Cancelled procedures in the English NHS: Evidence from the 2010 tariff reform

Graham Cookson[a], Simon Jones[b] and Ioannis Laliotis*[a]

[a]University of Surrey

[b]New York University

Abstract

This paper explores the role of incentives in the English NHS. Until financial year 2009/10, elective procedures that were cancelled after admission received a fixed reimbursement associated with a specific healthcare resource group code. We investigate whether this induced trusts to admit and then cancel, rather than cancel before admission and/or to cancel low fee over high fee work. As the tariff was ended in April 2010 we conduct an interrupted time series analysis to examine if their behaviour was affected after the tariff removal. The results indicate a small, yet statistically significant, decline in the probability of a last minute cancellation in the post-tariff period, especially for certain types of patients and diagnoses.

Keywords: Cancelled procedures; English NHS; Reforms; Incentives

JEL Classification: I10; L10

---

*Corresponding author: School of Economics, University of Surrey, Guildford, GU2 7XH, UK. E-mail: i.laliotis@surrey.ac.uk. We would like to thank the Editor and two anonymous referees for their constructive suggestions on an earlier draft of this paper. We have also received insightful comments from Jo Blanden, Sandra McNally and other attendees of the School of Economics Applied Micro Group workshop (University of Surrey) and from participants at the 2nd Annual Conference of the International Association for Applied Econometrics (IAAE) hosted by the University of Macedonia, Greece. The authors acknowledge financial support from The Leverhulme Trust (RL-2012-681). The usual disclaimer applies.

1 Introduction

During the last decade, approximately the 3.5% of elective procedures were cancelled after patients were admitted to NHS hospitals each year. Such cases are generally termed as "last minute cancellations" and they lead to waste of time and resources (Lacqua *et al.*, 1994). Moreover, they may have adverse impacts, emotional or economical, upon both patients and their families (Ivarsson *et al.*, 2002; Mangram, 1992; Tait *et al.*, 1997). There is an emerging strand of literature attempting to model the frequency of cancelled elective procedures. Two recent papers by Cookson *et al.* (2012) and McIntosh *et al.* (2012) have modelled the incidence of last-minute cancellations of elective procedures in the English National Health Service using patient-level information extracted from the Hospital Episode Statistics database. Both papers reported that age, gender, day of admission, socio-economic status and hospital characteristics were significant predictors of the probability of a procedure being cancelled after admission. An earlier observational study by Sanjay *et al.* (2007) investigated the incidence of all cancelled operations and reported that inconvenient appointments and list overruns were the most common reasons for cancelling.

Until the financial year 2009/10, each elective procedure that was cancelled after the patient was admitted, received a specific Healthcare Resource Group (HRG) code for a "planned procedure not carried out". This code was associated with a tariff of £469, known as the WA14Z tariff, and it was a part of the Payment by Results (PbR) system which was introduced in 2004/05 in order to help NHS providers to use their resources more efficiently and effectively, increase their activity, reduce their waiting times etc. (Department of Health, 2004).[1] Under this system NHS hospitals were reimbursed with a fixed amount of money every time they cancelled an elective procedure after the patient was being admitted. The tariff was paid for any planned elective admission (day case or overnight) that was cancelled on the last minute regardless of the type, cost or length of the procedure. Rather than compensating providers for a loss of revenue, it was intended

[1] Tariffs change every year responding to changing practices and market conditions. Before 2009/10 planned procedures not carried out were attached to the S22 HRG code and the associated tariff was £456 in 2008/09, £434 in 2007/08, £423 in 2006/07 etc.

to be a fixed fee to cover the cost of processing a patient (admission and discharge) when the provider would no longer receive a payment for that patient otherwise. However, unlike the other tariffs of the PbR system, the one for the cancelled operations did not reimburse providers for the cost of procedures they actually performed. On the contrary, its design could generate some adverse effects. As recently discussed by McIntosh *et al.* (2012), if the reason for a last-minute cancellation was clinical, then resources that were originally scheduled to be used to a given elective procedure were left idle and the opportunity cost of the cancellation exceeded the fixed reimbursement. Since other patients could have been treated instead and providers were reimbursed less than the actual cost of cancellation, this led to social waste. If providers should be reimbursed for not performing a procedure as planned in order to cover their revenue loss, they should be compensated with the actual cost of the last minute cancellation, which depends upon the particular procedure and patient characteristics.

However, the literature suggests that elective procedures are cancelled on the last minute mainly due to non-clinical reasons (Dexter *et al.*, 2005; Pollard *et al.*, 1996; van Klei *et al.*, 2002; Sanjay *et al.*, 2007). Whilst the WA14Z tariff was certainly lower than the total cost of aborting a procedure, since patients and their families were not reimbursed for their own costs, it was likely that the WA14Z tariff was greater than the hospital's last minute cancellation cost alone; the hospital was reimbursed for any other procedures it carried out as planned. Apart from being socially wasteful, the design of the WA14Z tariff provided hospitals with a disincentive to reduce their non-clinical last-minute cancellation rates since they were claiming reimbursements for both the last-minute cancellations and the other procedures they performed as scheduled. Moreover, the tariffs for some procedures were lower than the WA14Z one (McIntosh *et al.*, 2012). In fact, according to the 2009/10 mandatory prices information spreadsheet for admitted patient care, the WA14Z tariff was lying in the 28th percentile of the distribution of planned same day tariffs. As noted in the 2007-2008 NHS Admitted Patient Mandatory Tariff, minor endoscopic procedures on the bladder or diagnostic procedures on the stomach were attached to lower tariffs, so that providers could receive a greater reimbursement by cancelling them after admission rather than performing them as originally scheduled. Given all the above, the design of the WA14Z tariff was deemed problematic and ineffective. Furthermore, it had also substantial cost implications. Prior to mid-2010, more than 15,000 elective procedures were cancelled at the last minute each quarter on average. Hence, approximately £6.5 million each quarter was reimbursed to NHS providers for an activity they did not actually carry out, irrespective of other opportunity costs for hospitals and patients.

This paper investigates whether the design of this fixed tariff provided hospitals with an incentive to cancel procedures at the last minute and/or to cancel low tariff elective procedures in favour of higher tariff work. Fixed price reimbursements may often provide hospitals with an incentive to game the system by practising cost-driven patient selection (Propper *et al.*, 2004; Propper and Van Reenen, 2010). Cookson and Laucidella (2011) examined the possibility that providers may select against socio-economically disadvantaged hip-replacement patients, although they did not find strong evidence to support their hypothesis regarding incentives for patient selection. In the case of last-minute cancellations, however, the 2010 tariff removal offers a more appropriate setting which allows a thorough investigation regarding the existence of financial incentives to cancel a planned operation in the English NHS. More specifically, since financial year 2010/11 the HRG code associated with a last-minute cancellation was excluded from the scope of the mandatory tariff (Payments by Results Guidance 2010-11). This means that NHS providers can no longer claim a fixed reimbursement once they cancel a procedure after a patient's admission. Hence, if the tariff provided hospitals with an incentive to cancel an elective procedure at the last minute following a selective behaviour, its removal may have had a direct impact on their behaviour by decreasing the probability for a patient's procedure to be cancelled after his admission for non-clinical reasons. To explore this hypothesis, we make use of detailed patient-level data on elective procedures in all the English NHS trusts from January 1st of 2009 to December 31st of 2011. A before-and-after design is applied to evaluate whether the trend of last-minute cancellations has declined in the period after the tariff removal. Our results indicate a small, yet statistically significant decline in the probability of a last minute cancellation since the second quarter of 2010, especially for certain types of patients and diagnoses. Having controlled for patient and provider-level heterogeneity we interpret our findings as evidence favouring that removing a problematic tariff was effective.

The remainder of this article is structured as follows. Section 2 presents the estimation strategy. Section 3 outlines the data used in the analysis. Section 4 discusses the obtained results and Section 5 concludes.

