ISSN 1745-8587

Birkbeck Centre for Applied Macroeconomics

BCAM 1502

The Asymmetric Effects of Waiting Time Targets in Health Care

Sofia Dimakou
Technological Educational Institute of Athens

Ourania Dimakou
SOAS

Henrique S. Basso
Banco de España

July 2015

The Asymmetric Effects of Waiting Time Targets in Health Care*

Sofia Dimakou†, Ourania Dimakou‡, Henrique S. Basso§

Abstract

Waiting time targets have been a key policy intervention in many OECD countries, aimed at reducing persistent waiting times for healthcare. What is the impact of targets on the distribution of patients' waiting time? Do they affect healthcare outcomes? We address the first question by developing a theoretical model of healthcare provision and empirically assessing the entire distribution of patients' durations at the hospital level. Our model and empirical evidence identify two distinct admission patterns. Hospitals respond by either treating all patients faster or by 'substituting' among short and long waiters, indicating an asymmetric effect across patients. In order to address the impact of targets on healthcare outcomes (mortality, prolonged healthcare, delayed discharge at the patient level) we explore the identified heterogeneity of responses across hospitals. We find supportive evidence of a systematic difference in outcomes of patients treated in hospitals that exhibit asymmetric responses to targets.

JEL Classification: I18, I11, H51

Keywords: Waiting time targets, Hospitals, Prioritisation, Public Health Provision, Government Policy

*This paper is partially drawn from Sofia Dimakou's PhD dissertation at City University. The views expressed in this paper are those of the authors and do not necessarily coincide with those of the Banco de España and the Eurosystem. Ourania Dimakou and Henrique S. Basso are also affiliated with the Birkbeck Centre for Applied Macroeconomics (BCAM)

†Department of Business Administration, Technological Educational Institute of Athens, 12243, Aigaleo - Athens, Greece; e-mail: s.dimakou@gmail.com

‡Economics Department, School of Oriental and African Studies, University of London - Russell Square, London, WC1H 0XG, UK; e-mail: od1@soas.ac.uk

§Banco de España, Research Division, Alcalá 48, 28014 Madrid, Spain e-mail: henrique.basso@bde.es

Introduction

Large waiting lists and long waiting times for elective surgery have been at the core of policy concerns in many OECD countries. Within national health systems waiting lists function as non-price rationing and signaling devices that reconcile demand and supply, when healthcare provision is free and supply is constrained. Given the vital importance of providing good quality and prompt national healthcare, policy makers have extensively focused on waiting lists and times within a broader set of performance indicators. In many cases one of the main direct policies to tackle long waits has been the adoption of universal maximum waiting time targets and provision of direct financial incentives to reduce waiting times. Despite reports of reduced excessive and average waiting times, the ways by which hospitals manage their waiting lists and meet (or not) the corresponding targets have been less rigourously analysed.

This paper explores a more detailed representation of waiting lists to analyse the impact of waiting targets across different patients and to investigate whether healthcare outcomes are affected after their introduction. By analysing the different representations of the entire distribution of patients' waiting times our theoretical and empirical models provide unique insights on how hospital's admission patterns are affected by the implementation of targets. We identify heterogeneous effects of waiting time targets across different patients at the hospital level. We then explore those differences to measure the impact of targets on patients healthcare outcomes.

Our starting point is the development of a theoretical model that conceptualises the main characteristics of hospitals and allows us to determine the patterns by which elective patients are provided treatment. The core output of the model is a treatment plan that generates a distribution of waiting times across patients. Governments, concerned with excessive waiting times and the volume of treatment intervene by introducing maximum waiting times. We establish two distinct responses in hospitals treatment plans after the implementation of targets. First, we obtain a symmetric response, whereby all patients are treated faster while no one waits more than the maximum; government intervention is successful in decreasing waiting times and increasing societies' benefits from healthcare provision. Second, however, an asymmetric response may arise whereby, while excessive waiting times are eradicated, the waiting list is 'manipulated' and the prioritisation of treatment altered. Longer waiters are brought in before the target's limit (thus benefiting from the policy) at the expense of short waiters that now have to wait longer for treatment. The different responses depend on the structural conditions of the hospital (costs, benefits from treatment), as well as on the manager's effort in increasing treatment capacity while keeping hospital inputs (beds, operating rooms, medical and non-medical personnel) constant.

On the empirical side, we firstly employ the techniques of duration analysis and Hospital Episode Statistics (HES henceforth) data for the English NHS during 1997-2005 to estimate the whole waiting time distribution of elective patients at the hospital level. Our empirical results provide evidence of hospitals' efforts in reducing waiting lists and catching up with targets. Waiting time distributions display a typical 'bunching effect' whereby patients that

were treated after the target are re-ordered to the periods preceding the target for almost all hospitals in our sample. We also confirm the two key patterns identified in our theoretical model. There are NHS hospitals that manage to bring the whole waiting time distribution down. However, the most common response to targets (for around 60% of the hospitals in our sample) involves trade-offs between short and long waits (asymmetric/ 'shape' effects over time). This behaviour gets prominent as the time targets shorten. In this case, we observe not only a 'bunching effect' before the target's limit but a shift in the distribution affecting short duration patients. In our second empirical exercise we investigate whether targets affect healthcare provision. We utilise patient-level data, focusing on mortality, prolonged healthcare and delayed discharge as measures of healthcare outcomes and explore the identified heterogeneity of response across hospitals obtained from our duration analysis. We find supportive evidence of a systematic worsening of outcomes for patients treated in hospitals that exhibit asymmetric responses to targets compared to patient outcomes in those hospitals that do not. The odds ratio of suffering an adverse result, understood as the necessity of prolonged healthcare or death in hospital, in 2005 increased by 1.5 times for patients treated at hospitals where an asymmetric response across patients was observed after the introduction of targets, relative to hospitals in which responses were symmetric. The odds ratio of a delayed discharge (time to discharge is above the mean) increases by 1.25 times.

Related Literature

Although the literature on health care provision and waiting lists and times is vast, contributions that look at the overall distribution of waiting times are rare, particularly at the theoretical level. The closest analyses to our model come from Iversen (1993) and Siciliani (2006) with the later developing a continuous time dynamic framework. Dixon and Siciliani (2009) describe and map the distribution of patients already treated (HES data) with the distribution of patients waiting on the list (waiting list returns). However, while looking at the overall waiting time distributions, these are not derived from a model of hospital behaviour, but are rather based on ad hoc assumptions about hazard rates parametrisation. An important and novel aspect of our theoretical work is the emphasis on waiting time distributions, and the consequent insights we can draw regarding hospitals admissions patterns after the policy intervention.

On the empirical literature, duration analysis is only used in few studies. MacCormick and Parry (2003) applied it using data for one tertiary hospital in New Zealand and Levy, Sobolev, Hayden, Kiely, FitzGerald, and Schechter (2005) while looking at a subset of hospitals/ operations in Canada. For the UK national waiting targets, Dimakou, Parkin, Devlin, and Appleby (2009) use HES data for two years and focus at varying aggregation levels. Our work expands on the latter in several dimensions. By employing this technique for the UK, using a longer time span, and focusing at the hospital level we identify particular hospital-level patterns of admissions in response to the introduction of national targets. Unlike Dimakou et al. (2009) and guided by our theory, we concentrate attention away from peaks in hazards curves at target limits and towards the shape of survival and hazard curves at short durations.

Another empirical analysis that is related to ours is presented in Propper, Sutton, Whitnall, and Windmeijer (2010). In a different framework, they also assess the impact of waiting time targets on patients outcomes without however findings systematic effects. They establish the effects of targets on hospital behaviour by looking at the 'bunching effect' (percentage of patients close to breaching the target). However, guided by our theoretical and duration analysis contributions, we provide for a potentially better measure of hospitals' response to the policy intervention of targets. Besides the 'bunching effect', which is commonly observed across hospitals, exploring the heterogeneity with respect to short-waiters sheds more light on the potential effects of the targets policy and gives evidence of differentiated patients outcomes.

The rest of the paper is organised as follows. In Section 2 we present the methodology and results of our theoretical investigation of the effects of waiting targets at the hospital level. Section 3 proceeds in developing the two empirical exercises; by first confirming our theoretical predictions on hospital admission patterns after the introduction of the policy, and subsequently using these result to assess the impact of waiting time targets on patients healthcare outcomes. Section 4 concludes.

2 Theoretical Framework

We develop a framework to characterise hospital/manager's treatment plans of patients waiting on a list. Treatment plans describe the volume of patients treated at each period and the length of time each patient waited before treatment. While providing treatment, hospitals must allocate 'inputs' (beds, operating theaters, medical staff) for each patient, taking into account the desire to prioritise patients regarding the length of time they waited on the list, and the overall costs each treatment entails.

Treatment plans are determined by the hospital to maximise the benefits of healthcare provision. This decision is subject to three constraints. First, hospitals must abide by a budget constraint. Second, patients treated and changes in waiting lists must be consistent with the inflow of new patients each period. Finally, treatments supplied at each period must be consistent with the utilisation of resources and infrastructure of the hospital and its healthcare provision function, which determines the feasibility constraint of the hospital. We furthermore assume hospital's managers can increase the hospital's capacity of healthcare provision by exerting effort (organisation, overutilisation of resources) but may pay a utility cost in doing so.

The hospital's treatment plans resulting from this optimisation decision are described by the distribution of patient waiting times. We utilise this framework to study the effects of waiting time targets introduced by governments on this distribution. Comparing the benchmark model without the policy intervention and the hospital behaviour under waiting time targets yields predictions about the changes in distribution of waiting times after the introduction of targets depending on the implicit costs managers face to improve hospital's treatment capacity.

The two main elements of our model are the set of patients that are currently waiting

to be treated and the hospital that is the healthcare supplier. We explain each element in more detail next.

2.1 Patients

Patients currently in the waiting list, L_t, are characterised by the time they have been on the list, or their duration $d = 1, 2, ...q$. d denotes the period elapsed between joining the waiting list of a specialist and admittance for surgery at the hospital. The minimum possible waiting time is one period ($d = 1$) and the maximum time is bound by q (patients do not wait indefinitely). At each time t hospitals treat $k_{d,t}$ patients that have been in the list with duration d. Thus, total patients treated at time t is given by $k_t = \sum_d k_{d,t} \in L_t$.

We do not explicitly model the demand for health care, considering a reduced form relationship where the inflow of patients to the hospital is decreasing in expected duration. The higher the expected waiting time at the beginning of t is, the lower the demand for public health care.[1] Formally, the inflow of patients in the list, and equivalently, the demand for elective health care at the beginning of time t is given by

$$x_t = Z - \theta E_{t-1}(d)$$

where $E_{t-1}(d)$ denotes the duration patients, entering in the list at time t, expect at time $t-1$ (defined below), and Z is the potential demand for health care, being a function of a vector of exogenous demand factors. These may include socio-economic conditions and morbidity rates. Finally, the sensitivity of demand for healthcare to expected duration is captured by θ.

Before we describe the hospital's main features we briefly present the theoretical representations of the waiting time distribution. In our theoretical model waiting time is modeled as a discrete variable, where a period of time is equivalent to a month. The probability function (PF) of waiting time depicts the whole spectrum of the relative frequencies of patients having waited distinct periods of time until treatment at t, $f(d) = P(D = d)$. The cumulative function (CF) corresponds to the probability of having waited d periods or less, $F(d) = P(D \leq d)$. From here we obtain the two main representations of waiting time distributions used in our study, namely the survival and hazard functions. The survival function shows the probability of a person remaining (surviving) on the list beyond a given time and is indicative of cumulative rates of treatment. We derive the survival function as the complement of the cumulative function, that is $S(d) = 1 - CF = P(D > d)$. The hazard function is the risk of 'failure' at some time t. Essentially, it shows the rate at which patients leave the waiting list at a given time, conditional on having waited in the list up to that point. It thus approximates the conditional instantaneous probability of admission,

[1] This reduced form can be obtained by assuming that individuals' benefits from healthcare decrease while waiting for treatment and that they have a costly alternative (private providers) available, which is standard in the literature of waiting times. Expected waiting time acts as a rationing devise equilibrating demand and supply, similar to what prices do. See for instance Cullis, Jones, and Propper (2000), Goddard, Malek, and Tavakoli (1995), Iversen (1997), Besley, Hall, and Preston (1999), Martin and Smith (1999), Gravelle, Dusheiko, and Sutton (2002), Siciliani and Hurst (2005) and Siciliani (2006). Note that extensive expected waiting times can also reduce demand of elective surgeries by discouraging GPs from making referrals.

rather than the unconditional one (PF). Thus, $h(d) = P(D = d|D \geq d)$. Table 1 shows the different formats of the waiting time distribution.

Table 1: Theoretical Waiting Time Distribution

d	$f(d)$ $P(D = d)$	$F(d)$ $P(D \leq d)$	Survival Function $P(D > d)$	Hazard Function $P(D = d\|D \geq d)$
0	0	0	1	0
1	$\frac{k_{1,t}}{k_t}$	$\frac{k_{1,t}}{k_t}$	$1 - \frac{k_{1,t}}{k_t} = \frac{\sum_{d=2}^{q} k_{d,t}}{k_t}$	$\frac{k_{1,t}}{k_t}$
2	$\frac{k_{2,t}}{k_t}$	$\frac{k_{1,t}+k_{2,t}}{k_t}$	$1 - \frac{k_{1,t}+k_{2,t}}{k_t} = \frac{\sum_{d=3}^{q} k_{d,t}}{k_t}$	$\frac{k_{2,s,t}}{\sum_{d=2}^{q} k_{d,s,t}}$
.	.	.	.	.
.	.	.	.	.
$q-1$	$\frac{k_{q-1,t}}{k_t}$	$\frac{\sum_{d=1}^{q-1} k_{d,t}}{k_t}$	$\frac{k_{q,t}}{k_t}$	$\frac{k_{(q-1),t}}{k_{(q-1),t}+k_{q,t}}$
q	$\frac{k_{q,t}}{k_t}$	1	0	1

The expected waiting time at time t under rational expectations is given by

$$E_{t-1}^{RE}(d) = E_{t-1}\left(\sum_{d=1}^{q} d\,\frac{k_{d,t+d-1}}{x_t}\right) = E_{t-1}\left(1 \times \frac{k_{1,t}}{x_t} + 2 \times \frac{k_{2,t+1}}{x_t} + ... + q \times \frac{k_{q,t+(q-1)}}{x_t}\right).$$

2.2 Hospital

The three key features of the hospital in our model are the benefits of providing treatment (utility), its operational costs and its capacity (feasibility) constraint given resources and infrastructure.

