Spillover effects of financial incentives for providers onto non-targeted patients: daycase surgery in English hospitals

Philip Britteon1*, Søren Rud Kristensen1,2, Yiu-Shing Lau1, Ruth McDonald3 and Matt Sutton1,4

1Health Organisation, Policy and Economics, School of Health Sciences, The University of Manchester, Manchester, UK, 2Danish Centre for Health Economics, University of Southern Denmark, Odense, Denmark, 3Alliance Manchester Business School, The University of Manchester, Manchester, UK and 4Melbourne Institute: Applied Economics and Social Research, University of Melbourne, Melbourne, Australia

*Corresponding author: Philip Britteon; Email: philip.britteon@manchester.ac.uk

(Received 12 December 2021; revised 10 March 2023; accepted 25 March 2023; first published online 16 May 2023)

Abstract

Background. Incentives for healthcare providers may also affect non-targeted patients. These spillover effects have important implications for the full impact and evaluation of incentive schemes. However, there are few studies on the extent of such spillovers in health care. We investigated whether incentives to perform surgical procedures as daycases affected whether other elective procedures in the same specialties were also treated as daycases.

Data. 8,505,754 patients treated for 92 non-targeted procedures in 127 hospital trusts in England between April and March 2016.

Methods. Interrupted time series analysis of the probability of being treated as a daycase for non-targeted patients treated in six specialties where targeted patients were also treated and three specialties where they were not.

Results. The daycase rate initially increased (1.04 percentage points, SE: 0.30) for patients undergoing a non-targeted procedure in incentivised specialties but then reduced over time. Conversely, the daycase rate gradually decreased over time for patients treated in a non-incentivised specialty.

Discussion. Spillovers from financial incentives have variable effects over different activities and over time. Policymakers and researchers should consider the possibility of spillovers in the design and evaluation of incentive schemes.

Keywords: Multitasking; pay-for-performance; spillovers

1. Introduction

By rewarding or penalising performance, well-designed financial incentives can motivate providers to change their behaviour in line with the interests of the payer and population, leading to improvements in both quality and efficiency (Chalkley, 2018). However, incentives could also impact provider performance in areas of care that were not explicitly targeted by the incentives, including non-targeted tasks or the treatment of non-targeted patient groups (Kristensen et al., 2015). These wider effects are referred to as spillover effects in the literature on behaviour and incentives (Francetic et al., 2022).

Concerns for negative spillover effects in healthcare arise from the fact that only a subset of care provided by a professional or organisation can be observed and assessed by the government or an independent regulator (Chalkley and Malcomson, 1998). When this is the case, providers
may respond to an incentive scheme by reallocating effort and resources away from non-targeted dimensions of quality in order to bring about improvements in the incentivised area of care (Holmstrom and Milgrom, 1991; Kaarboe and Siciliani, 2011). This issue is particularly predominant in healthcare, where providers are required to perform multiple tasks and care for multiple patients under strict time and resource constraints. However, incentives may also have positive spillover effects when targeted and non-targeted tasks or patients share commonalities in production or agents learn from the incentivised task (Mullen et al., 2010; Kaarboe and Siciliani, 2011; Sherry et al., 2017). These beneficial spillover effects could outweigh any negative spillovers of the scheme.

Determining the sign and magnitude of spillover effects has important implications for the optimal design of an incentive scheme. In particular, theoretical evidence implies that incentives should be used with caution in situations where spillover effects are likely to have a negative affect non-targeted areas of care (Holmstrom and Milgrom, 1991; Kaarboe and Siciliani, 2011). In such a case, evidence suggests that targeted and non-targeted tasks should be split between different agents in order to reduce the potential for negative spillover effects and enable the use of higher-powered incentives (Holmstrom and Milgrom, 1991; Dewatripont et al., 2000). However, there exists little to no empirical evidence in healthcare in support of this theoretical finding.