2 Empirical strategy

Our analysis focuses on evaluating the impact, if any, of the tariff removal regarding the fixed reimbursement associated with last minute cancellations in the English NHS. Whether NHS providers had an incentive to admit a patient and then cancel the procedure or to cancel low fee work over high fee work, we seek to investigate if this incentive has weakened after the exclusion of that HRG code from the scope of the mandatory tariff. The tariff reform was introduced in the beginning of financial year 2010/11 (April) and it was applied to all NHS providers as it was part of the PbR system. Therefore there are no distinct treated and control providers so that a difference-in-differences or a propensity score matching approach can be used. However, given that our data are sufficiently distributed before and after the intervention, we can adopt another quasi-experimental research design based on a before-and-after approach.[2] Since there is no theoretical reason to expect a sharp change in the behaviour of providers right after the tariff removal we estimate interrupted time series models which allow for a more gradual adjustment of their behaviour over time. This method has a potentially high degree of internal validity and has been widely used in healthcare research (Bernal *et al.*, 2016; Cooper *et al.*, 2011; Huesch *et al.*, 2012; Linden, 2015).[3] In the absence of treated and comparison groups, the standard interrupted time series model assumes the following functional form:

[2]Some methodological concerns on measuring health policy interventions are compactly reviewed by Huesch *et al.* (2012).

[3]Cooper *et al.* (2011) also used interrupted time series models to estimate probabilities at the patient level. However, as they could define treated and control providers they performed a multiple-group analysis.

$$
\begin{aligned}
y_{ijt} = & \beta_0 + \beta_1 t + \beta_2 \mathbb{1}\{t \geq \tilde{t}\} + \beta_3 (t - \tilde{t}) \times \mathbb{1}\{t \geq \tilde{t}\} \\
& + \sum_{k=1}^{K} \gamma^k X_{ijt}^k + \delta_j + \eta_t + \epsilon_{ijt}
\end{aligned} \tag{1}
$$

where y_{itj} is an indicator for whether the procedure for the i-th patient admitted in the j-th NHS provider during month t was cancelled on the last minute (hence, the associated HRG code was WA14Z). Variable t is a running counter of months since January of 2009 which is the first month in our data and $\tilde{t}$ specifies the tariff removal introduced at the first month (April) of the 2010/11 financial year. Therefore, β_1 is the trajectory of the cancellation probability before the tariff removal, β_2 measures the step change of the outcome after the tariff removal and β_3 is the slope change in the post-tariff period. The latter is compared to the counterfactual, i.e. the pre-intervention trend that is assumed to had remained unchanged throughout the period in the absence of the tariff reform. Given the functional form of the model considered here this is the identification assumption of the analysis (Huesch *et al.*, 2012). Statistically significant p-values in β_2 or β_3 will indicate an immediate or a gradual treatment effect over time, respectively (Linden, 2015). Chow tests for the existence of a structural break in the series at the beginning of financial year 2010/11 are also performed (H_0: $\beta_2 = \beta_3 = 0$).

The model also controls for patient and provider-level heterogeneity although an advantage of the interrupted time series approach is that it is not affected from time-varying confounders which evolve slowly over time (Bernal *et al.*, 2016). More specifically, there are controls for gender, age, ethnicity, waiting time between referral and admission, transfer to another provider, diagnosis chapter heading, socioeconomic status of the area where the patient permanently lives into, discharge to the usual place of residence and controls about the day of the week the admission took place. At the provider level, the model controls for the hospital type, e.g. university teaching hospital, specialist hospital, foundation trust etc. However, the observed hospital type remains unchanged during the period considered here and these controls are omitted when we allow for provider-specific intercepts; δ is a set of provider-level fixed effects that capture all time invariant provider heterogeneity. Also, η denotes the month of admission and captures seasonal effects and ϵ is the error term. We model the probability of a last minute cancellation using OLS, although we probe the robustness of the results using logistic and complementary log-log regressions given that last-minute cancellations are rare (Cookson *et al.*, 2012).[4] The standard errors are corrected for heteroskedasticity and clustering at the provider level in order to allow for common error components for patients admitted to the same provider. We also conduct the analysis at the provider level taking into account the fractional nature of the response variable and controlling for provider fixed effects.

3 Data

We use the Hospital Episode Statistics (HES hereafter) database which contains anonymous administrative patient data. The original sample consisted of 22,005,931 observations of NHS patients who underwent elective procedures from January 2009 to December 2011 in various NHS facilities. After removing all the observations with missing data on key variables used in the analysis, we were left with 18,659,096 observations. A brief inspection did not indicate any correlation between missing data and particular provider or patient-level characteristics. The dependent variable a binary indicator of whether an elective procedure was cancelled after admission, i.e. whether a WA14Z HRG code was attached. There is no way of knowing what procedure had been originally planned since all cancelled operations are coded as WA14Z.

The patient characteristics include gender, age, ethnicity, waiting time, transferring to another provider, diagnosis chapter heading, socio-economic status, the Charlson comorbidity index, discharge to the usual place of residence and indicators regarding the day and month of admission. The diagnosis chapter headings follow the 10th Revision of the International Classification of Diseases and Related Health Problems (ICD-10). These diagnostic codes are used alongside with procedural codes (OPCS Classification of

[4]The estimation of linear probability models is quite common in the empirical literature and has been used in many healthcare studies (Angrist and Pischke, 2008; Cooper *et al.*, 2011).

Interventions and Procedures, not available in our HES extract) to capture every detail of a clinical event and determine the HRG codes. The admitted patients contained in our data can be classified into 1,574 3-digit alphanumeric diagnosis codes which can be aggregated to 20 chapters (see Table A1 in the Appendix for the full definition of the admissions examined here).[5] The socioeconomic status of patients is based on the socioeconomic quintile of their residence area and it is measured using the 2007 Index of Multiple Deprivation at the super output area (DCLG, 2007).[6] Waiting time measures the time between the referral of a patient and his admission for the procedure. Because of its skewed distribution, it is expressed in logarithms. The modified Charlson comorbidity index captures potential need and complexity of each procedure (Bottle and Aylin, 2011). The transfer indicator variable is included since hospitals may transfer a patient to another provider or admit a patient to a ward instead of cancelling the procedure (Cookson *et al.*, 2013). Individual time dummy variables indicate the day of week and the month of year the patient was admitted for a procedure. Provider characteristics include binary variables indicating the hospital type, i.e. whether it was a secondary care provider, a foundation trust, a university or a specialist hospital.[7] Table 1 displays the descriptive statistics on some key variables used in the analysis regarding the total period as well as before and after the tariff removal. Approximately 3% of the planned procedures are being cancelled on the last minute, with a small statistically significant decrease being observed in the post-tariff period. Regarding the other variables, their differences in means between the two periods are statistically significant in most cases, however, a brief inspection revealed that they have evolved rather slowly over time so they are not expected to violate the underlying assumptions of the segmented regression

[5]There were 4 patients classified into a special purposes code (U049) but they were excluded from the analysis due to missing information on other key variables.

[6]The index is constructed from 38 indicators across 7 weighted domains measuring an area's income, deprivation, employment deprivation, health deprivation and disability, education, skills and training, barriers to housing and service, crime and the local environment. The index is produced periodically for the Department of Communities and Local Government by the University of Oxford. The raw scores are meaningless and are categorized into quintiles since it is the relative deprivation that is relevant.

[7]Some patients were admitted to mental health and community care providers: 161,346 (0.73%) and 3,837 (0.02%) in the original sample, respectively. Moreover, due to the lack of information on key characteristics, they represent even smaller fractions in the utilized sample of non-missing values: 28,836 patients (0.15%) were admitted to mental health providers and only 9 (0.00%) to community care providers.

analysis design adopted here (Bernal *et al.*, 2016; Huesch *et al.*, 2012). This is important, especially for variables that could be considered as outcomes such as waiting time. However, given their small time variance over the period and because previous studies have used them as patient level controls when modelling last minute cancellations (Cookson *et al.*, 2013) they are not omitted from the explanatory vector and their inclusion or not does not affect the results. Moreover, controlling for the composition of the admitted population as well as for time and provider fixed effects will mitigate any biases.

