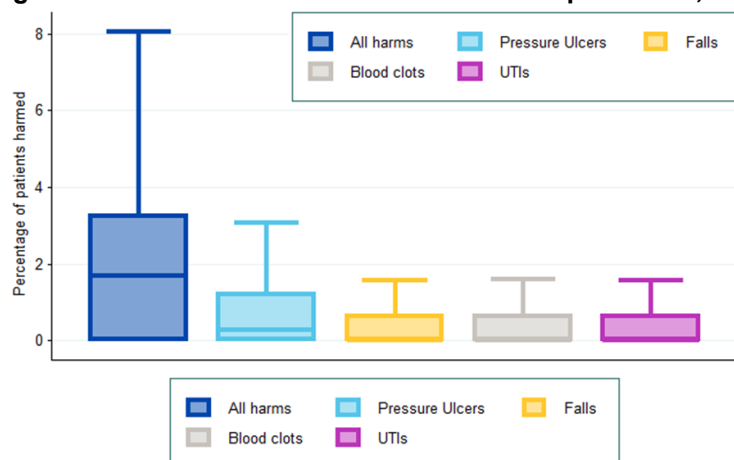
Figure 1 below shows the distribution of the combined harm indicator in the leftmost boxplot, and each subcomponent individually in the remaining boxplots. Whilst only about half of departments experienced any harm on any given subcomponent, only around 25% of departments were harm free on the combined measure.¹⁸ Further, on

¹⁷ There are also several other advantages. The measure more salient than others (such as PROMS) which exist in data files online but are not displayed on hospital wards nor included in hospital informational material. It is coded on an objective basis rather than based in self-reported assessments such as the Friends and Family test scores. Lastly, the measure is arguably more relevant: on many individual measure, negative health outcomes are fortunately quite rare. By comparison, harm on at least one of our four indicators is more common: the mean rate of harm is around 6 times more common than in-hospital mortality in our dataset. This greater prevalence increases the measure's suitability to standard econometric models

¹⁸ To ensure consistency, we define a 'harm' in the same way as the NHS Safety Thermometer (a key outlet for this data). We count all new UTIs and new venous thromboembolism (VTE). We discount falls with "no harm" but include all others including those with "low harm", and include all new pressure ulcers regardless of severity.

the combined measure, around 25% of departments saw 8% of their patients or more experience harm.

Figure 1: Distribution of harm rates across departments, for each subcomponent



Notes: Data is at department level, for all quarters 2013/14-2014/15. The whiskers represent the 1st and 99th percentiles of the data; the extremes of the box the 25th and 75th percentiles of the data, and the line the 50th percentile of the data. The indicator 'All Harms' is a measure of whether the patient received harm on ANY of pressure ulcers, falls, blood clots or UTIs. Sources: NHS Digital HES data, Harm Free Care data.

Hospitals which score well on one subcomponent of the indicator are likely to score well on others. Appendix 1 presents correlations between measures: although the magnitudes are not very high, the majority of correlations are significant and the signs are all positive.

Of course, our harm indicator is just one measure of quality: some highly regarded hospitals do not score highly on this measure, but do well on others. Nevertheless, measuring this new dimension of quality remains interesting. Our combined indicator has a correlation of 24% with unadjusted, in-hospital mortality,¹⁹ and this correlation is significant at the 1% level. Further, all subcomponents exhibit some correlation with mortality. The magnitude of the correlation implies that whilst our measure is not isolated, it is also measuring other (new) factors.

There appears to be significant variation in quality *within* trusts for the same harm indicator.²⁰ We calculated two measures of the similarity of harm rates across departments in each trust: mean absolute deviation (MAD) and the interquartile range (IQR). We found that around half of trusts' have an average deviation of their

¹⁹ For comparability with the existing literature, we exclude all patients with ages below 35 or over 75 in all mortality data. Mortality in these groups is less likely to be attributable to hospital care. We compute in-hospital mortality using HES data at patient level.

²⁰ A regression of harm rates aggregated to department level for Q1 2015 shows that including trust fixed effects explains 46% of the variation; the remaining variation is 'within' trusts.

departments' harm rates of over 0.8 percentage points from their mean. This is significant in the context of the overall mean harm rate of 1.9 percent.²¹ The IQR showed consistent results. Further analysis is shown in Appendix 1. This intra-trust variation is important for our estimation strategy, as we explain later.

5. Identification

Studies of market structure and conduct are notoriously hampered by endogeneity issues. The hospital competition literature has devoted significant attention to addressing these problems. This is both in terms of carefully isolating the variation used to identify the model, and in using instrumental variables (or similar) techniques to address simultaneous causality between concentration and quality. We discuss both these issues below.

Source of variation

It is generally preferable to exploit panel dimensions in data to find identifying variation, rather than conduct cross-sectional analysis. This is because using variation across different observations of the same subject (such as through time) allows the researcher to hold potentially confounding factors constant though the use of 'fixed effects'.²²

A natural approach would be to consider variation in concentration through time through mergers, entry and exit. However, there is not much meaningful timeseries concentration variation in the UK healthcare system over the period we are studying. New NHS hospitals are hardly ever built, and whilst they are sometimes 'downgraded', they are seldom closed.²³ In addition, since a wave of mergers in the 1990s, the number of NHS trusts has since remained relatively stable.

As a consequence, most studies have relied on a natural experiment to "induce" variation through time, namely the introduction of choice for first outpatient NHS care

²¹ The MAD across trusts is 0.74 percentage points

²² Several NHS concentration studies have nevertheless taken a cross-sectional approach. Bloom et al (2015) study cross sectional variation in fascia count across trusts to assess management quality of hospitals. Likewise, Skellern (2018) considers cross sectional variation in the HHI across hospital sites. However, conventional cross sectional studies are even more likely to be hampered by omitted variable arguments, because it is impossible to include trust fixed effects to address unobserved variation in factors such as sickness and urban density. As a consequence, these studies rely heavily on IV strategies to deal with omitted variable bias.

²³ Timeseries data on the number of sites associated with trusts has changed form through time, and is subject to changes in sites' classification. Our best estimate using HESF data is that around 4 sites per year were either closed or downgraded to a non-acute site between 2000 and 2016; and the same in the latter half of this period. This is a small fraction of the total number of sites.

in 2006-2008. In particular, the results presented in Cooper et al (2011) and Gaynor et al (2013) have been particularly widely cited. Although all hospitals were exposed to the introduction of choice, both papers argue that its effect would be greater for hospitals in less concentrated areas. By contrast, in the limiting case of a monopolistic market structure, patients would have no choices before or after the reform, so it would not make any difference.

Data on our quality measure was not fully established until 2013/14,²⁴ several years after the choice reform; this strategy is therefore unavailable to us. However, there are good reasons to prefer a new approach in any event: although the natural experiment so far used is undoubtedly useful, there is a risk that the literature relies heavily on this single policy change.

In this paper we introduce a new source of panel variation to the healthcare concentration literature. Rather than consider variation in concentration *within hospital trusts across time*, we exploit variation *within hospital trusts across departments*.²⁵ In particular, we utilise the fact that different departments within the same hospital trust commonly face different competitors. The differences arise because not all hospitals offer a full range of specialties at all sites, and because willingness to travel varies by area and department. This approach has two key advantages:

- (i) First, obtaining statistically and economically significant results may imply that hospitals optimise quality with respect to the competitive constraint *on each department* (rather than taking some average across all departments). This gives potential for new insight into the effects of competition *within* hospital trusts, thereby shedding light into what has always been something of a 'black box'.
- (ii) Second, not only can we introduce trust fixed effects to address unobservables and time fixed effects to address trends in harm rates, we can interact these fixed effects to give *trust-quarter combined* fixed effects. This means that our estimates are not sensitive to time-variant characteristics of individual trusts,

²⁴ Many hospitals began reporting in 2012/13, however coverage remained patchy until 2013/14

²⁵ We define a department as a provider's offering of a particular specialty group. This reflects the way hospitals in England commonly organise themselves. In particular, we study departments from 10 different specialties: Cardiology; Ear Nose and Throat (ENT); Gynaecology; Maternity Services; Medical and Clinical Oncology; Ophthalmology; Trauma and Orthopaedics; Urology; a group of medical specialties including General Medicine, Geriatric Medicine, Transient Ischaemic Attack and Stroke Medicine; and a group of surgical specialties including General Surgery, Breast Surgery, Vascular Surgery and Colorectal Surgery. Competition assessments of live merger cases have been conducted in all of these specialties: see Competition Commission (2013) and CMA (2015, 2017a, 2017b).

such as the quality of hospital buildings, deficits or management experience, unless these factors vary between different departments in the same trust.

Including data from several hospital departments is not new in itself: studies such as Cooper et al (2011) have included patients in a range of specialties in their analysis (such as AMI, hip replacements and cataract repairs, amongst others). Our innovation is to exploit variation between these departments. We address systematic differences in the probability of negative health outcomes across specialties by including a set of specialty fixed effects. Again, we interact these with the time fixed effects to address specialty-specific trends, such as all-NHS initiatives or guidance at specialty level.

In short, the only variation which we do not remove using fixed effects is that which varies across departments in the same trust, either cross-sectionally or through time. This could include ward layout and the case-mix that the department receives. We therefore need to introduce specific controls for these issues.

Regressors

Here, we explain how we measure concentration, before going on to specify our control variables.

Concentration variables

There is no single agreed way to measure the degree of competition faced by hospitals; however the two most common ways are by counting the number of competing organisations (“fascia”), and by calculating the Herfindahl-Hirschman Index (HHI) which is defined as the sum of squared market shares.

Fascia counts are not new to the healthcare literature, being the central measure in Bloom et al (2015). However, they are sometimes overlooked in favour of the HHI. This has its downsides, because the number of trusts in a given area is the policy-relevant measure for managers’ decisions to merge hospitals, and for competition authorities implementing merger control: for decision makers in these organisations, a key question is what would happen to quality were distinct trusts merged together. Study of fascia count means we are able to relate our results much more closely to this question than has been true in the existing literature, which has generally focussed more on whether the introduction of competition in the NHS as a whole was beneficial.

To compute fascia counts, we measure the number of trusts falling within each department’s catchment area. As such, catchment areas can vary across different specialties in the same trust as well as across trusts, depending on the distance

travelled by patients of each department. We count all NHS and ISTC trusts as alternatives, provided they treated a minimum number of patients in the relevant specialty in that quarter.^{26,27}

However, we also present results for the HHI. This index has the advantage that it can better account for the strength of alternatives, given that it assigns greater weight to trusts which take a greater share of the market. We do not treat different hospitals operated by the same trust as competitors, but otherwise follow the literature's approach to compute these.²⁸

Increasing concentration is unlikely to have a linear effect. Mergers to monopoly are more likely to have significant anticompetitive effects than mergers leaving many competitors. We therefore allow our quality measure to depend nonlinearly on both fascia count and the HHI. We estimate of the average effects of hypothetical mergers, conditional on existing concentration.²⁹

We show the distribution of our fascia count variable in Figure 2 below. This chart shows that a high proportion of departments have at most one other competitor. The chart also shows that rural areas tend to be more concentrated, a fact we pick up later in this paper.³⁰

²⁶ Our data lets us determine catchment area sizes by patient flows, and therefore to be specific to individual specialties (as well as quarters). This is in contrast to Bloom et al (2015), who use a uniform 15km radius.