2.2.1 Treatment capacity of the hospital

Each patient of duration d is treated utilising beds/operating theaters – hours[2] and hospital resources, which might simply be the amount of doctors/nurses (in hours worked), but could also include diagnostic tests being run using other hospital infrastructure. The hospital's treatment ('production') function is given by

$$f(b_{d,t}, r_{d,t}) = k_{d,t} = \chi_d(\delta_t) b_{d,t}^{\alpha} r_{d,t}^{\beta} \tag{1}$$

where $b_{d,t}$ are the beds allocated to the treatment of duration d patients, $r_{d,t}$ are the resources allocated to the treatment of duration d patients, and α, $\beta > 0$. Finally, χ_d determines the amount of patients treated of each duration for a given combination of beds and resources and δ_t the effort hospitals/managers make to increase treatment capacity. We assume

Assumption 1. $\frac{\partial \chi_d}{\partial d} > 0$, $\frac{\partial^2 \chi_d}{\partial d^2} < 0$ *and* $\frac{\partial \chi_d}{\partial \delta} > 0$.

[2]Given that each period consists of a month in our model, we interpret the amount of beds as the days/hours that a bed is occupied while patients of duration d are being treated.

Thus, treating a patient quicker requires more hours of beds and resources, since doctors might have to wait for diagnostic results and patients might have to wait for a time slot in operation theaters. Equivalently, with the same hours of beds/theaters and resources, a smaller number of patients can be treated with low duration.[3] However, the longer the duration, the lower the gain in decreased resource utilisation per treatment is, by further increasing the patient's wait. Finally, more manager's effort implies more treatment holding (hours of) beds and resources constant. In other words, more effort from the manager can lead to methods innovation, better organisation and more efficient allocation of given beds and resources, increasing treatment capacity. We define $B_t = \sum_d b_{d,t}$ as the total amount of bed-hours and $R_t = \sum_d r_{d,t}$ as the total amount of resources utilised by the hospital in period t.

2.2.2 The utility of the hospital

The hospital's utility from healthcare provision, or benefits from treatment, at any point in time t is given by

$$U_t = g(k_t) = \sum_d g(k_{d,t}) - c_\delta \delta_t. \tag{2}$$

The first term denotes the utility the hospital derives from treating patients at different durations. $g(k_{d,t})$ denotes the hospital's (monetary or non-monetary) gain from treating k patients of duration d. Recall that here the waiting time (d) is not a choice variable, but it is endogenously determined. The hospital chooses optimally the number of patients of each duration to be treated at time t, and this choice determines the waiting time implicitly. We make two general assumptions on the hospital's utility.

Assumption 2. *For a given number of patients treated (i.e. fixed k), the higher the waiting time, the lower the hospital's utility. That is,*

$$\frac{\partial g(k_{d,t})}{\partial d} < 0 \quad or \quad g(k_{d_1,t}) > g(k_{d_2,t}) \quad for\ d_2 > d_1.$$

The hospital prefers to treat as many people as possible sooner rather than later. Prioritisation by waiting times is equivalent to the assumption commonly done in the literature (see for instance Iversen (1993) and Siciliani (2006)) that the more a patient waits the lower are his/her benefits from treatment and thus the lower the hospital's utility.

Assumption 3. *For the same d, $g(k_{d,t})$ is concave in $k_{d,t} \in [0, k]$ and exhibits a turning point.*

Hospital's utility is increasing in the number of treatments until a threshold point. From that level of activity and onwards, utility declines as more patients (of the same d profile)

[3]The first statement in assumption 1 implies that it is hard for the hospital to treat patients quickly or equivalently some waiting allows the hospital to reduce costs of providing treatment, using resources more efficiently. Although this negative relationship is well established in the literature, both theoretically (Iversen (1993)) and empirically (Siciliani, Stanciole, and Jacobs (2009)), these contributions also suggest that there might be a level of duration beyond which beds/resources usage increases (due to higher administrative and medical resources required to manage a long waiting list). We assume that increased costs (in terms of resources/beds) due to long waits do not occur before q.

are treated. This assumption implicitly recognises that spreading treatment across different durations allows for a better management of capacity and resource utilitsation, increasing the hospital's gains from treatment (see also Iversen (1993)). Siciliani (2006) makes a similar assumption for average waiting time, while here we focus on the duration of each treatment.

The second term in the hospital utility denotes the manager's disutility of effort in increasing treatment capacity. $c_\delta > 0$ controls the relative weight of this utility loss, which might reflect manager's personal costs to ensure full maximisation of resources or losses derived from overutilisation of personnel. Note that c_δ may not merely reflect managerial ability, but could more likely be the result of current hospital conditions, which are related to the size of x-inefficiencies. This term introduces a wedge between society's benefits of healthcare ($\sum_d g(k_{d,t})$) and the hospitals' total utility of providing treatment.

2.2.3 The cost of the hospital

The hospital is, by construction, capacity constrained, and hence not able to treat all patients that require treatment at time t. Its cost from providing health care can be decomposed into four separable parts

$$C_t = c_B \bar{B} + c_R R_t + \tau (R_t - \bar{R})^2 + wt(k_{d,s}; \widehat{d}). \tag{3}$$

where c_B,c_R, $\tau > 0$. The first part denotes the cost of maintaining the total amount of bed-hours $\bar{B}$ available for each period t. The second and third relate to the costs of resources. c_R denotes the cost of utilising resources $\sum_d r_{d,t} = R_t$, while τ controls the additional costs of utilizing resources above an overall hospital resource availability, given by $\bar{R}$.[4] Finally, the last term in equation (3), $wt(k_{d,s}; \widehat{d})$, represents the waiting time target policy intervention, defined as

$$wt(k_{d,s}; \widehat{d}) = \begin{cases} 0 & \text{if } d \leq \widehat{d} \\ \phi k_{d,s} & \text{if } d > \widehat{d}. \end{cases}$$

$\widehat{d}$ is the universal waiting time target the government sets such that 'no elective patient should wait more than $\widehat{d}$ periods since added to the list'. ϕ is a scale parameter that ensures penalties from breaching the target are significant to alter hospitals management practices. This characterisation matches the unconditional maximum waiting time guarantee introduced by the NHS Plan in 2000, together with the penalties and rewards structure that accompanied it.[5]

2.3 Hospital's maximisation problem

In order to facilitate notation of the hospital's problem, we first describe the list of patients waiting to be treated. Let the number of patients of duration $d > 1$ currently waiting for treatment at time t be $\Psi_{d,t-1}$. This stock is equal to the inflow of patients in time $t - d + 1$

[4] R_t might be greater than $\bar{R}$, for instance, when hospitals require doctors/nurses to work over-time.

[5] Note that although we introduce targets as costs, we could equivalently set them as financial incentives for waiting time reduction (as it was in fact implemented in some OECD countries). Given that in both designs the hospital budget constraint is altered due to the policy, they generate similar implications to the optimal waiting list distribution.

minus all patients treated during periods $t-d+1$ until $t-1$. Formally, we define[6]

$$\Psi_{d,t-1} = x_{t-d+1} - \sum_{j=1}^{d-1} k_{d-j,t-j}.$$

The hospital maximises its utility function, selecting δ_t and $\{k_{d,t}, b_{d,t}, r_{d,t}\}$, for all d at time t subject to its constraints, thus

$$\max_{\delta_t,\{k_{d,t},b_{d,t},r_{d,t}\}_d} E_0 \sum_{t=0}^{\infty}\sum_{d=1}^{q} g(k_{d,t}) - c_\delta \delta_t$$

$$\begin{aligned}
\text{Subject to} \quad & \textit{(C1)} \quad c_B\bar{B} + c_R \sum_d r_{d,t} + \tau\left(\sum_d r_{d,t} - \bar{R}\right)^2 + wt(k_d;\widehat{d}) \leq M_t \\
& \textit{(C2)} \quad \sum_d b_{d,t} \leq \bar{B} \\
& \textit{(C3)} \quad k_{d,t} = \chi_d(\delta_t) b_{d,t}^{\alpha} r_{d,t}^{\beta} \\
& \textit{(C4)} \quad x_t = Z - \theta E_{t-1}(d) \\
& \textit{(C5)} \quad 0 \leq k_{d,t} \leq \Psi_{d,t-1}, \ \textit{(C6)} \quad \Psi_{d,t} = 0 \text{ for } d > q
\end{aligned}$$

With respect to the cost of health care provision, we assume that the hospital has a budget allocated for elective surgeries given by M_t, and the first constraint corresponds to the budget constraint of the hospital. Here and unlike in Ellis and McGuire (1986) the budget allocated to the hospital is exogenously given and thus our basic set-up (without government targets) is closely linked to the non-cooperative game of Iversen (1993). The second constraint, *(C2)*, ensures beds allocated are within hospital's infrastructure limit. *(C3)* states the treatment function. The forth constraint ensures the hospital takes the evolution of patients inflow into account. The fifth constraint states that the amount of patients of duration d treated at time t ($k_{d,t}$) must be between zero and the number of untreated patients in the list for that duration. In other words, the number of people selected for treatment at time t cannot exceed the corresponding number of people waiting. Lastly, using *(C6)*, we impose that the maximum waiting time is q. We solve this problem assuming a steady state has been reached (see Appendix A for details) and thus the number of entries to the list is equal to the number of patients treated at any point in time ($x_t = k_t$) and the optimal $k_{d,t}$ are time-invariant. At the steady state the expected waiting time becomes

$$E_{t-1}(d) = \bar{d} = \sum_{d=1}^{q} d\ f(d) = \sum_{d=1}^{q} d\,\frac{k_d}{k} = 1 \times \frac{k_1}{k} + 2 \times \frac{k_2}{k} + ... + q \times \frac{k_q}{k}.$$

2.4 Waiting Time Distribution and the Effects of Waiting Targets

The main output of our framework is the hospital's waiting time distribution. We concentrate on two key representations of that distribution, the survival curve and the hazard curve, as described in Table 1. In order to explore the mechanism and highlight our main

[6]The total list of patients at time t is given by the current inflow of new patients plus all untreated patients from previous periods, $L_t = x_t + \Psi_{2,t-1} + \Psi_{3,t-1} + \Psi_{4,t-1} + ... + \Psi_{q,t-1} = x_t + \sum_{d=2}^{q} \Psi_{d,t-1}$ and letting the inflow of patients at t $x_t = \Psi_{1,t-1}$, we can write $L_t = \sum_{d=1}^{q} \Psi_{d,t-1}$.

results we solve the model numerically and thus have to select functional forms for the utility $g(k_{d,t})$ and $\chi_d(\delta_t)$, and the main parameters of the model.[7] Functional forms are selected such that assumptions 1-3 are met, and details are presented in Appendix B. For simplicity we set $q = 24$, as the maximum duration possible being 24 months. In Dimakou et al. (2014) we explore how a similar model accounts for different hospital management practices and waiting time distributions when various functional forms and parameters are selected. The model there also allows for prioritisation by patient's level of severity. Here our focus is on analysing the effects of waiting time targets and thus we work with a benchmark specification that provides the best general fit for waiting time distributions prior to the government policy introduction.[8] The main theoretical implications stressed below, which we test in our empirical exercise, are generally robust to parameter changes.

In order to understand the effects of waiting targets we solve two versions of our model, (i) the *benchmark* model setting $\hat{d} = 24$, thus in this version there is no government intervention (no target imposed), and (ii) the *target* model in which we set $\hat{d} = 10 < d^*$, where d^* is the maximum duration observed in the benchmark model, thus the target binds. We perform this pre- and post- target analysis for two parameter specifications, one where c_δ is relatively small, thus managers may increase resource utilisation at lower utility cost, and one where c_δ is relatively high, thus it is very costly for managers to increase capacity, holding beds and resources constant. Recall that the size of this parameter does not solely reflect managerial ability; a high value of c_δ could be the result of the hospital already operating with a fully efficient allocation of beds and resources.

Figure 1 shows the graphical representations of the survival functions (upper graph) and the hazard functions (lower graph) for the *benchmark* and *target* versions of the model when costs of improving treatment capacity are low. The survival curves start from one, as all patients are waiting to be treated at duration zero, and then decrease monotonically while the hospital removes patients off the list. The hazard curves increase and reach one when all patients are treated.