[Table 1 about here]

[Figure 1a about here]

[Figure 1b about here]

The evolution of the cancellation rate is shown in Figures 1a and 1b. Figure 1a presents the average monthly cancelling rate over the total period. There seems to be a small decrease after the tariff removal however the time series plot is dominated by month effects. In order to remove them we regressed the incidence of last minute cancellation on month of admission indicators and plotted the mean residuals in Figure 1b where the decrease, although small, is more apparent. This can also be seen in Figure 2 where the relative frequency distribution of the cancellation rate across providers before and after the tariff change is depicted.[8] Moreover, the observed decline of the cancellation probability is statistically significant. Table 2 reports the cancellation rate before and after the tariff change for all admissions and for a series of sub-samples, namely by provider type and diagnosis chapter heading. There is descriptive evidence for a statistically significant incidence of last minute cancellations in the post-tariff period for patients admitted to foundation trusts and teaching hospitals as well as for certain ICD codes, e.g. infectious diseases, neoplasms, diseases of the musculoskeletal system, nervous system diseases, diseases of the digestive system, skin-related diseases, while the cancellation rate appears

[8]The upper and lower 1% tails of the distributions have been trimmed to avoid distortions caused from outliers (the mass of providers with a zero cancellation rate is higher in the post-tariff period). Results available upon request.

increased for some others, e.g. diseases of the eye and adnexa, diseases of the genitourinary system. Although small, the observed figures for last-minute cancellations are quite accurate because the HRG codes are tied to hospital reimbursement. Regarding some demographic characteristics, the cancellation probability has declined for females, those below 30 years old and those with a white ethnic background.[9]

[Figure 2 about here]

[Table 2 about here]

4 Results

Table 3 displays the results from modelling the incidence of a last minute cancellation at the patient level. In Panel A we performed the estimation on the full sample of patients with non-missing values on key variables. The model is progressively saturated to control for a variable indicating the removal of the tariff at the beginning of financial year 2010/11 (April 2010), period specific time trends, patient and provider characteristics, time fixed effects and time invariant unobserved heterogeneity at the provider level. As seen in the first two columns, the average cancellation rate is slightly above 3% and the probability of being cancelled seems to slightly decline after the tariff removal. Next, we condition the cancellation probability on a set of patient and provider-specific variables. Patient heterogeneity is captured using controls for gender (males are the base group), age (patients aged less than 9 years are the reference category), ethnic background (patients of Indian, Pakistani, Bangladeshi or any other Asian background are the base ethnic group classified as "Indian"), waiting time between referral and admission, a binary indicator of whether the patient was transferred to another provider, a binary indicator of whether the patient was discharged to his usual place of residence (rather to the temporary place of residence, a security institution, care home etc.), diagnosis chapter heading (certain infectious and parasitic diseases are the reference diagnosis category) and socioeconomic status of the patient's permanent residence area using quintiles of the 2007 IMD at the

[9]Results not shown here but are available from the authors upon request.

super output area (the lowest quintile is the base one). Provider characteristics include indicators regarding the observed hospital type, i.e. foundation trust, teaching hospital, specialist hospital, mental health hospital, community care provider and secondary care provider, but these are time invariant and they are removed from the specifications which control for provider-specific intercepts. Controlling for observed patient and provider characteristics in columns 3 and 4 does not seem to affect the magnitude and the sign of the estimated coefficient of interest (β_3, the post-tariff linear time trend) but there is also evidence for a statistically significant negative step change and a positive secular trend as well. However, controlling for time fixed effects (month and day of admission) and time invariant provider heterogeneity in columns 5 and 6 leads to the rejection of the hypothesis about the existence of a sharp change right after the tariff removal. Also the coefficient of the secular trend is now very low and marginally significant at the 10% level. The results from the full model specification seem to favour the existence of a slope change in the cancellation probability. Moreover a series of Chow tests performed in columns 2-6 also justify the existence of a structural break at the introduction of the 2010/11 PbR system (H_0: $\beta_2 = \beta_3 = 0$).

In panel B we probe the robustness of the baseline results by re-estimating the full model specification on various alternative sub-samples. In column 1, we restrict the sample to patients with length of stay less than a day since cases with greater length of stay are atypical and may confound the analysis (Cooper *et al.*, 2009). A similar point has been made by Cookson *et al.* (2013) who mentioned that any last-minute cancellations would normally occur shortly after admission. On the other hand, longer stays are more likely to result from complications or complexities not necessarily related to the elective procedure. The obtained results lead to the same conclusions regarding the existence of a slight but statistically significant slope change after the tariff removal. In column 2 of panel B we restrict the sample to patients admitted to secondary care providers as this is common when performing secondary care analysis (primary care providers, mental health and community trusts do not often perform the same types of procedures). Similarly, in column 3 of panel B we exclude patients admitted to mental health and community care providers and in column 4 we include only those patients who were discharged to their usual place of residence.[10] In column 4 we combine the restrictions about length of stay shorter than a day and discharge to the usual place or residence while in column 5 we impose all the aforementioned conditions simultaneously. Hence, the most restricted sample (column 6 of panel B) consists of patients with length of stay shorter than a day who were admitted to a secondary care provider (not to mental or community care providers) and who were discharged to their usual place of residence. The results are nearly identical to those reported in column 6 of panel A and indicate that the probability of being cancelled on the last minute declined by approximately 0.012% each month in the post-tariff period. Full results listing coefficients and t-statistics on the rest regressors of the models presented in column 6 of Table 3 are available in the Appendix Table A2.

[Table 3 about here]

Overall, the results indicate that providers seem to have responded to the tariff change, even if the estimated effect is small. The estimations on the full and the restricted samples suggest that the last minute cancellation probability for elective procedures has been declining by 0.012% each month, on average, during the post-tariff period, or approximately 3,000 less last-minute cancellations each month. In other words, removing a rather problematic tariff generated savings of about £14 million each month during the period examined here and prevented the adverse effects for a large number of patients. The same conclusions are drawn if we collapse our dataset and conduct the analysis at the provider level, given that we seek to test for any changes in the behaviour of providers in the post-reform period. In this case, given that the dependent variable is by construction bounded between 0 and 1, we estimate equation 1 using a fractional response (logit) model allowing for provider-specific intercepts (Hardin & Hilbe, 2007; Hausman & Leonard, 1997; Papke & Wooldridge, 2008). The results are provided in Table 4. In panel

[10] A cross tabulation over the total sample indicated the existence of substantial differences regarding the length of stay among patients with various destinations of discharge. Patients who were discharged to their usual place of residence had had a mean length of stay shorter than a day, while the mean length of stay of patients discharged elsewhere was approximately 19 days. More detailed results are available upon request.

A we have collapsed the dataset of all patients with non-missing values on key variables and in panel B we restrict the sample to secondary, non-mental health, non-community care providers. The results are quite similar to those obtained from the patient-level analysis and confirm the existence of a break, indicating a slightly declining trajectory of the cancellation rate after the removal of the tariff. Given the robustness of the finding across model specifications and levels of analysis, we interpret it as evidence of a slightly altered behaviour of providers regarding the last-minute cancellations of planned procedures. As a test to ensure that the model used here is not misspecified, we set some pre-policy counterfactual tariff removal dates (one, three and six months) before the official one in April 2010 and run a series of placebo regressions using the full model specification. Hence, we tested for any anticipation effects regarding the tariff removal although it was not something widely discussed before the introduction of the 2010/11 PbR system. However, the placebo regressions did not produce statistically significant evidence of a post-cancellation slope change.[11]

[Table 4 about here]

However, the tariff removal may have had a differentiated impact across types of patients. In accordance to the literature (Cookson *et al.*, 2013; McIntosh *et al.*, 2012) our estimations have shown that patients with specific demographic characteristics are more likely to be cancelled on the last minute. Also the cancellation probability varies by diagnosis chapter, therefore some procedures are more likely to be cancelled than others, for example if they are minor procedures or they are attached to tariffs lower than the WA14Z (or the S22 before the financial year 2009/10). Therefore, we estimated the full model specification (column 6, Table 3) on a series of sub-samples defined by demographic characteristics (age, gender, ethnic group, socioeconomic status), diagnosis code and provider type. Table 5 reports the results for the restricted sample, i.e. the one corresponding to panel B of Table 3, although the results are quite similar even when the

[11]More specifically, for the patient-level models the estimated coefficients of the post-tariff time trend were (t-stats in parentheses): -.00013 (-2.33), -.0008 (-1.14) and -.00011 (-1.22) for setting the tariff removal one, three and six months before the official date, respectively. For the provider-level regressions, the corresponding figures were -.00015 (-1.54), -.00006 (-0.53) and -.00017 (-1.11). The estimated coefficients of the period indicator were also zero. Results available upon request.

full sample of patients with non-missing values is used. The decline in the cancellation probability seems to be slightly greater for males. With respect to age it is mostly driven by patients aged between 60 and 90 years old. The decline is also stronger for patients with an Indian ethnic background and for those living in more deprived areas. Regarding the type of provider, there is weak evidence about a structural break in the case of specialist hospitals but in general there is not significant differentiation across provider types.