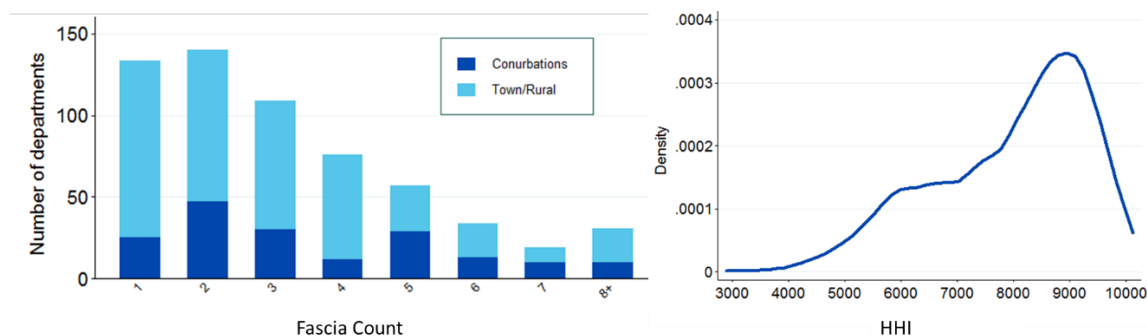
²⁷ Bloom et al (2015) note that given a catchment area around a hospital, any provider located up to twice this distance away will have a catchment area overlapping with the focal hospital, and as such could compete for some of that hospital's patients. For their main specification, the authors therefore construct a count of all hospitals located within a radius equal to twice that of the catchment area. Whilst clearly theoretically sound, this practice is not generally followed by Competition Authorities, including the CMA. This is because the proportion of patients such distant hospitals can compete for is very small. The authors address this criticism by introducing a battery of robustness tests in the distance over which fascia are counted. We likewise test various distances, but for our main specification adjust the multiplier down to 1.5 to bring the approach closer to how competition assessments are generally done.

²⁸ We follow Bloom et al in counting competition only from distinct trusts, rather than multiple hospital sites operated by the same trust. This is because we consider that hospitals have little or no incentive to increase quality to attract patients from their sister hospitals run by the same organisation.

²⁹ Although fascia count is a linear function of concentration, the HHI is by definition a nonlinear function of the number of (equal sized) competitors in the market. Nevertheless, including a squared term is a more flexible approach and allows the nonlinearity to be partly estimated by the model.

³⁰ Our HHI variable indicates a greater degree of concentration than previous work, such as Gaynor (2013). This is because we do not treat sites operated by the same trust as competitors to each other, and may also be due to the different time period and treatment mix studied.

Figure 2: Concentration measures: Fascia count and the HHI



Notes: Data is presented at department level (trust-specialty combinations) for 2015 quarter 1. For Panel A, Fascia count: Because multisite trusts can have different fascia counts at each site, but we observe harm at department level, the data shows the weighted average fascia count across sites in the same department, where the weights are the sum of the exogenous predicted probabilities that each hospital will be selected. For presentational purposes only, we round these to the nearest one. We split sites between ‘Conurbations’ and ‘Town/Rural’ using ONS classifications. For Panel B, HHI: we compute the HHI using the same formulae as Gaynor et al (2013), except we sum together market shares of all sites operated by the same trust. Sources: NHS Digital HES data; ONS NSPL data.

Control variables

Given that our data are at patient level, we are able to introduce controls which allow patients’ specific harm probabilities to differ. In particular, we control for patient age, gender, whether they have an old pressure ulcer or a long term catheter, and the (log of the) number of patients treated in the same ward.³¹

We combine our harm data with additional controls measured at department level:³² in particular, we control for average rates of comorbidities, emergency admissions,³³ the average resource intensity of procedures received,³⁴ an indicator for whether the

³¹ This differs from the number of patients treated in the department itself (which would be a function of quality), because hospitals have several wards. Whilst greater volumes are generally expected to improve the quality of many *primary interventions* due to greater experience gained by surgeons for example, the same is not necessarily true for broader measures of quality such as harm rates. This is because diseconomies of scale could affect the attention which any patient may receive. Adding this control could also proxy for resource constraints in periods of unexpectedly high demand. Indeed, we find that busier wards imply greater harm rates.

³² We calculate these measures only using patients who were being treated on the day each trust’s harm free care data was collected.

³³ We separate out emergency admissions from (i) the trust’s own Accident & Emergency (A&E) department and (ii) from all other sources, such requests for immediate admission from GPs, transfers from other hospitals, and transfers from other hospitals Accident & Emergency departments. Doing this allows for the possibility that pre-existing health status is more severe for patients who must be transferred for more specialist care than can be given in the initial location where they presented themselves, than those who can be treated in the same location they present themselves. Indeed, we find that harm rates are only higher (relative to elective patients) for the latter group, and exclude the control for admissions from A&E from the main specification.

³⁴ We measure this from NHS Digital Healthcare Resource Grouping data. In particular, each operation code is assigned a resource intensity which is ultimately used to match spells to payments; we average this hierarchy value to account for the fact that patients treated in departments undertaking more resource intensive surgery on average may be associated with greater pre-existing risks of harm.

department treats patients in the community,³⁵ and the percentages of patients for which our HES data lists a pre-existing diagnosis of pressure ulcers,³⁶ UTIs³⁷ or blood clots³⁸ respectively on admission. We find our conclusions are not sensitive to changes to these control variables.

There remain some unobservables. For example, although clear guidance exists for categorising harm, the comprehensiveness of the data recording could vary by department, and national initiatives (particularly around pressure ulcers) could have had varying local focus. Whilst specialty-quarter and trust-quarter fixed effects control for any systematic differences in these issues, some department specific variation may remain. Our assumption is that any residual variation is not correlated with concentration, conditional on all other covariates.

Simultaneity

Simultaneous causality describes the situation whereby one or more of the independent variables is jointly determined with the dependent variable. In our case, whilst concentration may influence quality, quality may also influence concentration. In order to estimate the former relationship reliably, we must be sure that our strategy is not confounded by the latter.

There are two main sources of simultaneity here. On the demand side, quality influences patient choices and willingness to travel, and therefore the number of competitors for each department. On the supply side, quality of incumbent departments may affect the number of providers of care through entry and merger decisions. We discuss each issue in turn.

Demand side

The literature has demonstrated that quality influences patient choices of hospital directly, and also influences patients' willingness to travel for care. As such, better hospital departments may have larger catchment areas than worse hospital departments, and therefore have more competitors. Alternatively, better departments might have smaller catchment areas because they capture a higher density of patients

³⁵ We exclude all community patients from the estimation itself, but control for the percentage of patients to partially address a selection issue where acute patients may be more complex and therefore harder to treat.

³⁶ We use the list of diagnosis codes under definition 2 in Ho et al (BMJ, 2017) to identify pre-existing pressure ulcers and risk factors.

³⁷ We use ICD10 code N39 as a proxy variable for UTIs and UTI risk factors

³⁸ We use the list of diagnosis codes for Pulmonary Embolism and Venous Thromboembolism provided in Sultan et al (BMJ, 2015) to identify pre-existing blood clots and risk factors.

than worse hospital departments: their quality overrides other factors which might make patients go elsewhere. This problem could cause us to identify a relationship between fascia count and quality, even if one did not exist.

Our solution is closely related to the existing literature's approach. In particular, Kessler & McClellan, Gaynor et al (2013) and subsequent authors construct a model of hospital demand which does not depend on quality. Patient flows are then predicted from this model, and used to compute concentration measures which are, by construction, exogenous to hospital quality. We follow Moscelli et al (2017) in using the resulting measures as *instruments* for the endogenous variables, rather than including the measures directly into the regression equation itself. We also make some modifications to the precise definition of the instrument, as discussed below, although find that our results are directionally robust without these.³⁹

Supply side

The current number of hospitals in any given area can be thought of as a combination of the historical locations of hospitals, which was primarily determined by (i) accident and population-based targets in the mid-1900s, which are likely as good as randomly assigned for purposes of studying today's quality, and (ii) and subsequent changes to this number through entry, exit and mergers. If recent changes are related to quality, this would cause an endogeneity problem. Is this the case?

Entry and exit of NHS hospitals are very rare. This is because NHS hospital *sites* are rarely built or demolished. Whilst occasionally new departments may open or be reconfigured, inspection of our data confirmed a very small number instances of this over three years.⁴⁰ Mergers between NHS hospitals present a slightly bigger issue. These reduce the number of competitors and have often taken place where one provider was not performing well. Although the number of NHS trusts remained fairly

³⁹ As an alternative, Bloom et al (2015) define each hospital's catchment area based on an exogenous, fixed radius of 15km. As a consequence, the fascia counts they use as their concentration measure are not influenced by patient flows. A downside of this approach is that it is likely to introduce measurement error. In our case, this could be severe because a large part of our identifying variation comes from differences in department-specific catchment areas within the same trust. We report a sensitivity using this approach in the appendix.

⁴⁰ Over the period 2010-2015, we find using NHS England Beds data that the number of trust-specialty combinations fell by just 0.45% per year overall, accounting for 0.03% of beds. This does not tell us about consolidation of specialties on particular sites with trusts; however using NHS Digital HESF data we find that only around 1.5% of sites close or are reclassified as not primarily acute sites per year (4.3, annually, from 2000-2015). In combination, entry and exit of NHS sites through time appears too low to have a significant impact on our analysis. Further, most 'exit' occurs for central capacity planning reasons, rather than due to local quality issues, which significantly mitigates any residual concern.

static in the years immediately before our data begins, the persistence of quality in the NHS may mean there is a residual endogeneity concern.

This issue has received some attention in the literature. Most significantly, Bloom et al (2015) instrument for concentration using the degree of political marginality in the parliamentary constituencies in the vicinity of each hospital trust. The paper is however almost entirely focussed on NHS hospitals, arguing that in 2006 ISTCs provided care was negligible in size; was restricted to “elective services for which there [were] long waiting lists in the NHS”, and was predominantly privately rather than publicly funded. Other papers also omit consideration of ISTC endogeneity. Given that the existing literature does not extend beyond 2010, this is perhaps reasonable as the ISTC program remained in its infancy over this period.

However, by the start of our particular data series, ISTCs accounted for 4.2% of admissions.⁴¹ Further, ISTCs generally entered near NHS hospitals which had high waiting times. These waiting times will be related to the quality of the NHS hospital, but the direction is ambiguous: either the hospital is of high quality and waiting times are high to ration ensuing demand, or it is of lower quality and waiting times are symptomatic of managerial and resourcing issues. In any event, the number of ISTCs within a given radius is likely to be endogenous, and cannot be ignored in our case.

The instrument

Our instrument is computed as the weighted average of the number of NHS sites available to each potential patient of a given department, where the weights are an exogenous probability that patient selects that department.⁴² For each department, we calculate the following:

$$IV_{stq} = \frac{\sum_i^I (NumNHSsites_{istq} \cdot Pr(choice)_{istq})}{\sum_i^I Pr(choice)_{istq}}$$

Where subscript i denotes individuals, s denotes specialties, t denotes trusts, and q denotes quarters. We construct the measure first by computing two component parts:

⁴¹ Up to 17% of departments in our data were ISTCs, although this varies significantly by specialty: Trauma & Orthopaedics had by far the highest proportion of ISTCs.