Before we focus on the effects of targets on treatment plans we provide a general discussion on the main mechanisms that shape survival and hazard curves. The mechanisms that drive the hospital's admission behaviour depend mainly on its utility, production structure and inflow interactions. The hospital would prefer to treat as many patients as possible immediately ('front-loading'), increasing benefits. However this comes at a higher cost, since it forces the hospital to allocate more bed-hours and resources for each patient. Moreover, hospitals refrain from treating too many patients up front since that would reduce expected waiting time, increasing the demand for health care in the future and generating longer waiting lists as hospitals are capacity constrained. Therefore, the hospital also takes into account the impact of its own behaviour in the future flow of patients. Finally, due to congestion (assumption 3) benefits also decrease when the number of treated patients of

[7] We obtain the solution by employing a constrained nonlinear optimisation routine in Matlab. Although it is fairly easy to determine the first and second order conditions of our maximisation problem, these involve many Kuhn-Tucker equations. Thus, it is easier to solve the optimisation problem directly instead of using the resulting system of equations.

[8] Differentiating patients by the severity of their case would not change qualitatively our results. However, we only consider prioritisation by duration here.

the same duration are too high, given incentive for hospitals to smooth treatment across duration.[9] As a result, hospitals treat most patients that are in the list for the first 3 months, and continue treating patients such that hazard rates are monotonically increasing for longer durations.

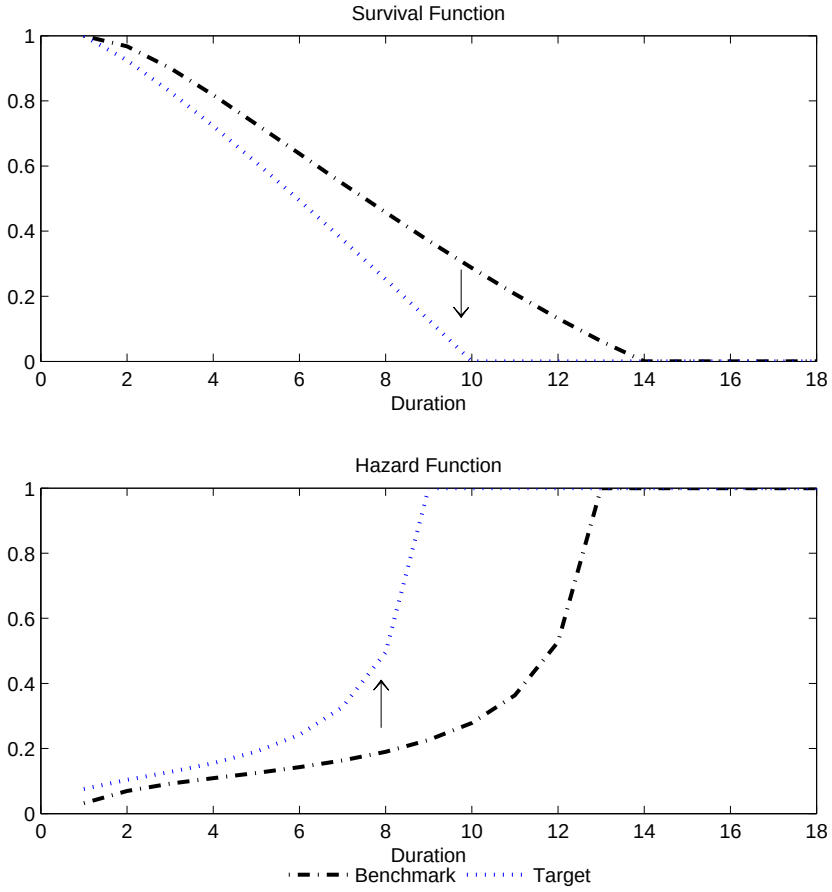


Figure 1: Effect of Waiting Targets - Low Costs (c_δ) - *Case 1*

As the survival curves (Figure 1) demonstrate the introduction of waiting time targets are successful in ensuring that no patient waits more than 10 months ($\hat{d} = 10$) and moreover, the probability of being treated at any duration conditional on still being in the list (hazard curve) also increase across all durations. The survival curve shifts leftwards, closer to the origin and the hazard curve shifts upwards, while becoming steeper. The introduction of the waiting time target not only benefits long waiters (i.e. patients that in the pre-target case were waiting for 10 to 14 months), since those now wait less for treatment but also shorter waiters, since now they are more likely to wait less. This symmetric improvement across all patients is possible since managers find it optimal to improve treatment capacity holding. With a re-organisation of beds/resources or through methods innovation, they manage to continue treating as many patients of low duration as possible - costs of investing in δ are

[9]This feature essentially prevents the hospital from treating the maximum amount of patients possible in the first period and leaving the remaining patients to be treat at the limit of 24 months duration to ensure expected duration is not too low - an extreme case of front loading.

small relative to the incentive to front-load. Finally, having assumed that government's sole care is about the benefits of treatment ($\sum_{d=1}^{q} g(k_d)$) and not managerial costs of improving capacity ($c_\delta \delta$), the introduction of targets also leads to higher benefits from healthcare provision.

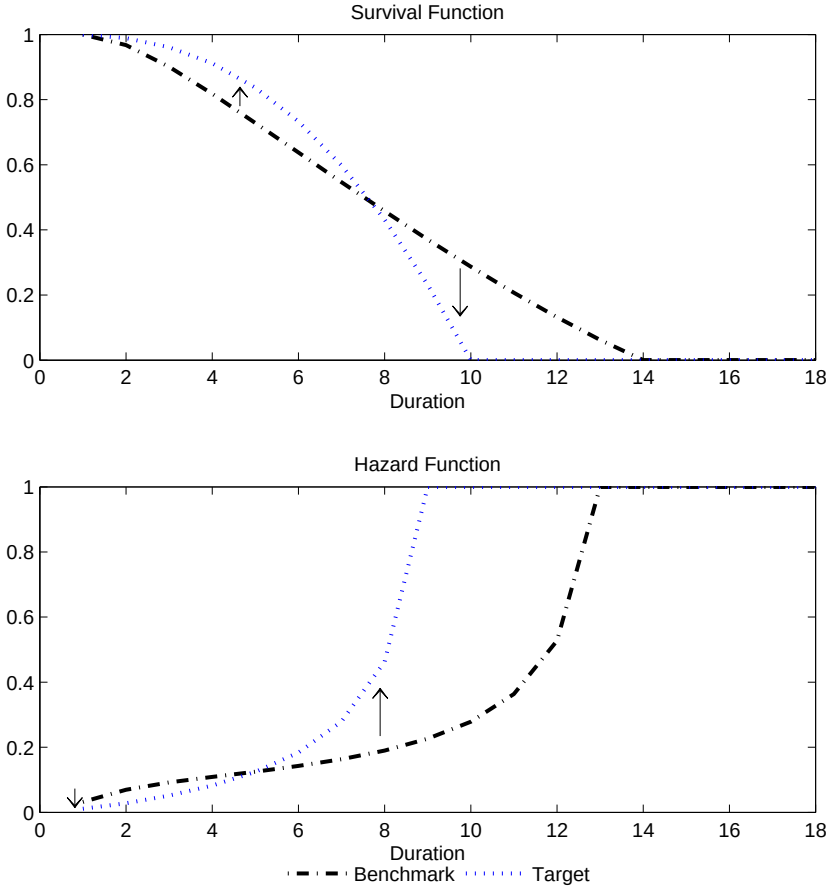


Figure 2: Effect of Waiting Targets - High Costs (c_δ) - *Case 2*

Figure 2 shows the graphical representations of the survival functions (upper graph) and the hazard functions (lower graph) for the benchmark and *target* versions of the model when costs of improving treatment capacity (c_δ) are high. Again, the waiting target is met, as no patient is treated with duration greater than $\hat{d} = 10$ but in this case the hospital manages to eliminate the long waiters (i.e. patients previously treated after the set target) by reducing the amount of very short waiters and at the same time increasing the amount of medium waiters (increased treatments in the periods prior to the target). This shift in treatment plans can be identified in two equivalent ways. First, while the hazard curve, again, becomes steeper (the probability of being treated before the target increases in both *Case 1* and *Case 2*), it now decreases for short durations; patients are less likely to be treated during the first four period of wait. Second, or equivalently, patients are more likely to survive in the list for longer since few are treated up front. Thus, the survival curve pivots up for short duration. In order to ensure all patients are treated before the target, the probability of treatment of middle duration patients (5 till 10) must increase.

Therefore, the model identifies an asymmetric effect of the targets on long and short waiters resulting from a 'manipulation' of waiting list that alters the prioritisation of treatment. The introduction of targets does not induce managers to improve treatment capacity, since the 'front loading' motive is weak relative to the disutility of ensuring a re-organisation of inputs that would yield more treatments, holding beds and resources constant. As a result this 'manipulation' of the waiting time distribution is necessary in order to reduce costs and keep the steady state expected duration and overall number of treatments controlled. Finally, we observe that due to the lack of improvement in capacity, the total benefit from treatment ($\sum_{d=1}^{q} g(k_d)$) is reduced. Thus, in this case while the policy intervention managed to ensure no patient waits too long to be treated, it did not increase the total benefits of healthcare provision after the introduction of targets.

3 Empirical Analysis

Based on our theoretical model, the introduction of waiting targets is effective in eliminating long-waiters but can generate a positive and symmetric effect across patients (all patients are treated relatively more quickly than before) or it can generate an asymmetric effect across patients, with long-waiters benefiting in detriment of short-waiters. The first aim of our empirical analysis is to explore whether these two responses, identified theoretically, are observed when waiting time targets are introduced. The second aim is to investigate whether this asymmetric response, when occurring, affects patients healthcare outcomes. Before detailing the empirical methodology used for each empirical exercise we briefly describe our dataset.

The HES is the database employed. This covers all NHS hospital patients treated in a given financial year in England and Wales, recording both the date the patient was placed on the waiting list of a specialist and the treatment date. The difference between the two serves as the measure of waiting time (or duration). We evaluate data on three specialties (general surgery, trauma and orthopaedics, and ophthalmology) consisting of more than 50% of patients waiting for elective surgery. The time coverage is nine years from 1997/98 until 2005/06. The majority of procedures are general surgery, followed by orthopaedics and ophthalmology, and there is a steady increase of admission numbers of all specialities through the years. After excluding trusts with missing data over the nine years from 1997/98 until 2005/06, a set of 52 hospitals remains.[10]

3.1 Empirical Waiting Time Distributions

Methodology

[10]Our sample includes all hospitals, identified by their NHS code in the HES, for which there is available data for all years. Some NHS code changes might occur due to mergers or other organisational reforms, thus for some NHS codes we have data for a subset of periods within the 9 years. Although it would be possible to construct a time series by linking different identifiers that relate to a known hospital, we exclude them since, depending on the degree of reform, comparisons of waiting distributions might be misleading.

We employ duration (also known as time-to-event or survival) analysis to obtain empirical representations of patients' waiting time patterns.[11] Duration analysis, by exploring conditional probabilities of treatment and the cumulative density function, is a robust and informative approach, allowing for an in-depth exploration and comparison of distinct admission behaviours. Following our theoretical model closely, we estimate survival and hazard functions using the non-parametric Kaplan-Meier or product limit estimator (Kaplan and Meier, 1958). Comparisons are then performed using both graphical techniques and log-rank statistical tests to ensure the survival curves obtained are statistically different.

The key general characteristics of survival curves that guide our analyses, using the terminology of Weon and Je (2012), are their variation in terms of 'shape' and 'scale'. Scale refers to changes in the position of the curve. Survival curves closer to the origin imply faster admission rates, since a smaller proportion of patients is left waiting on the list at each duration. Scale changes after the introduction of targets are therefore directly linked to a symmetric response in hospital admission behaviour. Shape refers to changes in the slope (size and sign of second derivative) of the survival curve. Shape changes after the introduction of targets are therefore indicative of potential asymmetric effect across patients.

Results

In 2000 the UK government introduced a national NHS inpatient waiting time target for all public hospitals of 18 months, which was gradually reduced by three months at every year from 2002 until 2005/2006 (see Table 2 below).[12] For most of our empirical analysis we focus on the effects of the policy on hospitals admission patterns, comparing waiting time distributions of the period prior to the policy to the ones observed in 2005/06 after all target changes have been introduced. Nonetheless, we start by presenting the evolution of admission patterns over the nine years (1997-2005) for Hammersmith hospital as it shares common characteristics with many of the other trusts.

Table 2: Waiting time targets - timeline

	Year				
	2000/01	2002/03	2003/04	2004/05	2005/06
Targets	18m(546 days)	15m(456 days)	12m(365 days)	9m(273 days)	6m(182 days)

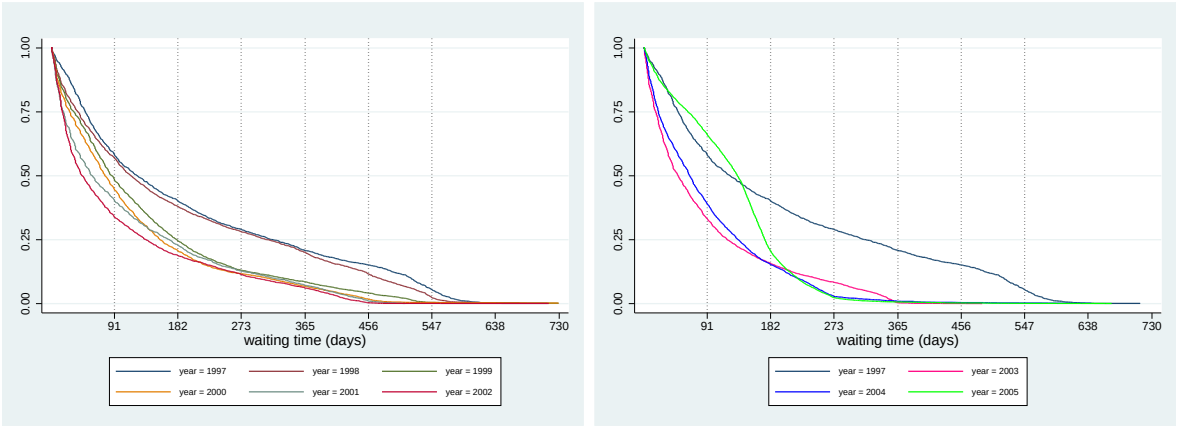
The trends in the survival curves of Hammersmith differ markedly across the nine year period (Figure 3). Admission rates for the first two years are quite slow, but improve gradually as time passes; the survival curves shift leftwards for the whole range of durations, implying a proportional decrease in the waiting time of all patients (scale/symmetric effect). It is also evident that much effort is devoted to reducing extremely long waiters; while 20% of patients had to wait more than a year in 1997 and 1998, from 2003 onwards there were no

[11] In our context, the 'event' of interest is admittance to hospital, 'survival' corresponds to remaining on the list, and 'time' is that between being placed on a waiting list until admitted for surgery.