Moreover, the reduction in the last-minute cancellation probability seems to be greater for certain types of diagnoses, namely for infectious and parasitic diseases (ICD-I), diseases of the blood and certain disorders involving the immune mechanism (ICD-III), diseases of the ear (ICD-VIII) and symptoms and signs (ICD-XVIII). This provides some justification about last minute cancellations of minor procedures that are attached to lower tariffs (prior to its removal, the WA14Z tariff was lying in the 28th percentile of the 2009/10 distribution of planned same day tariffs). According to the 2009/10 tariff information spreadsheet there were several HRG codes attached to tariffs lower than the WA14Z one and were associated with ICD-10 diagnosis chapters for which the cancellation probability followed a downward slopping post-reform trajectory, e.g. thalassaemia, pleurisy, disorders of immunity without HIV/AIDS, thrombocytopenia, minor and intermediate ear disorders, etc. Ideally, we would be able to know what kind of operation would have been carried out, however, there is no way to acquire this information because all cancelled operations are coded as WA14Z rather than as the originally planned procedure (McIntosh *et al.*, 2012).[12] Knowing this information could enable us to test the extent to which providers responded to the tariff removal on an individual basis. Also, a difference-in-differences style estimator comparing procedures with high versus low tariffs (relative to WA14Z) before and after the reform could have been used. Moreover, we do not know if the procedure was cancelled for clinical or non-clinical reasons, as the latter

[12]A reasonable proxy would be to use the HES ID of patients who were cancelled on the last minute and check if they had a further elective admission in the same ICD-10 diagnosis chapter within a short period after the cancellation. However, lack of Information Governance Clearance to hold HES ID (or an encrypted version of it) restricts us from devising such a proxy. Moreover, using data from April 2013 to November 2016 we note that there are 172,062 elective admissions with WA14Z, 65.5% of which had a subsequent elective admission within 90 days and 54% of them had the same ICD 10 diagnosis chapter.

could be considered of greater relevance since the literature suggests that the reasons of last minute cancellations of elective procedures are mainly non-clinical (e.g. Dexter *et al.*, 2005; Sanjay *et al.*, 2007). However, given the design of the WA14Z tariff and its relative position over the tariff distribution one can argue about the existence of an incentive to cancel a planned procedure after admission and test this hypothesis using the timing of the tariff removal.

[Table 5 about here]

5 Conclusions

In this article we attempted to evaluate a recent tariff reform for cancelled elective procedures in the English NHS. Until financial year 2009/10 providers were able to claim a fixed reimbursement of £469 for procedures that were cancelled after the patient was admitted. Given the fact that the tariff associated with last minute cancellations was abolished in April 2010, we seek to explore whether it acted as an incentive for providers to cancel at the last minute and/or cancel low tariff procedures in favour of receiving the fixed reimbursement or higher tariff work. In order to do so, we compared their behaviour regarding last minute cancellations before and after the tariff elimination. A main assumption we make is that if the tariff did operate as an incentive for trusts, its removal would result in a decreasing probability of a procedure being cancelled on the last minute in the post-tariff period. However, due to the fact the the tariff removal was applied to all NHS providers under the 2010/11 Payment by Results framework, there was no way to define treated and control providers. Moreover, as there was no information on what procedure was originally planned to be carried out, a difference-in-differences style estimator could not be adopted either. Instead, as common in the healthcare literature we followed an interrupted time series approach to examine if there was a change in the behaviour of providers regarding last minute cancellation in the post-reform period. Based on an extract of the Hospital Episode Statistics data between January of 2009 and December of 2011 we present evidence of a statistically significant decline of the last minute cancellation probability during the post-tariff period. The results were consistent across several model specifications, sub-samples and levels of analyses, indicating a slight decrease in the incidence of last minute cancellations after controlling for patient and provider level heterogeneity. Moreover, the decline was found to be more pronounced for specific groups of patients and diagnosis types. Although there is no way of knowing what operation was originally planned to be carried out or whether the reason of the cancellation was clinical or non-clinical, we interpret that finding as evidence of hospitals' response to the tariff change. The magnitude of the estimates presented here indicate that even if the incentive to cancel was not very strong, removing a problematic tariff resulted into fewer cancellations, preventing the adverse effects for a large number of patients, and generated additional savings.

References

[1] Angrist, J.D. and Pischke, J.S. (2008), "Mostly Harmless Econometrics: An Empiricist's Companion", Princeton University Press, 2008, ISBN 978-1-4008-2982-8.

[2] Bernal, J.L., Cummins, S. and Gasparrini, A. (2016), "Interrupted Time Series Regression for the Evaluation of Public Health Interventions: A Tutorial", *International Journal of Epidemiology*, forthcoming.

[3] Bottle, A. and Aylin, P. (2011), "Comorbidity Scores for Administrative Data Benefited from Adaptation to Local Coding and Diagnostic Practices", *Journal of Clinical Epidemiology*, 64(12): 1426–1433.

[4] Cookson, G., Jones, S. and McIntosh, B. (2013), "Cancelled Procedures: Inequality, Inequity and the National Health Service Reforms", *Health Economics*, 22(7): 870–876.

[5] Cookson, R. and Laucidella M. (2011), "Do the Poor Cost More? The Relationship Between Small Area Income Deprivation and Length of Stay for Elective Hip Replacement in the English NHS from 2001 to 2008", *Social Science & Medicine*, 77(2): 173–184.

[6] Cooper, Z., Gibbons, S., Jones, S. and McGuire, A. (2011), "Does Hospital Competition Save Lives? Evidence from the English NHS Patient Choice Reform", *Economic Journal*, 121(554): F228–F260.

[7] Cooper, Z.N., McGuire, A., Jones, S. and Le Grand, J. (2009) , "Equity, Waiting Times and NHS Reforms: Retrospecive Study", *British Medical Journal*, 339: b3264.

[8] Department of Communities and Local Government, 2007. Indices of Deprivation, 2007.

[9] Department of Health, 2004. NHS reference costs 2003 and National tariff 2004 (Payment by results core tools 2004), 2004.

[10] Dexter, F., Marcon, E., Epstein, R.H. and Ledolter, J. (2005), "Validation of Statistical Methods to Compare Cancellation Rates on the Day of Surgery", *Anesthesia & Analgesia*, 101(2): 465–473.

[11] Hardin, J.W. and Hilbe, J.M. (2007), "Generalized Linear Models and Extensions", Stata press.

[12] Hausman J.A. and Leonard, G.K. (1997), "Superstars in the National Basketball Association: Economic Value and Policy", *Journal of Labor Economics*, 15(4): 586–624.

[13] Huesch, M.D., Østbye, T. and Ong, M.K. (2012), "Measuring the Effect of Policy Interventions at the Population Level: Some Methodological Concerns", *Health Economics*, 21(10): 1234–1249.

[14] Ivarsson, B. Kimbald, P.O., Sjoberg, T. and Larsson, S. (2002), "Patient Reactions to Cancelled or Postponed Heart Operations", *Journal of Nursing Management*, 10: 75–81.

[15] Linden, A. (2015), "Conducting Interrupted Time Series Analysis for Single and Multiple Group Comparisons", *Stata Journal*, 15(2): 480–500.

[16] Lacqua M. and Evans, J. (1994), "Cancelled Elective Surgery: An Evaluation", *American Surgeon*, 60(3).

[17] Mangram, J.L. (1992), "Total Joint Replacement: Implication of Cancelled Operations of Hospital Costs and Waiting List Management", *Quality in Health Care*, 1: 34–37.

[18] McIntosh, B., Cookson, G. and Jones, S. (2012), "Cancelled Surgeries and Payment by Results in the English National Health Service", *Journal of Health Services Research and Policy*, 17(2), 79–86.

[19] Papke, L.E. and Wooldridge, J.M. (2008), "Panel Data Methods for Fractional Response Variables with an Application to Test Past Rates", *Journal of Econometrics*, 145(1-2): 121–133.

[20] Payment by Results Guidance for 2010-11, Department of Health, Payment by Results Team.

[21] Pollard, J.B., Zboray, A.L. and Mazze, R.I. (1996), "Economic Benefits Attributed to Opening a Preoperative Evaluation Clinic for Outpatients, *Anesthesia & Analgesia*, 83(2): 407–410.

