RESEARCH ARTICLE

Health Economics WILEY

The effect of hospital spending on waiting times

Callum Brindley[1] | **James Lomas[2]** | **Luigi Siciliani[3]**

[1]Erasmus School of Health Policy and Management, Erasmus University, Rotterdam, The Netherlands

[2]Centre for Health Economics, University of York, York, UK

[3]Department of Economics and Related Studies, University of York, York, UK

Correspondence
Callum Brindley, P.O. Box 1738, Rotterdam 3000 DR, The Netherlands.
Email: brindley@eshpm.eur.nl

Abstract

Long waiting times have been a persistent policy issue in the United Kingdom that the COVID-19 pandemic has exacerbated. This study analyses the causal effect of hospital spending on waiting times in England using a first-differences panel approach and an instrumental variable strategy to deal with residual concerns for endogeneity. We use data from 2014 to 2019 on waiting times from general practitioner referral to treatment (RTT) measured at the level of local purchasers (known as Clinical Commissioning Groups). We find that increases in hospital spending by local purchasers of 1% reduce median RTT waiting time for patients whose pathway ends with a hospital admission (admitted pathway) by 0.6 days but the effect is not statistically significant at 5% level (only at the 10% level). We also find that higher hospital spending does not affect the RTT waiting time for patients whose pathway ends with a specialist consultation (non-admitted pathway). Nor does higher spending have a statistically significant effect on the volume of elective activity for either pathway. Our findings suggest that higher spending is no guarantee of higher volumes and lower waiting times, and that additional mechanisms need to be put in place to ensure that increased spending benefits elective patients.

KEYWORDS
elective treatment, hospital spending, rationing, waiting times

JEL CLASSIFICATION
C26, I11, I18

1 | INTRODUCTION

Waiting times are a significant policy issue for health systems across several OECD countries (OECD, 2020). The combination of health care provision that is free at the point of use and capacity constraints generates an excess demand, which translates into a waiting list. In effect, waiting time serves as a means to ration care in the absence of a price (Iversen & Siciliani, 2011). As a result, some patients have to wait for weeks or months before receiving treatment. Waiting times generate dissatisfaction amongst patients due to forgone health benefits while waiting, and the risk of health deterioration (Nikolova et al., 2016). Waiting times can differ substantially across countries due to differences in funding, capacity, workforce and payment systems. Due to the COVID pandemic, much elective care was postponed and waiting times and waiting lists have risen sharply as a result. The question of how to best reduce this backlog has received widespread attention (Maringe et al., 2020; O'Dowd, 2021; Stoye et al., 2022; Sud et al., 2020).

One approach to reduce waiting times is to increase health spending and allocate more resources to the hospital sector, in particular with the view of increasing volume and reducing waiting times. However, increases in spending require funds to be taken away from other publicly-funded activities or increases in taxation or salary contributions, which can be politically unpopular and may result in the policy not being cost-effective. In addition, it is possible that the additional spending may not be effective. That is, waiting times are not necessarily reduced as a result of increased spending because the increase in supply can be offset by increases in demand, or because increased spending does not translate into higher supply. The extent to which this occurs is an empirical question that has been examined in previous studies where supply and demand elasticities with respect to waiting time are estimated separately.

Evidence from England suggests that the demand elasticity with respect to waiting time is around −0.1 or −0.2 (Dusheiko et al., 2004; Gravelle et al., 2003; Martin et al., 2007; Martin & Smith, 2003; Sivey, 2012). The estimates for the supply elasticity vary, with recent studies finding a lower elasticity of around 0.1 compared to earlier studies finding a higher elasticity of around 3–5 (Gravelle et al., 2003; Martin et al., 2007; Martin & Smith, 1999, 2003; Windmeijer et al., 2005). There is only limited evidence from other countries (Australia and Italy) about the demand and supply elasticities, with estimates likely to differ based on institutional arrangements and funding levels (Riganti et al., 2017; Stavrunova & Yerokhin, 2011). A study looking specifically at the part of the United States' healthcare system that provides public care to veterans found that a 10% increase in capacity reduces primary care waiting times by 2.1% (Yee et al., 2022). Other studies have investigated whether waiting times for publicly-funded patients differ by socioeconomic status for a given level of need (Landi et al., 2018; Monstad et al., 2014; Moscelli et al., 2018; Siciliani, 2016), or whether patients with higher need are prioritized on the list (Gravelle & Siciliani, 2008; Gutacker et al., 2016).

This study uses the Referral-to-Treatment (RTT) data collection in England to analyze the causal effect of hospital spending on waiting times and elective volumes in England. It exploits variations in hospital spending at the local purchaser level (Clinical Commissioning Groups or CCGs) between 2014 and 2019. As shown in the theory section below, this can be seen as a reduced form of the structural equations estimated in the previous literature. To address possible omitted variable bias at the local purchaser level, we employ a first-differences panel approach to account for time-invariant unobserved factors at the CCG level. Health spending varies significantly across CCGs but also over time with the *within* variation in spending accounting for 41% of overall spending variation. To further account for time-varying factors potentially correlated with waiting times and spending, we also use an instrumental variable strategy. Hospital spending at CCG level is determined by a formula that uses historical data related to the demand for and cost of providing health care—for example, age profile, morbidity and mortality information and local labor market conditions, and is not related to waiting times. We instrument hospital spending with an index that exogenously determines the overall annual spending growth based on the difference between previous years' funding and a target spending allocation, which we refer to as "distance to target" index (see Data section for more details) and also exhibits significant within variation (about 36% of overall variation). We complement our main analysis with data available from the Hospital Episodes Statistics (HES) on inpatient waiting time and more refined measures of volume (emergency, elective, planned, day cases) and measures related to intensity of care (length of stay and bed days).

We make three contributions to the literature. First, we exploit variation in spending measured at local purchaser level, which is distinct from previous literature that is at provider or small-area level. Our unit of aggregation matches well with our focus on spending that is allocated at the purchaser level. Second, we measure waiting times from general practitioner (GP) referral to treatment, known as the Referral-To-Treatment (RTT) waiting time, therefore going beyond the typical focus on inpatient waiting time, from specialist consultation to treatment. The outpatient waiting time, from GP referral to specialist visit, can be a significant component of the duration of the wait along the patient pathway (OECD, 2016). Moreover, there may be gaps between the outpatient and the inpatient waiting time, which are included in the RTT waiting time but may be missed if measuring outpatient and inpatient waiting time separately. Third, we make use of a novel instrument in the context of the waiting time literature that is based on changes in the allocation formula for commissioners. We argue this is a source of exogenous variation in hospital funding that affects waiting times only indirectly through changes in spending.

Our results are as follows. We find that the elasticity of RTT waiting time for those patients whose pathway ends with a hospital admission (admitted pathway) to hospital spending is −0.87 but it not statistically significant at 5% level (only at the 10% level). The effect on waiting times appears quantitatively small. Given a median waiting time of 10.6 weeks, a 1% increase in hospital spending reduces median RTT waiting time for an admitted patient by 0.6 days.[1]

We find no statistically significant effect of hospital spending on RTT waiting time for those patients whose treatment ends with a specialist consultation (labeled in the official data as "non-admitted pathway").[2]

Neither do we find that hospital spending increases RTT volume of elective activity for either the admitted or non-admitted pathway. The results are corroborated by additional analysis using the alternative HES data. We find that hospital spending did not affect inpatient waiting time (from specialist addition to list to admission to treatment), nor did it affect elective admissions,

episodes of care, emergency admissions, or day cases. Higher spending is positively associated with length of stay, potentially suggesting higher intensity of care, but this is only statistically significant at the 10% level, while the effect on bed days is not significant.

The rest of the paper is organized as follows. Section 2 provides the institutional background. Section 3 outlines our theoretical framework. Section 4 presents the data, and Section 5 explains the methodology. Section 6 provides the results, and Section 7 discusses and concludes.

2 | INSTITUTIONAL BACKGROUND

The United Kingdom has a predominantly public health care system that has undergone a series of pro-market reforms since the 1990s. The National Health Service (NHS) retains a strong gatekeeping, referral system (Cylus et al., 2015). The NHS is funded out of general taxation and health care is generally free at the point of access. In 2012 the Conservative government passed the Health and Social Care Act, which removed barriers for the NHS to purchase some services from the private sector and restructured the NHS purchasing bodies, introducing CCGs (Peedell, 2011; Speed & Gabe, 2013).

In the NHS, patients register with a local GP who they must consult for a referral to see a specialist (NHS England, 2019b). Patients have a right to choose which hospital to attend for the first specialist appointment (NHS England, 2019a). The NHS Constitution establishes a legal right for patients to receive specialist care within 18 weeks of referral, or for the NHS to offer a range of alternative providers if this is not possible. Since the beginning of the COVID19 pandemic in March 2020, a sustained disruption of routine health care has led to larger waiting lists and longer waiting times (NHS England, 2022e). In February 2022, the Government announced recovery funding of £8 billion for elective care and set targets of an 30% increase in activity to address a backlog of referrals (Committee of Public Accounts, 2022; NHS England, 2022a).