⁴² We compute this measure at site level, and then aggregate to trust level as the weighted average across each site, where the weights used are the sum of the exogenous predicted probabilities of each site.

- (i) We model patient flows as a function of distance and distance interacted with patient urban status only.⁴³ For each patient, we obtain the predicted probability $Pr(choice)$ that they select any given department in their choice set, but this probability does not depend on the quality of that department.⁴⁴
- (ii) We count the number of competitors which each patient has in their choice set. We then weight this number by the predicted probabilities above. Using these weights excludes patients who do not have that department in their choice set, and captures the idea that departments are unlikely to place much weight on competition in distant areas from which they are unlikely to attract notable volumes.

There are two conditions for the validity of an instrument: exclusion and relevance. To satisfy the exclusion restriction, again conditional on all covariates, the instrument cannot be correlated with the error term. This will be true if both components (i) and (ii) of the measure are not functions of quality:

The discussion in the previous section shows that the number of fascia could be endogenous due to mergers and ISTC entry. For this reason, our competitor count is the number of NHS sites. Specifically, we count sites rather than trusts, because the number of acute sites does not usually change following a merger. We exclude ISTCs from the count because the presence of an ISTC may be related to local NHS quality. The number of NHS sites is highly correlated with the number of trusts patients' choice sets, but is exogenous.

We then weight this count by the 'exogenous' predicted probabilities calculated in (i). These are not a function of hospital quality. This will be true provided that distance between patients and hospitals and its interaction term with patient urban status are unrelated to quality. Contrary to schooling, it is very rare for people to move to live close to good hospitals. Although ultimately an assumption, this is a common plank of a wide range of studies.⁴⁵

Because none of the component parts of the instrument depend on quality of the local hospitals, the instrument is plausibly exogenous, and therefore satisfies the exclusion

⁴³ We divide all patients between those living in a conurbation (33% of our data), and those in a more rural area. Our model uses 40% of available patients in order to reduce the computational burden of the exercise.

⁴⁴ For reasons that will become clear below exclude independent hospitals and do not aggregate together different sites of the same NHS trust in computing these probabilities.

⁴⁵ The choice set is defined as all hospitals located within a circle with the radius equal to the 80th percentile distance travelled by patients to their chosen hospital, considering all patients with the same rural/urban status in the country. Since this radius is a national average, it does not vary by local area, and therefore local quality. On the latter, there is potential for further endogeneity. We discuss this more in the next section

restriction. Conditional on all covariates, the instrument is also highly correlated with concentration, suggesting instrumental relevance. We present the standard tests later in this paper.

6. Data and Estimation Strategy

Our primary source of data is an extract of Hospital Episode Statistics (HES) data. We match in various additional publicly available data. This includes UK Office of National Statistics (ONS) information at postcode or lower super output area (LSOA) level, such as rural/urban status; NHS Digital Organisation Data Services data on the locations of NHS Trusts, Sites and GP Practices in England; NHS Digital Hospital Estates and Facilities data covering further characteristics of all NHS Trusts (such as number of beds); and data from the NHS Harm Free Care Team on harm rates, wards and further patient characteristics.

Summary statistics for the variables which enter our main specification are shown in Table 1 below. Additional summary statistics can be found in the appendix.

Table 1: Summary statistics

	(1) N	(2) Mean	(3) Sd	(4) Min	(5) Max
Patient acquired at least one new harm	1,196,000	0.0282	0.165	0	1
Fascia Count	1,176,000	3.005	1.898	1	10
HHI	1,181,000	8,528	1,291	4,936	10,000
Patient had Old Pressure Ulcer	1,196,000	0.038	0.192	0	1
Patient has Long Term Catheter	1,196,000	0.0079	0.090	0	1
Patient is Age 70+	1,196,000	0.60	0.49	0	1
Patient is Female	1,196,000	0.58	0.49	0	1
Number of patients in same ward	1,196,000	73.64	25.18	1	246
Department treats community patients	1,196,000	0.076	0.26	0	1
Department mean procedure resource intensity	1,196,000	5.195	1.69	2.08	8.52
Department % high comorbidities	1,165,000	0.36	0.15	0	0.91
Department % transferred non-elective admissions	1,196,000	0.18	0.13	0	0.60
Department % admitted with pressure ulcer	1,192,000	0.0473	0.0413	0	0.172
Department % admitted with UTI	1,194,000	0.0685	0.0667	0	0.256
Department % admitted with blood clot	1,194,000	0.0119	0.0119	0	0.0479

Notes: Summary statistics for variables included in the main specification only. The data is at patient level. Therefore, for variables measured at department level such as fascia count, this table shows the patient-

weighted averages of those variables. Several variables enter our regression in logs but to aid interpretation we show the untransformed variables. Sources: NHS Digital HES data; Harm Free Care Data; ONS NSPL.

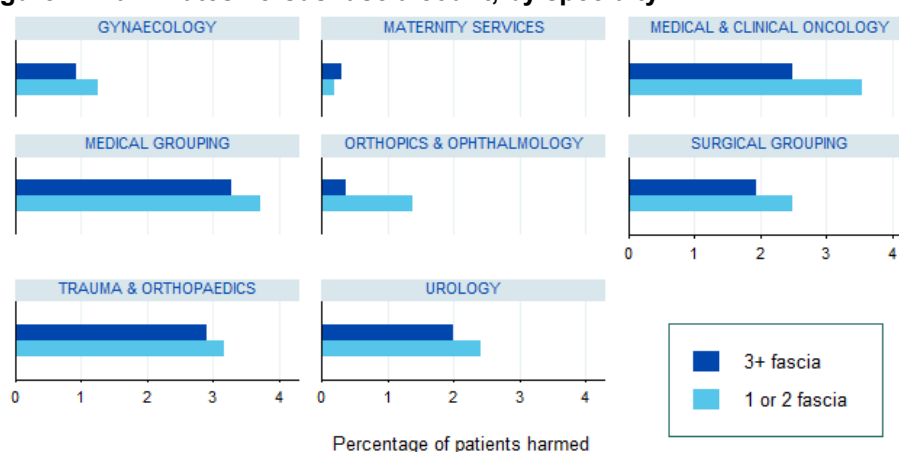
As we explained above, our data contains 8 specialty groupings.⁴⁶ The largest comprises 19% of department-quarter combinations in our data.⁴⁷ We use eight quarters of data in the financial years 2013/14 and 2014/15, each comprising just over 12% of the data.

Descriptive comparisons

Before undertaking any statistical work, it is instructive to check whether a relationship already exists in the data. Whilst these statistics do not allow us to make any causal interpretation, they provide context.

A key contribution of this paper is to relate analysis to specialties, the actual way in which hospitals organising themselves. Figure 2 shows the average harm rate for each specialty, split by whether the department has at most one other competitor. The chart shows that in departments more concentrated areas almost always have higher harm rates on average.

Figure 2: Harm rates versus fascia count, by specialty



Notes: Each bar represents the mean rate of harm across departments, split by specialty and concentration. We measure concentration using fixed radius fascia counts: i.e. the number of distinct NHS Trusts and ISTCs within a fixed radius catchment area determined, for each specialty, as 1.5 times the distance over which 80% of

⁴⁶ Our list of 8 specialties is formed of 18 different “treatment specialties”. We make the following groupings: Obstetrics and Midwifery (collectively termed “Maternity Services”); Medical and Clinical Oncology; Geriatric Medicine, Transient Ischaemic Attack, Stroke Medicine and General Medicine (collective, “Medical Specialties”); and Breast Surgery, General Surgery, Colorectal Surgery and Vascular Surgery (collectively, “Surgical Specialties”). We undertake this grouping exercise because each individual treatment specialty is frequently cross-coded others in this list, because they often draw on similar resources and staff, and because it is common for hospitals to locate these specialties in the same physical locations. We draw on previous views expressed by hospitals and economic analysis from CMA merger cases in order to determine which specialties should be grouped together (CMA, 2015 and 2017). Including treatment specialty fixed effects alleviates concern about within-group heterogeneity in underlying riskiness of patients potentially generated in this process.

⁴⁷ The number of patients is skewed towards the grouped specialties; we address this using fixed effects and a sensitivity in which we run regressions at department level, therefore assigning each equal weight.

patients travel to hospital: this radius is common to all departments in the same specialty. We then split the data by whether each department has at most one other competitor in this radius, or more. Harm rates are the percentage of patients experiencing harm from a fall, pressure ulcer, blood clot or UTI. Sources: NHS Digital HES; Harm Free Care data; ONS NSPL data.

Estimation strategy

We use OLS to estimate a simple linear probability model (LPM) as follows.

$$Pr(harm)_{itqs} = \alpha + \beta_1 Conc_{tqs} + \beta_2' D_{tqs} + \beta_3' P_{itqs} + \mu_{itqs} + \gamma_{sq} + \delta_{tq} + \varepsilon_{itqs}$$

where X is a vector of explanatory variables, α is a constant term; $Conc_{tqs}$ is concentration, and D_{tqs} is a vector of control variables measured at the trust-quarter-specialty level with associated coefficient vector β_2 ; and P_{itqs} is a vector of controls at *patient*-trust-quarter-specialty level with coefficients β_3 . We also include three sets of fixed effects for treatment groups μ_{itqs} , specialty-quarter combinations γ_{ts} , and trust-quarter combinations δ_{tq} . Subscripts are denoted as follows: i individual; q quarter; s specialty; and t trust. For all regressions except where specified, we use heteroskedasticity-robust standard errors clustered at the trust-quarter level.

Regressions of this form will only yield unbiased estimates of β_1 if all explanatory variables are uncorrelated with the error term. However, for the reasons set out above, this is unlikely to be true because concentration is simultaneously determined with quality. For our main results, we therefore use two stage least squares, instrumenting for concentration.

There are well known disadvantages of using a LPM, in which the effect of explanatory variables is assumed to have a linear effect on probabilities. In particular, the model permits the predicted probability of harm to be outside the [0,1] interval. However, it also has several advantages. In particular, we are able to use the standard two stage least squares estimator which relies less strong assumptions than some standard IV techniques for probit regression.⁴⁸ In any event, we find that a small proportion of predicted values are negative, and our estimates do not change significantly when we implement probit regression (as shown in the appendix).⁴⁹

We expect introducing more competitors will reduce the rate of harm. However, we do not necessarily expect this effect of be linear: that is, adding competitors will reduce rates of harm, but the size of this reduction is likely get smaller as the number of

⁴⁸ Further, the choice of nonlinear regression function is ultimately fairly arbitrary: therefore even if the function should be nonlinear, there is no guarantee that probit, logit or any other function is the right nonlinear one.