[22] Propper, C., Burgess, S. and Green, K. (2004), "Does Competition Between Hospitals Improve the Quality of Care? Hospital Death Rates and the NHS Internal Market", *Journal of Public Economics*, 88(7-8): 1247–1272.

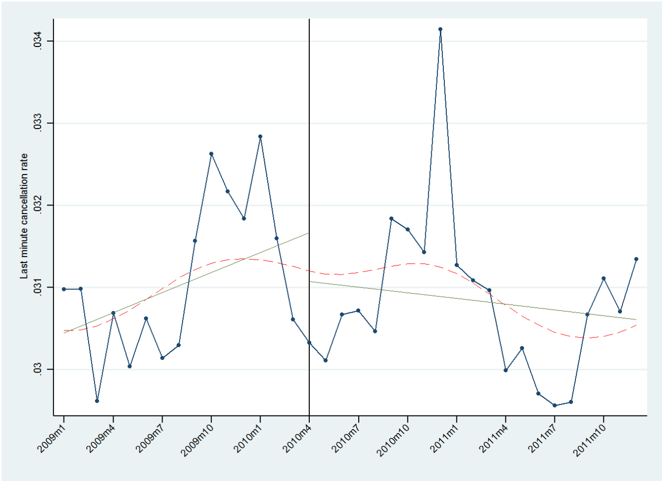
[23] Propper, C. and Van Reenen, J. (2010), "Can Pay Regulation Kill? Panel Data Evidence of the Effect of Labor Markets on Hospital Performance", *Journal of Political Economy*, 118(2): 222–273.

[24] Sanjay, P., Dodds, A., Miller, E., Arumungam, P. and Woodward, A. (2007), "Cancelled Elective Operations: An Observational Study From a District General Hospital", *Journal of Health Organisation and Management*, 21(1): 54–58.

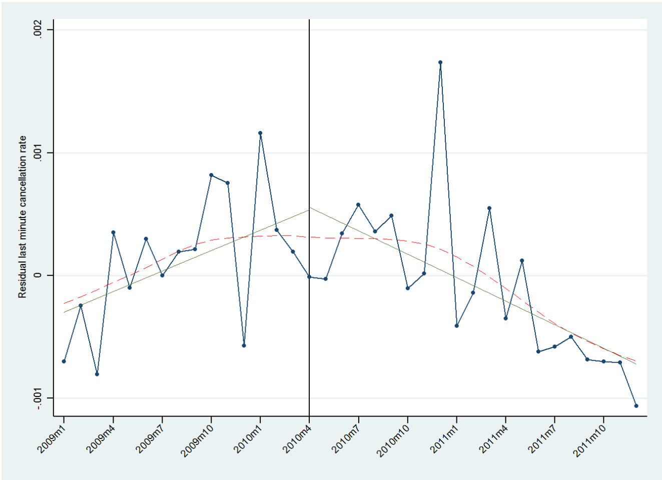
[25] Tait, A.R., Voepel-Lewis, T. and Munro, H.M. (1997), "Cancellation of Pediatric Outpatient Surgery: Economic and Emotional Implications of Patients and Their Families", *Journal of Clinical Anaesthesiology*, 9: 213–219.

[26] van Klei, W.A., Moons, K.G.M., Rutten, C.L.G., Schuurhuis, A., Knape, J.T.A., Kalkman, C.J. and Grobbee, D.E., "The Effect of Outpatient Preoperative Evaluation of Hospital Inpatients on Cancellation of Surgery and Length of Hospital Stay, *Anesthesia & Analgesia*, 94(6): 644–649.

Figures & Tables



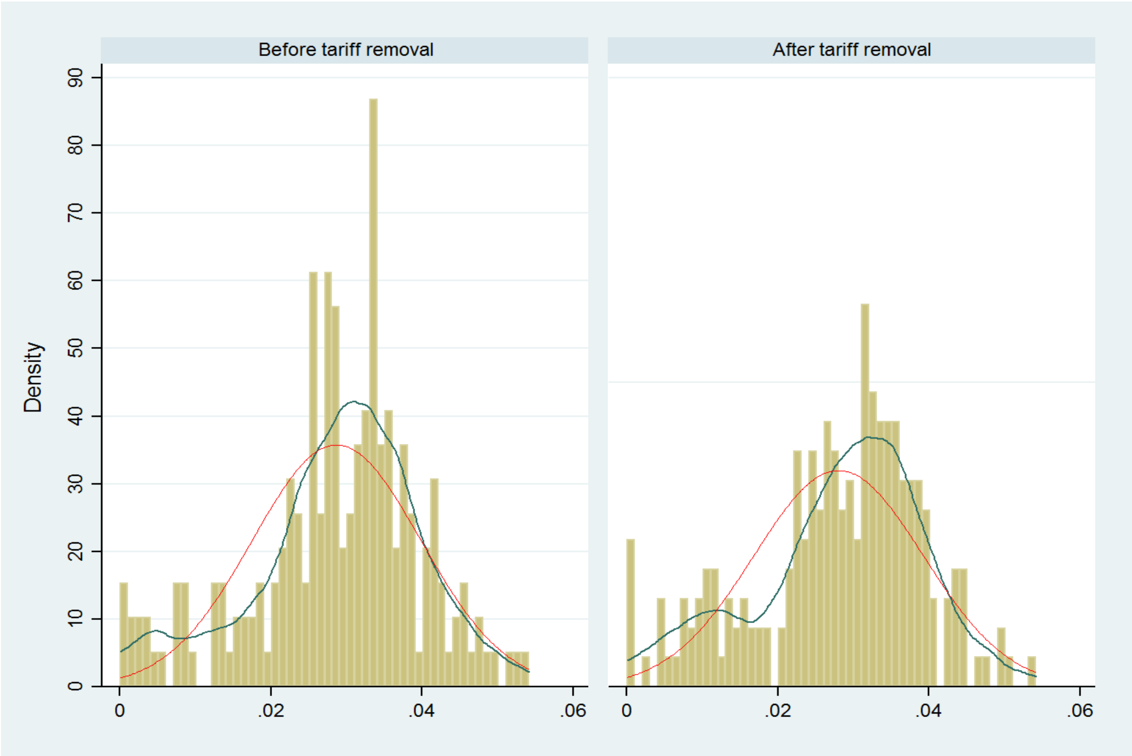
(a)



(b)

Figure 1: *Source:* Hospital Episode Statistics, 2009m1-2011m12.
Notes:(a) Tariff removal is indicated by the black vertical line. Dashed red line is running mean lowess smoothing line. Green lines are period-specific linear fit lines. (b) Mean monthly cancellation after removing the month effect.

Figure 2: Cancellation rate distribution across providers before and after the tariff removal.



Source: Hospital Episode Statistics, 2009m1-2011m12.
Notes: Navy lines are Epanechnikov kernel density estimates of the cancellation rate distributions. Red lines are overlaid normal density curves.

Table 1: Summary statistics overall and by period.