3 | THEORETICAL FRAMEWORK

Our approach builds on previous studies that have estimated supply and demand functions for elective care (Martin & Smith, 1999, 2003). The demand and supply functions can be described as:

$$Y_i^D = \alpha_0 + \alpha_1 w_i + \alpha_2 x_i^D + \alpha_3 z_i + e_i^D \tag{1}$$

$$Y_i^S = \beta_0 + \beta_1 w_i + \beta_2 x_i^S + \beta_3 z_i + e_i^S \tag{2}$$

where Y_i^D and Y_i^S denote respectively the (log of) demand and supply of health care in area i (or, alternatively, for provider i), and w_i is the (log of) waiting time (Siciliani & Iversen, 2012). We expect $\alpha_1 < 0$ and $\beta_1 > 0$. On the demand side, longer waiting times can induce some patients to opt for the private sector if they are able to pay out of pocket or are privately insured, or seek alternative (less invasive) medical treatments, therefore reducing demand. On the supply side, longer waiting can induce providers to increase volume if providers are altruistic and have a disutility from patients having to wait longer. In addition, in the presence of waiting time targets, longer waiting times will translate into poorer provider performance increasing the chance of regulatory interventions. Vector x_i^D contains variables that affect demand (e.g., the health needs and overall size of the population in the area). Vector x_i^S contains variables which affect supply (e.g., number of doctors, number of beds, hospital spending, and type of hospital). Vector z_i contains variables which affect both the supply and the demand of services.

Under the equilibrium assumption $Y_i^D = Y_i^S = Y_i$, we can also directly write the waiting time as a function of demand and supply shifters. We obtain:

$$w_i = \gamma_0 + \gamma_1 x_i^D + \gamma_2 x_i^S + \gamma_3 z_i + e_i^w \tag{3}$$

where $\gamma_0 = (\alpha_0 - \beta_0)/(\beta_1 - \alpha_1)$, $\gamma_1 = \alpha_2/(\beta_1 - \alpha_1)$, $\gamma_2 = -\beta_2/(\beta_1 - \alpha_1)$, $\gamma_3 = (\alpha_3 - \beta_3)/(\beta_1 - \alpha_1)$. Therefore, the effect of health spending on reducing waiting times is given by γ_2, which gives the supply elasticity with respect to spending (β_2) divided by sum of the supply elasticity and the demand elasticity (in absolute value) with respect to waiting times, $(\beta_1 - \alpha_1) > 0$. It is γ_2 that we are able to identify in our econometric model and is the main focus of this paper. Higher spending reduces waiting times, but its effect is attenuated when the demand or supply elasticity is higher. This is because higher spending increases supply, which reduces waiting times, but in turn this can encourage a higher demand of treatments and also reduce provider marginal disutility from waiting (due to altruism or targets) therefore reducing supply.

4 | DATA

4.1 | Waiting times and activity

We use data on RTT waiting times for elective (non-emergency) treatment published by NHS England for 182 CCGs from 2014 to 2019 (NHS England, 2022d) (see Table A1 in the Appendix for CCG list). Waiting time is defined as the period from when a GP refers a patient to a hospital until the time a specialist treats the patient (e.g., a consultation or hospital admission if the latter is required), and it is measured in weeks.

Three main measures of RTT waiting times are available, known as (i) the admitted pathway, (ii) the non-admitted pathway, and (iii) the incomplete pathway. The admitted RTT waiting time captures the full (completed) duration of the patient's journey from RTT, for all patients requiring a hospital admission (either as inpatient or day case). The non-admitted RTT instead measures the (completed) waiting time for those patients whose pathway ended with a specialist consultation (it is "non-admitted" since it does not require a hospital admission). Both measures include all patients whose waiting finished in a given month. Lastly, the incomplete RTT pathway measures the waiting time of the patients on the list at a census date. The complete pathway is more representative than the incomplete pathway, which tends to oversample patients with long waiting times (Dixon & Siciliani, 2009). Our study concentrates on the completed pathway therefore using the RTT waiting time measure on the admitted and non-admitted pathway.

NHS England collects waiting time data from health care providers and publishes it on a monthly basis at the CCG level.[3] It provides the number of patients waiting by week (<1 week, between 1 and 2 weeks, etc.). To obtain an annual measure of waiting time that is aligned with hospital spending and to smooth seasonality (e.g., emergency care in winter tends to delay elective procedures), we aggregate the monthly records of each year for every CCG. We next calculate the cumulative distribution function of waiting times and obtain the waiting time at each decile of the distribution (i.e., 10th, 20th, …, 80th, 90th).

We also calculate the total number of admitted and non-admitted RTTs each year for each CCG, which is the amount of waiting list activity or volume.

Our main measure of waiting time is the median RTT waiting time (in weeks) of the admitted pathway for each CCG and year across all specialties. We focus on the median waiting time because the distribution of waiting times is generally right skewed, and it is also used by the NHS to monitor performance. As a supplementary analysis, we consider the waiting times of patients at each decile of the waiting time distribution (i.e., 10th, 20th, …, 80th, 90th). We also perform a disaggregated analysis of RTT waiting times for specific specialties with the largest volumes and longest average waiting times, namely general surgery, trauma and orthopedics, ear nose and throat, ophthalmology and gynecology. We also perform these analyses for the non-admitted pathway. Our annual measure of waiting time is aligned with the English financial year and the sample period ends in March 2019 just before the beginning of COVID-19 pandemic, which severely affected health services and waiting times.

4.2 | Alternative measures of waiting times, volumes and other outcomes

We complement the analysis on waiting time and activity measures from the RTT database with additional outcomes from the HES, which are also available by CCG and year (NHS England, 2022c): (i) total number of hospital episodes; (ii) total number of hospital admissions; (iii) number of elective admissions, which includes planned admissions and admissions from a waiting list; (iv) number of emergency admissions; (v) number of daycases; (vi) number of bed days; (vii) mean inpatient length of stay; and (viii) median inpatient waiting time from addition to the waiting list to admission for treatment.

We use these alternative dependent variables to check whether our analysis is robust to alternative data sources of waiting times and volumes. They also allow us to investigate possible mechanisms through which changes in hospital spending might have an effect on the waiting list (e.g., changes in the number and type of admissions, or length of stay). The RTT and HES databases are independently collected: HES is from routine hospital records, while RTT is a separate database used to monitor waiting times. HES describes episodes of continuous spells of admitted care. For example, in situations where responsibility for a patient's care is transferred to a second or subsequent specialist (known in England as a "consultant"), there will be two or more HES episodes relating to the patient's stay in hospital. This is why the total number of hospital (consultant/specialist) episodes is greater than the total number of hospital admissions. Episodes and admissions are recorded in the year in which the care ended.

The total number of hospital admissions is defined as the sum of different types of HES admissions: emergency admissions, elective admissions and admission as day cases. Length of stay, measured in the days, is the duration of an episode from hospital admission to discharge. Number of bed days is the sum of the episode duration of all hospital episodes within the year. Day cases, which imply that the patient is discharged on the same day as admission and therefore has a length of stay of 0 days, are

excluded from the calculation of length of stay. The inpatient waiting time is the time from the specialist decision to add the patient on the waiting list to admission for treatment and is measured in weeks. By definition, this HES inpatient waiting time is shorter than the RTT admitted waiting time because the latter measures the complete care pathway from GP RTT, therefore including both outpatient and inpatient waiting time.

4.3 | Hospital spending

Our data on hospital spending by CCG for financial years 2014–2019 are from NHS England's allocations of funding (NHS England, 2020a). These annual allocations of funding to each CCG consist of three streams of healthcare services: "core services," "primary care" and "specialized services." Our study uses the funding for "core services" of which approximately two-thirds is assigned to hospital-based care and one-third assigned to community-based care and mental health (NHS England, 2020b). We refer to this stream as "hospital spending" for short. We therefore exclude the separately reported "primary care" funding stream, as this does not directly affect the supply of elective hospitalizations. We also exclude "specialized services" for uncommon conditions, for which there are few providers and costs are very high. In 2019 "core services" represented 75% of CCG funding, while "primary care" and "specialized services" accounted respectively for 8% and 17% of the funding.

Each CCG's allocation of funding per person is determined by (i) a target funding allocation, which is based on a weighted capitation model, and (ii) a "distance from target (DFT)" index which determines the annual growth from the previous year's allocation toward the target. The CCG funding allocation is given by the product of (i) and (ii), which are described in more detail below.

The weighted capitation model for the target funding allocation (i) is based on a CCG's population size, which is risk adjusted for the demographic profile of the local population using an *age index*, and for the level of health inequalities and unmet need using an *additional needs index* that captures local deprivation and the standardized mortality ratio (i.e., excess deaths). The capitation model is also adjusted using an *input prices index* that reflects unavoidable differences in the costs of delivering health services including remoteness and input prices (e.g., staff, buildings and equipment). Each of these three indices is a relative index with the national average given a value of one.[4]

Finally, the target funding allocation (i) is adjusted using the *DFT index* (ii). This index is based on (a) the percentage difference between the previous year's funding allocation and the target funding allocation, and (b) a pace of change policy (NHS England, 2020b). The pace of change policy is set nationally and aims to move CCGs toward their target allocations over time smoothing potentially large fluctuations in funding. Those furthest below (above) target will see the largest annual increase (decrease) in allocations. A separate DFT index is calculated annually for each CCG and each funding stream. As explained below in the Methods section, we use the distance to target index for CCG "core services" as the basis for an instrumental variable for hospital spending.

Our hospital spending variable is measured per capita to allow comparisons across CCGs with different population size.

4.4 | Demand-side and supply-side control variables

We control for a range of demand-side factors associated with need that potentially influence waiting times. We use NHS England's age group demographics for each CCG from 2014 to 2019 (percentage of total population including 0–9, 10–19, 20–29, …, 70–79, >80 years old) (Office for National Statistics, 2020). We also control for socioeconomic status using six out of seven English Indices of Deprivation, including the domains for income, employment, education, crime, living environment, and barriers to housing and services (McLennan et al., 2019; Ministry of Housing, Communities and Local Government, 2019). We exclude the health domain because it contains health outcomes that may themselves be affected by hospital expenditure. We also exclude the overall Index of Multiple Deprivation (IMD), which is a weighted aggregate of the separate domains, to avoid multicollinearity.