Failure to consider spillover effects when evaluating an incentive scheme could also result in incorrect conclusions about its true overall effect. In particular, studies using non-targeted areas of care as a control group to estimate the targeted effect of a scheme may be downwardly or upwardly biased if the non-targeted area of care was positively or negatively affected by the reform (Rubin, 1990). This underlying assumption of common policy evaluation methods has been largely overlooked in studies assessing the intended impact of a financial incentive schemes (Eijkenaar et al., 2013; Kondo et al., 2015; Cattel and Eijkenaar, 2020; Zaresani and Scott, 2021).

To address these concerns, we evaluated the spillover effects of a national pay-for-performance scheme in England, the Best Practice Tariff (BPT) scheme, using an interrupted time series model. The BPT scheme incentivised hospitals to treat certain elective procedures within six specialities as a daycase but not others, despite other procedures within the same and different specialities also being recommended for treatment as a daycase by a national group of leaders in day and short stay surgery: the British Association of Day Surgery (BADS). This provided us with a natural experiment to examine the presence of spillover effects from targeted to non-targeted procedures and to test whether spillovers on the proportion of patients treated as a daycase differed between specialities based on their exposure to the incentives.

In doing so, we contribute to literature in two ways. First, we build on a handful of studies investigating the extent to which spillover effects occur in health care. Previous studies have investigated whether spillover effects on non-incentivised tasks differed between targeted and non-targeted patient groups (Mullen et al., 2010; Sutton et al., 2010; Doran et al., 2011; Binyaruka et al., 2015). However, very little is known about the extent to which spillover effects may vary between non-targeted areas of care based on their proximity to the targeted dimension of care. We address this shortcoming by investigating the extent to which spillover effects differed between patients undergoing a non-targeted procedure across different specialities depending on their exposure to the incentives. In doing so, we provide new empirical evidence on the extent to which incentivised and non-incentivised tasks should be split or shared between different areas of care and staff within an organisation, supporting theoretical evidence on the optimal design of incentive schemes (Holmstrom and Milgrom, 1991; Dewatripont et al., 2000).

Second, we develop previous evaluations of the BPT daycase incentive scheme by extending their focus to include the impact of the scheme on all procedures recommended for day surgery but not targeted by the scheme (Allen et al., 2016; Gaughan et al., 2019). In doing so, we also test an underlying assumption of the estimation strategies adopted by the two previous studies. Specifically, both studies estimated the targeted effect of the policy by exploiting information on patients that underwent a non-targeted procedure to construct a control group, using a difference-in-differences model or the synthetic control method. Estimates from these analyses
may have been biased if changes in response to the incentives also impacted the treatment of non-targeted patients (Angrist and Pischke, 2008; Abadie, 2021). We test the robustness of these findings using an interrupted time series model under the alternative assumption that daycase trends throughout the period before the incentives would have continued without interruption in the absence of the scheme. This approach avoided imposing assumptions on the spillover effects of the scheme and was preferred given the absence of an alternative uncontaminated control group.

2. Methods

2.1 Incentive scheme

The BPT scheme was a large pay-for-performance scheme that was first incorporated into national reimbursement system for hospitals in England in 2010 (Darzi, 2008). The scheme incentivised the performance of hospitals across a range of clinical areas by adjusting the amount that hospitals were reimbursed by the National Health Service (NHS) for treating certain patients (Gershlick, 2016). In this study, we focus on a subset of these incentives that aimed to increase the proportion of patients that were treated as a daycase (Leroy et al., 2017); where a daycase defined a period of care in which a patient was admitted to hospital for a planned surgical procedure and then successfully discharged in the same calendar day.

When appropriately applied, treating a patient as a daycase is generally cheaper than admitting the patient overnight (Street and Maynard, 2007) and can free up beds for other patients without requiring additional capacity (Alderwick et al., 2015). There is also evidence to suggest that daycase surgery has quality advantages. Studies have shown that patients have higher satisfaction when treated as a daycase, have reduced risk of hospital acquired infection from spending less time in hospital, and have the same complication and readmission rates than if they were admitted overnight (for overview see: Jarrett and Staniszewski, 2006).