	Total period		Before tariff removal		After tariff removal		t-test
Variable name	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	p-value
Cancelled procedure	.0309	.1731	.0311	.1736	.0308	.1728	$.0001^a$
Female	.5283	.4992	.5291	.4991	.5277	.4992	$.0000^a$
Age: 0-9 years	.0411	.1984	.0411	.1986	.0410	.1983	.3093
Age: 10-19 years	.0380	.1912	.0390	.1936	.0373	.1895	$.0000^a$
Age: 20-29 years	.0637	.2443	.0650	.2465	.0628	.2427	$.0000^a$
Age: 30-39 years	.0840	.2773	.0866	.2813	.0821	.2745	$.0000^a$
Age: 40-49 years	.1311	.3376	.1323	.3389	.1303	.3366	$.0000^a$
Age: 50-59 years	.1524	.3594	.1518	.3588	.1528	.3598	$.0000^a$
Age: 60-69 years	.2026	.4019	.1998	.3998	.2045	.4034	$.0000^a$
Age: 70-79 years	.1840	.3875	.1829	.3866	.1847	.3881	$.0000^a$
Age: 80-89 years	.0935	.2912	.0926	.2899	.0942	.2921	$.0000^a$
Age: $\geq$90 years	.0096	.0976	.0089	.0937	.0101	.1002	$.0000^a$
Ethnic group: Indian	.0378	.1908	.0361	.1866	.0390	.1936	$.0000^a$
Ethnic group: Black	.0203	.1411	.0208	.1427	.0200	.1401	$.0000^a$
Ethnic group: Chinese	.0023	.0477	.0022	.0471	.0023	.0481	$.0000^a$
Ethnic group: White	.8183	.3856	.8089	.3932	.8248	.3801	$.0000^a$
Ethnic group: Other	.0172	.1298	.0165	.1273	.0176	.1316	$.0000^a$
Ethnic group: Unknown	.1041	.3054	.1155	.3196	.0962	.2949	$.0000^a$
Waiting time (log)	3.3143	1.3148	3.3099	1.3020	3.3173	1.3236	$.0000^a$
Transferred	.0029	.0542	.0032	.0568	.0028	.0524	$.0000^a$
Discharged home	.9923	.0876	.9919	.0894	.9925	.0864	$.0000^a$
Deprivation index: I	.1956	.3966	.1941	.3955	.1966	.3974	$.0000^a$
Deprivation index: II	.2168	.4121	.2166	.4119	.2170	.4122	$.0314^b$
Deprivation index: III	.2104	.4076	.2110	.4080	.2100	.4073	$.0000^a$
Deprivation index: IV	.1949	.3961	.1955	.3966	.1945	.3958	$.0000^a$
Deprivation index: V	.1823	.3861	.1829	.3866	.1818	.3857	$.0000^a$
Charlson score	1.7050	4.1191	1.4989	3.8665	1.8478	4.2795	$.0000^a$
ICD chapter: I	.0023	.0483	.0023	.0480	.0024	.0486	$.0087^b$
ICD chapter: II	.1728	.3781	.1688	.3746	.1755	.3804	$.0000^a$
ICD chapter: III	.0200	.1401	.0198	.1393	.0202	.1406	$.0000^a$
ICD chapter: IV	.0142	.1182	.0138	.1165	.0144	.1193	$.0000^a$
ICD chapter: V	.0014	.0372	.0015	.0386	.0013	.0362	$.0000^a$
ICD chapter: VI	.0257	.1583	.0251	.1565	.0261	.1595	$.0000^a$
ICD chapter: VII	.0824	.2749	.0833	.2764	.0817	.2739	$.0000^a$
ICD chapter: VIII	.0103	.1010	.0108	.1033	.0100	.0994	$.0000^a$
ICD chapter: IX	.0572	.2322	.0583	.2343	.0564	.2307	$.0000^a$
ICD chapter: X	.0229	.1496	.0234	.1512	.0226	.1486	$.0000^a$
ICD chapter: XI	.1810	.3850	.1788	.3832	.1826	.3863	$.0000^a$
ICD chapter: XII	.0222	.1475	.0234	.1513	.0214	.1447	$.0000^a$
ICD chapter: XIII	.1364	.3432	.1365	.3433	.1363	.3432	.3925
ICD chapter: XIV	.0837	.2769	.0868	.2816	.0815	.2737	$.0000^a$
ICD chapter: XV	.0111	.1049	.0118	.1081	.0106	.1025	$.0000^a$
ICD chapter: XVI	.0003	.0166	.0002	.0157	.0003	.0172	$.0000^a$
ICD chapter: XVII	.0116	.1071	.0117	.1074	.0115	.1068	$.0107^b$
ICD chapter: XVIII	.0633	.2434	.0643	.2453	.0625	.2421	$.0000^a$
ICD chapter: XIX	.0184	.1344	.0177	.1320	.0188	.1359	$.0000^a$
ICD chapter: XXI	.0628	.2426	.0615	.2403	.0637	.2427	$.0000^a$
Foundation trust	.5009	.5000	.5020	.5000	.5001	.5000	$.0000^a$
University hospital	.3815	.4858	.3778	.4848	.3841	.4864	$.0000^a$
Specialist hospital	.0400	.1959	.0404	.1969	.0397	.1951	$.0000^a$
Secondary care provider	.9569	.2031	.9591	.1980	.9553	.2066	$.0000^a$
Observations	18,659,096		7,635,636		11,023,460		

Source: Hospital Episode Statistics (HES), 2009m1-2011m12. Full sample of patients with non-missing values on key variables. $^a p < .001$; $^b p < .05$; $^c p < .01$

Table 2: Cancellation rate by various subgroups before and after the tariff removal.

Subgroup	Total period	Before tariff removal	After tariff removal	t-test
Total sample	.0309 (.1731)	.0311 (.1736)	.0308 (.1728)	.0001^a
Foundation trust	.0294 (.1689)	.0297 (.1690)	.0291 (.1683)	.0000^a
University hospital	.0327 (.1778)	.0331 (.1789)	.0324 (.1771)	.0000^a
Specialist hospital	.0291 (.1680)	.0292 (.1683)	.0289 (.1677)	.6410
ICD chapter: I	.0174 (.1309)	.0191 (.1368)	.0163 (.1267)	.0315^b
ICD chapter: II	.0193 (.1375)	.0199 (.1398)	.0188 (.1360)	.0000^a
ICD chapter: III	.0197 (.1389)	.0193 (.1376)	.0199 (.1397)	.1670
ICD chapter: IV	.0280 (.1649)	.0284 (.1662)	.0277 (.1641)	.2468
ICD chapter: V	.0134 (.1149)	.0126 (.1113)	.0141 (.1177)	.2984
ICD chapter: VI	.0167 (.1281)	.0173 (.1302)	.0163 (.1267)	.0143^a
ICD chapter: VII	.0337 (.1805)	.0328 (.1780)	.0344 (.1822)	.0000^a
ICD chapter: VIII	.0375 (.1899)	.0376 (.1903)	.0374 (.1897)	.7633
ICD chapter: IX	.0326 (.1776)	.0326 (.1777)	.0326 (.1776)	.8644
ICD chapter: X	.0376 (.1906)	.0383 (.1919)	.0374 (.1897)	.1283
ICD chapter: XI	.0242 (.1536)	.0244 (.1542)	.0240 (.1532)	.0560^c
ICD chapter: XII	.0351 (.1840)	.0360 (.1863)	.0344 (.1823)	.0053^a
ICD chapter: XIII	.0306 (.1724)	.0316 (.1750)	.0299 (.1705)	.0000^a
ICD chapter: XIV	.0371 (.1889)	.0367 (.1880)	.0374 (.1897)	.0243^b
ICD chapter: XV	.0127 (.1120)	.0139 (.1169)	.0118 (.1081)	.0000^a
ICD chapter: XVI	.0215 (.1451)	.0224 (.1479)	.0210 (.1435)	.7519
ICD chapter: XVII	.0403 (.1968)	.0406 (.1973)	.0401 (.1964)	.6523
ICD chapter: XVIII	.0508 (.2197)	.0494 (.2168)	.0519 (.2217)	.0000^a
ICD chapter: XIX	.0367 (.1879)	.0374 (.1897)	.0362 (.1868)	.0719^c
ICD chapter: XXI	.0554 (.2288)	.0543 (.2266)	.0562 (.2303)	.0000^a

Source: Hospital Episode Statistics (HES), 2009m1-2011m12. Full sample of patients with non-missing values on key variables. Standard deviations in parentheses. For the t-test, the p-values are reported. $^a p < .001$; $^b p < .05$; $^c p < .01$.

Table 3: Modelling last-minute cancellations at the patient level.