Regarding the included indices of deprivation, the income domain measures the proportion of the population in an area who experience deprivation relating to low income. The definition of low income includes people who are out-of-work, as well as those who are in-work, but with low earnings. The education domain measures the proportion of the working age population in an area who are involuntarily excluded from the labor market. This includes people who would like to work but are unable to do so because of unemployment, sickness, disability, or caring responsibilities. The education domain and crime domain respectively measure the lack of attainment of skills in the local population and the risk of personal and material victimization. The living environment domain measures the quality of the local environment including both the quality of indoor living such as housing, and the quality of outdoor living such as air pollution and road traffic accidents. Lastly, the barriers to housing and services domain measures the physical proximity and financial accessibility of housing and local services. The deprivation scores of local areas are population-weighted and aggregated to obtain an average score for the overarching CCG. While income and employment have a common interval scale, the other domains have unique ordinal scales. For interpretation and comparability, we use each CCG's percentile rank based on its domain score (0 = most deprived CCG, 0.5 median CCG score,

1 = least deprived CCG). The values for 2014 were equated to those from the 2011 IMD survey, and the values for 2016–2018 were estimated using linear interpolation between the surveys of 2015 and 2019.

On the supply side, we control for geographic and temporal differences in input prices using the NHS England market forces factor (MFF) index. Specifically, MFF reflects differences in average health care provider costs for staff, buildings and equipment between CCGs. We estimate by linear interpolation the values between 2014 and 2019 published by NHS England.

5 | METHODS

Our empirical strategy aims at estimating the causal effect of hospital spending by CCGs on waiting times for elective treatments. As highlighted by our theoretical framework, there may be omitted factors that affect both the demand for and the supply of health services and therefore correlate with hospital spending and waiting times. For example, a sicker population will translate into a higher demand, increasing waiting time, but also into a higher supply through higher allocation of hospital spending. Urban areas may have a higher cost of delivering services, potentially reducing supply, and higher demand of services due to environmental factors (e.g., pollution) therefore increasing waiting times. Many of these factors will be persistent across geographical areas and commissioners. To control for time-invariant differences in demand and supply factors, we therefore use a first-differences approach[5]:

$$\Delta W_{it} = \beta \Delta S_{it} + \Delta X_{it}^D \gamma^D + \Delta X_{it}^S \gamma^S + \Delta \delta_t \gamma^t + \Delta \varepsilon_{it} \tag{4}$$

where $\Delta W_{it} = W_{it} - W_{it-1}$ is the difference in median RTT waiting time in CCG i between year t and $t-1$, $\Delta S_{it} = S_{it} - S_{it-1}$ is the difference in nominal hospital spending per capita between year t and $t-1$.

Some demand and supply factors may be time-varying. On the demand side, we control for time-varying factors related to demographics and socioeconomic status, X_{it}^D. Similarly, on the supply side, we control for variations over time of hospital input costs, X_{it}^S.

$\Delta \delta_t = \delta_t - \delta_{t-1}$ is the difference between year effects to control for common trends that are uniform across CCGs in waiting time and spending over time, for example, due national price inflation or changes in the financial climate and generosity of funding for the health sector. ε_{it} is an idiosyncratic error term. Both waiting times and hospital spending are measured in logs, while all other variables are untransformed.

The coefficient of interest is β. We expect that CCGs that had sharper increases in hospital spending over time experienced a larger reduction in waiting times relative to CCGs with smaller increases in hospital spending. We therefore expect β to be negative. We estimate model (1) in Equation (4) by OLS using CCG population weights and robust standard errors that are clustered at CCG level.

The coefficient β will be unbiased if variations in hospital spending over time are exogenous, while possible bias may arise in the presence reverse causality and time-varying omitted variable bias. Hospital spending per capita for a given CCG is determined by the formula described in the Data section. The formula is calculated using historical data relating to the demand for and cost of providing health care—for example, age profile, morbidity and mortality information and local labor market conditions, and is not related to waiting times. We therefore do not consider reverse causality to be an issue in our analysis. However, there may be time-varying factors that are omitted from our econometric model that affect the level of hospital spending. While the first-differences approach eliminates the possibility of time-invariant unobserved factors, omitted variable bias could still occur through time-varying unobserved supply- or demand-side factors that are associated with spending allocation and may also influence waiting times.

We therefore complement our main analysis with an instrumental variable approach. We instrument hospital spending with the *DFT index*. Recall that this index determines the overall annual spending growth from the previous year's allocation toward the target. The DFT index directly and exogenously affects hospital spending at the CCG level by determining the growth of budget allocations. An index greater (less) than one means a CCG's current allocation is above (below) the target allocation so that hospital spending needs to decrease (increase) over time to reach the target. Unobserved hospital demand and supply shocks at the CCG level could affect waiting times and hospital spending, but will not affect the DFT index. More formally, the first stage regression of our first difference model is:

$$\Delta S_{it} = b \Delta Z_{it} + \Delta X_{it}^D g^D + \Delta X_{it}^S g^S + \Delta \delta_t g^t + \Delta \varepsilon_{it} \tag{5}$$

where Z_{it} is an instrument, the *DFT index*. We therefore claim that the DFT index directly and exogenously affects hospital spending at the CCG level by determining the growth of budget allocations without directly affecting waiting times (only indirectly through increased spending).

For our IV strategy to be valid, the instrument has to be a strong exogenous predictor of hospital spending (relevance) and the instrument should not be correlated with the error term in the second-stage regression, conditional on the covariates (i.e., the exclusion restriction). We discuss these two conditions in turn.

In terms of relevance, the distance to target index is correlated to hospital finding through (a) the percentage difference between the previous year's funding allocation and the target funding allocation, and (b) the pace of change policy (NHS England, 2020b). The formula itself, which is used to estimate health care need for the target allocation, is periodically revised (although not during the period covered in our analysis). If allocations were revised strictly in line with the targets, there would be substantial re-allocations of funding that would potentially result in major disruptions in services. As such, CCG allocations are only revised partially so that they transition toward target over time. This is operationalized through the pace-of-change policy. The pace of change policy determines how quickly CCGs should move toward their target allocations and is informed by policy considerations such as the maximum decrease and maximum increase in spending that can be implemented without unduly affecting the availability, quality and efficiency of services, rather than outcomes such as waiting times and mortality (NHS England, 2020b). The difference between the allocation given to a CCG and the target allocation is known as the DFT index, which we use as our instrumental variable.

In addition to being valid, our instrument needs also to satisfy the exclusion restriction. Our IV strategy is based on changes in the DFT index, which in turn depends upon the pace of change policy. We argue that the pace of change policy, which is a key determinant of distance to target, is mostly politically determined and therefore not affected by waiting times, volumes or other outcomes. Moreover, we control for a range of time-varying demand variables that could potentially lead to both higher hospital spending and longer waiting times (e.g., population growth, aging, increasing health needs). However, it could still be the case that the pace of change policy is related to past changes in spending that may have a lagged effect on waiting times, the importance of which we assess through sensitivity analysis (by including lagged hospital spending as an additional variable).

Although there is no formal test for the exclusion restriction, that is, the instrument should not be correlated with the error term in the second-stage regression, we can indirectly test for it by investigating the correlation between the instrument and our observed covariates. The logic is that if there is a strong statistical relationship between the IV and observed covariates, then it is likely that such relationship may exist also with unobserved variables.

In Table A2 in the Appendix, we regress the DFT index against our vector of control variables. None of the coefficients is statistically significant at the 5% level. Similarly, in Table A3, column (2), we regress each of the control variables (as dependent variable) against the instrument (DFT index) and the remaining control variables (as independent variables, not reported). Again, the instrument is not statistically significant at the 5% level in any of the multivariate regressions. For completeness, in column (1) of Table A3, we also report the univariate regression of each control variable against the instrument. These findings give us confidence that the exclusion restriction is satisfied.

Our second-stage regression is:

$$\Delta W_{it} = \beta \widehat{\Delta S_{it}} + \Delta X_{it}^D \gamma^D + \Delta X_{it}^S \gamma^S + \Delta \delta_t \gamma^t + \Delta \varepsilon_{it} \tag{6}$$

which is estimated by two-stage least squares (2SLS), using the Stata command xtivreg. We use CCG population weights and robust standard errors that are clustered at CCG level.

In an analysis of heterogeneity, we estimate the effect of hospital spending at different points along the distribution of waiting times. Specifically, we use the waiting time at each decile in the cumulative distribution function of waiting times as the dependent variable (in the same manner as the median outlined above). We also conduct a separate disaggregated analysis of median RTT waiting time by speciality (e.g., general surgery, gynecology, trauma and orthopedics).

We also perform a volume analysis with waiting list activity performed in each year. This is similar to the analysis of hospital spending on waiting times outlined above, but where the waiting time is replaced with volume while using the same independent variables in the first-difference model in [4]. Similarly, we use as alternative dependent variable alternative measures of waiting times and volume (including emergency admissions, day cases and bed days) from the HES, and length of stay.

6 | RESULTS

6.1 | Summary statistics

Table 1 reports the descriptive statistics of waiting time for elective treatment, volume of hospital activity and other variables from 2014 to 2019. On average, CCGs have a median RTT waiting time of 10.63 and 6.35 weeks for the admitted and non-admitted pathways respectively (i.e., inpatient and outpatient treatment). The median HES elective admission inpatient waiting time, which covers a shorter period from the decision to add the patient to the list to hospital admission, was 4.83 weeks. Each year, an average of 15,737 patients started a treatment ending with an admission and 53,410 patients started a treatment not ending with a hospital admission. On average, CCGs had 103,785 hospital admissions, which consisted of approximately 40% elective admissions, 30% emergency admissions and 30% day cases. The average number of total bed days was 221,144

TABLE 1 Descriptive statistics. Years 2014–2019.