[12] Together with the introduction of targets other policies increasing hospital capacity and promoting productivity gains were implemented. Nonetheless, our empirical results indicate that targets have been the key drivers of hospital's waiting lists management (hazard rates peak at the targets).

patients with that duration. During the introduction of the laxer targets, the hospital was able to improve treatment capacity and hence results match the theoretical prediction for *Case 1* (Fig. 3 (a)). However, in the last two years of our sample the Hammersmith hospital increased the waiting times of people with duration less than 6 months relative to previous years, particularly so in 2005/06. This is reflected by a rightwards pivot of the first (0-182 days) segment of the survival curve, together with a noticeable change in the curvature (significant shape effect). In fact, the results indicate that in 2005 the hospital treated relatively less patients during the first 4 months of wait comparing to years 1997 and 1998, before the targets were set (Fig. 3 (b)). As a result, facing stricter targets, the hospital is unable to increase treatment capacity further and hence manipulates the treatment plans, reducing the number of long-waiters in the detriment of short-waiters as predicted by *Case 2* of our theoretical framework.

Figure 3: Evolution of survival curves of Hammersmith from 1997 to 2005.

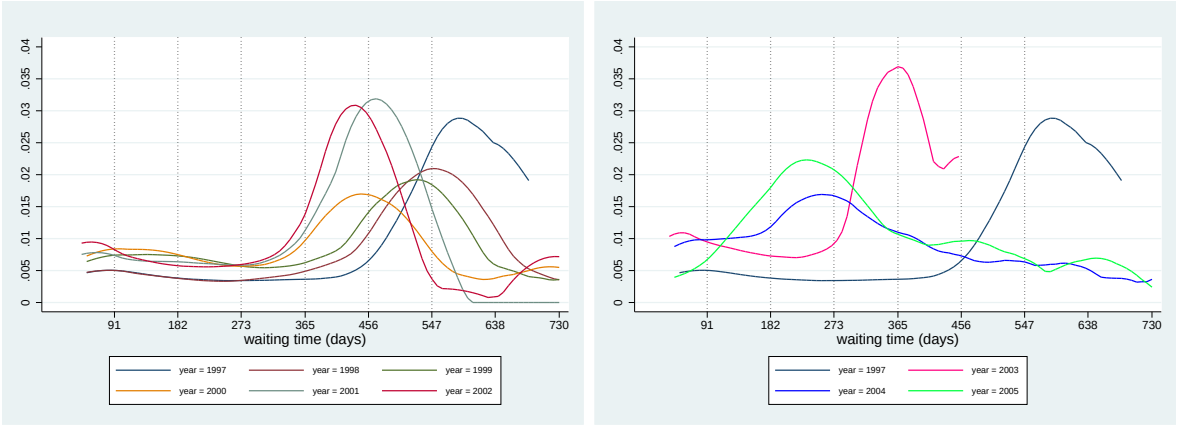


(a) (b)

Figure 4 reports the evolution of the hazard curves for the same hospital and time frame. A key difference between the theoretical hazard curves and their empirical counterpart is that the first always approach one as the list is eliminated, while this is not so for the second. The empirical curves increase sharply when the list gets substantially reduced but as a residual amount of patients may remain in the list for substantially long periods, we observe a fall after the sharp increase instead of a cumulative curve that approaches one. Nonetheless, as in the theoretical model those peaks indicate the point where survival curves approach zero. Note that we observe peaks even for the first three years (1997-99), where no national targets were rigourously imposed. Such a behaviour is consistent with hospitals' having their own 'internal' targets. With the introduction of decreasing national targets the peaks in the hazard rates move leftwards, broadly in line with these targets. Thus, it is evident that hospitals are consistently responding to this government policy. From 2000 until 2004 the universal targets have decreased from 18 to 9 months and the peak of the hazard curve occurred before or at the target in all years. However, in 2005, with the decrease of the target to six months, the peak occurred after the target, already indicating the hospital's difficulty in meeting a stricter target. Once again the estimated changes in waiting time distribution match well our theoretical predictions. In some cases hazard

curves move up, targets are effective in eliminating long waiters but have positive effects across all patients given that hazard rates shift up across all durations. However, for years 2004 and 2005 we observe that while the peak moves leftwards, the lower duration portion moves downwards, indicating list manipulation and asymmetric effects across patients.

Figure 4: Evolution of hazard curves of Hammersmith from 1997 to 2005.



(a) (b)

To sum up, hazard curves increase around the target duration for each of the years from 2000 until 2005, reflecting a natural shift of patients that are now treated before the target and consistent with our model predictions (for both *Case 1* and *Case 2*). However this 'bunching' in the distribution of waiting times when a threshold is introduced does not reflect whether 'manipulation' of waiting lists occurs or not. Rather, the indication of manipulation of waiting lists and suboptimal outcomes, suggesting hospitals alter clinical priorities due to the policy intervention, is reflected by the effects on the left-hand side of the distribution. An effective intervention increases the probability of treatment for short duration patients, while a suboptimal intervention leads to a decrease in the probability of treatment of short duration patients. We explore this heterogeneity of responses to analyse the effects of targets on health care outcomes of patients in section 3.2.

A brief look at the overall effect of the waiting targets introduced in the UK (comparing survival curves from 1997 and 2005) for all hospitals in our sample indicates the consistency of our findings. Figure 5 illustrate the results for a subset of 9 hospitals while the remaining 43 hospitals, for which comparison is possible, are presented in Appendix C. For 33% of the hospitals a symmetric effect across patients is observed (positive scale effect with survival curves shifting down, e.g. Bradford and St. George hospitals), matching *Case 1* of our theoretical framework, while for 58% of the hospitals a trade-off between short and long waiters emerges; the short-end of survival curves move up indicating short-waiter are worse off (e.g. Hamstead and Nuffield and South Manchester), in accordance with *Case 2* of our model.[13]

[13]Three hospitals in our sample exhibit negative scale effects, whereby all patients are waiting longer after the introduction of the targets.

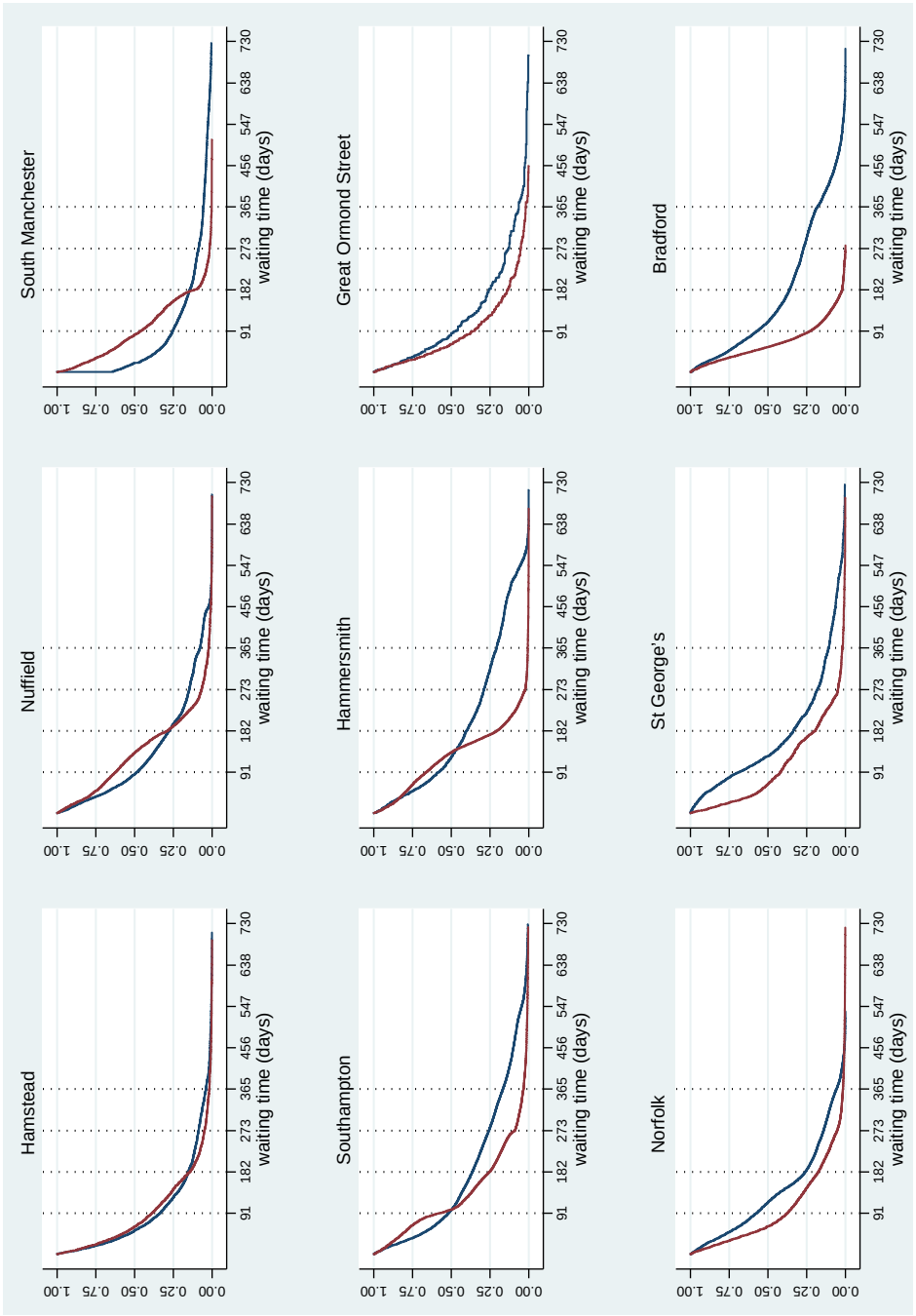


Figure 5: Patterns of survival curves in 1997 *(blue line)* and 2005 *(red line)*.

3.2 Impact of waiting time intervention on healthcare outcomes

Methodology

In this section we consider the effect of the introduction of waiting time targets on patients' healthcare outcomes. The first step to perform such analysis is to define how to identify and measure when targets are potentially altering the quality of patient treatment in a NHS hospital, aside from their role in reducing average waiting times. Our theoretical and empirical analysis suggests that the impact of targets is efficient and improves benefits of healthcare when short duration patients are better off, while potential clinical distortions might be present when the treatment of short duration patients is delayed relative to the pre-target waiting time distribution to accommodate for the faster treatment of long duration patients. We thus classify hospitals in our sample according to their response in the treatment of short duration patients. As such we set group 1 hospitals as the ones that responded to the target in an asymmetric manner (shape effect) and group 2 hospitals the ones that responded following a symmetric pattern across patients (scale effect). We then characterise patients according to each hospital they were treated. The benchmark classification we use is done comparing the waiting time distributions for 1997 (the first year in our sample) and 2005 (the last change in waiting targets - 6 months). We verify the robustness of our results by using waiting time distributions for 1999 (the year before the targets were imposed) and 2004, when targets were set at 9 months.

The second step is to define how to measure healthcare outcomes or the quality of patient treatment. One possibility is to use measures such the quality-adjusted life-year (QALY) for each healthcare intervention, however as in many studies that use HES, lack of QALY data at the patient level prevents us from doing so. The level of detail in our dataset allow us to construct three distinct measures of outcomes. The first, and generally accepted measure, is patient *mortality* rate. Our dataset only includes death in hospital, or before discharge, while mortality until 30 days after the discharge is also commonly used.[14] As a result our first measure is quite restrictive and as such in our second measure we include both mortality and discharge to a general ward/ another hospital, which we named *prolonged care or mortality*. Finally, we select a third measure of health care outcome that explores the period from admission until discharge, identifying patients with above average length of stay (*delayed discharge*).

The third and final step is to define the empirical model. We employ a cross section logit regression model at the patient level to measure the impact of the targets on a set of patient outcomes controlling for patient and other hospital characteristics. More precisely we estimate the following specification

$$o_i = \beta_0 + \beta_1 shape_i + \beta_2 X_{1i} + \beta_3 X_{2h} + u_i \tag{4}$$

where i refers to patients and h to hospitals. $o_i \in \{$*mortality*, *prolonged care or mortality*,

[14] In recent years the HES databased provide the 30 day mortality variable by using death records in the UK from the Office of National Statistics (ONS). However, ONS data was not available to us for our sample period.

delayed dischrage} refers to the measure of healthcare outcomes, $shape_i$ takes the value of one for patients treated in hospitals in group 1 (asymmetric response) and zero for patients treated in hospitals in group 2 (symmetric response). X_{1i} includes the following set of patient and episode level controls; age, sex, index of multiple deprivation, main specialty of episode and number of complications/comorbidities per case, while X_{2h} refers to hospital level controls, namely type and size of hospital, hospital resources (total number of employees, beds), total capital investment and total number of complaints regarding facilities. The analysis is performed at the cross-sectional level with outcome and control variables evaluated at the year of study (2005). Table 3 describes all variables and sources.