	[1]	[2]	[3]	[4]	[5]	[6]
			Panel A: Results for the full sample			
Post-tariff indicator	-.00028	-.00082^b	-.00086^a	-.00080^b	.00009	-.00001
	(-0.81)	(-2.44)	(-2.63)	(-2.47)	(0.25)	(0.03)
Time trend	-	.00012^a	.00012^a	.00012^a	.00006	.00007^c
		(3.21)	(3.17)	(3.36)	(1.67)^c	(1.81)
Post-tariff time trend	-	-.00014^a	-.00013^a	-.00014^a	-.00011^a	-.00011^a
		(-3.39)	(-3.21)	(-3.40)	(-2.77)	(-2.77)
Patient controls	No	No	Yes^a	Yes^a	Yes^a	Yes^a
Provider controls	No	No	No	Yes^a	Yes^a	No
Time fixed effects	No	No	No	No	Yes^a	Yes^a
Provider fixed effects	No	No	No	No	No	Yes^a
R-squared	.0000	.0000	.0055	.0061	.0063	.0080
Chow test (F-stat)	-	7.23^a	7.01^a	7.33^a	5.36^b	4.26^b
Patients			18,659,096			
Providers			301			
			Panel B: Results for alternative restricted samples			
Post-tariff indicator	.00039	-.00008	-.00001	-.00003	.00031	.00030
	(0.90)	(-0.22)	(-0.03)	(-0.07)	(0.71)	(0.67)
Time trend	.00007	.00006	.00007^c	.00007^c	.00006	.00006
	(1.63)	(1.51)	(1.80)	(1.80)	(1.28)	(1.24)
Post-tariff time trend	-.00014^a	-.00010^b	-.00011^a	-.00011^a	-.00012^b	-.00012^b
	(-2.75)	(-2.49)	(-2.75)	(-2.75)	(-2.44)	(-2.40)
Patient controls	Yes^a	Yes^a	Yes^a	Yes^a	Yes^a	Yes^a
Provider controls No	No	No	No	No	No	
Time fixed effects	Yes^a	Yes^a	Yes^a	Yes^a	Yes^a	Yes^a
Provider fixed effects	Yes^a	Yes^a	Yes^a	Yes^a	Yes^a	Yes^a
R-squared	.0104	.0074	.0080	.0080	.0098	.0098
Chow test (F-stat)	5.38^a	3.39^b	4.22^b	4.14^b	4.24^b	4.04^b
Patients	14,844,816	17,854,568	18,630,260	18,514,651	14,202,494	14,156,288
Providers	265	213	253	297	197	171

Source: Hospital Episode Statistics (HES), 2009m1-2011m12.
Notes: OLS estimates. t-statistics in parentheses have been calculated using standard errors corrected for heteroskedasticity and clustering by provider. $^a p < .001$; $^b p < .05$; $^c p < .01$ (for groups of variables they indicate the results of a joint significance test). Chow test H_0: $\beta_2 = \beta_3 = 0$. In Panel B the estimation samples are as follows: Length of stay<1 day (Column 1), admitted to a secondary care provider (Column 2), excluding those admitted in mental or community care providers (Column 3), discharged home (Column 4), Length of stay<1 day & admitted to a secondary care provider (Column 5), all previous restrictions imposed simultaneously (Column 6).

Table 4: Modelling last-minute cancellations at the provider level.

	[1]	[2]	[3]	[4]
	Panel A: Results for the full sample			
Post-tariff indicator	.00082 (1.48)	.00007 (0.10)	-.00042 (-1.05)	.00007 (0.16)
Time trend	-	.00013 (1.93)^c	.00014 (3.19)^a	.00011 (2.28)^b
Post-tariff time trend	-	-.00014 (-1.75)^c	-.00012 (-2.18)^b	-.00011 (-2.14)^b
Provider fixed effects	Yes^a	Yes^a	Yes^a	Yes^a
Patient controls	No	No	Yes^a	Yes^a
Time fixed effects	No	No	No	Yes^a
Pseudo R-squared	.0647	.0647	.0746	.0749
Chow test (F-stat)	-	3.11	5.13^c	4.95^c
Observations	8,763	8,763	8,763	8,763
Providers	301	301	301	301
	Panel B: Results for the restricted sample			
Post-tariff indicator	.00080 (0.21)	-.00107 (-3.01)^a	-.00141 (-3.86)^amc	-.00072 (-1.73)^c
Time trend	-	.00014 (3.63)^a	.00011 (2.55)^b	.00006 (1.28)
Post-tariff time trend	-	-.00013 (-2.78)^a	-.00010 (-2.04)^b	-.00010 (-1.86)^c
Provider fixed effects	Yes^a	Yes^a	Yes^a	Yes^a
Patient controls	No	No	Yes^a	Yes^a
Time fixed effects	No	No	No	Yes^a
Pseudo R-squared	.0105	.0105	.0112	.0113
Chow test (F-stat)	-	14.15^a	18.11^a	5.69^c
Observations	6,606	6,606	6,606	6,606
Providers	176	176	176	176

Source: Hospital Episode Statistics (HES), 2009m1-2011m12.
Notes: Average marginal effects obtained from a fixed effects fractional logit model. z-statistics in parentheses have been calculated using standard errors obtained via the Delta method. $^ap < .001$; $^bp < .05$; $^cp < .01$ (for groups of variables they indicate the results of a joint significance test). Chow test H_0: $\beta_2 = \beta_3 = 0$.

Table 5: Modelling last-minute cancellations across different groups of patients.

Patient group	$\hat{\beta}_1$ (t-stat)	$\hat{\beta}_2$ (t-stat)	$\hat{\beta}_3$ (t-stat)	R-sq.	Chow F-stat
Males	.00008(1.34)	.00043(0.75)	-.00014(-2.22)^b	.0105	3.48^b
Females	.00004(0.88)	.00018(0.40)	-.00010(-2.09)^b	.0095	2.83^c
Age: 0-9 years	-.00002(-0.11)	.00018(0.14)	-.00019(-1.19)	.0159	0.72
Age: 10-19 years	-.00015(-1.29)	.00134(1.10)	-.00012(-0.90)	.0115	1.47
Age: 20-29 years	-.00010(-1.16)	-.00021(-0.23)	.00008(0.89)	.0146	0.48
Age: 30-39 years	.00007(0.79)	-.00064(-0.73)	-.00009(-0.97)	.0113	0.54
Age: 40-49 years	.00002(0.23)	.00072(0.96)	-.00010(-1.36)	.0107	1.61
Age: 50-59 years	.00006(1.09)	-.00031(-0.44)	-.00008(-1.36)	.0103	0.74
Age: 60-69 years	.00002(0.38)	.00060(1.04)	-.00011(-1.68)^c	.0099	2.43^c
Age: 70-79 years	.00015(2.27)^b	.00055(0.89)	-.00023(-3.08)^a	.0113	6.51^a
Age: 80-89 years	.00022(2.66)^a	.00018(0.19)	-.00017(-1.82)^c	.0159	2.00
Age: $\geq$90 years	.00003(0.14)	.00298(1.22)	.00011(0.39)	.0310	0.77
Ethnic group: Indian	.00044(3.16)^a	-.00241(-1.75)^c	-.00045(-3.11)^a	.0145	5.16^b
Ethnic group: Black	.00011(0.51)	-.00095(-0.42)	.00001(0.04)	.0142	0.11
Ethnic group: Chinese	.00002(0.05)	-.00442(-0.80)	.00020(0.42)	.0201	0.70
Ethnic group: White	.00005(1.19)	.00039(0.84)	-.00013(-2.50)^b	.0090	4.61^b
Ethnic group: Other	-.00008(-0.40)	.00284(1.47)	-.00012(-0.53)	.0145	1.60
Ethnic group: Unknown	-.00005(-0.64)	-.00013(-0.18)	.00001(0.15)	.0139	0.03
Deprivation index: I	-.00006(-1.05)	.00082(1.37)	-.00002(-0.33)	.0089	1.19
Deprivation index: II	.00007(1.40)	.00050(0.89)	-.00016(-2.78)^a	.0086	4.97^b
Deprivation index: III	.00009(1.70)^c	-.00014(-0.21)	-.00014(-2.16)^b	.0089	2.35^c
Deprivation index: IV	.00002(0.36)	.00050(0.73)	-.00006(-0.85)	.0097	1.01
Deprivation index: V	.00016(1.94)^c	-.00031(-0.46)	-.00022(-2.56)^b	.0115	3.27^b
ICD chapter: I	.00080(2.75)^a	-.00628(-1.77)^c	-.00081(-2.74)^a	.0141	4.08^b
ICD chapter: II	-.00004(-0.55)	.00114(1.84)^c	-.00015(-1.35)	.0128	2.58^c
ICD chapter: III	.00028(2.28)^b	-.00093(-0.71)	-.00028(-2.20)^b	.0142	2.45^c
ICD chapter: IV	.00075(1.86)^c	-.00047(-0.08)	-.00112(-1.62)	.0623	1.42
ICD chapter: V	-.00111(-1.04)	.01392(1.31)	.00032(0.25)	.0588	0.88
ICD chapter: VI	-.00021(-1.69)^c	.00242(2.06)^b	.00010(0.65)	.0089	2.17
ICD chapter: VII	-.00003(-0.25)	.00122(1.00)	.00013(0.89)	.0222	1.25
ICD chapter: VIII	.00035(1.53)	-.00220(-0.86)	-.00051(-1.84)^c	.0147	1.79
ICD chapter: IX	-.00003(-0.31)	.00180(1.57)	-.00002(-0.15)	.0095	1.27
ICD chapter: X	-.00043(-1.65)	.00137(0.49)	-.00007(-0.24)	.0258	0.18
ICD chapter: XI	.00002(0.31)	-.00027(-0.40)	-.00001(-0.14)	.0062	0.08
ICD chapter: XII	.00011(0.68)	.00002(0.01)	-.00029(-1.59)	.0078	1.37
ICD chapter: XIII	-.00019(-1.68)^c	-.00033(-0.31)	.00014(1.13)	.0109	0.89
ICD chapter: XIV	.00006(0.47)	.00174(1.38)	-.00013(-0.92)	.0124	1.61
ICD chapter: XV	-.00023(-1.41)	.00079(0.46)	.00021(1.17)	.0489	0.70
ICD chapter: XVI	-.00037(-0.43)	.00745(0.65)	-.00011(-0.11)	.0721	0.21
ICD chapter: XVII	-.00005(-0.20)	.00189(0.70)	-.00005(-0.14)	.0119	0.24
ICD chapter: XVIII	.00051(3.00)^a	.00041(0.26)	-.00069(-3.17)^a	.0161	5.63^b
ICD chapter: XIX	-.00001(-0.03)	.00223(0.69)	-.00028(-1.00)	.0150	0.94
ICD chapter: XXI	.00048(3.30)^a	-.00202(-1.29)	-.00034(-1.64)	.0183	2.00
Foundation trust	-.00003(-0.63)	.00031(0.57)	-.00003(-0.44)	.0089	0.35
University hospital	.00002(0.26)	.00042(0.59)	-.00012(-1.36)	.0097	1.38
Specialist hospital	.00018(1.67)	-.00207(-1.53)	-.00020(-1.25)	.0131	4.31^b