Prior to the BPT incentives, hospitals were reimbursed the same amount per admission regardless of whether a patient was treated as a daycase or not. The amount that hospitals received for each admission was determined under the prospective payment system, known as Payment by Results (PbR) (Department of Health, 2012). Under PbR, hospitals were paid a fixed tariff per admission to cover the costs incurred throughout each patient’s spell in hospital. Tariffs were set at a price equal to the national average cost of treating patients that had previously undergone the same procedure and required similar levels of resources to treat. This incentivised hospitals to reduce costs to a level below the national average in order to reduce deficits and remain financially viable (Shleifer, 1985). Since overnight admissions cost more on average, the payment system also meant that hospitals were incentivised to treat patients as a daycase prior to the introduction of the BPT incentives (Street and Maynard, 2007).

After the introduction of the BPT daycase incentives, hospitals were reimbursed more money if a patient undergoing a targeted procedure was treated as a daycase and less if they were admitted overnight (Table A1 in the supplementary material). This increased the incentives for hospitals to treat patients with a targeted condition within the same day. Differences between the daycase and overnight tariff varied between procedures from £150 to £325 (NHS England, 2016). This equated to an extra 8.8% to 40.8% in revenue for treating a patient as a daycase rather than as an overnight admission. The new daycase and overnight tariffs were set based on a predetermined target daycase rate at which hospital would have received the same average revenue per patient as in the period before the incentives. This meant that hospitals were rewarded for producing above the target daycase rate but were penalised for producing below. The target daycase rate varied between procedures dependent on the recommended rate of best practice (BADS, 2016) and the extent to which this rate was deemed achievable.

The scheme targeted procedures with daycase rates that varied substantially between hospitals and on average, fell below the national recommended rate (NHS England, 2016). Procedures were selected from a directory produced by the BADS (BADS, 2016). The directory listed procedures
that were considered, by a panel of experts, to be suitable for treatment as a daycase. In total, 24 procedures listed in the BADS directory were targeted in the rollout of the BPT incentives between April 2010 and April 2013.

In April 2010, hospitals were only incentivised to treat patients undergoing a laparoscopic cholecystectomy as a daycase (Figure 1). The incentives were then expanded to include an additional 13 procedures in April 2011, six procedures in April 2012 and four procedures in April 2013. A further 18 procedures were introduced in April 2017. Information on the rollout of the BPT initiative was published in January each year. However, hospitals were able to familiarise themselves with the detail of the new proposed tariff structure approximately 6 months before it was introduced, during the ‘road testing’ stage of its release (Department of Health, 2012).

2.2 Potential spillover mechanisms

Despite the rollout of the BPT incentives, the set of elective procedures targeted by the scheme only made up a small proportion of the total procedures that were recommended for treatment as a daycase by the BADS (BADS, 2016). Patients undergoing a procedure that was recommended for treatment as a daycase but was not incentivised under the scheme may have been affected by the BPT incentives if induced changes in the behaviour of hospitals had an impact on the care that they received. We refer to these procedures and patients admitted for one of these procedures as being ‘non-targeted’.

There are different mechanisms through which the incentives may have impacted patients undergoing a non-targeted procedure that was also recommended for treatment as a daycase. On the one hand, the BPT incentives may have had a negative spillover effect on the proportion of non-targeted patients that were treated as a daycase if hospitals reallocated effort and resources towards the incentivised area of care (Holmstrom and Milgrom, 1991). For example, daycase units and physicians may have delayed the treatment of non-targeted patients within the same calendar day to ensure that targeted patients were treated first and had time to recover before discharge.

On the other hand, there may have been a positive spillover effect of the scheme if targeted and non-targeted patients shared commonalities in production as a daycase (Mullen et al., 2010; Sherry et al., 2016). For example, patients undergoing different procedures may have shared the same clinical staff, specialist equipment and treatment rooms within a hospital. In such a case, improvements in the targeted area of care may have also led to improvements in another through two potential mechanisms.
First, improvements in the proportion of targeted patients that were treated as a daycase may have been achieved through investment in the daycase capacity of a hospital. For example, hospitals may have invested in larger daycase facilities and hired additional staff. In return, such investments may have led to an increase in the proportion of non-targeted patients that were treated as a daycase if additional resources were also used to treat these patients.