Table 3: Variables in estimation results

Variable	Description
Patient outcome variables (source: HES)	
mortality	1 if patient died in hospital; 0 otherwise
prolonged care/ mortality	1 if patient was discharged to another hospital facility (general ward) or another hospital, or if patient died; 0 otherwise
delayed discharge	1 if patient stayed above average from admission for surgery until discharge. Evaluated at speciality level.
prolonged care	same as *prolonged care/ mortality*, but patients that died are excluded (missing values). Used for robustness
overly delayed discharge	1 if patient stayed one standard deviation above the average by specialty stay until discharge. Used for robustness
Main independent variable (source: HES)	
shape	1 if hospital responded in an asymmetric way to waiting time targets (shape effect); 0 if scale (hospital reduced waiting times throughout the scale). Comparison made between: 1997-2005; 1999-2005; 1997-2004; 1999-2004. Cases identified as negative scale (treating all patients slower) are included in the shape effect.
Patient level control variables (source: HES)	
groupage	each patient classified in one of 11 ten-year groups
sex	1 if female
imd_j	index of multiple deprivation: $j \in \{hd, i\}$, for health and disability and income components respectively.
ortho	1 if main specialty of the episode is orthopedics; 0 if ophthalmology or general surgery
ophtha	1 if main specialty of the episode is ophthalmology; 0 otherwise
CC_n	set of categorical variables showing the number of patient diagnoses (primary and three secondary) classified under complications and comorbidities in HRGv4. CC_1= 1 if patient has one CC amongst his/her diagnosis; CC_2 =1 if patient has two CCs, and so on
Hospital level control variables (source: HEFS)	
large, medium, small, teaching, specialist	a set of categorical variables for large, medium, small acute, teaching and specialist types of hospitals. Takes value 1 for each type; 0 otherwise
beds	available number of beds
totempl	total (medical and not) number of employees (in WTE)
totcapinv	total capital investment (in £)
complaints	total number of complaints about facilities

All data are from HSCIC: HES, hospital episode statistics; HEFS, hospital estates and facilities statistics. The assignment of CCs to patient episodes in done using the methodology of HRGv4, as outlined in http://www.hscic.gov.uk/article/2322/HRG4-200708-Reference-Costs-Grouper-Documentation. Hospital level variables from HEFS are at the trust level, since not available at hospital (site) level.

As discussed in the introduction Propper et al. (2010) also assess the impact of the targets policy on patients outcomes, employing a different methodology than ours and finding no systematic differences in outcomes. One fundamental difference stems from the identifica-

tion of the targets effect on hospitals (first step discussed above). This is done by employing census data and looking at the 'bunching effect', that is, the percentage of patients at the end of each quarter that risk breaching the target if not treated in the following quarter. We, on the other hand, define a measure that focuses on the short-side of the waiting distribution (of HES data). Our theoretical and duration analysis confirms that 'bunching' around the time of the target (humps in hazard curves) is a common phenomenon, hence what differentiates hospitals' responses is not whether (previously) long-waiters are treated at the margin of the target limit, but whether this is done at the expense of short-waiters that now have to wait longer. Our measure, thus, provides for a potentially better and simpler identification of hospitals' heterogeneity affecting patient's i treatment.[15]

Results

Results, presented in Table 4, are obtained employing our logit estimation and are depicted in odds ratios.[16] Our findings provide supportive evidence for a systematic difference in outcomes among patients treated in hospitals that responded asymmetrically to targets. For the first outcome measure (column *(1)*), the impact is insignificant but this can be largely attributed to the fact that patient mortality in hospital is of extremely low frequency (about 0.12% of patients in 2005 died while in hospital). For the second outcome measure (column *(2)*) our findings are stronger. We confirm that for those patients treated in hospitals with an asymmetric response to targets the odds of dying in hospital or being discharged to other facility/hospital for continued care are almost 1.5 times more relative to the odds for patients treated in hospitals that reduced waiting times throughout the distribution.

The majority of control variables at patient level are also significant. A unit increase in age group (i.e. a decade increase) more than doubles the odds of the worse outcome occurring, while health and disability deprivation is also increasing the odds. Being a female patient also seems to increase the odds of prolonged care, while they decrease for mortality. The odds of the worse outcome occurring are also increasing in the number of complications in each episode.[17] The type of hospital is also shown to be a significant factor, with large acute hospitals having higher odds relative to all others. All other characteristics of hospitals do not seem to be altering the odds on patient outcomes.

The last four columns depict results for the third patient outcome. The odds of staying

[15] Also note that Propper et al. (2010) use quarterly frequency of a shorter time span (2000/01-2004/05), employing a panel data analysis and using the ONS-HES linked mortality data. At the hospital level, they also investigate different types of list manipulation, such as whether the type of admission for elective surgery has changed (from waiting list to booked or planned, with the later not included in the target policy). Our focus is only on waiting list admissions.

[16] Estimated parameters have been adjusted such that a parameter of 1 indicates the variable of interest does not affect the oddsratio of observing the outcome.

[17] It could be argued that patients' episodes with complications could be the outcome of prolonged wait prior to admission. That is, complications may be the consequence of prolonged waiting times, and this outcome could be worsened in those hospitals that substitute among long and short waiters. However, results whereby 'patient with complications' is treated as the dependent variable actually show that the odds of developing complications among the hospitals with scale admissions patterns are higher, thus validating our choice of using CCs as a control variable.

Table 4: Estimation Results

outcome:	*(1)* mortality	*(2)* prolonged care or mortality	*(3)* delayed discharge	*(4)*	*(5)*	*(6)*
shape	1.036	1.471	1.241	1.452	0.968	1.760
	(0.24)	(7.36)***	(16.02)***	(16.08)***	(-1.89)*	(10.78)***
groupage	2.597	2.120	1.244	1.344	1.494	0.917
	(16.72)***	(40.50)***	(71.87)***	(51.09)***	(89.07)***	(-8.59)***
sex (female)	0.638	1.225	1.208	1.408	1.099	1.149
	(-3.44)***	(4.39)***	(16.95)***	(17.88)***	(6.16)***	(3.16)***
imd	1.194	1.203	1.184	1.993	0.692	3.294
	(2.13)**	(6.38)***	(3.12)***	(7.32)***	(-5.00)***	(6.20)***
ortho	0.271	3.417				
	(-8.41)***	(21.37)***				
ophtha	—	0.062				
		(-13.11)***				
No. of complications (CCs):						
CC_1		1.642	1.873	1.493	1.428	2.440
		(9.25)***	(51.93)***	(18.31)***	(22.19)***	(17.52)***
CC_2		2.755	2.610	2.116	1.790	4.790
		(15.22)***	(52.41)***	(26.53)***	(21.25)***	(15.55)***
CC_3		5.866	3.809	2.976	2.755	8.366
		(19.90)***	(39.50)***	(24.51)***	(16.89)***	(6.50)***
CC_4		11.201	5.714	4.204	4.813	—
		(16.85)***	(21.95)***	(15.41)***	(9.21)***	
medium		0.775	0.619	0.695	0.907	0.601
		(-3.42)***	(-20.47)***	(-9.44)***	(-3.00)***	(-4.17)***
small		0.228	0.738	0.788	0.952	1.263
		(-12.81)***	(-10.02)***	(-4.59)***	(-1.21)	(1.54)
teaching		0.397	0.858	0.790	0.762	0.142
		(-9.41)***	(-6.83)***	(-6.86)***	(-8.44)***	(-12.96)***
specialist		0.277	1.230	—	1.241	10.471
		(-9.47)***	(4.97)***		(4.26)***	(10.18)***
complaints		0.995	1.000			
		(-8.71)***	(0.31)			
totempl	1.000	0.998	1.000	1.000	1.000	1.000
	(0.96)	(-9.34)***	(-8.02)***	(-1.27)	(2.43)**	(5.50)***
totcapinv	1.000	1.000	1.000	1.000	1.000	1.000
	(-0.19)	(1.72)*	(7.55)***	(3.70)***	(3.87)***	(-0.41)
Obs.	165,118	193,442	193,986	67,475	94,081	35,552

Note: t-ratios (based on robust standard errors) in parentheses; * = 10%, ** = 5%, and *** = . 1% levels of significance. In column *(2)* hospital size/resources is measured by number of beds; in all others by total number of employees. *imd* is measured by the health and disability domain in columns *(1)-(2)* and by the income one in columns *(3)-(6)*. A dash line indicates that the outcome measure was not identified in a particular category (e.g. no patient died from ophthalmology in col. *(1)*

in hospital longer than average until discharge are 1.25 times higher for patients treated in hospitals that exhibit an asymmetric response than patients treated in hospitals with a scale effect. In the last three columns of the table, we disentangle the effects across the specialties at which the average stay has been evaluated. Among general surgery, the odds of *delayed discharge* are 1.45 times more for patients treated at 'shape' hospitals than for those at 'scale' hospitals, and they increase to 1.8 for ophthalmology. For orthopedics, the odds are reversed with prolonged stay somewhat more likely among patients at 'scale' hospitals, although the results are barely significant. As before, the effect of control variables stays unchanged.

Summarising, our results indicate that patients treated in hospitals in which the in-

troduction of targets generated 'manipulation' of lists to the detriment of short duration patients are more likely to observe a worse outcome after the policy intervention. This includes facing a higher mortality rate, and particularly, the likelihood of being in need of prolonged healthcare after discharge or facing a prolonged stay after admission. We perform a number of robustness exercises and selected results are depicted in Table 7 in Appendix D. First, we employ different year comparisons to construct the $shape_i$ variable (1999) and a different year for assessing patient outcomes (2004). Second, we deploy two different proxies for healthcare outcomes, namely prolonged care (excluding mortality) and delayed stay (for those that wait much longer until discharge). Findings are qualitatively unchanged.[18]

The *shape* variable used in our regressions has been constructed based on the identification of the heterogenous effects of targets on the empirical waiting time distributions. However, the differences across hospitals that based our classification, in line with a possible interpretation of our theoretical model, might be the result of hospital's management quality, which are potentially related to quality of healthcare. As such, our classification identifies the effect of targets at the hospital level but may also be related to the differences in quality or performance that existed before the introduction of the waiting targets. In order to explore to which extent this is the case we perform a 'placebo' (falsification) exercise whereby we attempt to investigate whether our *shape* measure would also affect patients' outcomes for the years prior to 2000. In Table 5 below we show estimation results for the coefficient and significance of the *shape* variable (all other control variables, although not reported, are the same as in Table 4).

Table 5: Results for years prior to the targets

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
	mortality		prolonged care or mortality		delayed discharge	
	1997	1999	1997	1999	1997	1999
shape	0.948	1.189	0.870	0.936	1.145	1.299
	(-0.47)	(1.48)	(-3.39)***	(-1.65)*	(11.88)***	(24.61)***
Obs.	234163	268543	234065	268581	233132	267200

For both years prior to the implementation of targets, the impact of being treated in a hospital that later displayed shape effects plays no role for mortality, while for the second outcome we find that the odds of prolonged care/ mortality are lower among the patients in hospitals that later exhibit asymmetric reductions to waiting times. This result indicates that with the policy intervention these hospitals, due to changes in treatment plans, have delivered relatively worse outcomes. Finally, for *delayed discharge*, we see that while in 1997 the effect is significant, albeit lower than in 2005, this is not the case for 1999, suggesting that, even from before the targets, in hospitals that asymmetric effects of targets were later observed, patients were more likely to face prolonged stays. However, given that in 1997 this effect was smaller and that in 1999 some form of anticipation of the introduction of targets

[18]We also estimate all specifications of Table 4 using robust standard errors clustered at the hospital level. The *shape* variable becomes insignificant only for column *(2)*. We found that when specialist hospitals are included clustering standard errors are significantly higher. As a result we exclude patients from these hospitals (patients from 41 large, medium, small and teaching hospitals remain in the sample), also dropping the control for specialty from specification *(2)*, and find that the *shape* variable is again significant. We conclude that our main results are also robust to clustering standard errors at the hospital level (results are available from the authors upon request).

might have occurred, we cannot rule out that targets might also have affected patient's length of stay.

4 Conclusion

Waiting time targets have been widely used to improve the provision of healthcare in several OECD countries. Many contributions to the literature have analysed the effectiveness of such policy interventions, particularly looking at their effect on average (and excessive) waiting times and concluded that targets have been successful in reducing patients waiting times. Using data from the UK National Health System we also analyse the impact of waiting targets on healthcare provision, but as opposed to other studies, we focus on the entire distribution of patients waiting times, being able to identify different effects across patients.

We show that the great majority of hospitals responded to the targets policy by setting treatment plans such that no patient is treated after the target's limit, and thus average waiting time is reduced. However, this observation does not necessarily indicate that the policy intervention increased the benefits of healthcare provision. While a proportion of hospitals alter treatment plans such that all patients are treated fasters and hence healthcare provision improves, the majority of hospitals manage to eliminate patients with duration greater than the target limit, by decreasing treatment of patients who have just entered on the waiting list. As such a subset of patients that were treated quickly before the policy intervention are now worse off. This is particularly so as the targets tighten.

We then explore this heterogeneity across hospitals to identify whether waiting time targets have altered a set of outcomes of healthcare interventions. We find that in hospitals in which targets have produced trade-off across patients, indicating 'manipulation' of waiting lists and potential clinical distortions, quality of healthcare provision become systematically worse relative to hospitals where the prioritisation of short duration patients has remained the same. Delayed treatment of short-waiters is associated with increased likelihood of patient mortality or greater chance of need for prolonged healthcare. It also increases the probability the the length of stay from admission until discharge gets higher. As a result, although we find evidence that waiting time targets corrected the problem of large waiting times in elective surgery in the UK, we also find empirical support for the hypothesis that targets also generated a negative effect on patient welfare.