Source: Hospital Episode Statistics (HES), 2009m1-2011m12.
Notes: OLS estimates. t-statistics in parentheses have been calculated using standard errors corrected for heteroskedasticity and clustering by provider. $^a p < .001$; $^b p < .05$; $^c p < .01$. Chow test H_0: $\beta_2 = \beta_3 = 0$.

Table A1: Sample distribution and definitions of principal diagnoses by ICD-10 codes.

Chapter	Diagnosis block	Frequency	%	ICD-10 definition
ICD-I	A00-B99	43,699	0.23	Certain infectious and parasitic diseases
ICD-II	C00-D48	3,223,788	17.28	Neoplasms
ICD-III	D50-D89	373,562	2.00	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
ICD-IV	E00-E90	264,282	1.42	Endocrine, nutritional and metabolic diseases
ICD-V	F00-F99	25,840	0.14	Mental and behavioural disorders
ICD-VI	G00-G99	479,868	2.57	Diseases of the nervous system
ICD-VII	H00-H59	1,536,648	8.24	Diseases of the eye and adnexa
ICD-VIII	H60-H95	192,323	1.03	Diseases of the ear and mastoid process
ICD-IX	I00-I99	1,066,880	5.72	Diseases of the circulatory system
ICD-X	J00-J99	427,662	2.29	Diseases of the respiratory system
ICD-XI	K00-K93	3,377,888	18.10	Diseases of the digestive system
ICD-XII	L00-L99	414,972	2.22	Diseases of the skin and subcutaneous tissue
ICD-XIII	M00-M99	2,545,103	13.64	Diseases of the musculoskeletal system and connective tissue
ICD-XIV	N00-N99	1,561,740	8.37	Diseases of the genitourinary system
ICD-XV	O00-O99	207,510	1.11	Pregnancy, childbirth and the puerperium
ICD-XVI	P00-P96	5,158	0.03	Certain conditions originating in the perinatal period
ICD-XVII	Q00-Q99	216,449	1.16	Congenital malformations, deformations and chromosomal abnormalities
ICD-XVIII	R00-R99	1,180,386	6.33	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified
ICD-XIX	S00-T98	343,108	1.84	Injury, poisoning and certain other consequences of external causes
ICD-XXI	Z00-Z99	1,172,230	6.28	Factors influencing health status and contact with health services
Total		18,659,096	100.00	

Source: Hospital Episode Statistics, 2009m1-2011m12.

Table A2: Full set of results on the rest covariates.

	Full sample		Restricted sample	
Variable	Coeff.	t-stat.	Coeff.	t-stat.
Female	-.00397^a	(-16.74)	-.00397^a	(-13.33)
Age: 10-19 years	-.00288^b	(-2.53)	-.00230	(-1.50)
Age: 20-29 years	.00117	(0.96)	.00231	(1.54)
Age: 30-39 years	.00026	(0.22)	.00137	(0.96)
Age: 40-49 years	.00057	(0.47)	.00162	(1.11)
Age: 50-59 years	.00112	(0.90)	.00214	(1.43)
Age: 60-69 years	.00158	(1.26)	.00307^b	(2.03)
Age: 70-79 years	.00368^a	(2.94)	.00551^a	(3.62)
Age: 80-89 years	.00706^a	(5.30)	.00909^a	(5.56)
Age: $\geq$90 years	.01391^a	(6.65)	.01587^a	(6.14)
Ethnic group: Black	.00287^a	(3.02)	.00356^a	(3.03)
Ethnic group: Chinese	-.00100	(-0.88)	-.00186	(-1.30)
Ethnic group: White	-.00512^a	(-8.34)	-.00629^a	(-7.68)
Ethnic group: Other	.00018	(0.27)	.00367	(0.44)
Ethnic group: Unknown	-.00790^a	(-10.86)	-.01027^a	(-10.38)
Waiting time (log)	.00337^a	(13.71)	.00518^a	(16.58)
Transferred	-.00111	(-0.45)	-	-
Discharged home	.00526^b	(2.47)	-	-
ICD chapter: II	.00523^a	(3.35)	.00823^a	(4.49)
ICD chapter: III	.00469^b	(3.09)	.00495^a	(2.98)
ICD chapter: IV	.01187^a	(3.46)	.01752^a	(3.82)
ICD chapter: V	.01525^a	(5.07)	.02635^a	(6.55)
ICD chapter: VI	.00169	(1.27)	.00281^c	(1.83)
ICD chapter: VII	.01342^a	(5.64)	.01194^a	(4.68)
ICD chapter: VIII	.01896^a	(11.49)	.02662^a	(13.05)
ICD chapter: IX	.01547^a	(9.86)	.01654^a	(9.88)
ICD chapter: X	.02114^a	(10.25)	.04748^a	(13.95)
ICD chapter: XI	.00756^a	(6.40)	.00835^a	(6.26)
ICD chapter: XII	.01869^a	(13.36)	.01961^a	(12.51)
ICD chapter: XIII	.01435^a	(11.12)	.02354^a	(13.88)
ICD chapter: XIV	.02048^a	(13.68)	.03127^a	(16.55)
ICD chapter: XV	.00687^a	(3.19)	.01106^a	(4.65)
ICD chapter: XVI	.00516	(1.32)	.00910^b	(2.01)
ICD chapter: XVII	.02266^a	(11.90)	.03433^a	(13.71)
ICD chapter: XVIII	.03561^a	(16.83)	.03942^a	(17.14)
ICD chapter: XIX	.02239^a	(12.01)	.04352^a	(15.85)
ICD chapter: XXI	.03674^a	(19.69)	.04079^a	(19.41)
Deprivation index: II	.00119^a	(8.70)	.00146^a	(8.60)
Deprivation index: III	.00256^a	(15.79)	.00318^a	(15.60)
Deprivation index: IV	.00557^a	(23.18)	.00678^a	(23.14)
Deprivation index: V	.00938^a	(25.20)	.01135^a	(25.82)
Charlson score	-.00028^a	(-2.96)	-.00017	(-1.46)

Source: Hospital Episode Statistics (HES), 2009m1-2011m12.
Notes: OLS estimates. Models correspond to those in column 6 of Table 3. t-statistics in parentheses have been calculated using standard errors corrected for heteroskedasticity and clustering by provider. $^a p < .001$; $^b p < .05$; $^c p < .01$.