Second, there may have been a ‘learning effect’ (or knowledge spillover) that benefited the treatment of targeted and non-targeted patients (Chandra and Staiger, 2007). This effect may have occurred without any investment in additional resources. For example, a physician may have learnt from performing minimal access techniques on targeted patients and applied this skill when performing similar non-targeted procedures. Moreover, physicians that tended to admit patients overnight may have switched their treatment style towards daycase surgery after the incentives were introduced. Knowledge could have also spread between physicians that regularly interacted within the same network (Landon et al., 2012). Similarly, daycase units may have become more efficient and organised at admitting and discharging patients within the same calendar day.

Whether the overall spillover effect of the BPT scheme was positive or negative is dependent on which of these prevailing mechanisms caused the spillover to occur. Moreover, the interaction between these mechanisms is likely to have varied between different non-targeted patients dependent on the exposure to the incentives of the physician or surgical team responsible for the treatment of the patient.

On the one hand, positive spillover effects from investment and learning were likely to have been larger for patients within the same speciality as a targeted procedure (Mullen et al., 2010). For example, whilst investments in general daycase facilities (e.g. admin staff, operation theatres) are likely to have been wide reaching, investments in specialist resources may have only benefitted the treatment of patients whose treatment required the same equipment or was conducted in the same location within the hospital. Similarly, positive spillover effects through learning may have been specific to certain types of procedure and larger for patients treated by physicians who were also treating more patients as daycases in response to the incentives.

By the same logic, hospitals were also more likely to have substituted effort away from non-targeted patients within the same area as a targeted procedure if they shared the use of the same specialist equipment or were treated by the same pool of specialist physicians (Holmstrom and Milgrom, 1991). Understanding the extent to which spillover effects differed between patients within an organisation is therefore ultimately an empirical question.

### 2.3 Data

To evaluate the spillover effects of the BPT incentives, we used patient-level data from Hospital Episode Statistics (HES). The same data were used to reimburse providers under PbR and the BPT scheme. The data contained clinical, personal and geographical information on all NHS-funded patients that were admitted to a hospital trust in England.

We evaluated the effect of the incentives on a binary variable from HES that measured whether the hospital both planned to and successfully discharged the patient within the same calendar day. This matched the definition of a daycase used to reimburse hospitals under the BPT scheme (Department of Health, 2013). Whether the hospital planned to treat the patient within the same day was recorded by the hospital prior to the patient’s admission. Whether the patient was treated in the same day was determined using information on the date of admission and the date of discharge from hospital.

In addition, we used information from HES to construct control variables in the analysis. These variables included: the age of the patient (coded in 10-year bands from 19–30 to ≥80 years), the sex of the patient (male = 1), 31 binary variables that indicated whether the patient had a specific Elixhauser comorbidity or not (yes = 1) (Elixhauser et al., 1998), whether the
patient was admitted from their usual place of residence or not (yes = 1), the income deprivation
score of the patient’s area of residence in 2010 (McLennan et al., 2011), and the average drive
time to hospital from the population weighted centroid of the local geographical area in which
the patient lived (Gibbons et al., 2019).

We used HES data on admissions to hospital from the start of April 2007 to end of March
2016. This period included: three years before the introduction of the first incentives; the three-
year rollout period; and three years after the final set of incentives were introduced (Figure 1).

We focused on patients undergoing an elective procedure that was recommended for treatment
as a daycase but never targeted by the BPT incentives during the period of study. We refer to these procedures as non-targeted procedures. In total, we considered patients admitted for one of 92 different non-targeted procedures. We identified non-targeted procedures from the procedure codes recorded in the second to fifth edition of the BADS directory of procedures (BADS, 2016). The directories listed procedure codes deemed suitable for treatment as a daycase by leaders in day and short stay surgery and from a review of national performance data. The same directories were used in the development of the BPT daycase scheme and in the reporting of day surgery rates on a quarterly basis on the ‘Better Care Better Values’ website.

We grouped the 92 non-targeted procedures into incentivised and non-incentivised special-
ities. We defined specialties based on the surgical groupings in the BADS directories. Incentivised specialties contained both targeted and non-targeted procedures and included: general surgery; gynaecology; urology; breast surgery; orthopaedic surgery; and ear, nose and throat (ENT). 71 of the 92 non-targeted procedures belonged to one of these six categories. The remaining 21 non-targeted procedures belonged to one of three non-incentivised specialties that contained no targeted procedures: vascular surgery; head and neck; and ophthalmology. A full list of the non-targeted procedures analysed in the study is included in Table A2 in the supplementary material.