References

Besley, T., Hall, J., Preston, I., 1999. The demand for private health insurance: Do waiting lists matter? Journal of Public Economics 72(2), 155–181.

Cullis, J. G., Jones, P. R., Propper, C., 2000. Waiting lists and medical treatment: Analysis and policies. In: Culyer, A. J., Newhouse, J. P. (Eds.), Handbook of Health Economics, volume 1, chapter 23, pp. 1201–1249, Elsevier, Amsterdam.

Dimakou, S., Dimakou, O., Basso, H., 2014. Waiting time distribution in public health care: Empirics and theory. Mimeo, SOAS - University of London.

Dimakou, S., Parkin, D., Devlin, N., Appleby, J., 2009. Identifying the impact of government targets on waiting times in the NHS. Health Care Management Science 12(1), 1–10.

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.

Ellis, R. P., McGuire, T. G., 1986. Provider behavior under prospective reimbursement: Cost sharing and supply. Journal of Health Economics 5(2), 129–151.

Goddard, J., Malek, M., Tavakoli, M., 1995. An economic model of the market for hospital treatment for non-urgent conditions. Health Economics 4(1), 41–55.

Gravelle, H., Dusheiko, M., Sutton, M., 2002. The demand for elective surgery in a public system: Time and money prices in the UK national health service. Journal of Health Economics 21(3), 423–449.

Iversen, T., 1993. A theory of hospital waiting lists. Journal of Health Economics 12(1), 55–71.

Iversen, T., 1997. The effect of a private sector on the waiting time in a National Health Service. Journal of Health Economics 16(4), 381–396.

Kaplan, E., Meier, P., 1958. Nonparametric estimation from incomplete observations. Journal of the American statistical association 53(282), 457–481.

Levy, A., Sobolev, B., Hayden, R., Kiely, M., FitzGerald, J., Schechter, M., 2005. Time on wait lists for coronary bypass surgery in British Columbia, Canada, 1991 - 2000. BMC Health Services Research 5(22).

MacCormick, A., Parry, B., 2003. Waiting time thresholds: Are they appropriate? ANZ Journal of Surgery 73(11), 926–928.

Martin, S., Smith, P., 1999. Rationing by waiting lists: An empirical investigation. Journal of Public Economics 71(1), 141–164.

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.

Siciliani, L., 2006. A dynamic model of supply of elective surgery in the presence of waiting times and waiting lists. Journal of Health Economics 25(5), 891–907.

Siciliani, L., Hurst, J., 2005. Tackling excessive waiting times for elective surgery: a comparative analysis of policies in 12 OECD countries. Health policy 72(2), 201–215.

Siciliani, L., Stanciole, A., Jacobs, R., 2009. Do waiting times reduce hospital costs? Journal of Health Economics 28(4), 771–780.

Weon, B., Je, J., 2012. Trends in scale and shape of survival curves. Scientific Reports 2.

A Steady state hospital's maximisation problem

Here we show in more detail the steady state hospital's maximisation problem.

$$\begin{aligned}
&\max_{\delta,\{k_d,b_d,r_d\}_d} E_0 \sum_{d=1}^{q} g(k_d) - c_\delta \delta \\
&\text{Subject to} \quad c_B \bar{B} + c_R \sum_d r_d + \tau \left(\sum_d r_d - \bar{R} \right)^2 + wt(k_d; \widehat{d}) \leq M \\
&\qquad \sum_d b_d \leq \bar{B},\ k_d = \chi_d(\delta) b_d^\alpha r_d^\beta \\
&\qquad \sum_d k_d = Z - \theta E(d) \\
&\qquad 0 \leq k_d \leq \Psi_d,\ \Psi_d = 0 \text{ for } d > q
\end{aligned}$$

Recall that $k = \sum_d k_d$, the steady state expected duration is defined as $E(d) = \sum_d d \frac{k_d}{k}$ and $\Psi_d = k - \sum_{h=1}^{d-1} k_h$. At the steady state the restrictions that $k_d \leq \Psi_d$ are satisfied as long as k_d is non-negative for all d. Thus, the Lagrange function reads:

$$\begin{aligned}
\max_{\delta,k,\{k_d,b_d,r_d\}_d} \mathfrak{L} = &\sum_d g(k_d) - c_\delta \delta + \lambda \left(M - c_B \bar{B} - c_R \sum_d r_d - \tau \left(\sum_d r_d - \bar{R} \right)^2 - wt(k_d; \widehat{d}) \right) \\
&+ \sum_d v_{d,s} k_d + + \mu \left(Z - \theta E(d) - k \right) + \nu \left(\bar{B} - \sum_d b_d \right) + \varsigma \left(-k_d + \chi_d(\delta) b_d^\alpha r_d^\beta \right)
\end{aligned} \tag{5}$$

where λ is the lagrangian multiplier of the hospital budget constraint, $v_{d,s}$ is the lagrange multiplier of the Kuhn-Tucker constraint $k_{d,s} \geq 0$, and μ is the multiplier for the condition that ensures that the steady state inflow and outflow are equal. ν relates to the constraint on stock of beds and ς to the treatment function constraint.

Solving the hospital's problem gives rise to $3(d) + 5$ Karush–Kuhn–Tucker (KKT) con-

ditions. For each k_h where $h = 1, 2, ...q$

$$\frac{\partial \mathfrak{L}}{\partial k_h} = \frac{\partial g(k_h)}{\partial k_h} + \frac{\partial wt(k_h; \widehat{d})}{\partial k_h} + v_h - \mu \left(\theta \frac{\partial E(d)}{\partial k_h} + 1 \right) = 0$$

$$\frac{\partial \mathfrak{L}}{\partial r_h} = -c_R - \tau \left(\sum_d r_d - \bar{R} \right) + \varsigma \left(\beta \frac{\chi_h b_h^\alpha r_h^\beta}{r_h} \right) = 0$$

$$\frac{\partial \mathfrak{L}}{\partial b_h} = -\nu + \varsigma \left(\alpha \frac{\chi_h b_h^\alpha r_h^\beta}{b_h} \right) = 0$$

$$\frac{\partial \mathfrak{L}}{\partial v_h} = k_h \geq 0,\ v_h \geq 0 \quad \text{and} \quad v_h k_h = 0$$

$$\frac{\partial \mathfrak{L}}{\partial \delta} = -c_\delta + \delta \left(\sum_d \frac{\partial \chi_d}{\partial \delta} b_d^\alpha r_d^\beta \right) = 0$$

$$\frac{\partial \mathfrak{L}}{\partial \lambda} = \left(M - c_B \bar{B} - c_R \sum_d r_d - \tau \left(\sum_d r_d - \bar{R} \right)^2 - wt(k_d; \widehat{d}) \right) \geq 0,\ \lambda \geq 0 \quad \text{and} \quad \lambda \frac{\partial \mathfrak{L}}{\partial \lambda} = 0$$

$$\frac{\partial \mathfrak{L}}{\partial \nu} = \bar{B} - \sum_d b_d \geq 0,\ \nu \geq 0 \quad \text{and} \quad \nu \frac{\partial \mathfrak{L}}{\partial \nu} = 0$$

$$\frac{\partial \mathfrak{L}}{\partial \mu} = Z - \theta E(d) - k = 0$$

$$\frac{\partial \mathfrak{L}}{\partial \varsigma} = -k_d + \chi_d b_d^\alpha r_d^\beta = 0$$

From this we can derive the optimal number of patients treated after having waited d durations as a function of all the structural parameters (denoted $\mathfrak{z}$) of the model, $\forall \{d\} \quad k_d^* = k_d^*(\mathfrak{z})$.

B Functional and Parameter Specification - Basic benchmark Case

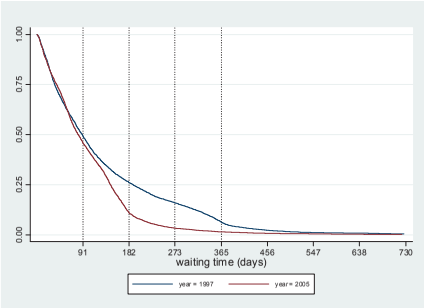
Table 6 shows the parameters values used in the numerical solution of the model.

C Effects of Waiting Time Targets: Full Sample

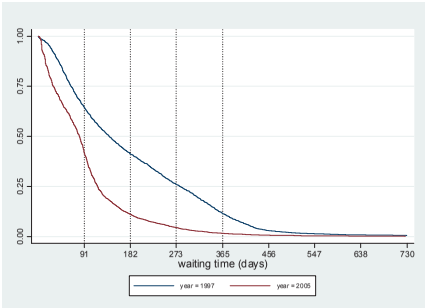
Here we show the comparison of survival curves of 1998/99 and 2005/06 for all 52 hospitals in our sample (Figures A1 to A6). For 30 hospitals we observe asymmetric effects, for 17, positive scale effects (shift inwards), for 4, negative scale effects (shift outwards) and for 1 hospital the survival curves remain largely unchanged.

Table 6: Benchmark functional specifications and parameter values

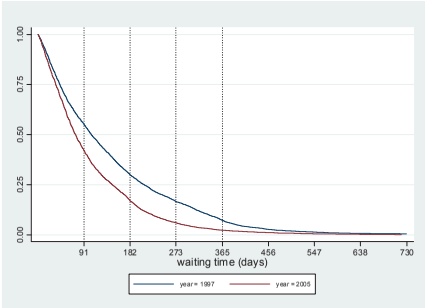
$g(k_d) = a_d k_d^3 + b_d k_d^2 + c_d k_d$	Utility from treating k patients with duration d
where $a_d = -0.0002 + \frac{0.0001}{d}$ $b_d = 0.02 - \frac{0.01}{d}$ $c_d = 2 + \frac{5}{d}$	parameters of the cubic utility function
$\alpha = 0.5$, $\beta = 0.1$, $\chi_d(\delta) = 0.1 + \log((d+1)^2) + \delta$	Production Function Terms
$c_\delta = 10000$	High Cost Case
$c_\delta = 500$	Low Cost Case
$c_B = 1$	Monetary cost of maintaining beds
$c_R = 1$	Monetary cost of operating resources
$\tau = 0.33$	Additional monetary cost of operating resources above limit
$M = 3000$	Hospital's budget
$\bar{B} = 1000$	Hospital's beds
$\bar{R} = 1000$	Hospital's resources
$Z = 1000$	Potential demand for healthcare
$\theta = 50$	Sensitivity of inflow to expected waiting time
q=24	Maximum allowed waiting time



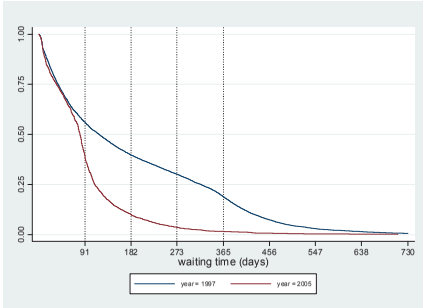
(a) Frimley Park Hospital



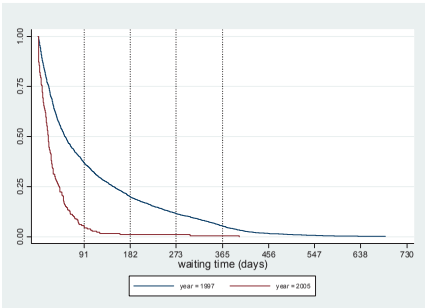
(b) Bromley Hospitals NHS Trust



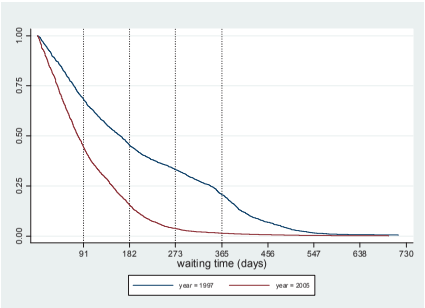
(c) Peterborough Hospitals NHS Trust



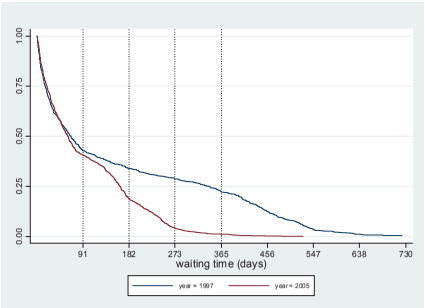
(d) Countess of Chester Hospital NHS Trust