	Obs.	Mean	Standard deviation			Min	Max
			Overall	Between	Within		
Elective waiting time							
Admitted pathway median waiting time (weeks), RTT	1089	10.63	1.93	1.64	1.02	5.00	17.00
Nonadmitted pathway median waiting time (weeks), RTT	1089	6.35	1.16	0.93	0.69	3.00	10.00
Elective admission median waiting time (weeks), HES	1089	4.83	0.70	0.60	0.37	2.14	8.57
Hospital activity (volume) and other outcomes							
Admitted pathway waiting list activity, RTT	1089	15,737	9074	8787	2429	1486	75,155
Nonadmitted pathway waiting list activity, RTT	1089	53,410	27,349	25,836	9496	3178	192,219
Total episodes, HES	1089	94,345	46,936	46,890	6531	24,943	307,810
Total admissions, HES	1089	103,785	51,926	51,855	7385	28,351	360,140
Elective admissions, HES	1089	41,170	20,763	20,746	2827	11,406	144,351
Waiting admissions, HES	1089	30,143	14,591	14,510	2264	7944	96,679
Planned admissions, HES	1089	11,028	7310	7205	1624	881	50,005
Emergency admissions, HES	1089	28,572	14,528	14,421	2523	7335	100,245
Day cases, HES	1089	34,042	17,283	17,213	2827	8474	120,382
Bed days, HES	1089	221,144	115,974	115,117	20,041	64,725	1,050,072
Length of stay mean (days), HES	1089	4.81	0.62	0.50	0.37	2.75	7.98
Hospital spending, IV and controls							
Hospital spending per capita (£)	1089	1229	135	123	56	878	1693
Distance from target index	1089	1.004	0.050	0.047	0.018	0.880	1.339
Market forces factor index	1089	0.999	0.066	0.066	0.001	0.927	1.160
Population 0–9 years (%)	1089	0.117	0.015	0.015	0.002	0.075	0.181
Population 10–19 (%)	1089	0.111	0.010	0.010	0.001	0.076	0.166
Population 20–29 (%)	1089	0.133	0.031	0.031	0.003	0.088	0.265
Population 30–39 (%)	1089	0.139	0.033	0.032	0.004	0.088	0.263
Population 40–49 (%)	1089	0.137	0.011	0.009	0.006	0.110	0.174
Population 50–59 (%)	1089	0.133	0.015	0.014	0.003	0.070	0.161
Population 60–69 (%)	1089	0.106	0.022	0.022	0.003	0.041	0.163
Population 70–79 (%)	1089	0.077	0.022	0.021	0.005	0.022	0.142
Population 80 plus (%)	1089	0.047	0.013	0.013	0.001	0.015	0.084
IMD income	1089	0.499	0.290	0.290	0.016	0.000	1.000
IMD employment	1089	0.499	0.290	0.290	0.019	0.000	1.000
IMD education	1089	0.500	0.291	0.290	0.024	0.000	1.000
IMD crime	1089	0.500	0.290	0.284	0.063	0.000	1.000
IMD living environment	1089	0.501	0.288	0.286	0.041	0.000	1.000
IMD barriers to housing & services	1089	0.499	0.292	0.288	0.048	0.000	1.000

Note: Values are measured for each year and CCG. RTT is the time from when a GP refers a patient to a hospital until the time a specialist treats the patient. RTT is also the main database on elective waiting times and waiting list activity. HES is a complementary database from routine hospital records, which are collected independently from RTT data. Age groups are measured as a percentage of a CCG's total population. The IMD are measured as a CCG's percentile ranking (i.e., equal to 0 for the most deprived CCG and equal to 1 for the least deprived CCG for a given index and year). Sample period includes six financial years 2014–2019 (e.g., the 2014 financial year is April 2014 to March 2015).

Abbreviations: CCG, Clinical Commission Group; HES, Hospital Episode Statistics; IMD, indices of multiple deprivation; RTT, referral-to-treatment.

and the mean length of stay was 4.81 days. Mean hospital spending per capita was £1229. The mean DFT index, which recall determines the overall annual spending growth from the previous year's allocation toward the target, is equal to one.

Waiting time, volumes and hospital spending exhibit large within and between variation. The within variation for waiting times is at least 50% of the overall variation. For volume, it is 27% and 35% of the overall variation for the admitted and

FIGURE 1 Median waiting time from referral to treatment (RTT), admitted and non-admitted pathways.
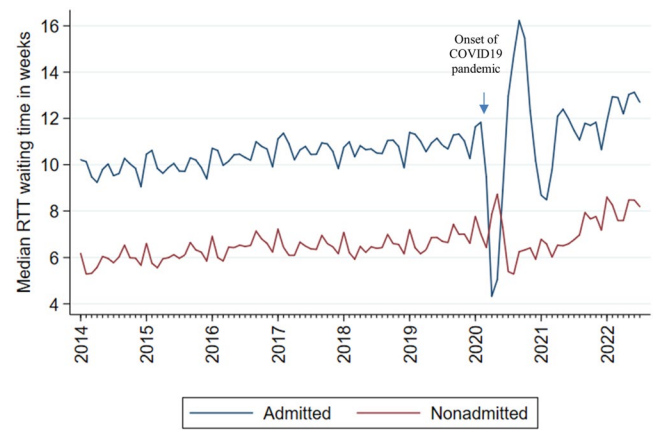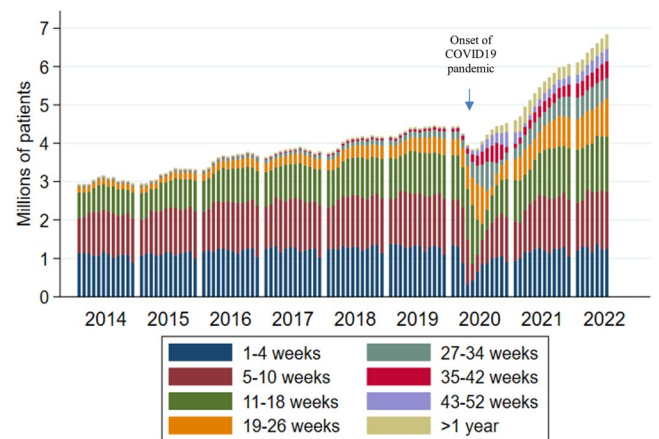

FIGURE 2 Total number of referral to treatment patients on the waiting list, incomplete pathway.

non-admitted pathway, respectively. For health spending it is about 41% of overall variation. In relation to the IV, we observe a comparable within variation. For the distance to target index, the within variation accounts for 36% of the overall variation.

Figure 1 shows the median RTT waiting time for the admitted and non-admitted pathways from 2014 to the middle of 2022. Over the period of analysis from April 2014 to March 2020, the waiting time for the admitted (inpatient) pathway steadily increased from 9.2 to 11.8 weeks (an increase of 28%). The waiting time for the non-admitted (outpatient) pathway also rose from 5.6 to 7.0 weeks (an increase of 25%). The distribution of median waiting times across CCGs is close to be normally distributed although the waiting time of individual patients tends to be right-skewed (i.e., most patients receive treatment within 3 months but some wait much longer up to over a year). Figure 2 shows the number RTT patients on the waiting list (i.e., the incomplete pathway). Over the period of analysis from April 2014 to March 2020, the waiting list increased on average by 7% per month and the number of people waiting <18 weeks fell from 93.7% to 79.6%. The COVID19 pandemic, which is beyond the analysis of this paper, has exacerbated prior trends.

6.2 | Elasticity of waiting times to hospital spending

Table 2 compares the first-difference OLS and 2SLS regression results for median RTT waiting time. The OLS estimated elasticity of the RTT waiting time for patients whose pathway ends with a hospital admission is −0.73 and statistically significant at the 1% level (see column [1]). The IV elasticity in column (2) of Table 2 is −0.87 and is not statistically significant at 5% level (but only at the 10% level). These coefficients imply local purchasers that increase hospital spending by 1% reduce the median RTT waiting time experienced by their patients by 0.73%–0.87%.

The quantitative effect on waiting time is small. Given a median waiting time of 10.6 weeks at the sample mean and using our IV estimate, a 1% increase in hospital spending would reduce the median RTT waiting time for an admitted patient by 0.6 days.[6] None of the demand or supply side controls are statistically significant while all of the year effects are significant showing that waiting times have gradually increased over time, which is consistent with Figure 1.

Table 3 provides the results from the first-stage regression (see Table A4 in the Appendix for reduced form). Our instrument, DFT, has an elasticity of 0.518 on health spending and it is statistically significant at the 1% level. The Kleibergen-Paap test for under-identification implies that DFT is "relevant" and a strong predictor of hospital spending (p-value = 0.000) (H_0: IV is

T A B L E 2 Regression results. RTT waiting time, admitted pathway.

| | (1) | | (2) | |
| | Ln(Admitted median RRT waiting time) | | Ln(Admitted median RRT waiting time) | |
	First difference		IV	
Ln(Hospital spending per capita)	−0.731***	(0.277)	−0.869*	(0.463)
Market forces factor	6.052	(7.037)	6.198	(6.935)
Population 0–9	0.816	(7.739)	0.575	(7.671)
Population 10–19	−2.742	(7.202)	−3.086	(7.261)
Population 20–29	−0.238	(6.899)	−0.635	(6.881)
Population 30–39	−1.444	(7.17)	−1.761	(7.119)
Population 40–49	2.645	(7.324)	2.215	(7.324)
Population 50–59	−0.703	(9.073)	−1.216	(8.953)
Population 60–69	0.828	(8.093)	0.32	(8.08)
Population 70–79	0.163	(8.754)	−0.222	(8.752)
Population ≥80 (ref. group)				
IMD income	−0.053	(0.357)	−0.069	(0.356)
IMD employment	0.017	(0.34)	0.013	(0.336)
IMD education	0.01	(0.195)	0.01	(0.193)
IMD crime	−0.131	(0.095)	−0.133	(0.093)
IMD living environment	−0.102	(0.135)	−0.099	(0.134)
IMD housing & services	0.014	(0.099)	0.021	(0.101)
Year 2015	0.039**	(0.016)	0.041***	(0.016)
Year 2016	0.155***	(0.033)	0.166**	(0.039)
Year 2017	0.178***	(0.043)	0.190***	(0.048)
Year 2018	0.212***	(0.053)	0.226***	(0.06)
Year 2019	0.262***	(0.064)	0.278***	(0.074)
Observations	908		908	
R^2	0.054		0.053	

Note: Robust standard errors in parenthesis.