We identified admissions where a patient underwent a non-targeted procedure in HES using
the procedure codes recorded throughout the duration of their hospital spell. We did not consider admissions where the patient underwent more than one non-targeted procedure in the same hospital spell because daycase surgery was unlikely to have been suitable for patients that required more than one significant procedure. Similarly, we did not consider admissions where a patient underwent both a targeted and non-targeted procedure in the same hospital spell as hospitals were rewarded for treating these patients as a daycase under the BPT scheme.

In total, we identified 10,478,334 hospital admissions to 204 hospital trusts for a single non-targeted procedure during the study period. Out of these admissions, we excluded 1,890,489 admissions to 76 hospital trusts that did not perform at least one non-targeted procedure from both an incentivised and non-incentivised specialty in every month. This ensured that we estimated the spillover effect of the incentives on a balanced panel and for the same set of hospital trusts when stratifying the estimation sample. Additionally, we excluded 82,091 admissions to a single hospital trust that did not perform at least one targeted procedure in every month in order to exclude hospitals that were not regularly incentivised by the BPT incentives during the study period.

The final estimation sample included 8,505,754 (81.2%) patients that were admitted for a single non-targeted procedure to 127 different hospital trusts between April 2007 and March 2016 (Figure 2). These patients were similar to the population of patients, but the characteristics of hospitals included in the sample differed as result of the inclusion criteria (Table A3 in the supplementary material).

2.4 Estimation strategy
2.4.1 Main analysis
We estimated average changes in the probability that a non-targeted patient was treated as a day-

case using an interrupted time series (ITS) model (Wagner et al., 2002). We opted for this
approach because unlike other policy evaluation methods, the ITS method did not impose any assumptions on the potential spillover effects of the scheme. Instead, the approach exploited changes in the probability of being treated as a daycase throughout the before period to predict what the daycase rate would have been in the post-policy period had the BPT incentives not been introduced.

Specifically, the method relied on the assumption that the trend throughout the before period would have continued without interruption had the BPT incentives not been introduced (Wagner et al., 2002). This assumption may not have held if: (i) the shape of the trend over time was incorrectly specified in the model; (ii) other policy changes or technological advances affected the daycase trend in the post-policy period; or (iii) changes in the daycase trend were caused by changes in other variables that were associated with the probability of being treated as a daycase (Lopez Bernal et al., 2018).

We estimated the ITS model as follows:

\[
\Pr [Y_{ijht} = 1] = \alpha + \beta A_t + \sum_{k=2010}^{2015} [ \gamma_k D_k + \delta_k (A_t \times D_k)] + X_t \varphi + \sum_{s=1}^{12} \sigma_s M_s + \tau_j + \eta_h + \epsilon_{ijht}
\]  

where, \( Y_{ijht} \) was a binary variable that indicated whether patient \( i \) was treated as a daycase for non-targeted procedure \( j \) at hospital trust \( h \) in month \( t \); \( A_t \) was a continuous variable that indicated the time in months from the start of the study period; \( D_k \) was a binary variable equal to one in post-policy year \( k \), and zero otherwise; \( X_t \) were a vector of patient characteristics (as described in Figure 2. Volume of non-targeted patients in the estimation sample, by speciality. Parentheses indicate the number of non-targeted procedures within each speciality.)

https://doi.org/10.1017/S1744133123000063 Published online by Cambridge University Press
above); \( \tau_j \) were procedure fixed effects; \( M_s \) were month dummies; and \( \eta_h \) were hospital trust fixed effects.

We included \( D_k \) in the model to allow for shifts in the level of the daycase trend in each post-policy year. Similarly, we included an interaction term between \( A_t \) and \( D_k \) to allow for changes in the shape of the daycase trend in each post-policy year (Karaca-Mandic et al., 2012). This approach allowed us to investigate whether spillover effects of the incentives varied in the short- and long-term periods as a result of the yearly rollout of the scheme.