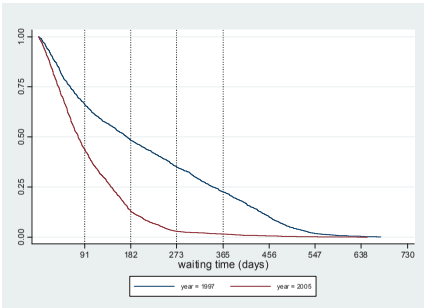
(e) East Somerset NHS Trust



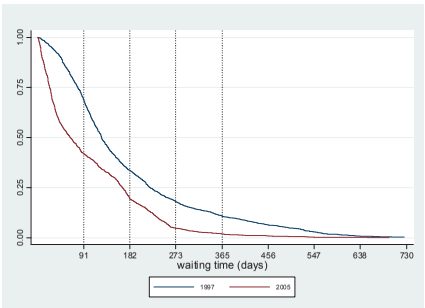
(f) Bedford Hospitals NHS Trust



(g) Ealing Hospital NHS Trust

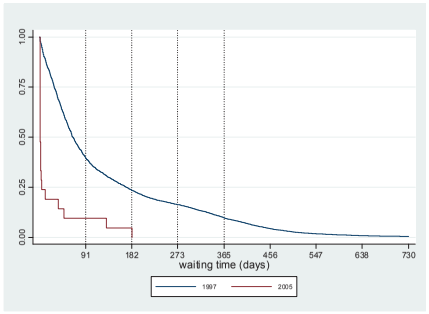


(h) Milton Keynes General Hospital

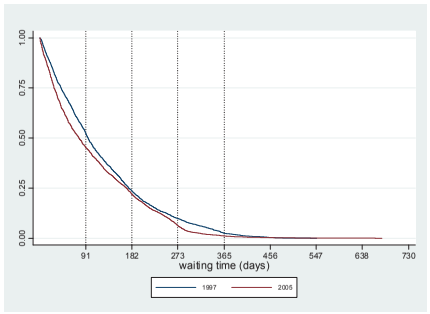


(i) St George's healthcare NHS Trust

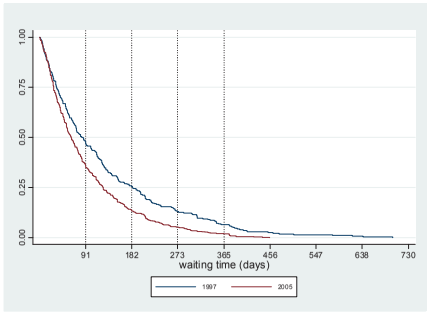
Figure 6: Survival Curves - 1997/98 vs 2005/06 - Part 1



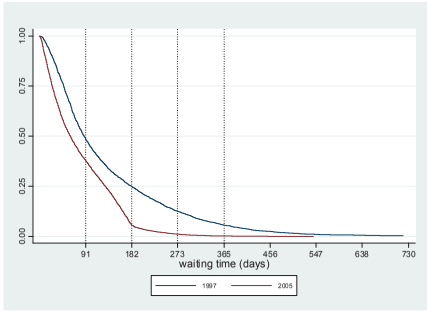
(a) The United Bristol healthcare



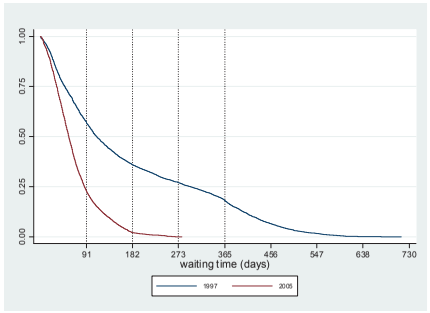
(b) The Royal National Orthopaedic



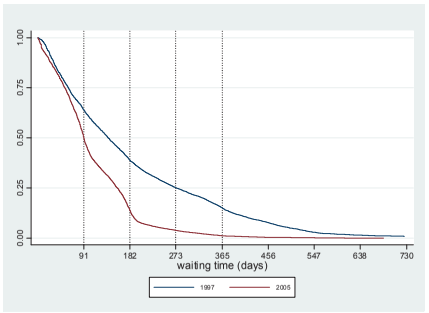
(c) The Great Ormond Street Hospital



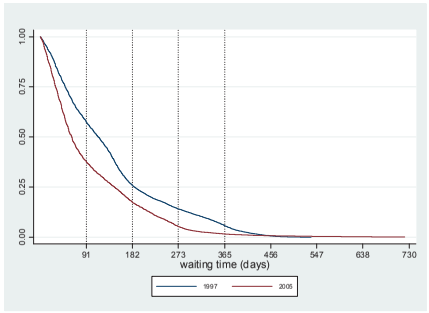
(d) Royal Orthopaedic Hospital NHS Trust



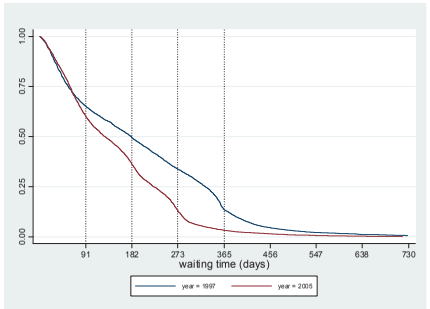
(e) Bradford Hospitals NHS Trust



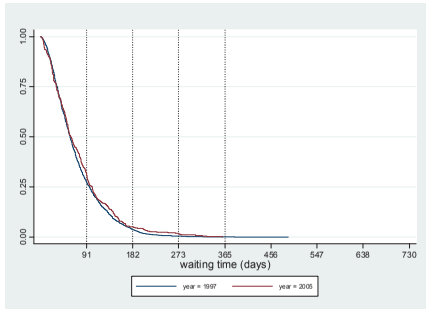
(f) Royal Berkshire Battle Hospital



(g) Norfolk Norwich Healthcare

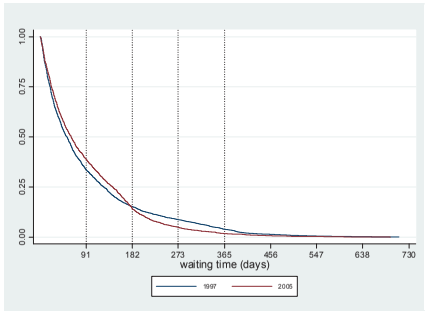


(h) Royal United Hospital Bath NHS

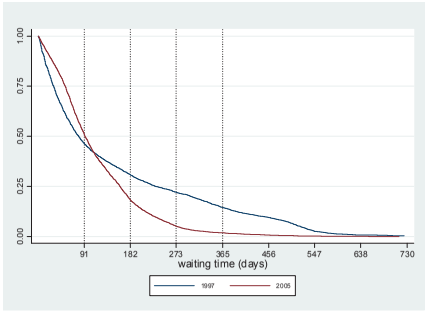


(i) Poole Hospitals NHS Trust

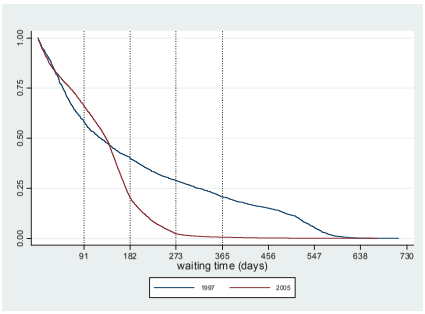
Figure 7: Survival Curves - 1997/98 vs 2005/06 - Part 2



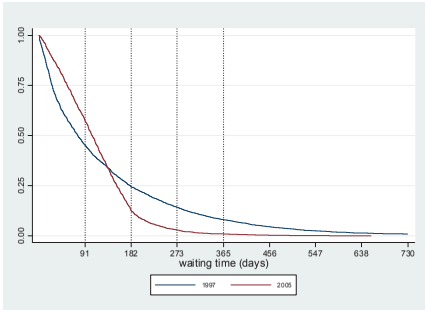
(a) The Royal Free Hampstead Hospital



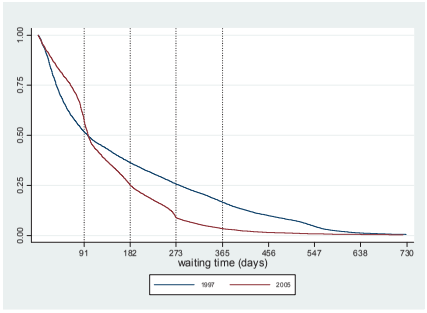
(b) Guy's & St Thomas NHS Trust



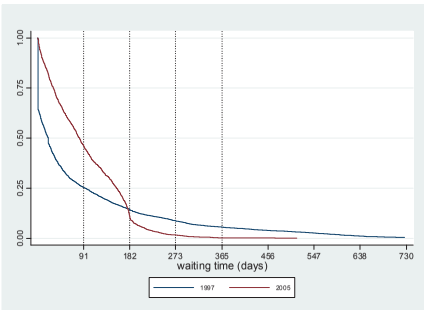
(c) The Hammersmith hospitals



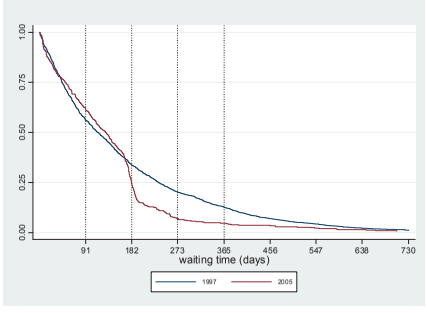
(d) Nottingham City Hospital NHS Trust



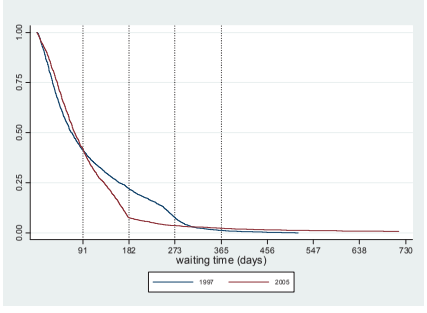
(e) Southampton University Hospital



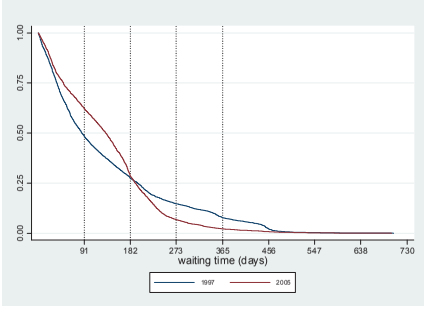
(f) SouthManchester University Hospitals



(g) Salford Royal Hospitals NHS Trust

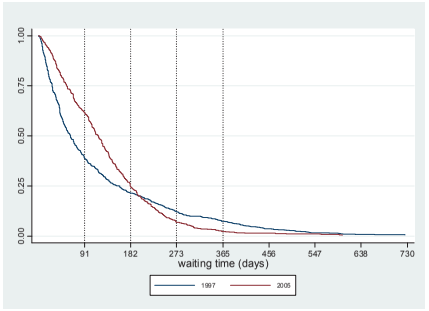


(h) University Hospital Birmingham

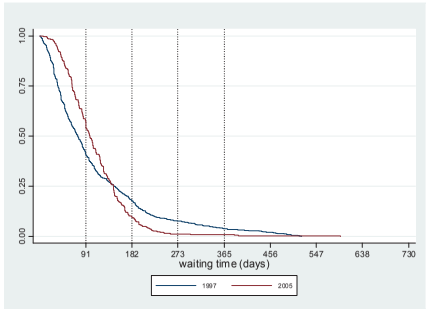


(i) Nuffield Orthopaedic NHS Trust

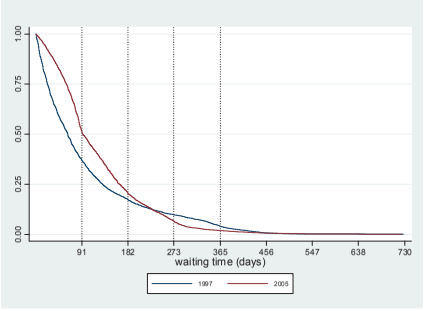
Figure 8: Survival Curves - 1997/98 vs 2005/06 - Part 3



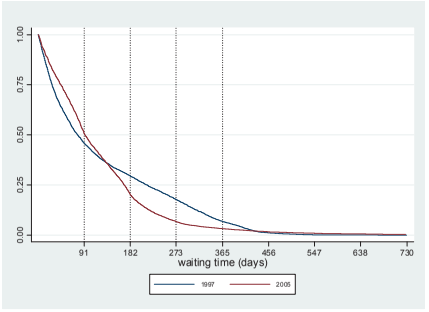
(a) Sheffield Childrens Hospital



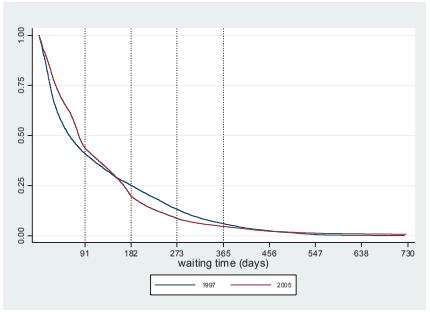
(b) The Queen Victoria Hospital NHS Trust