Abbreviations: IMD, indices of multiple deprivation; RTT, referral-to-treatment.

Significance levels *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

weak). The first stage F-statistic is above 130, which is considerably higher than the "rule of thumb" threshold of 10. The result of the Hausman test suggests that hospital spending is not endogenous (p-value = 0.531) (H_0: spending is exogenous). While this suggests that OLS is more efficient than the IV, both estimated waiting time elasticities are quantitatively similar. We emphasize the IV result because it addresses potential endogeneity from residual time-varying omitted variables. The Ramsey RESET test suggests that there are no neglected nonlinearities in the model (p-value = 0.531) (H_0: functional form well specified).

Table 4 provides the first-difference OLS and 2SLS regression results for the non-admitted (outpatient) pathway. Unlike the admitted pathway, we find a positive elasticity of hospital spending on median RTT waiting time for the non-admitted pathway, which has the opposite sign to the one we hypothesize. However, the OLS estimate of 0.599 is only significant at the 10% level and the 2SLS estimate of 0.288 is instead not significant.

6.3 | Heterogeneity along the distribution of waiting times and by speciality

Figure 3 shows the IV estimated marginal effects of hospital spending on RTT waiting times for the admitted pathway, evaluated at each decile of the distribution (i.e., from shortest to longest wait). Specifically, we use the waiting time at different points along the cumulative distribution function of waiting times as the dependent variable in the same manner as the median discussed above. We find that the elasticity of waiting times to hospital spending ranges from −0.59 at the 70th percentile to −1.24 at the 20th percentile (see Table A5 in the Appendix). Evaluated at the average median waiting time, we find that the marginal effect of

TABLE 3 First stage regression.

	Ln(Hospital spending per capita)	
Distance from target	0.518***	(0.045)
Market forces factor	0.187	(0.708)
Population 0–9	−2.187**	(0.885)
Population 10–19	−3.08***	(0.918)
Population 20–29	−2.724***	(0.835)
Population 30–39	−3.119***	(0.87)
Population 40–49	−3.481***	(0.876)
Population 50–59	−3.599***	(1.03)
Population 60–69	−3.312***	(0.932)
Population 70–79	−3.309***	(0.998)
Population ≥80 (ref. group)		
IMD income	−0.095**	(0.042)
IMD employment	−0.007	(0.029)
IMD education	0.016	(0.026)
IMD crime	−0.0003	(0.01)
IMD living environment	0.012	(0.011)
IMD housing & services	0.026**	(0.013)
Year 2015	0.013***	(0.001)
Year 2016	0.076***	(0.003)
Year 2017	0.09***	(0.004)
Year 2018	0.105***	(0.005)
Year 2019	0.12***	(0.007)
Observations	908	
First-stage F stat	130.9	
Kleibergen-Paap test for relevance	p-value < 0.001 H_0: IV is weak	
Hausman test for endogeneity	p-value = 0.690 H_0: Spending is exogenous	
Ramsey RESET test for nonlinearities	p-value = 0.531 H_0: Functional form well specified	

Note: Robust standard errors in parenthesis.

Abbreviation: IMD, indices of multiple deprivation.

Significance levels *p < 0.1, **p < 0.05, ***p < 0.01.

a 1% increase in hospital spending would lead to a 0.6 days reduction in waiting time for the median admitted patient. The 95% confidence interval of the estimated elasticity gives a range of 1.3 to 0 days decrease in the median RTT waiting time.

For the admitted pathway (i.e., ending with an inpatient admission), we observe that the marginal effect increases (in absolute value) along the distribution, which implies that patients who wait longest would see the largest reduction in waiting time from increased hospital spending. These results are significant at the 5% only at the 60th and 80th decile with an elasticity respectively equal to −0.995 and −0.775 (while at the 10% level for the 20th, 40th, 50th and 70th decile). For the non-admitted pathway, we find that the IV estimated elasticity of hospital spending is not significant at the 10% level for any of the deciles along the waiting time distribution (see Table A6). The one exception is the 30% percentile, where we find an elasticity of −1.89 at the 5% significance level, and a marginal effect of a 0.4 days reduction in waiting time (see Figure A1).

Table 5 shows the first difference IV estimated elasticity of hospital spending on median RTT waiting time by speciality for the admitted (inpatient) and non-admitted (outpatient) pathways. For the admitted pathway, we find that the majority of specialities have a negative elasticity ranging from −0.29 (trauma and orthopedics) to −3.32 (thoracic medicine). However, most are not significant at the 10% level or less. In the opposite direction, cardiothoracic surgery has a positive elasticity of 3.05 that is significant at the 10% level, but it is a speciality that represents <1% of all waiting list activity (see Table A7). For the non-admitted pathway, we also find mixed results with close to half of the specialities showing a negative elasticity ranging from −0.11 (plastic surgery) to −2.23 (general medicine), and the other half of the specialities having a positive elasticity ranging from 0.34 (neurology) to 2.62 (neurosurgery). Most of the results are not significant at the 10% level or less. Gynecology

TABLE 4 Regression results. RTT waiting time, non-admitted pathway.

	(1)		**(2)**	
	Ln(Non-admitted median RRT waiting time)		**Ln(Non-admitted median RRT waiting time)**	
	First difference		**IV**	
Ln(Hospital spending per capita)	0.599*	(0.305)	0.228	(0.416)
Market forces factor	2.168	(3.508)	2.562	(3.442)
Population 0–9	−2.22	(8.768)	−2.869	(8.597)
Population 10–19	−4.839	(9.056)	−5.764	(8.916)
Population 20–29	−1.562	(8.555)	−2.632	(8.386)
Population 30–39	−4.654	(8.796)	−5.509	(8.695)
Population 40–49	−0.334	(8.571)	−1.49	(8.418)
Population 50–59	−6.977	(10.18)	−8.36	(10.05)
Population 60–69	−3.118	(9.986)	−4.486	(9.855)
Population 70–79	−8.384	(10.347)	−9.423	(10.179)
Population ≥80 (ref. group)				
IMD income	0.283	(0.344)	0.242	(0.344)
IMD employment	0.194	(0.319)	0.183	(0.314)
IMD education	0.126	(0.276)	0.126	(0.272)
IMD crime	−0.086	(0.095)	−0.091	(0.092)
IMD living environment	−0.283**	(0.132)	−0.275**	(0.13)
IMD housing & services	−0.218*	(0.113)	−0.201*	(0.113)
Year 2015	0.067***	(0.015)	0.071***	(0.016)
Year 2016	0.109***	(0.032)	0.137***	(0.039)
Year 2017	0.11**	(0.042)	0.143***	(0.05)
Year 2018	0.131**	(0.055)	0.168***	(0.063)
Year 2019	0.216***	(0.07)	0.259***	(0.078)
Observations	908		908	
R^2	0.110		0.110	

Note: Median RTT waiting time. Robust standard errors in parenthesis.

Abbreviations: IMD, indices of multiple deprivation; RTT, referral-to-treatment.

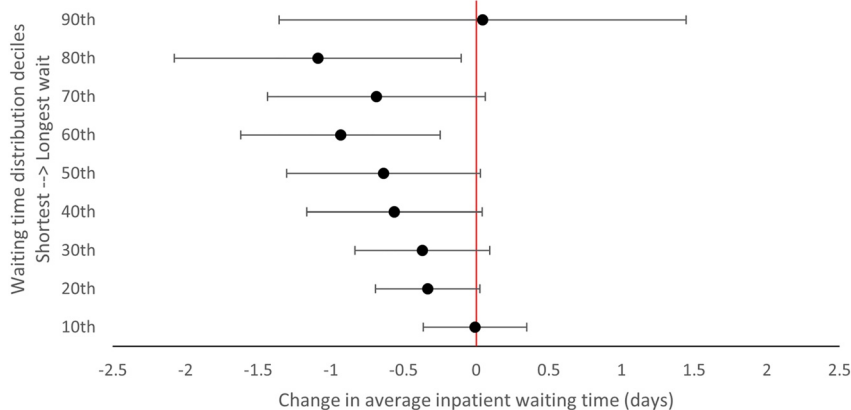Significance levels *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

FIGURE 3 Marginal effect of 1% increase in hospital spending on referral to treatment waiting times along the waiting time distribution, admitted pathway.

and "other" specialties are the only specialities that have an elasticity that is significant at the 10% level or less and constitute a large share of non-admitted (outpatient) waiting list activity that is >5% (see Table A7).

TABLE 5 Regression results. IV estimated elasticity of waiting time with respect to hospital spending by speciality.