We estimated the model using a logit regression with cluster-robust standard errors at the hospital trust level. This approach allowed us to adjust for potential floor or ceiling effects under the assumption that the probability of being treated a daycase followed an S-shaped trend over time that flattened as the probability reached 0 or 1 (Lopez Bernal et al., 2018). Using the output from the logit model, we then calculated the spillover effect of the policy as the average marginal effect of being treated in post-policy year \( k \), when setting the continuous time variable \( (A_t) \) equal to the midpoint of year \( k \). This estimated the average difference between the level of the extrapolated daycase trend from the before period and the level of the actual daycase trend in each post-policy year.

We estimated the model separately for patients undergoing a non-targeted procedure in either an incentivised speciality or a non-incentivised speciality to investigate whether the average spillover effect of the incentives differed between non-targeted patients based on their proximity to the targeted area of care. In addition, we also stratified the analysis by the speciality in which a patient was treated. This allowed us to further examine whether spillover effects varied between different incentivised specialities depending on when they became targeted by the policy.

### 2.5 Sensitivity analysis

We also tested the sensitivity of the results to the assumption that the trend in the daycase rate throughout the before period would have continued without interruption in the absence of the incentive scheme. This assumption may have failed if trends throughout the before period did not follow an S-shaped path as modelled by the logit regression (Verbyla et al., 1999). To investigate this further, we estimated the ITS model using a linear probability model.

Estimates from the main analysis may have also been biased if hospital trusts responded to the announcement of the incentives up to six months before they were implemented, as suggested in a previous evaluation of the BPT scheme (Allen et al., 2016). As an additional sensitivity test, we therefore also estimated changes in the post-policy trend every six months from October 2009 to the end of the study period.

### 3. Results

#### 3.1 Descriptive statistics

During the period before the incentives were introduced, patients undergoing a non-targeted procedure from an incentivised speciality were less likely to be treated as a daycase than patients undergoing a non-targeted procedure from a non-incentivised speciality (61.9% vs 87.9%) (Table 1). This difference was in part due to differences in the complexity of treating patients as a daycase between the incentivised and non-incentivised specialities considered in the study; reflected by the difference in the BADS recommended daycase rate (90.0% vs 97.8%). Throughout the study period, the average daycase rate increased for non-targeted patients in both incentivised and non-incentivised specialities. Relative to the extrapolated daycase trend from the before period, however, the trend in the daycase rate increased for patients in an incentivised speciality but decreased for patients in a non-incentivised speciality (Figure 3).

Over the same period, the severity of patients in the estimation sample also increased (Table 1). The average number of comorbidities for patients undergoing a non-targeted procedure in an incentivised speciality was higher from 2013/14 onwards (0.8 comorbidities) than in the
Table 1. Characteristics of non-targeted patients over time, by type of speciality

<table>
<thead>
<tr>
<th>Observations</th>
<th>Non-targeted patients</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Incentivised speciality</td>
<td>Rollout</td>
<td>After</td>
<td>Non-incentivised speciality</td>
<td>Rollout</td>
<td>After</td>
</tr>
<tr>
<td>Number of admissions</td>
<td>1,487,832</td>
<td>1,478,185</td>
<td>1,514,446</td>
<td>1,349,344</td>
<td>1,290,151</td>
<td>1,385,796</td>
</tr>
<tr>
<td>Number of hospital trusts</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
</tr>
<tr>
<td>Daycase rate (%)</td>
<td>61.9</td>
<td>66.2</td>
<td>69.3</td>
<td>87.9</td>
<td>89.9</td>
<td>91.5</td>
</tr>
<tr>
<td>BADS recommended rate in 2016 (%)</td>
<td>90.0</td>
<td>89.9</td>
<td>90.2</td>
<td>97.8</td>
<td>97.6</td>
<td>97.6</td>
</tr>
</tbody>
</table>