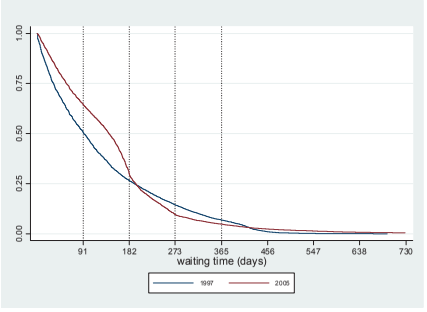
(c) Wirral Hospital NHS Trust



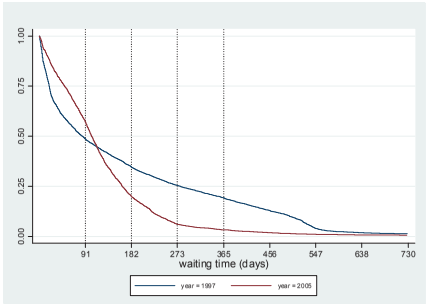
(d) Royal Cornwall hospitals



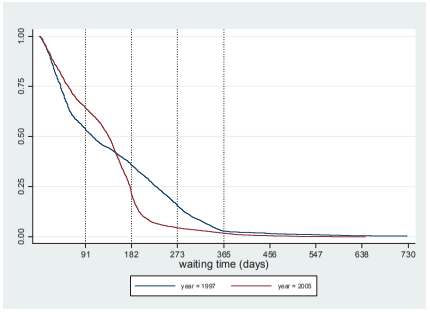
(e) Royal Devon & Exeter Healthcare



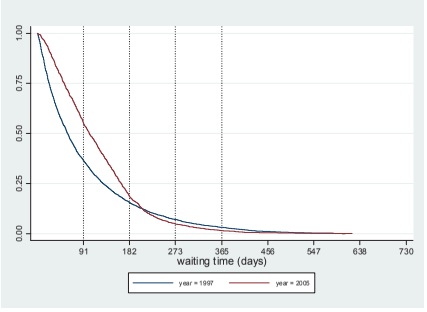
(f) Plymouth Hospitals NHS Trust



(g) Royal Surrey County Hospital NHS Trust

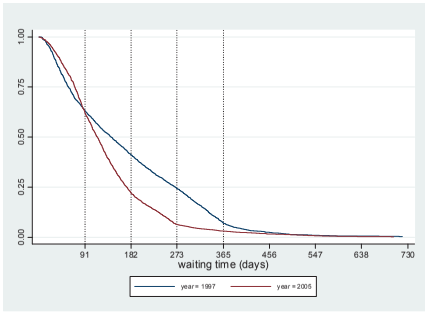


(h) Walsall Hospitals NHS Trust

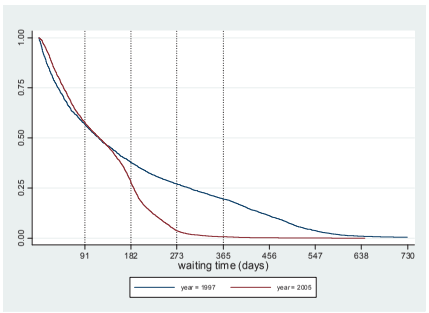


(i) Basildon & Thurrock General Hospital

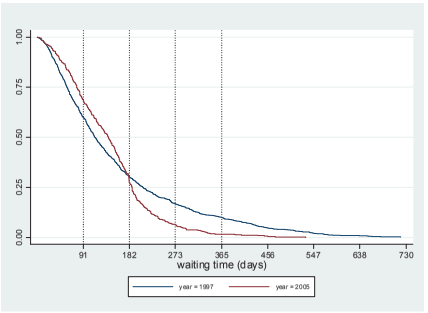
Figure 9: Survival Curves - 1997/98 vs 2005/06 - Part 4



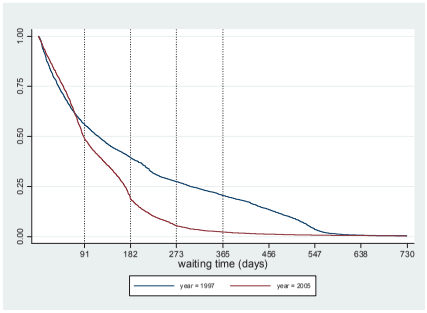
(a) James Paget Hospital NHS Trust



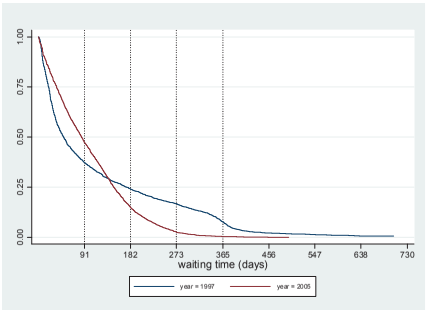
(b) Swindon & Marlborough NHS Trust



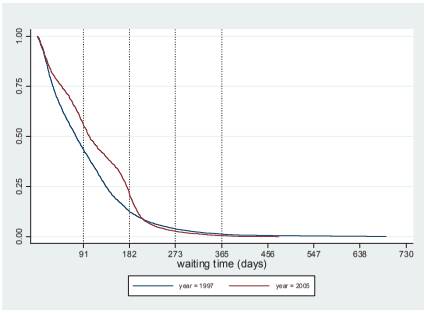
(c) Newham Healthcare NHS Trust



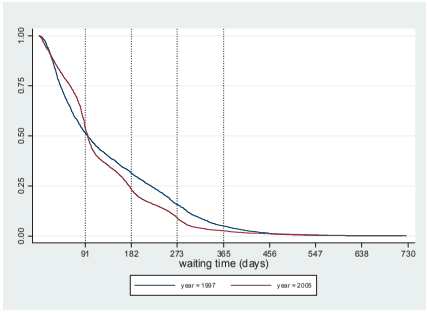
(d) Worthing Southlands Hospital



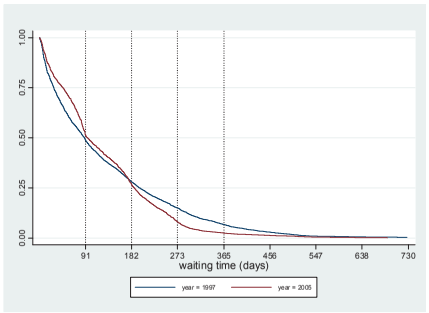
(e) Weston Area Health NHS Trust



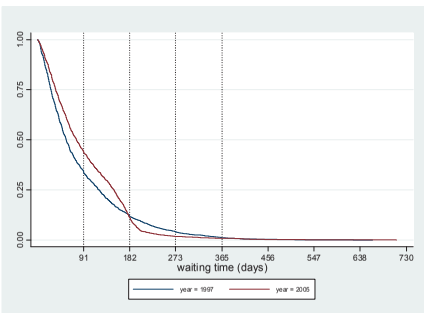
(f) West Dorset General Hospitals



(g) Burton Hospitals NHS Trust

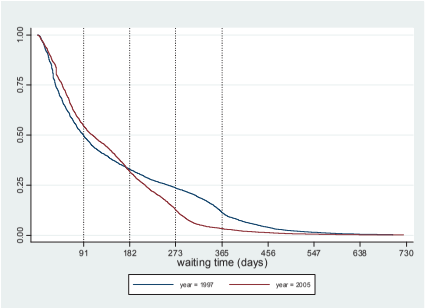


(h) Hereford Hospitals NHS Trust

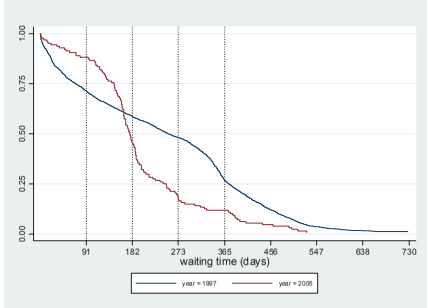


(i) Tameside & Glossop Acute Services

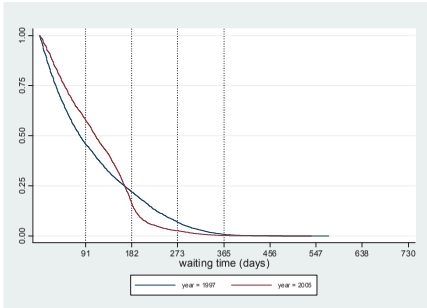
Figure 10: Survival Curves - 1997/98 vs 2005/06 - Part 5



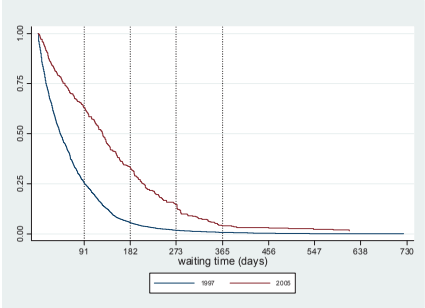
(a) Kettering General Hospital NHS Trust



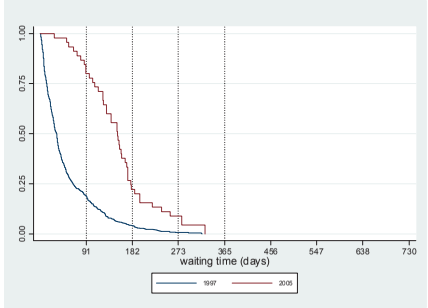
(b) The Royal West Sussex NHS Trust



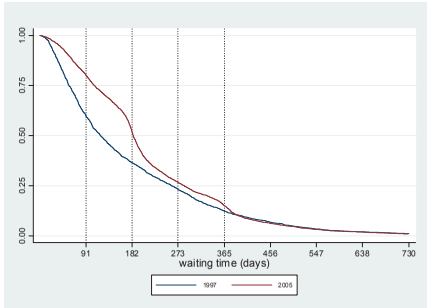
(c) Birmingham Heartlands & Solihull



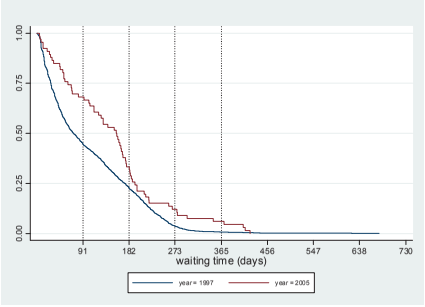
(d) Chelsea & Westminster healthcare



(e) Royal Liverpool Childrens NHS Trust



(f) Robert Jones & Agnes Hunt Orthopaedic Hospital



(g) South Warwickshire General Hospital

Figure 11: Survival Curves - 1997/98 vs 2005/06 - Part 6

D Robustness Tests

Table 7: Estimation Results - Robustness Exercises

	(1) 99-05 mortality	*(2)* 97-04	*(3)* 99-05 prolonged care/ mortality	*(4)* 97-04	*(5)* 99-05 delayed discharge	*(6)* 97-04	*(7)* 97-05 prolonged care	*(8)* 97-05 overly delayed discharge
shape	0.997	1.236	1.546	2.016	1.151	1.061	1.582	1.122
	(-0.02)	(1.40)	(8.87)***	(12.11)***	(11.74)***	(4.60)***	(8.29)***	(5.85)***
groupage	2.598	1.846	2.127	2.152	1.245	1.251	2.102	1.401
	(16.72)***	(11.48)***	(40.58)***	(41.90)***	(72.16)***	(73.57)***	(38.20)***	(66.65)***
sex	0.637	0.884	1.231	1.375	1.207	1.201	1.318	1.214
	(-3.45)***	(-0.95)	(4.48)***	(6.95)***	(16.88)***	(16.32)***	(5.56)***	(11.58)***
imd	1.201	0.977	1.195	1.286	1.183	1.035	1.207	1.524
	(2.20)**	(-0.27)	(6.14)***	(8.77)***	(3.10)***	(0.63)	(6.11)***	(5.23)***
ortho	0.271	0.425	3.450	3.093			4.765	
	(-8.40)***	(-5.43)***	(21.58)***	(19.60)***			(23.44)***	
ophtha	—	0.016	0.064	0.123			0.085	
		(-4.10)***	(-12.99)***	(-15.17)***			(-11.47)***	
totempl	1.000	1.000	0.998	0.998	1.000	1.000	0.998	1.000
	(0.90)	(0.96)	(-9.53)***	(-13.88)***	(-7.98)***	(-2.09)**	(-9.80)***	(-4.91)***
totcapinv	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(-0.09)	(0.47)	(1.93)*	(13.09)***	(9.33)***	(-12.98)***	(2.85)***	(3.83)***
No. of complications (CCs):								
CC_1		2.089	1.635	1.448	1.857	2.003	1.630	1.930
		(2.85)***	(9.17)***	(7.24)***	(51.34)***	(57.52)***	(8.94)***	(35.28)***
CC_2		10.748	2.750	2.384	2.585	2.999	2.549	3.188
		(9.91)***	(15.20)***	(13.11)***	(51.95)***	(57.27)***	(13.25)***	(46.69)***
CC_3		33.735	5.896	4.475	3.771	5.111	3.923	5.799
		(14.36)***	(19.96)***	(16.13)***	(39.24)***	(43.24)***	(12.60)***	(44.48)***
CC_4		109.335	11.345	13.412	5.680	11.314	4.188	10.911
		(17.40)***	(16.96)***	(17.73)***	(21.89)***	(23.38)***	(6.41)***	(29.28)***
medium			0.802	0.873	0.645	0.753	0.796	0.739
			(-3.03)***	(-1.57)	(-18.63)***	(-11.33)***	(-2.90)***	(-9.99)***
small			0.230	0.551	0.782	0.845	0.189	1.168
			(-12.88)***	(-5.25)***	(-8.10)***	(-5.29)***	(-13.43)***	(3.90)***
teaching			0.403	0.386	0.886	1.245	0.280	0.851
			(-9.40)***	(-9.05)***	(-5.45)***	(10.07)***	(-11.04)***	(-4.94)***
specialist			0.274	0.486	1.257	1.523	0.254	1.484
			(-9.57)***	(-5.08)***	(5.44)***	(10.29)***	(-9.70)***	(7.40)***
complaints			0.995	0.995	1.000	1.000	0.992	1.000
			(-8.54)***	(-6.49)***	(3.17)***	(-0.55)	(-8.03)***	(-3.04)***
	(-25.24)***	(-24.88)***	(-44.20)***	(-46.51)***	(-50.99)***	(-56.66)***	(-42.54)***	(-71.03)***
Obs.	165,118	198,938	193,442	198,976	193,986	196,410	193,203	190,065

Note: t-ratios (from robust standard errors) are in parentheses; * = 10%, ** = 5%, and *** = 1% levels of significance. Columns (3), (4), (7), (8) use beds instead of total employment as a measure of hospital's size. Columns (1)-(4) and (7) use the health and disability domain of the imd; the rest use the income domain.

In columns *(1), (3), (5)* we assess patient outcomes in 2005 (as in Table 4) using however hospital classification based on the 1999-2005 comparison of survival curves at short durations. Columns *(2), (4), (6)* perform the same exercise for outcomes in 2004. In the last two columns we display results for the two additional patients' outcomes, namely *prolonged care* and *overly delayed discharge*. Results are in line with the ones obtained in Table 4.