	Elasticity admitted pathway		Elasticity non-admitted pathway	
All specialities	−0.869*	(0.463)	0.228	(0.416)
General surgery	−0.948	(0.715)	−1.184	(0.846)
Urology	0.673	(0.532)	−0.908	(0.683)
Trauma & orthopedics	−0.294	(0.388)	−1.053	(0.803)
ENT	−0.883	(0.735)	1.053	(0.803)
Ophthalmology	0.221	(0.660)	−0.879	(1.037)
Oral surgery	–		–	
Neurosurgery	−0.549	(1.546)	2.617***	(0.977)
Plastic surgery	−1.438	(1.110)	−0.107	(1.733)
Cardiothoracic surgery	3.051*	(1.585)	2.149	(3.135)
General medicine	6.138	(3.910)	−2.293	(1.772)
Gastroenterology	−0.461	(0.988)	−1.914*	(1.049)
Cardiology	−0.633	(0.887)	0.486	(0.976)
Dermatology	−1.991	(1.327)	1.878*	(0.963)
Thoracic medicine	−3.322*	(1.819)	−1.758*	(0.959)
Neurology	−2.137	(2.009)	0.337	(1.043)
Rheumatology	−2.345	(3.453)	0.115	(0.926)
Geriatric medicine	–		–	
Gynecology	−1.232	(0.949)	1.868**	(0.885)
Other	−1.782*	(1.000)	1.142*	(0.664)

Note: Median RTT waiting time. Elasticity based on first difference IV estimate. Robust standard errors in parenthesis.

Abbreviation: RTT, referral-to-treatment.

Significance levels *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Estimates for oral surgery and geriatric medicine omitted due to small annual patient volumes.

6.4 | Elasticity of waiting list activity (volume) with respect to hospital spending

Table 6 shows the first difference IV estimated elasticity of hospital spending on RTT waiting list activity (i.e., volumes of patients) for both the admitted (inpatient) and non-admitted (outpatient) pathways. We find an elasticity in the expected direction of 0.41 and 0.88 for the admitted and non-admitted pathways respectively, with a 1% increase in spending increasing volume by 0.4% or 0.9%. Although the order of magnitude appears to be plausible, these results are not significant at the 10% level. Change in the volume of hospital activity is a key mechanism through which spending can impact on waiting times. However, our analysis suggests that CCGs that had relatively higher spending over time did not have significantly higher volumes. This can help explain the very limited or no effect of spending on waiting times.

6.5 | Elasticity of other outcomes and alternative measures of waiting times and volumes with respect to hospital spending

Table 7 shows the elasticity of various alternative measures of waiting times, volume and other outcomes from HES data with respect to hospital spending and also compares these additional results with the main results using RTT data. We comment the results from our instrumental-variable specification (but also report the first-difference results, which are qualitatively similar).

In relation to inpatient waiting time, we find that elasticity of median inpatient waiting time (as measured in HES) with respect to hospital spending is −0.619 but is not statistically significant. This is largely consistent with the elasticity of RTT admitted-pathway waiting time with respect spending of −0.869, presented above, which is statistically significant only at 10% level.

In relation to volume, we have several additional measures. Starting with elective admissions, the elasticity of total elective admissions with respect to hospital spending is 0.185, suggesting that higher spending is positively associated with volume but the effect is not statistically significant. The coefficient is larger for elective admissions from the waiting list and equal to 0.237, and smaller for planned elective admissions and equal to 0.073. Again, the coefficients are not statistically significant.

TABLE 6 Regression results. IV estimated elasticity of hospital spending on waiting list activity per capita.

	(1)		(2)	
	Ln(Waiting list activity per capita) admitted pathway		Ln(Waiting list activity per capita) non-admitted pathway	
Ln(Hospital spending per capita)	0.412	(0.677)	0.876	(0.834)
Market forces factor	31.463***	(10.63)	34.71*	(18.26)
Population 0–9	12.16	(17.81)	21.17	(19.63)
Population 10–19	−26.58	(17.52)	−14.25	(20.68)
Population 20–29	0.223	(14.94)	11.09	(17.64)
Population 30–39	1.695	(17.36)	18.987	(20.79)
Population 40–49	10.43	(16.94)	26.13	(21.47)
Population 50–59	6.542	(19.62)	30.218	(23.61)
Population 60–69	−8.611	(17.57)	−3.449	(20.67)
Population 70–79	−4.278	(16.98)	1.18	(20.38)
Population ≥80 (ref. group)				
IMD income	1.444	(1.109)	0.705	(1.287)
IMD employment	−1.203	(0.989)	−0.626	(1.04)
IMD education	0.638	(0.573)	0.691	(0.734)
IMD crime	0.083	(0.17)	0.07	(0.193)
IMD living environment	0.082	(0.217)	0.075	(0.29)
IMD housing & services	−0.125	(0.239)	0.375	(0.282)
Year 2015	−0.063***	(0.022)	−0.028	(0.03)
Year 2016	−0.049	(0.064)	0.031	(0.074)
Year 2017	−0.122	(0.086)	0.019	(0.09)
Year 2018	−0.069	(0.117)	0.096	(0.114)
Year 2019	−0.104	(0.155)	0.115	(0.149)
Observations	908		908	
R^2	0.112		0.0851	

Note: Robust standard errors in parenthesis.

Abbreviation: IMD, indices of multiple deprivation.

Significance levels *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

For a given volume of admissions, higher spending could translate into more episodes of care for each individual patient, given that each patient admission involves one or more episodes of care (e.g., different visits by different specialists). This is however not the case. The elasticity of total episodes of care with respect to hospital spending is 0.078 and again not statistically significant.

Higher hospital spending could also be associated with a higher propensity to admit emergency patients following an A&E attendance. We do not find that this is the case. Spending is negatively associated with emergency admissions and again it is not statistically significant. Our findings are consistent with the demand for emergency care being less responsive than the demand for elective care. The result is also reflected in the effect on spending on total admissions (the sum of elective and emergency admissions and daycases) as the coefficient for total admissions is 0.078, and therefore smaller than for elective admissions, as expected.

Last, if higher spending does not affect volume, it could still affect the intensity of care to treat patients as measured by patient length of stay in hospital. We find that the elasticity of mean length of stay with respect to hospital spending is 0.519 and statistically significant at the 10% level. Similarly, the elasticity of spending on bed days (i.e., the product of volume and length of stay) with respect to spending is slightly smaller and equal to 0.434 and not significant even at 10% level.

6.6 | Additional robustness checks

We conduct several robustness checks. We find that our main results are similar whether we use controls with a log or untransformed scale, and that the precision increases by adding all of the IMD controls and population age groups. We also obtain

TABLE 7 Regression results. Estimated elasticity of hospital spending on other outcomes and alternative measures of waiting times and volumes.

	(1) Ln(Hospital spending per capita) First difference		(2) Ln(Hospital spending per capita) IV	
Elective waiting time				
Ln(Admitted pathway median waiting time in weeks, RTT)	−0.731***	(0.277)	−0.869*	(0.463)
Ln(Nonadmitted pathway median waiting time in weeks, RTT)	0.599*	(0.305)	0.228	(0.416)
Ln(Elective admission median waiting time in weeks, HES)	−1.351	(1.163)	−0.619	(1.810)
Hospital activity (volume) and other outcomes				
Ln(Admitted pathway waiting list activity per capita, RTT)	−0.328	(0.556)	0.412	(0.677)
Ln(Nonadmitted pathway waiting list activity per capita, RTT)	−0.338	(0.688)	0.876	(0.834)
Ln(Total episodes per capita, HES)	0.079	(0.082)	0.078	(0.115)
Ln(Total admissions per capita, HES)	0.080	(0.091)	0.117	(0.139)
Ln(Elective admissions per capita, HES)	0.013	(0.105)	0.185	(0.169)
Ln(Waiting admissions per capita, HES)	0.088	(0.120)	0.237	(0.198)
Ln(Planned admissions per capita, HES)	−0.258	(0.301)	0.073	(0.465)
Ln(Emergency admissions per capita, HES)	0.163	(0.142)	−0.09	(0.194)
Ln(Day cases per capita, HES)	0.079	(0.135)	0.189	(0.217)
Ln(Bed days per capita, HES)	0.267	(0.232)	0.434	(0.322)
Ln(Length of stay mean, HES)	0.210	(0.237)	0.519*	(0.294)

Note: Year effects and all demand-side and supply-side controls included. Robust standard errors in parenthesis.

Abbreviations: HES, Hospital Episodes Statistics; RTT, referral-to-treatment.

Significance levels *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

similar elasticity estimates for RTT waiting time after adjusting for nonlinearity using Duan smearing factors, which remain relatively small <1.05 for all deciles along the RTT waiting time distribution (see Table A8). The Duan smearing factors for admitted and non-admitted waiting list activity were also close to one.

We also investigated whether lagged expenditures had an effect on waiting time and volumes by adding 1- and 2-year lagged hospital spending as additional covariates in Equation (4). An effect might be observed, for example, if resources were used to build new capacity that took time to take effect. Patients with long waiting times that start and finish in different years might also be impacted by spending over multiple years. We found, however, that 1-year lagged hospital spending was generally highly insignificant (see Tables A9 and A10). We obtained less precise estimates and observed multicollinearity with a reversal of the main coefficient when adding both a 1- and 2-year lag.