⁴⁹ 6% of predicted probabilities are negative in our main specification; we drop these in postestimation

competitors increases. We therefore compare models in which concentration enters linearly with nonlinear transformations which give this property. For fascia count we test the reciprocal of fascia count, and for the HHI, which is *decreasing* in the number of competitors, we follow the literature in using the log of the HHI. For both, we also test a square term.⁵⁰⁵¹

7. Main estimates

Table 2 shows the OLS, first stage, reduced form and LPM IV estimates for these models using our main set of controls.

Table 2: Effect of concentration on rates of harm

Regression	Fascia Count		HHI	
	Linear	Nonlinear (Reciprocal)	Linear	Nonlinear (Log)
OLS	-0.0011*** (0.00026)	0.003* (0.0017)	0.00000024 0.00000050	0.00069 (0.0037)
IV	-0.0023 (0.0019)	0.05*** (0.017)	0.0000047 (0.0000033)	0.061** (0.027)
First stage	0.43*** (0.076)	0.21*** (0.046)	-241*** (38)	-0.066*** (0.011)
Reduced form	-0.0011 (0.00077)	0.011*** (0.0031)	-0.0011 (0.00077)	-0.0041** (0.0017)

Notes standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. The dependent variable is a dummy which takes value 1 if the patient experienced a fall, a new pressure ulcer, a blood clot or a UTI in the care of the hospital. All regressions include our main set of control variables to account for confounding factors such as differing pre-existing health status and department characteristics, as well as trust-quarter and quarter-specialty fixed effects. Concentration is measured using both the count of competitors in the hospital's catchment (Fascia Count), and the HHI (which is measured on a scale between 0 and 10,000). Standard errors are clustered at department-quarter level.

The estimated coefficients in linear probability models are interpreted as the *percentage point* change in probability of the outcome, in response to a unit change in the independent variable. Note that a one unit increase in the reciprocal of fascia count is in fact an *increase* in concentration, rather than a decrease, and so the expected sign is different from that on the linear model. Because the magnitude of the effects

⁵⁰ Note that different nonlinear transformations are required across concentration measures: it is not appropriate to take the reciprocal of the HHI or the log of fascia count. This is because the former would transform the HHI from a nonlinear function of the number of (equal size) competitors to a linear one, and logging fascia count would require the impact of adding fascia to be *greater* for higher values of fascia count, not smaller.

⁵¹ We also test a square term: the results are broadly consistent, and are reported in the appendix. For reasons of simplicity, we select the reciprocal of fascia count and the log of the HHI as our headline results, rather than use quadratic terms: dealing with one excluded instrument is simpler than two excluded instruments.

are hard to interpret from the regression output itself, we focus immediately on the sign and significance levels of the coefficients, before transforming the estimated effects to allow more natural interpretation.

The results show that higher concentration is associated with higher probability of harm for patients, and that this effect is statistically significant.⁵² The concentration coefficients in the linear model are not significant, but become so in the models with nonlinear transformations, indicating that the data is more consistent with the benefits of competition ‘petering out’ once there is a sufficient number of alternatives. Both measures of concentration give similar results. The IV estimates are larger in magnitude than the OLS estimators, which indicates that the IV strategy is necessary to produce unbiased results. The instruments pass conventional tests: in particular, the f-test on the excluded instrument is 22.09 for the reciprocal fascia count and 36.46 for the log of the HHI. Both exceed the Stock & Yogo (2005) critical values, such that we can reject the hypothesis of weak instruments based on a relative size of 10%.

Table 3 shows the IV coefficients for the specifications with non-linearly transformed concentration variables below for all control variables. Controls variables generally have the expected signs. Further sensitivity checks to the set of controls are presented below.

⁵² It is less straightforward to perform tests on the strength of the first-stages of IV regressions for two excluded instruments under clustered standard errors, than it is for just one excluded instrument. Therefore, for simplicity, we choose the transformations where we only require one excluded instrument.

Table 3: Effect of concentration on rates of harm – IV regression tables showing controls

	(1)	(2)
Reciprocal fascia count	0.05*** (0.017)	
Log HHI		0.061** (0.027)
(Dummy) Patient had old pressure ulcer	0.012*** (0.0013)	.011*** (0.0012)
(Dummy) Patient has long term catheter	0.039*** (0.0031)	0.039*** (0.003)
(Dummy) Patient is age 70+	0.015*** (0.00046)	0.015*** (0.00045)
(Dummy) Patient is female	-0.0034*** (0.00041)	-0.0035*** (0.00041)
(Log) Number of patients in same ward	0.0017** (0.00075)	0.0017** (0.00075)
(Dummy) Department treats community patients	0.0027 (0.0019)	0.0049** (0.0021)
(Log) Department mean procedure resource intensity	0.011** (0.0048)	0.0074* (0.0039)
(Log) Department % high comorbidities	-0.0062 (0.012)	-0.016* (0.0092)
(Log) Department % transferred non-elective admissions	0.012** (0.0056)	0.0029 (0.0047)
(Log) Department % admitted with pressure ulcer	-0.0035 (0.017)	-0.0074 (0.016)
(Log) Department % admitted with UTI	0.018 (0.018)	-0.0013 (0.014)
(Log) Department % admitted with blood clot	0.14** (0.064)	0.1* (0.057)
Constant	-0.064* (0.033)	-0.82** (0.34)
Trust-Quarter Fixed Effects	YES	YES
Quarter-Specialty Fixed Effects	YES	YES
Treatment Specialty Fixed Effects	YES	YES
Observations	1,057,236	1,059,728

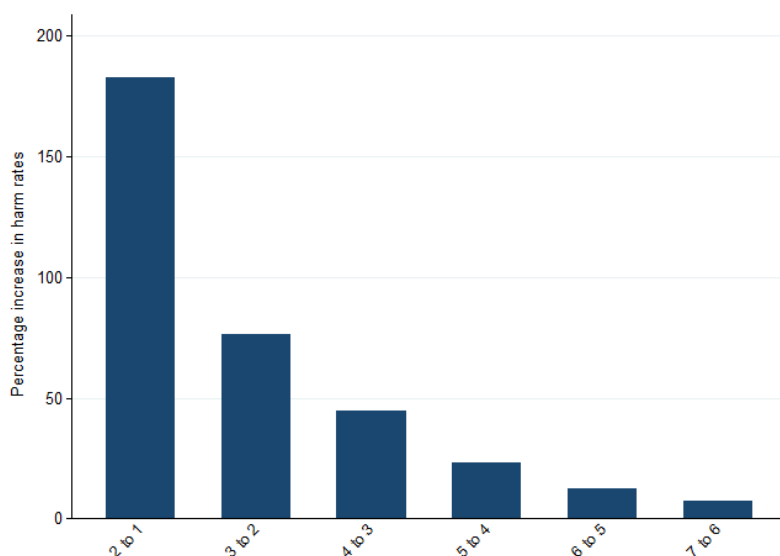
Notes: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. The dependent variable is a dummy which takes value 1 if the patient experienced a fall, a new pressure ulcer, a blood clot or a UTI in the care of the hospital. Both regressions instrument for concentration using the same nonlinear transformation of the instrument as is used for the concentration variable. The controls for Old Pressure Ulcers, Long Term

Catheters, Age, and Gender are all dummy variables at patient level, and the number of patients being treated in the same ward also enters into the regressions at patient level. The other controls enter at department level: we control for whether the department treats some patients in the community, the mean procedure hierarchy (a measure of the intensity of operations performed), as well as the percentage of patients treated which have above average comorbidities, are non-elective patients (admitted through sources other than the hospital's own A&E department), and are overnight stay patients. Standard errors are clustered at department-quarter level.

Interpreting the magnitude of the percentage point changes in the above tables is not straightforward due to the nonlinear transformation for fascia count, and because even comparatively small percentage point increases in probabilities can be economically significant where the event is relatively rare to start with.⁵³

It is therefore useful think instead about *percentage* changes in the dependent variable when concentration increases. For the fascia count regression (1) above we compute this percentage change in response to a one unit *reduction* in fascia count (i.e. as a semi-elasticity); for the HHI in response to a 10% *increase* in the index (i.e. as an elasticity). We compute the average (semi-)elasticity within each hospital department,⁵⁴ and present the median across departments, conditional on the degree of remaining competition after a hypothetical merger.⁵⁵ The relevant formulae are provided in Appendix 2.

Figure 3: Estimated percentage increase in harm from a merger: main specification



Notes: The chart above shows the median percentage harm increase (semi-elasticity) across departments following a one unit reduction in fascia count. These are calculated as follows: first, we find the patient level

⁵³ That is, adding an additional 0.3 percentage points onto the probability of harm might seem like a small effect of mergers, but this is equivalent to a 10 percent increase in harm rates (which start from a base of around 3%).

⁵⁴ We provide further detail in how these are calculated in the appendix.

⁵⁵ We do not test the statistical significance of the semi-elasticities, but note that the marginal effects are all significant at the 1% level.

semi-elasticity using the coefficients from model (1) and its predicted harm rates; we then take the mean semi-elasticity within the department to get department level semi-elasticities. We plot the median semi-elasticity (to account for potential outliers) for all departments conditional on the (rounded) number of competitors faced, omitting departments facing no competitors for which there can be no further decrease.

Figure 3 shows that under our assumptions on the form of nonlinearity, the estimated percentage increase in harm rates is significantly higher when there are few remaining competitors to when there are many. In particular, a hypothetical merger to monopoly would result in a 182% increase in the number of patients experiencing harm.

At the median department, a one unit increase in fascia count is estimated to lead to a 41% increase in harm rates. The estimated effects for the (log) HHI are consistent with those for fascia count, if not larger: at the median department, a 10% increase in the HHI is estimated to lead to a 34% increase in harm rates.⁵⁶

8. Robustness checks

To test whether our results were driven by any particular modelling decisions or data issues, we undertook a range of sensitivity tests. The results for eight of these are presented in Table 4 below: further tests are shown in the appendix.

Table 4: Control and specification sensitivities

	Reciprocal Fascia Count		Log HHI	
	IV Coefficient	Median Semi-elasticity	IV Coefficient	Median Elasticity (10% increase in HHI)
1. Add further controls	0.046*** (0.014)	39%	0.087** (0.036)	46%
2. Add further VTE controls	0.05*** (0.017)	45%	0.061** (0.027)	35%
3. Add interaction terms for age	0.051*** (.017)	43%	0.061** (.027)	36%
4. Include day of week dummies	0.05*** (0.017)	42%	0.061** (.027)	34%
5. Exclude casemix severity variables	0.061*** (0.02)	52%	0.076*** (0.028)	43%
6. Exclude pre-existing rates of pressure ulcers, UTIs and VTE	0.062*** (0.023)	52%	0.065** (0.03)	35%

⁵⁶ We report the median rather than the mean because dividing the marginal effect by the probability of harm leads to some very large (semi)-elasticities where the predicted rate of harm is very low for some departments. The median is not affected by large outlier values.