**Patient characteristics**

| Male (%)                     | 47.2                  | 47.0      | 46.5      | 41.9      | 43.0      | 43.2 |
| Age (Years)                  | 53.1                  | 54.2      | 55.2      | 65.1      | 65.3      | 65.9 |
| Coexisting conditions (No. per patient) | 0.4                   | 0.6       | 0.8       | 0.6       | 0.9       | 1.0 |
| Admitted from home (%)       | 99.8                  | 99.8      | 99.7      | 99.8      | 99.8      | 99.5 |
| Drive time (mins)            | 24.0                  | 24.4      | 24.6      | 24.0      | 24.5      | 24.8 |
| Income deprivation score in 2010 (%) | 21.7                  | 21.7      | 21.5      | 21.6      | 21.8      | 21.3 |

**Before Period** is from April 2007 to March 2010.
**Rollout Period** is from April 2010 to March 2013.
**After Period** is from April 2013 to March 2016.

Incentivised specialities are specialities that also included patients undergoing a targeted procedure.
Non-incentivised specialities are specialities that did not include any patients undergoing a targeted procedure.

BADS – British Association of Day Surgery.
before (0.4 comorbidities) period. The same patients were also older on average in the after period (55.2 years) than in the before (53.1 years) and rollout (54.2 years) periods. Similar patterns were observed for non-targeted patients in a non-incentivised speciality.

3.2 Main analysis

The BPT incentives had an initial positive spillover effect on patients undergoing a non-targeted procedure within an incentivised speciality (Table 2). In 2010/11, the probability that these patients were treated as a daycase was 1.04 percentage points (SE: 0.30) higher as a result of the BPT incentives. This effect increased to 1.78 percentage points (SE: 0.49) by the midpoint of 2011/12. The spillover effect declined each year after 2012/13 but remained positive in magnitude until the end of the study period. However, the effect was no longer statistically different from zero from 2014/15 onwards.

Conversely, the BPT incentives had no initial spillover effect on the treatment of patients within a non-incentivised speciality in 2010/11 (0.01 percentage points, SE: 0.16). However, there was evidence of a gradual reduction in the probability of being treated as a daycase throughout the rollout of the incentives. By 2013/14, patients were 0.97 percentage points (SE: 0.41) less likely to be treated as a daycase as a result of the policy. This statistically significant effect persisted until the end of the study period.

Over the whole study period, the estimates from the ITS model suggest that an additional 41,898 non-targeted patients were treated as a daycase in an incentivised speciality as a result of the BPT incentives (Figure 4). Conversely, there was a reduction of approximately 15,218 in the number of patients that were treated as a daycase in a non-incentivised speciality. This equated to an additional 26,680 patients being treated as a daycase between April 2010 and March 2016. The majority of these additional daycases were provided in the first few years of the scheme, when the positive spillover effect on patients in an incentivised speciality was large and the negative spillover effect on patients in a non-incentivised speciality was small.

Estimates of the spillover effect of the BPT scheme also differed between specialities (Table 3). Estimates of the spillover effect on each of the six incentivised surgical were positive in each post-
policy year. However, a statistically significant positive effect was only observed for patients undergoing a procedure in general surgery, orthopaedics and urology. In each of the specialties where there was a significant result, the spillover effect developed around the time that the incentives were introduced. The largest increase in the daycase rate was observed in general surgery; the first specialty to be incentivised by the BPT scheme.

In comparison, the effect on non-targeted patients in a non-incentivised specialty varied between each specialty. There was only evidence of a statistically significant negative effect on patients undergoing a head and neck operation. The probability of being treated as a daycase reduced over time for these patients and was significantly less than zero from 2013/14 onwards (−1.42 percentage points, SE: 0.54).

### 3.3 Sensitivity analysis

Estimates from the sensitivity analyses are included in the supplementary material provided with the paper. There was no evidence of a spillover effect in response to the announcement of the BPT incentives between October 2009 and April 2010 on the daycase rate of patients in incentivised (0.11 percentage points, SE: 0.22) and non-incentivised (0.11 percentage points, SE: 0.18) specialities (Table A4 in the supplementary material).