Only in one instrumental-variable specification we find that higher spending reduces waiting times for patients with an inpatient admission which is significant only at the 10% level. Using our theoretical framework, the elasticity of waiting time with respect to spending is $dw_i / d\, x_i^S = \gamma_2 = -\beta_2 / (\beta_1 - \alpha_1) = -0.869$. Similarly, substituting (3) into (1) and differentiating with respect to x_i^S in equilibrium, the elasticity of volume to spending is $dY_i / d\, x_i^S = \alpha_1\, \gamma_2 = 0.412$ (though this is not statistically significant). Comparing the two coefficients, we can infer that the elasticity of demand to waiting time is $\alpha_1 = 0.412/(-0.869) = -0.474$. This elasticity is higher than the one reported in the literature for England (around −0.1 or −0.2) though these lower elasticities relate to a time period where waiting times were much longer than in our sample period (Dusheiko et al., 2004; Gravelle et al., 2003; Martin et al., 2007; Martin & Smith, 2003; Sivey, 2012).

7 | DISCUSSION AND CONCLUSIONS

This study has investigated the extent to which higher hospital spending by local purchasers in England reduces waiting times for elective treatment. We find that increases in hospital spending by local purchasers of 1% reduce median RTT waiting time for patients whose pathway ends with a hospital admission by 0.6 days but the effect is not statistically significant at 5% level (only at the 10% level). We also find that higher hospital spending does not affect the RTT waiting time for patients whose

pathway ends with a specialist consultation. Nor does higher spending have a statistically significant effect on the volume of elective activity for either pathway. In this section, we discuss the implications of our findings for possible policy interventions at the national and international level.

Siciliani et al. (2013) review different policy initiatives to reduce waiting times across 12 OECD countries with long waiting times. They find that maximum waiting time guarantees can be effective in reducing waiting times but only if accompanied by penalties for providers through regulatory interventions (as in England[7] or Finland) or patient choice policies where patients have the option of going private if the maximum waiting time guarantee has not been satisfied (as in Denmark and Portugal). They also find that the Netherlands experienced a significant reduction in waiting times by combining maximum waiting time guarantees with an increase in health spending and by switching hospital reimbursement toward activity-based financing and fee-for-service for specialists. It also concludes that additional funding should be made conditional on additional activity and maximum waiting times to ensure that the additional activity is not offset by additional demand. Another systematic review of policies highlights that there is limited high-quality evidence evaluating policy interventions aimed at reducing waiting times (Ballini et al., 2015). Only eight studies met the inclusion criteria, of which seven related to restructuring the intake assessment/referral process. The review found no studies evaluating interventions to increase capacity or reduce demand. It concluded that, although no firm conclusions could be reached, policy interventions which improved accessibility of services (through open access or direct booking/referral) showed some promise.

A more recent report (OECD, 2020) documents that following the financial crisis the main types of policies to reduce waiting times (e.g., various forms of maximum waiting-time guarantees) have been maintained but that health spending growth has slowed down in several countries. Between 2007 and 2019, waiting times have been relatively stable in some countries (e.g., Finland) or slowly increasing (e.g., in England, and the Netherlands) possibly as a result of a tighter financial environment and reduced investments.

Our findings echo one of the key policy findings of Siciliani et al. (2013) that additional funding is no guarantee of success in reducing waiting times. For additional funding to work, it has to be contingent on providers committing to additional activity and maximum waiting-time targets. Our analysis shows that volume did not significantly increase. One reason to explain this is that hospitals may be already working close to full capacity with very limited ability to expand in the short run. There is evidence that bed occupancy rates has been steadily increasing in the years before COVID-19 toward high levels, above 85% occupancy rates (Bosque-Mercader & Siciliani, 2022). This is consistent with beds per 1000 population in England (2.4 beds) being below the EU average in 2020 (OECD and European Union, 2022). Moreover, the health workforce is limited, and providers may struggle in the short run to recruit additional staff, especially nurses and anesthesiologists, or to extend the working hours of current health workers (OECD, 2023). Third, investing in additional capacity can be a slow process requiring investment and disruption of current practices.

The results have also implications for the elective backlog that was accumulated during the COVID-19 pandemic. As suggested by Figure 2, the waiting list has rapidly grown from <4.5 million at the beginning of 2020 to more than 6 million in 2022. The current Delivery Plan for England involves additional funding and several initiatives to boost supply, such as increasing the proportion of procedures performed in outpatient departments rather than in inpatient surgical theaters, making temporary staffing banks more attractive to facilitate staff taking on extra shifts and paying promptly for these, proactively supporting temporary staff, including offering more permanent employment or development opportunities, and the creation of dedicated elective facilities both within hospitals or new ones, which are ring fenced against disruptions from emergency care (NHS England, 2022a, 2022b). Our study warns policymakers that the link between funding and activity will have to be carefully monitored for this delivery plan to succeed. Policy options to boost supply include additional reliance on activity-based financing, greater conditionality of funding to additional supply and lower waiting times, and using existing capacity in the private sector to treat publicly-funded patients. Given that the health workforce is key to boost supply, retention policies that improve working conditions and remuneration are likely to play an important role to address workers' shortages (OECD, 2023).

Our results are not incompatible with previous literature, which has focused on estimating the demand and supply elasticity to waiting times. A key finding in this literature is that the demand for health care is inelastic to waiting times, with an elasticity of around −0.1 or −0.2 (Dusheiko et al., 2004; Gravelle et al., 2003; Martin et al., 2007; Martin & Smith, 2003). In turn, this implies that a higher supply of health care will be effective in reducing waiting times, as the demand response to a reduction in waiting time will have a limited effect in boosting demand. Our results show that the link between spending and supply can be a tenuous one, and the extent to which higher spending translates into higher supply is crucial for policies aimed at reducing waiting times.

Last, our study focuses on England. Future work could investigate the relation between spending and waiting times in other health systems that also exhibit long waiting times, such as Australia, Italy, Ireland and Spain (OECD, 2020), and test whether different health system infrastructure and capacity leads to a stronger relationship between spending and waiting times. Our

analysis uses data pre COVID-19. As time passes and new databases become available, future work could investigate the relation between spending and waiting times in the post-COVID period.

AUTHOR CONTRIBUTIONS

Callum Brindley and James Lomas conceptualized the study. Callum Brindley wrote the first draft with input from James Lomas and Luigi Siciliani. All authors contributed to subsequent revisions and approved the manuscript prior to its submission. All authors are accountable for the research presented.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Analysis was performed on publicly available data. Code is available upon request.

ETHICS STATEMENT

This study did not receive nor require ethics approval, as it does not involve human and animal participants.

PATIENT CONSENT FOR PUBLICATION

Not applicable.

PATIENT AND PUBLIC INVOLVEMENT

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

ORCID

Callum Brindley https://orcid.org/0000-0002-5478-3974
James Lomas https://orcid.org/0000-0002-2478-7018
Luigi Siciliani https://orcid.org/0000-0003-1739-7289

ENDNOTES

[1] There is a statistically significant effect (at 5% level or less) only for the 60th and 80th percentiles of the admitted pathway waiting time distribution, and the effect is again quantitatively small. The effect is heterogeneous across specialties.

[2] If a patient's pathway ends with a specialist consultation, then this implies that the patient does not require an inpatient admission. Otherwise, the patient's wait would continue and the patient would be included amongst those with an admitted pathway.

[3] This study focuses on NHS wait times, which primarily concerns public providers but also includes situations where a patient is offered an appointment with a private provider as part of an outsourcing arrangement; waiting times in the private sector where patients have bypassed the NHS are outside the scope of this study and represent a relatively small share of consultations and procedures. In 2014, around 13% of all elective surgery on UK residents was privately funded (The King's Fund, 2014).

[4] In short, the target funding allocation works by applying a funding stream target share and three indices to the national budget per capita: target allocation = (national budget per person) × (funding stream target share) × (age index) × (additional needs index) × (input price index).

[5] We favor a first difference specification over a fixed effects one because of the presence of serial correlation, which makes the first difference estimator more efficient (Baltagi, 2005).

[6] Note that this result is obtained without application of a smearing factor. This is because the smearing factor makes very little difference (see Supporting Information S1: Table A4).

[7] Using a difference-in-difference approach with Scotland as control group, Propper et al. (2008) show that the introduction of maximum waiting times in England reduced waiting times by about seven percentage points, without any adverse effects on quality of care (Propper et al., 2010). However, Askildsen et al. (2011) show that introducing maximum waiting times could come at the cost of lower priority for the more severe patients in Norway.

REFERENCES

Askildsen, J. E., Holmås, T. H., & Kaarboe, O. (2011). Monitoring prioritisation in the public health-care sector by use of medical guidelines. The case of Norway. *Health Economics*, *20*(8), 958–970. https://doi.org/10.1002/hec.1659

Ballini, L., Negro, A., Maltoni, S., Vignatelli, L., Flodgren, G., Simera, I., Holmes, J., & Grilli, R. (2015). Interventions to reduce waiting times for elective procedures. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD005610.pub2

Baltagi, B. H. (2005). *Econometric analysis of panel data* (3rd ed.). John Wiley & Sons.

Bosque-Mercader, L., & Siciliani, L. (2022). The association between bed occupancy rates and hospital quality in the English National Health Service. *The European Journal of Health Economics*, *24*(2), 209–236. https://doi.org/10.1007/s10198-022-01464-8

Committee of Public Accounts. (2022). *NHS backlogs and waiting times in England*. United Kingdom House of Commons.

Cylus, J., Richardson, E., Findley, L., Longley, M., O'Neill, C., & Steel, D. (2015). United Kingdom Health System Review (2015) (No. 17(5)), Health Systems in Transition.