7. Cross-sectional variation only	0.047* (0.025)	52%	0.078* (0.045)	56%
8. Drop first 6 months of data	0.058** (0.023)	53%	0.062** (0.031)	37%
9. Cluster SEs across trusts	0.05* (0.03)	57%	0.061 (0.046)	55%

*Notes: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. The dependent variable is a dummy which takes value 1 if the patient experienced a fall, a new pressure ulcer, a blood clot or a UTI in the care of the hospital. All regressions instrument for concentration using the same nonlinear transformation of the instrument as used for the concentration variable. All regressions except (8) cluster standard-errors at trust-quarter level. The leftmost column describes the differences between the regression run and our main regressions (1) for fascia count and (2) for the HHI. The median (semi-)elasticity is the percentage increase in harm rates at the median. We use median to account for potential outliers.*

Row 1 shows the results when we add controls for rates of patients living in above median unemployment or deprivation areas, A&E admission status, overall volumes, and whether the ward only treated overnight patients. Row 2 also controls for the percentage of patients risk-assessed for VTE, and the percentage of patients receiving VTE prophylaxis. Row 3 interacts patient age with all other patient level case-mix controls (female, old pressure ulcers, long term catheter and the treatment specialty fixed effects). Row 4 includes the weekday on which the hospital collected the data to account for potential systematic variation in case-mix (i.e. complexity of treatment) across different times of the week. Row 5 excludes case-mix controls which could be picking up sicker patients going to better hospitals: co-morbidities; number of patients treated in the ward; procedure intensity. Row 6 excludes the controls for pre-existing rates of pressure ulcers, UTIs and blood clots. Row 7 averages each department's concentration values across quarters, thereby using only cross-sectional (within trust) variation to estimate the concentration coefficient. Row 8 drops the first 6 months of data (our series was only 12 month old at the start of the time period, so we check whether the results are driven by data teething issues early on). Row 9 clusters at trust rather than trust-quarter level. In all cases, the sign is the same and the magnitude of the estimated effects remains appreciable. The effects remain statistically significant at (or below) the 5% level in most specifications.

We also test the sensitivity of our results to the decision to use a linear probability model over nonlinear alternatives. In particular, we estimate our main specifications (1) and (2) using instrumental variables probit regression via maximum likelihood. Table 5 shows the coefficients of these regressions for each concentration measure.

Table 5: Probit regressions

Regression	Fascia Count		HHI	
	Linear	Nonlinear (Reciprocal)	Linear	Nonlinear (Log)
Not instrumented	-0.015*** (0.0044)	0.018 (0.031)	0.0000085 (0.0000087)	0.047 (0.064)
IV	-0.047 (0.037)	0.84*** (0.31)	0.000099 (0.000067)	1.1** (0.52)
First stage	0.43*** (0.076)	0.21*** (0.046)	-239*** (38)	-0.065*** (0.01)
Reduced form	-0.024 (0.015)	0.18*** (0.056)	-0.024 (0.015)	-0.075** (0.033)

Notes standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. The dependent variable is a dummy which takes value 1 if the patient experienced a fall, a new pressure ulcer, a blood clot or a UTI in the care of the hospital. All regressions include our main set of control variables, including trust-quarter and quarter-specialty fixed effects. The IV estimates are computed using stata's `-ivprobit-` command.

These regressions show a very similar picture to those presented by our main LPM estimates: in particular, the IV estimates are larger in magnitude than those for the model which does not attempt to address endogeneity issues, and the IV coefficients on the nonlinearly-transformed concentration variables are statistically significant. The t-statistics on the first stage coefficients are also high, likely indicating instrumental relevance.

The reciprocal fascia count model implies a median 20% increase in harm rates following a hypothetical merger, and the log HHI model implies a median 41% increase following a 10% increase in the HHI: these magnitudes are reasonably comparable to our main results.⁵⁷

Our conclusions are also robust to other changes in specification; in particular, we obtain broadly similar results when we use an alternative non-linear transformation for concentration (inclusion of a quadratic term); compute fascia counts over different distances; use different radii definitions in computing the fascia counts; use individual sub-indicators of the harm measure as the dependent variable; and when we aggregate all data to department level.

⁵⁷ The estimated semi-elasticity for hypothetical 2 to 1 mergers is 274% again this magnitude is very large.

Application to other quality variables

The existing literature has focussed on rarer but more severe health outcomes, usually related to the primary healthcare intervention the patient receives.⁵⁸ It is possible to extend our methodology to consider one often studied such measure, in-hospital mortality. This serves as a way to link our ‘harm’ results to the pre-existing literature, and is also of interest in itself.

Table 6 below shows the concentration coefficients for our mortality regressions. We use site (rather than trust) level data for these regressions, since the quality measure can be constructed from our HES data itself. We include similar although not identical controls to those used in our harm indicator regressions: some changes are warranted due to the different nature of the outcome we are modelling (casemix measures which are predictive of our four harms are likely to be different from those which are predictive of mortality). Some others are required by the different sources of data we use.⁵⁹

Table 6: Mortality regressions

Regression	Linear Probability Model	
	Reciprocal FC	Log HHI
OLS	.00074 (.00063)	.0047*** (.0011)
IV	.011** (.0041)	.021** (.0094)
First stage	0.28*** (0.041)	-0.058*** 0.0090
Reduced form	.003*** (.001)	-.0012** (.00053)

⁵⁸ One approach in the literature is to risk-adjust the dependent variable by generating the ratio of the observed probability to the predicted probability of death at hospital level, where predicted probabilities are computed in a regression of mortality on a set of patient characteristics and fixed effects. Rather than undertake this two-step procedure, we simply include patient characteristics and fixed effects in our main regression.

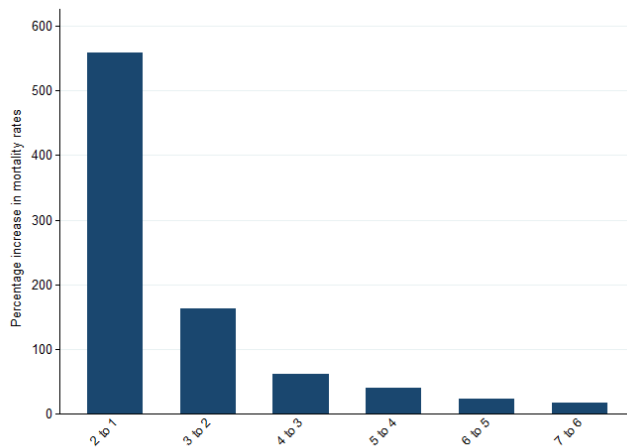
⁵⁹ In particular, we exclude controls for old pressure ulcers and long term catheters since these are unlikely to be good proxies for mortality risk. We exclude the number of patients treated in the same ward since this information is not available in HES data, and because the quality of the primary intervention may be more important for mortality than the care received on wards. We include controls for the broad diagnosis (ICD chapter) of the patient; the day of the week they were admitted, the nationally-determined price paid to the hospital for the treatment, the patient’s neighbourhood (LSOA) deprivation index (IMD), and whether the patient was admitted from the trust’s A&E (separately from other non-elective admission methods). We replace the dummy for patients over 70 years old with a (logged) continuous variable for age. We control directly for overnight status, non-elective status, and procedure hierarchy at the patient level rather than using department level averages. Our results do not change when we use as close to the same set of controls as the harm regressions as the data allows.

Notes standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. The dependent variable is a dummy which takes value 1 if the patient died within 30 days of admission in hospital (we do not observe deaths outside of hospital).

Consistent with much of the existing literature, greater concentration appears to increase mortality. The mortality results are larger although directionally consistent with the harm results, and the f-statistics on the first-stage regressions again exceed the critical values: 49.21 for the reciprocal fascia count, and 42.9 for the log specification.

Figure 4 below shows the semi-elasticities generated by the reciprocal fascia count LPM. The median department sees a harm rate increase of 55% following a hypothetical one unit decrease in fascia, and 129% following a 10% increase in the HHI. The larger figures may in part reflect a smaller base: in-hospital mortality is very rare for elective care.

Figure 4: Estimated percentage increase in mortality from a merger



Notes: The bar chart shows the median semi-elasticities across departments from our main in-hospital mortality specification. This uses similar controls as our main harm specification. These are calculated as follows: first, we find the patient level semi-elasticity using the reciprocal FC LPM regression and its predicted harm rates; we then take the mean semi-elasticity within the department to get department level semi-elasticities. We plot the median semi-elasticity (to account for potential outliers) for all departments conditional on the (rounded) number of competitors faced, omitting departments facing no competitors for which there can be no further decrease.

9. Policy implications and conclusion

Our paper contributes further empirical evidence that competition ultimately benefits patients. We have used a new methodology in this literature to exploit variation in the degree of competition faced by the same hospital trusts across their different departments in order to estimate the impact of competition on quality. We have studied a new quality measure: the incidence of patient harm through falls, pressure ulcers,

blood clots and UTIs. Our analysis covered two years of data in 2013/14 and 2014/15, and eight specialties.

Our main estimate is that a hypothetical future merger between two geographically proximate hospitals would, on average and assuming no offsetting clinical benefits are unlocked by the merger, result in a 41% increase in harm rates. The magnitude of the effect is significantly greater where few competitors remain post-merger, and smaller when several alternatives for patients remain. This effect is robust to a number of alternative specifications and holds also when we consider in-hospital mortality.

The primary cost of a fall in quality is to patients' well-being. If valued in quality adjusted life years, for example, the magnitudes would be high. This is consistent with existing literature. There is also a financial cost because the harms we have studied must then be treated by the NHS. Research placed the financial impact of these harms at around £430 million per year across the whole of the English NHS (Department of Health, NHS's Harm Free Care team, 2010). Taking the 139 trusts in our data, this implies an average cost of £3.1 million per trust per year, assuming the total cost of harm has not changed significantly. If the harm rate were to increase by 41% as a result of a representative merger affecting all specialties in the hospitals, across both trusts involved, this would give an additional annual cost of £2.5 million per year counting only the four harm types.⁶⁰ For a merger to monopoly, the cost increase could be much higher. This simple calculation does not include the costs associated with other harms, such as increased mortality risk, nor lost revenue from reduced patient numbers caused by quality-conscious patients deciding to go elsewhere. This figure is clearly only indicative. However, it does demonstrate that competition in the NHS can also be beneficial in financial terms. Competition is likely to bring benefits through time; as a result, any such savings are likely to compound.

Clearly, there are a range of other factors which are central to efforts to drive up and maintain quality in healthcare systems. The validity of our conclusions regarding concentration to the sector in the future depends on how conditions and priorities in the NHS change, relative to the period studied. For example, deficits rose significantly in the years immediately following our data series, which may represent pressures on the supply side that reduce the salience of competition as a driver of quality. Nevertheless, our empirical work suggests that patient choice and competition between public hospitals can be an important driver of quality.

⁶⁰ That is, a cost of £430m/139 per trust, times two trusts in the data, multiplied by 41% to account for the effects of the merger. Note that the aggregate cost of harm is likely to have fallen since 2010 in line with lower overall harm rates; the magnitude of any such fall is not clear however.

10. References

- Beckert, W., Christensen, M., & Collyer, K. (2012). 'Choice of NHS-funded hospital services in England', *The Economic Journal*, vol.122(560), pp. 400-417.
- Bloom, N., Propper, C., Seiler, S., & Van Reenen, J. (2015). 'The impact of competition on management quality: evidence from public hospitals', *The Review of Economic Studies*, vol.82(2), pp. 457-489.
- Cooper, Z., Gibbons, S., Jones, S., & McGuire, A. (2011). 'Does hospital competition save lives? Evidence from the English NHS patient choice reforms', *The Economic Journal*, vol.121(554), pp. F228-F260.
- Competition Commission (2013). 'The Royal Bournemouth and Christchurch Hospitals NHS Foundation Trust/Poole Hospital NHS Foundation Trust: A report on the anticipated merger of The Royal Bournemouth and Christchurch Hospitals NHS Foundation Trust and Poole Hospital NHS Foundation Trust: Final Report', <https://www.gov.uk/cma-cases/royal-bournemouth-and-christchurch-hospitals-nhs-foundation-trust-poole-hospital-nhs-foundation-trust-merger-inquiry-cc> (last accessed: 16 December 2018).
- CMA (2015). 'Ashford and St Peter's and Royal Surrey County A report on the anticipated merger of Ashford and St Peter's Hospitals NHS Foundation Trust and Royal Surrey County Hospital NHS Foundation Trust: Final Report', <https://www.gov.uk/cma-cases/ashford-st-peter-s-nhs-foundation-trust-royal-surrey-county-nhs-foundation-trust> (last accessed: 16 December 2018).
- CMA (2015). 'Central Manchester University Hospitals and University Hospital of South Manchester: A report on the anticipated merger between Central Manchester University Hospitals NHS Foundation Trust and University Hospital of South Manchester NHS Foundation Trust: Final Report', <https://www.gov.uk/cma-cases/central-manchester-university-hospitals-university-hospital-of-south-manchester-merger-inquiry> (last accessed: 16 December 2018).
- Department of Health (2012). 'Delivering the NHS Safety Thermometer CQUIN 2012/13: A Preliminary Guide to Measuring "Harm Free" Care', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216534/dh_134329.pdf (last accessed 20 August 2018).
- Department of Health (2015). 'Handbook to the NHS Constitution', <https://www.gov.uk/government/publications/supplements-to-the-nhs-constitution-for-england>, (last accessed 20 August 2018).
- Gaynor, M., Moreno-Serra, R., & Propper, C. (2013). 'Death by market power: reform, competition, and patient outcomes in the National Health Service', *American Economic Journal: Economic Policy*, vol.5(4), pp. 134-66.
- Gaynor, M., Propper, C., & Seiler, S. (2016). 'Free to choose? Reform, choice, and consideration sets in the English National Health Service', *American Economic Review*, vol.106(11), pp. 3521-57.

- Gaynor, M. and Town R.J. (2012). 'Competition in Health Care Markets', in (M.V. Pauly, T.G. McGuire, P.P. Barros, eds.), *Handbook of Health Economics Vol. 2*, pp. 499-637, Amsterdam: Elsevier Science B.V.
- Harm free care (2014). 'NHS Safety Thermometer CQUIN 2014/15: Frequently Asked Questions', http://harmfreecare.org/wp-content/files_mf/HFC-FAQ-19.02.14.pdf, (last accessed 3 February 2018).
- Kessler, D. P., & McClellan, M. B. (2000). 'Is hospital competition socially wasteful?' *The Quarterly Journal of Economics*, vol.115(2), pp.577-615.
- Monitor (2014), 'NHS foundation trusts: consolidated accounts 2013/14', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/331166/Consolidated_Accounts_2013-14.pdf (last accessed 16 December 2018).
- Monitor (2015), 'NHS foundation trusts: consolidated accounts 2014/15', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/447099/50702_HC_238_WEB.pdf (last accessed 16 December 2018).
- Moscelli, G., Gravelle, H., Siciliani, L. (2016). 'Market structure, patient choice and hospital quality for elective patients', Working Paper, Centre for Health Economics, University of York, 2016.
- NHS England (2017). 'RTT Annual Report', <https://www.england.nhs.uk/statistics/statistical-work-areas/rtt-waiting-times/> (last accessed 14 September 2018).
- NHS Improvement (2016). 'Making mergers work: Improvements NHS providers have achieved through merger', https://improvement.nhs.uk/documents/104/Mergers_factors.pdf (last accessed 10 August 2018).
- NHS Improvement and NHS England (2016), 'Outpatient appointments summary', <https://www.england.nhs.uk/wp-content/uploads/2015/09/monitor-nhse-outpatient-appointments-summary.pdf> (last accessed 10 August 2018).
- NHS Improvement (2018). 'NHS foundation trusts: consolidated accounts 2017/18', https://improvement.nhs.uk/documents/3017/Consolidated_FT_accounts_201718.pdf (last accessed 16 December 2018).
- Skellern, M. (2018). 'The Effect of Hospital Competition on Value-Added Indicators of Elective Surgery Quality', Working Paper, London School of Economics.
- Stock J., Yogo M. (2005). 'Testing for Weak Instruments in Linear IV Regression', in (Andrews D.W.K., ed), *Identification and Inference for Econometric Models*, pp. 80-108, New York: Cambridge University Press.
- Sultan A. A., West J., Grainge M. J., Riley R.D., Tata L.J., Stephansson Olof et al. (2016) 'Development and validation of risk prediction model for venous thromboembolism in postpartum women: multinational cohort study', *BMJ* 2016; pp.355:i6253

APPENDIX

A1. Additional statistics and explanation

Quality measures

We tabulate the correlations between all subcomponents of our combined ‘harm rates’ measure of quality below. We include the correlations of each subcomponent with in-hospital mortality.

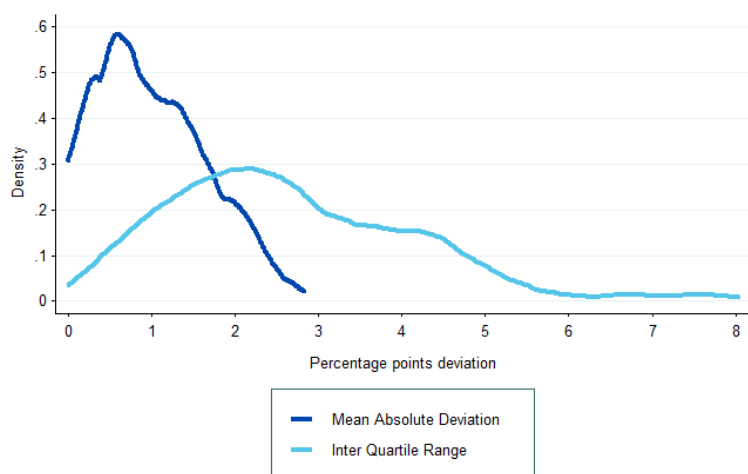
Table A1: Department level correlations between individual quality indicators

<i>Indicator name</i>				
Falls	Falls			
Pressure Ulcers	26%***	Pressure Ulcers		
Blood Clots	20%***	18%***	Blood Clots	
UTIs	4%	9%**	2%	UTIs
In-Hospital Mortality	14%***	17%***	16%***	7%*

Notes: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. Data is at department level for 2015 quarter 1 only. Source: NHS Digital HES, Harm Free Care data.

As explained in the main body, the mean absolute deviation statistic indicates that there is significant intra-trust variation in quality. We also find consistent results when we use the IQR: around half of trusts had a difference of 2 percentage points between departments situated at the bottom of their first and third quartiles. We plot both the MAD and IQR distributions below.

Figure A1: Intra-trust variation in quality: MAD and IQR



Instrumental variables approach:

Here, we explain how we have computed the instrumental variables in greater detail, and present additional statistics and results relevant to the critical assessment of the strength of our IV approach. As set out above, our approach is an adjusted of the approaches used in Kessler & McLellan (2000) and Gaynor et al (2013).

Our adjustments reflect the fact that we have used fascia counts rather than the HHI as the dependent variable; that we have an IV approach rather than the ‘reduced form’; and that we have accounted for possible simultaneity arising from ISTC entry decisions. A core part of our approach remains the calculation of exogenous patient flows however. We explain our specific approach, contrasted to the existing approach, in what follows.

As in the conventional approach, we begin by constructing a model of patient choice. We treat utility as an additive function of a systematic component \bar{u}_{ihsq} and a random component ε_{ihsq} . We are then able to use a conditional logit model in which patient i chooses hospital h to maximise utility u_{ij}

$$\begin{aligned}
 u_{ihsq} &= \bar{u}_{ihsq} + \varepsilon_{ihsq} \\
 &= \gamma_1 dist_{ihsq} + \gamma_2 (dist_{ihsq} \times urban_{ihsq}) + \varepsilon_{ihsq}
 \end{aligned}
 \tag{1}$$

where $dist_{ihsq}$ is the distance between patient i and hospital h in specialty s and quarter q , $urban_{ihsq}$ is a patient-specific variable denoting the urban density in the patient’s LSOA of residence, and ε_{ij} is an error term. Interacting distance and urbanicity allows for the fact that patients in built up areas may have different costs of travel to those in more rural areas.

We allow patients to select any hospital offering their specialty of treatment in that quarter, provided that hospital is located within radius r . This radius is defined as the 80th percentile of the distribution of patient journeys nationally, for all patients of the same urban status treated in that specialty and quarter.⁶¹

⁶¹ That is, for a given specialty and quarter, patient choice sets are permitted to be larger for patients living in a rural area than in an urban area. Being defined on the national rather than the local distribution of distances, the choice set is not specific to the area around any given hospital and therefore does not depend on any individual hospital’s quality.

Assuming that the error term is i.i.d. extreme-value distributed, equation (1) can be estimated as a conditional logit model via maximum likelihood. We use a 40% sample of the data to reduce the computational burden of the model; this still contains millions of patients. We estimate separate models for each specialty and quarter: both distance and the interaction term are highly significant. Under the relevant statistical assumptions on ε_{ij} specified above, the conditional probability that individual i chooses hospital h (denoted s_{ihsq}) is given by:

$$s_{ihsq} = \frac{\exp(\bar{u}_{ihsq})}{\sum_{m=0}^M \exp(\bar{u}_{imsq})} \quad (2)$$

where M is the set of all hospitals. These predicted probabilities are plausibly exogenous, since they depend only on the definition of the choice set, and two regressors: distance and patient urban density. As a consequence, concentration measures (such as the HHI) which are purely a function of these probabilities are also exogenous.

Our choice model differs from that taken in the literature, in two main ways, explained below.

The first difference is that our demand model is much simpler: we do not define choice sets to ensure that patients are able to choose from a range of hospital types (e.g. large and small hospitals), nor do we account for the potential influence of hospital characteristics (such as whether the hospital has teaching status, is a private hospital or is a large hospital) on patient choice.

However, because we follow Moscelli et al (2017) in using the concentration measures constructed from our predicted probabilities *as instruments*, rather than including them directly in the regression as in Kessler & McClellan (2011) and Gaynor et al (2014), this is not concerning. Provided the constructed concentration measures are sufficiently correlated with observed concentration, and the exclusion restriction continues to hold in our setting, the greater simplicity of our demand model generating patient flows will not influence the analysis.⁶² Indeed, as we explained in the main body, our instruments are highly correlated with observed concentration, which is required for instrumental relevance.