Dixon, H., & Siciliani, L. (2009). Waiting-time targets in the healthcare sector: How long are we waiting? *Journal of Health Economics*, *28*(6), 1081–1098. https://doi.org/10.1016/j.jhealeco.2009.09.003

Dusheiko, M., Gravelle, H., & Jacobs, R. (2004). The effect of practice budgets on patient waiting times: Allowing for selection bias. *Health Economics*, *13*(10), 941–958. https://doi.org/10.1002/hec.937

Gravelle, H., & Siciliani, L. (2008). Is waiting-time prioritisation welfare improving? *Health Economics*, *17*(2), 167–184. https://doi.org/10.1002/hec.1262

Gravelle, H., Smith, P., & Xavier, A. (2003). Performance signals in the public sector: The case of health care. *Oxford Economic Papers*, *55*(1), 81–103. https://doi.org/10.1093/oep/55.1.81

Gutacker, N., Siciliani, L., & Cookson, R. (2016). Waiting time prioritisation: Evidence from England. *Social Science & Medicine*, *159*, 140–151. https://doi.org/10.1016/j.socscimed.2016.05.007

Iversen, T., & Siciliani, L. (2011). Non-price rationing and waiting times. In S. Glied & P. C. Smith (Eds.), *The Oxford handbook of health economics* (pp. 648–670). Oxford University Press.

Landi, S., Ivaldi, E., & Testi, A. (2018). Socioeconomic status and waiting times for health services: An international literature review and evidence from the Italian National Health System. *Health Policy*, *122*(4), 334–351. https://doi.org/10.1016/j.healthpol.2018.01.003

Maringe, C., Spicer, J., Morris, M., Purushotham, A., Nolte, E., Sullivan, R., Rachet, B., & Aggarwal, A. (2020). The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: A national, population-based, modelling study. *The Lancet Oncology*, *21*(8), 1023–1034. https://doi.org/10.1016/S1470-2045(20)30388-0

Martin, S., Rice, N., Jacobs, R., & Smith, P. (2007). The market for elective surgery: Joint estimation of supply and demand. *Journal of Health Economics*, *26*(2), 263–285. https://doi.org/10.1016/j.jhealeco.2006.08.006

Martin, S., & Smith, P. C. (1999). Rationing by waiting lists: An empirical investigation. *Journal of Public Economics*, *71*(1), 141–164. https://doi.org/10.1016/S0047-2727(98)00067-X

Martin, S., & Smith, P. C. (2003). Using panel methods to model waiting times for National Health Service surgery. *Journal of the Royal Statistical Society: Series A*, *166*(3), 369–387. https://doi.org/10.1111/1467-985X.00282

McLennan, D., Noble, S., Noble, M., Plunkett, E., Wright, G., & Gutacker, N. (2019). *English indices of deprivation 2019: Technical report*. Ministry of Housing, Communities and Local Government.

Ministry of Housing, Communities and Local Government. (2019). *The English indices of deprivation 2019: Statistical release*. Ministry of Housing, Communities and Local Government.

Monstad, K., Engesæter, L. B., & Espehaug, B. (2014). Waiting time and socioeconomic status—An individual-level analysis. *Health Economics*, *23*(4), 446–461. https://doi.org/10.1002/hec.2924

Moscelli, G., Siciliani, L., Gutacker, N., & Cookson, R. (2018). Socioeconomic inequality of access to healthcare: Does choice explain the gradient? *Journal of Health Economics*, *57*, 290–314. https://doi.org/10.1016/j.jhealeco.2017.06.005

NHS England. (2019a). Guidance - NHS hospital services [WWW document]. Retrieved April 22, 2022, from https://www.nhs.uk/using-the-nhs/nhs-services/hospitals/about-nhs-hospital-services/

NHS England. (2019b). Guidance - referrals for specialist care [WWW document]. Retrieved April 22, 2022, from https://www.nhs.uk/using-the-nhs/nhs-services/gps/referrals-for-specialist-care/

NHS England. (2020a). Budget allocations [WWW document]. Retrieved September 12, 2020 https://www.england.nhs.uk/allocations/

NHS England. (2020b). Fair shares - a guide to NHS allocations.

NHS England. (2022a). Delivery plan for tackling the covid-19 backlog of elective care [WWW document]. Retrieved April 22, 2022, from https://www.england.nhs.uk/coronavirus/publication/delivery-plan-for-tackling-the-covid-19-backlog-of-elective-care/

NHS England. (2022b). Elective recovery programme update [WWW document]. Retrieved September 12, 2022, from https://www.england.nhs.uk/wp-content/uploads/2022/03/BM2208Pu-elective-recovery-update-march-2022.pdf

NHS England. (2022c). Hospital admitted patient care activity [WWW document]. NHS Digital. Retrieved April 3, 2023, from https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity

NHS England. (2022d). NHS RTT waiting times [WWW document]. Retrieved September 12, 2022, from https://www.england.nhs.uk/statistics/statistical-work-areas/rtt-waiting-times/

NHS England. (2022e). Statistical press notice - NHS referral to treatment (RTT) waiting times data - July 2022 [WWW document]. Retrieved September 12, 2022, from https://www.england.nhs.uk/statistics/statistical-work-areas/rtt-waiting-times/rtt-data-2022-23/

Nikolova, S., Harrison, M., & Sutton, M. (2016). The impact of waiting time on health gains from surgery: Evidence from a national patient-reported outcome dataset. *Health Economics*, *25*(8), 955–968. https://doi.org/10.1002/hec.3195

O'Dowd, A. (2021). NHS waiting list hits 14 year record high of 4.7 million people. *BMJ*, *373*, n995. https://doi.org/10.1136/bmj.n995

OECD. (2016). Waiting times for elective surgery. In *Health at a glance: Europe 2016* (pp. 174–175). OECD.

OECD. (2020). *Waiting times for health services: Next in line, OECD health policy studies*. OECD.

OECD. (2023). *Ready for the next crisis? Investing in health system resilience, OECD health policy studies*. OECD. https://doi.org/10.1787/1e53cf80-en

OECD, European Union. (2022). *Health at a glance: Europe 2022: State of health in the EU cycle, health at a glance: Europe*. OECD. https://doi.org/10.1787/507433b0-en

Office for National Statistics. (2020). *Population estimates by output areas, electoral, health and other geographies, England and Wales: Statistical bulletin*. Office of National Statistics.

Peedell, C. (2011). Further privatisation is inevitable under the proposed NHS reforms. *BMJ*, *342*(may17 3), d2996. https://doi.org/10.1136/bmj.d2996

Propper, C., Sutton, M., Whitnall, C., & Windmeijer, F. (2008). Did "targets and terror" reduce waiting times in England for hospital care? *The B.E. Journal of Economic Analysis & Policy*, *8*(2). https://doi.org/10.2202/1935-1682.1863

Propper, C., Sutton, M., Whitnall, C., & Windmeijer, F. (2010). Incentives and targets in hospital care: Evidence from a natural experiment. *Journal of Public Economics*, *94*(3–4), 318–335. https://doi.org/10.1016/j.jpubeco.2010.01.002

Riganti, A., Siciliani, L., & Fiorio, C. V. (2017). The effect of waiting times on demand and supply for elective surgery: Evidence from Italy. *Health Economics*, *26*, 92–105. https://doi.org/10.1002/hec.3545

Siciliani, L. (2016). Waiting times: Evidence of social inequalities in access for care. In B. Sobolev, A. Levy, & S. Goring (Eds.), *Data and measures in health services research, health services research* (pp. 1–17). Springer US. https://doi.org/10.1007/978-1-4899-7673-4_17-1

Siciliani, L., Borowitz, M., & Moran, V. (Eds.). (2013). *Waiting time policies in the health sector: What works?, OECD health policy studies*. OECD. https://doi.org/10.1787/9789264179080-en

Siciliani, L., & Iversen, T. (2012). Waiting times and waiting lists. The Elgar companion to health economics (2nd ed.). Edward Elgar Publishing. https://www-elgaronline-com.eur.idm.oclc.org/display/edcoll/9781849802673/9781849802673.00036.xml

Sivey, P. (2012). The effect of waiting time and distance on hospital choice for English cataract patients. *Health Economics*, *21*(4), 444–456. https://doi.org/10.1002/hec.1720

Speed, E., & Gabe, J. (2013). The Health and Social Care Act for England 2012: The extension of 'new professionalism'. *Critical Social Policy*, *33*(3), 564–574. https://doi.org/10.1177/0261018313479010

Stavrunova, O., & Yerokhin, O. (2011). An equilibrium model of waiting times for elective surgery in NSW public hospitals. *The Economic Record*, *87*(278), 384–398. https://doi.org/10.1111/j.1475-4932.2011.00726.x

Stoye, G., Warner, M., & Zaranko, B. (2022). *The NHS backlog recovery plan and the outlook for waiting lists*. The IFS.

Sud, A., Torr, B., Jones, M. E., Broggio, J., Scott, S., Loveday, C., Garrett, A., Gronthoud, F., Nicol, D. L., Jhanji, S., Boyce, S. A., Williams, M., Riboli, E., Muller, D. C., Kipps, E., Larkin, J., Navani, N., Swanton, C., Lyratzopoulos, G., … Turnbull, C. (2020). Effect of delays in the 2-week-wait cancer referral pathway during the COVID-19 pandemic on cancer survival in the UK: A modelling study. *The Lancet Oncology*, *21*(8), 1035–1044. https://doi.org/10.1016/S1470-2045(20)30392-2

The King's Fund. (2014). Commission on the Future of Health and Social Care in England: The UK private health market.

Windmeijer, F., Gravelle, H., & Hoonhout, P. (2005). Waiting lists, waiting times and admissions: An empirical analysis at hospital and general practice level. *Health Economics*, *14*(9), 971–985. https://doi.org/10.1002/hec.1043

Yee, C. A., Barr, K., Minegishi, T., Frakt, A., & Pizer, S. D. (2022). Provider supply and access to primary care. *Health Economics*, *31*(7), 1296–1316. https://doi.org/10.1002/hec.4482

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.