Omitting some hospital characteristics such as teaching hospital status is unlikely to cause our instrument to violate the exclusion restriction: the model just depends on

⁶² By contrast, a demand model with good predictive power is necessary to generate concentration measures that can be included directly into a regression, if the 'exogenous' concentration measures are to be interpreted on the same scales as the 'observed' or endogenous concentration measures.

fewer parameters which need to be justified as exogenous. Indeed, there is good reason to think that some of the characteristics which were previously included are potentially endogenous: higher quality hospitals are more likely to be teaching hospitals. Therefore, a simpler model may be better.

A second difference is our treatment of distance. Whilst we use the distance from patients' addresses to hospitals as a regressor in the demand model,⁶³ some previous papers in the literature have modelled choice probability as a function of *differential distance* (that is, the distance between hospitals). This is not a problem for instrumental relevance. For the exclusion restriction, we need to make slightly stronger assumptions however.

Specifically, conditional on all covariates, there cannot be any residual correlation between patients' distances from hospital and their underlying health status. We have mitigated concern by including a wide range of fixed effects and controls for underlying patient health status. Further in the UK people rarely relocate in order to increase their proximity to secondary acute healthcare. Therefore, this assumption is likely to be satisfied.

⁶³ We follow papers such as Gaynor, Propper and Seiler (2016) in taking this approach.

Summary statistics

Table A2: Summary statistics for the mortality dataset

	(1) N	(2) Mean	(3) Sd	(4) Min	(5) Max
Patient died within 30 days of admission	5,727,000	0.01	0.07	0	1.00
Fascia Count	5,727,000	3.04	1.97	1.00	10.00
HHI	5,727,000	8,370.00	1,180.00	5,111.00	10,000.00
Age	5,727,000	57.28	11.54	35.00	74.92
Patient is Female	5,727,000	0.57	0.18	0.07	1.00
Department treats community patients	4,668,000	0.05	0.21	0	1.00
Department mean procedure resource intensity	4,370,000	5.69	2.53	0	15.00
Patient neighbourhood has above median unemployment	5,727,000	0.49	0.50	0	1.00
Patient neighbourhood has above median deprivation	5,727,000	0.47	0.50	0	1.00
Patient has above median charlson comorbidity score	5,727,000	0.31	0.46	0	1.00
Patient non-elective, admitted via own A&E	5,727,000	0.21	0.41	0	1.00
Patient non-elective, transferred patient	5,727,000	0.11	0.31	0	1.00
Patient had overnight stay	4,664,000	0.76	0.28	0	1.00
Tariff payment for recorded care	4,809,000	0.50	0.50	0	1.00

A2. Calculation of (semi-)elasticities

As trailed in the main body, we transform the regression coefficients arising from our main regressions to aid interpretation. The statistics we calculate differ slightly depending on whether fascia count or the HHI is the underlying independent variable, due to convention in how these variables are usually interpreted. We take both in turn, first for the statistics following the linear probability models, and then the probit models.

In the case of fascia count, we compute the median (across departments) percentage increase in harm in response to a one unit decrease in fascia count. We use the median to account for the fact that dividing by small predicted probabilities makes the mean sensitive to specification.

In effect, we are “undoing” the non-linear transformation we applied to estimate the regression: it is harder to immediately visualise what harm linearly decreasing in the *reciprocal* of fascia count actually looks like, than it is to visualise a relationship between harm and fascia count itself. We are also relativizing the change in harm to the level of harm actually observed, by using percentage changes rather than percentage point changes (as estimated in the regression).

The semi-elasticity is calculated as follows, where the data is structured as combinations of individuals i (of I total individuals in a department), and departments d ; ${}^{64} Pr(\widehat{harm})_{its}$ is the predicted harm; $recip(\cdot)$ the reciprocal function, $\beta_{recip(FC)}$ the estimated coefficient from the regression, and $p50_{ts}[\cdot]$ the median function across departments. Note that we multiply by minus one because we want a measure of concentration which is decreasing in the number of competitors.

$$\begin{aligned}
 median\ semielasticity &= -p50_d \left[\frac{1}{I} \sum_{i \in d} \frac{\partial Pr(\widehat{harm})_{id}}{\partial FC_d} \cdot \frac{1}{Pr(\widehat{harm})_{id}} \right] \\
 &= -p50_d \left[\frac{1}{I} \sum_{i \in d} \frac{\partial Pr(\widehat{harm})_{id}}{\partial recip(FC)_d} \cdot \frac{\partial recip(FC)_d}{\partial FC_d} \cdot \frac{1}{Pr(\widehat{harm})_d} \right] \\
 &= -p50_d \left[\frac{1}{I} \sum_{i \in d} \beta_{recip(FC)} \cdot \left(-\frac{1}{FC_d^2} \right) \cdot \frac{1}{Pr(\widehat{harm})_d} \right]
 \end{aligned} \tag{1}$$

In the calculation, we use the chain rule to calculate the marginal effect of a one unit increase in *fascia count* at the individual level). To get the *percentage* rather than *percentage point* increase in harm rates, we compute the predicted harm probability of the patient (that is the pre-existing probability of harm implied by our model absent any changes in concentration), and divide the marginal effect by the pre-existing probability. This gives us a semi-elasticity. We then compute the mean semi-elasticity across all patients treated in the same department (that is, in the same specialty-trust combination, regardless of the point in time in which they were treated). In doing this, we assume the department’s semi-elasticity is equal to that of the representative patient. We then find the median of the resulting statistic across departments. Because the independent variable (fascia count) is still interpreted in

⁶⁴ A department is a trust-specialty combination. Concentration may vary for departments through time: we simulate an increase concentration across all periods, and so omit the time subscript.

units, but the dependent variable is interpreted as a percentage, this is a semi-elasticity.

In the case of the HHI, it is generally conventional to express concentration as an elasticity rather than a semi-elasticity: see Gaynor et al (2013) who use a log-log specification, for example. This is because the HHI does not, in itself, have a tangible interpretation.

We calculate this as follows, where HHI_i is the HHI which varies across each individual i , $\log(\cdot)$ denotes the natural logarithm, and all other notation is as above. To relate our results more closely to the magnitude of changes in concentration which could conceivably give rise to concern, we multiply the result by 10 to approximate the effect of changing the HHI by 10 percent, rather than just one percent (which is of course a tiny change).

$$\begin{aligned}
 \text{median elasticity} &= p50_{ts} \left[\frac{1}{I} \sum_{i \in d} \frac{\partial Pr(\widehat{harm})_{id}}{\partial HHI_d} \cdot \frac{HHI_d}{Pr(\widehat{harm})_{id}} \right] \\
 &= p50_{ts} \left[\frac{1}{I} \sum_{i \in d} \frac{\partial Pr(\widehat{harm})_{id}}{\partial \log(HHI)_d} \cdot \frac{\partial \log(HHI)_d}{\partial HHI_{ts}} \cdot \frac{HHI_d}{Pr(\widehat{harm})_{id}} \right] \\
 &= p50_{ts} \left[\frac{1}{I} \sum_{i \in d} \frac{\partial Pr(\widehat{harm})_{id}}{\partial \log(hhi)_d} \cdot \frac{1}{HHI_d} \cdot \frac{HHI_d}{Pr(\widehat{harm})_{id}} \right] \\
 &= p50_{ts} \left[\frac{\beta_{\log(HHI)}}{Pr(\widehat{harm})_{id}} \right]
 \end{aligned} \tag{2}$$

In the calculation, we compute again the marginal effect of log HHI which is equal to the estimated β . We divide by the mean pre-existing predicted probability of harm for each patient. This already gives us the elasticity, rather than the semi-elasticity because independent variables in regressions of linear-log form are already interpreted in percentage terms. We then find the mean elasticity within departments, and take the median of the resulting distribution.

To relate our results more closely to the magnitude of changes in concentration which could conceivably give rise to concern, we multiply the result by 10 to approximate the effect of changing the HHI by 10 percent, rather than just one percent (which is of course a tiny change).

For the probit sensitivities, we compute the same statistics in much the same way. For fascia count, the calculation is as follows, where φ represents a standard normal

probability density function and $\mathbf{x}_i' \boldsymbol{\beta}$ the matrix formed by multiplying the coefficient vector by the vector of explanatory variables for each observation its ; that is, the linear prediction of the model, and all other notation is as above.

$$\begin{aligned}
 \text{median semielasticity} &= -p50_d \left[\frac{1}{I} \sum_{i \in d}^{I \in d} \frac{\Pr(\widehat{\text{harm}})_{id}}{\partial FC_d} \cdot \frac{1}{p(\widehat{\text{harm}})_{id}} \right] \\
 &= -p50_d \left[\beta_{\text{recip}(FC)} \cdot \left(-\frac{1}{FC_d^2} \right) \cdot \varphi(\mathbf{x}_{id}' \boldsymbol{\beta}) \cdot \frac{1}{\Pr(\widehat{\text{harm}})_{id}} \right]
 \end{aligned} \tag{3}$$

Again, we express the effect of fascia count in a (negative) semi-elasticity and of the HHI in an elasticity.

$$\begin{aligned}
 \text{median elasticity} &= p50_d \left[\frac{1}{I} \sum_{i \in d}^{I \in d} \frac{\partial \Pr(\widehat{\text{harm}})_{id}}{\partial HHI_d} \cdot \frac{HHI_d}{\Pr(\widehat{\text{harm}})_{id}} \right] \\
 &= p50_d \left[\frac{\beta_{\log(HHI)} \cdot \varphi(\mathbf{x}_{id}' \boldsymbol{\beta})}{p(\widehat{\text{harm}})_{id}} \right]
 \end{aligned} \tag{4}$$

In computing these statistics from the linear probability model, we have to drop all patients which have predicted probabilities of zero or lower: we cannot divide by zero, and negative predicted probabilities introduce sign errors into the end results. Fortunately, only a small proportion of patients need to be dropped. None need to be dropped in the post-estimation following the probit regressions, which yield similar results. Identical formulae were used for the mortality regressions: we simply replaced the dependent variables.

A3. Further sensitivities

Here, we present the results of additional sensitivities we have conducted. First, we demonstrate an alternative nonlinear transformation to our concentration variables, Second, we flex the radius over which we count fascia. Lastly, we adjust the weighting of departments in the data by running regressions on department-level data.

Alternative non-linear form

Our main results use nonlinear transformations of our concentration variables, specifically chosen such that we only have one endogenous variable and excluded

instrument whilst retaining plausible functional form. Other non-linear transformations are of course also possible. We test one other commonly used candidate approach: introducing a quadratic term alongside the linear term. This has the advantage of flexibility but also introduces some additional complexity given that we now require multiple first-stage regressions for each IV estimate.

Table A3: Concentration coefficients (using quadratic terms)

	Fascia Count		HHI	
	Linear term	Square term	Linear term	Square term
OLS	-0.00087** (0.00035)	-0.000054 (0.000065)	0.00000091 -0.00000067	0.00000000040* 0.00000000023
2SLS	-0.0081*** (0.0023)	0.0013** (0.00051)	0.000013*** (0.0000047)	0.0000000052*** (0.0000000016)
Linear First Stage	0.56*** (0.1)	0.77 (0.53)	-211*** (48)	-22 (26)
Quadratic First Stage	-0.13** (0.065)	-0.11* (0.057)	-148831 (120063)	397883*** (72679)
Reduced form	-0.0036*** (0.001)	0.0019*** (0.00051)	-0.0036*** (0.001)	0.0019*** (0.00051)

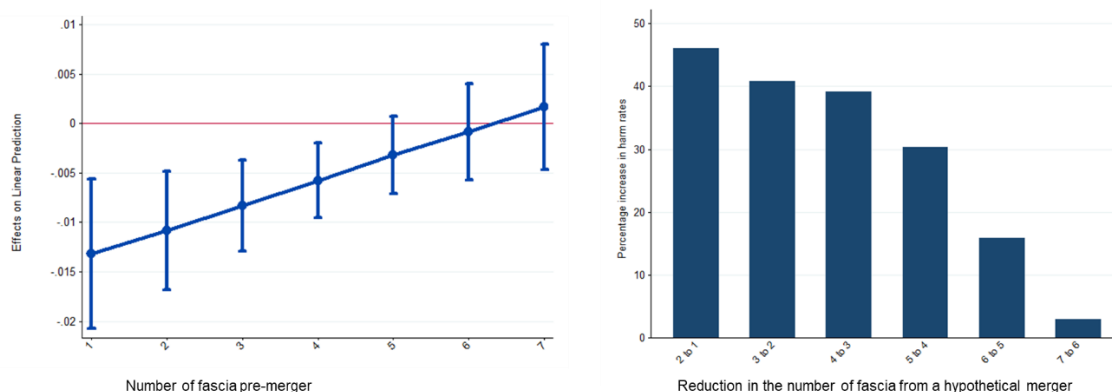
*Notes standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. The dependent variable is a dummy which takes value 1 if the patient experienced a fall, a new pressure ulcer, a blood clot or a UTI in the care of the hospital. All regressions include our main set of control variables to account for confounding factors such as differing pre-existing health status and department characteristics, as well as trust-quarter and quarter-specialty fixed effects. Concentration is measured using both the count of competitors in the hospital's catchment (Fascia Count), and the HHI (which is measured on a scale between 0 and 10,000). Standard errors are clustered at department-quarter level.*

The results in Table A1 indicate that the precise nonlinear transformation used does not influence the interpretation of the results: the coefficients also imply similar magnitude semi-elasticities in our main results.

Examination of the marginal effects implied by the fascia count regression allows us to consider the two concentration coefficients in this regression together. Further, the (patient weighted) marginal effect plot in Figure A2 Panel A, below is also informative of the number of competitors at which quality begins to fall: the estimated effect of adding one more fascia is statistically significant only where there are 5 or fewer fascia in a given area. The effect of changes in the number of competitors is also much more sizable when there are very few alternatives than when there are already

several.⁶⁵ In Panel B we show the median department weighted semi-elasticities for comparability with the charts in the main body: these tell a similar story.

Figure A2: Mean (patient level) marginal effects and median department weighted semi-elasticities implied by the quadratic term fascia count regression



Notes: Panel A in the above shows the average (mean) marginal effects (i.e. percentage point increase in harm rates) of adding one more competitor, conditional on the number of existing fascia. The whiskers represent 95% confidence intervals found using the delta method. Panel B shows the median percentage harm increase (semi-elasticity) across departments following a one unit reduction in fascia count. These are calculated as per the main body, and as follows: first, we find the patient level semi-elasticity using the coefficients from model (1) and its predicted harm rates; we then take the mean semi-elasticity within the department to get department level semi-elasticities. We plot the median semi-elasticity (to account for potential outliers) for all departments conditional on the (rounded) number of competitors faced, omitting departments facing no competitors for which there can be no further decrease. Note that we reverse the sign of Panel B for comparability with the main text, but do not do so for Panel A for comparability with the regression coefficients above.

Fascia count computation sensitivities

As explained in our identification section, there is no one way of computing a count of competitors. Here we test the impact of changing the way we do this. The results in Table A2 show that the significance of our main estimates is not sensitive to changes in the radius over which we count competitors.

Table A4: Fascia count sensitivities

	Reciprocal Fascia Count	
	IV Coefficient	Median Semi-elasticity
1. Use narrower radius (1x the hospital 80% catchment area)	0.035***	27%

⁶⁵ Patient level marginal effects are a component of the department level semi-elasticity plots we have produced in the main body. We focus here on marginal effects largely because it is much simpler to assess the statistical significance of these relative to semi-elasticities at department level.

	(0.011)	
2. Use wider radius (2x the hospital 80% catchment area)	0.083**	65%
	(0.038)	
3. Use standardised catchment area (1.5x the average 80% catchment area)	0.026***	20%
	(0.0074)	

Notes: Standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. The dependent variable is a dummy which takes value 1 if the patient experienced a fall, a new pressure ulcer, a blood clot or a UTI in the care of the hospital. All regressions instrument for concentration using the same nonlinear transformation of the instrument as is used for the concentration variable. The leftmost column describes the differences between the regression run and our main regressions (1) for fascia count and (2) for the HHI. The median (semi-)elasticity is the percentage increase in harm rates at the median. We use median to account for potential outliers.

Department weighting

Our analysis is conducted at patient level. We cluster standard errors at department-quarter level, the level at which we measure fascia count. Nevertheless, there could be a residual concern that some departments have more patients than others. We address this by regressing quality on fascia count at the aggregate level, where each observation represents a department-quarter combination.

For controls entering into our main specification as patient level dummies, we use (logged) percentages at department quarter level. We remove the patient-level treatment specialty fixed effects, given that each department treats a mix of treatment specialties. The results on department-aggregated data are shown in Table A4 below.

Table A5: Regressions on department-aggregated data

	Nonlinear (Reciprocal)	Nonlinear (Log)
OLS	0.00086 (0.0018)	0.0026 (0.0031)
IV	0.028*** (0.0093)	0.077*** (0.027)
First stage	0.36*** (0.042)	-0.065*** (0.0097)
Reduced form	0.01*** (0.0035)	-0.005*** (0.0019)

Notes standard errors in parentheses: * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. The dependent variable is a continuous variable containing the proportion of patients in the department-quarter combination which experienced a fall, a new pressure ulcer, a blood clot or a UTI whilst under the care of the trust. We constructed this dependent variable from the same patient level dataset as is used in our main regressions. The regressions are run using

OLS and two stage least squares with OLS first and second stages. We use the same controls as in our main regression, although also aggregate them to department level. We include trust-quarter and quarter-specialty fixed effects. Concentration is measured using both the count of competitors in the hospital's catchment (Fascia Count), and the HHI (which is measured on a scale between 0 and 10,000). Standard errors are clustered at department-quarter level.

The results are broadly consistent with our main estimates. The implied percentage change in harm rates following a one-unit reduction in fascia count is 15.8%, and following a 10% increase in the HHI is 35.3%. The F-statistic on the first-stage regression is 74.44 for fascia count, and 44.66 for the HHI.

ADDITIONAL INFORMATION potentially for data README file

Information on Dataset construction

The primary dataset used in our analysis is the Hospital Episodes Statistics (HES) dataset held by the CMA, provided by NHS Digital. This extract contains three years of administrative data at the level of individual healthcare interventions received by each patient in the UK between April 2012 and March 2015. This dataset is available from [NHS Digital](#) under license.

Initial Dataset cleaning

We run a set of cleaning rules over the raw dataset. In particular, the cleaning code aggregates individual healthcare interventions (or ‘episodes’) received by individual patients as part of their stay in hospital into single hospital ‘spells’ by selecting the most relevant, which is typically the episode which uses the most resources.

Missing or invalid data is recovered where possible by using information present in episodes other than the final selected one or by derivation from other variables. A small fraction of records are dropped where the data indicates the patient’s care was very unusual, such as spells in hospital of over one year which are more likely to be due to coding errors. Patients receiving specialised services under NHS definitions are also dropped, alongside any patients whose care is not funded by the NHS.

Additionally, various local area characteristics are imported from [Office of National Statistics \(ONS\) data](#) based on patients’ Lower Super Output Areas (LSOA) of residence, tariff information made available by NHS Digital / The [National Casemix Office](#) is imported, and the Charlson comorbidity index is computed.

Regression Dataset Construction, and Analysis

We make additional processing steps to the HES dataset, in order to extract the key information from richer variables: we create indicator variables for patients together based on whether their admission was through elective, non-elective, or emergency department routes; for characteristics such as gender and above-average comorbidities; for mortality outcomes using the discharge method variable (for patients aged between 35 and 75 only, for consistency with the literature); and for the [International Classification of Diseases, Tenth Revision, Clinical Modification](#)

(ICD) chapter of their diagnoses; and to create specialty grouping variables as set out in the main body. We drop all patient spells from the 2012/13 financial year.

For the mortality data, we keep only a 50% sample of patients treated to reduce computational intensity. For the harm regressions, we compute mean rates of patient characteristics in departments in each quarter, including only patients who were in hospital on the days when the harm data was collected by the relevant department.

We derive additional data from a range of other sources. In particular, we use the ONS [National Statistics Postcode Lookup directory](#) to merge in further information based on patients' LSOA level and hospitals' postcodes: this allows us to identify the LSOA's degree of urbanicity, deprivation (using the index of multiple deprivation) and the longitude and latitude of the population-weighted LSOA centroid.

Further [ONS data](#) allows us to identify population characteristics in the LSOA. We use NHS Digital [Organisation Data Services \(ODS\) data](#) to match identifiers in the HES data for trusts, sites and GP practices to postcodes; we use the ODS data to validate the information in the HES data, and to geocode the locations in which patients are treated. We use [Hospital Estates and Facilities \(HESF\)](#) data to obtain information on the categorisation of sites and trusts, to ensure that our analysis covers only trusts offering acute care (rather than, for example, mental health care for which rules may differ), as well as some additional information such as bed volumes which is used in our descriptive statistics. We further drop any department which treats fewer than 100 patients in a quarter; is identified as being part of an NHS site not recorded in the ODS data; is not located in England, or is located in London.

We construct radial catchment areas around each site using the geocoded site locations and LSOA population-weighted centroids, and from these derive the fascia counts within each given radius, and also the choice sets available to patients for the demand model used in the construction of our instrument. We also construct HHIs using observed patient flows.

Our quality data uses patient level extracts made available through the [NHS safety thermometer portal](#). The datasets contain anonymised returns from internal censuses taken by trusts which contains data at patient level on the types of any harm received, including identifiers for trusts and wards, as well as the dates of collection. The data also contains basic patient information such as whether the patient is under 18 or over 70 years old, the presence of any pressure ulcers on initial inspection, the usage of catheters as part of treatment, the risk assessment or prophylaxis for venous thromboembolism, the specialty in which they were treated,

and whether the patient is classified as acute or community. We make basic cleaning steps to extract the key information from above, to remove patients aged under 18 years of age, and to compute the number of patients treated in each ward.

We match together all these datasets, create further derived variables (such as logging controls), and drop outliers. For all control variables, we define an outlier as a record which has a value which are more than three standard deviations away from the mean of the variable. For our central independent variable, we are somewhat more cautious and inspect several cases where the above process indicates an outlier. We took the decision to drop records whose values exceeded the 99th percentile, and in the case of the HHI, additionally where records' values were below the 5th percentile.