In addition, the results from the main analysis for patients in an incentivised specialty did not differ based on the choice of model (Table A5 in the supplementary material). However, the negative spillover effect on patients in a non-incentivised specialty was larger when estimating the analysis using a linear probability model. The difference between estimates from the linear

---

**Table 2. Spillover effect of the BPT incentives on the probability that a non-targeted patient was treated as a daycase, by type of speciality**

<table>
<thead>
<tr>
<th>Year</th>
<th>Incentivised speciality</th>
<th>Non-incentivised speciality</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010/11</td>
<td>1.04***</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>2011/12</td>
<td>1.78***</td>
<td>−0.09</td>
</tr>
<tr>
<td></td>
<td>(0.49)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>2012/13</td>
<td>1.76**</td>
<td>−0.32</td>
</tr>
<tr>
<td></td>
<td>(0.62)</td>
<td>(0.32)</td>
</tr>
<tr>
<td>2013/14</td>
<td>1.46*</td>
<td>−0.97*</td>
</tr>
<tr>
<td></td>
<td>(0.74)</td>
<td>(0.41)</td>
</tr>
<tr>
<td>2014/15</td>
<td>1.28</td>
<td>−0.98*</td>
</tr>
<tr>
<td></td>
<td>(0.81)</td>
<td>(0.41)</td>
</tr>
<tr>
<td>2015/16</td>
<td>1.09</td>
<td>−0.98*</td>
</tr>
<tr>
<td></td>
<td>(0.90)</td>
<td>(0.41)</td>
</tr>
<tr>
<td>Providers</td>
<td>127</td>
<td>127</td>
</tr>
<tr>
<td>Procedures</td>
<td>71</td>
<td>21</td>
</tr>
<tr>
<td>Patients</td>
<td>4,480,463</td>
<td>4,025,291</td>
</tr>
</tbody>
</table>

Estimates are the average marginal effect of the BPT incentives on the probability that a patient undergoing a non-targeted procedure was treated as a daycase at the midpoint of each year, measured in percentage points. Parentheses include cluster robust standard errors at the hospital trust level.

* p < 0.05, ** p < 0.01, *** p < 0.001.
The largest spillover effects were observed in the initial year of the study period, with a positive spillover effect on patients in incentivised specialities increasing over time. In the second year of the scheme, the spillover effect further increased but then slowly reduced until the end of the study period. Similarly, the negative spillover effect on patients in non-incentivised specialities worsened over time and was statistically significant from the fourth year of the policy. Overall, the magnitude of the spillover effect in percentage points was smaller than the targeted effect observed in previous evaluations of the scheme (Allen et al., 2016; Gaughan et al., 2019). This was expected given that the incentives were likely to have had the largest effect on the group of patients that were directly targeted by the scheme. Nonetheless, the spillover effect on the volume of daycases was relatively large given the total number of non-targeted patients in the population.
Table 3. Spillover effect of the BPT incentives on the probability that a non-targeted patient was treated as a daycase, by speciality

<table>
<thead>
<tr>
<th></th>
<th>General surgery</th>
<th>Gynaecology</th>
<th>Orthopaedic</th>
<th>Urology</th>
<th>Breast</th>
<th>ENT</th>
<th>Vascular surgery</th>
<th>Head and neck</th>
<th>Ophthalmology</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010/11</td>
<td>1.68***</td>
<td>0.11</td>
<td>0.95**</td>
<td>1.27</td>
<td>1.39</td>
<td>1.32</td>
<td>0.10</td>
<td>−0.09</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
<td>(0.47)</td>
<td>(0.34)</td>
<td>(0.75)</td>
<td>(1.86)</td>
<td>(0.78)</td>
<td>(0.97)</td>
<td>(0.26)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>2011/12</td>
<td>2.52**</td>
<td>0.53</td>
<td>1.87***</td>
<td>2.19*</td>
<td>0.41</td>
<td>1.39</td>
<td>−0.90</td>
<td>−0.53</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>(0.87)</td>
<td>(0.79)</td>
<td>(0.56)</td>
<td>(1.00)</td>
<td>(3.77)</td>
<td>(1.35)</td>
<td>(1.67)</td>
<td>(0.37)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>2012/13</td>
<td>2.83*</td>
<td>1.42</td>
<td>1.57*</td>
<td>1.70</td>
<td>0.72</td>
<td>2.29</td>
<td>0.60</td>
<td>−0.88</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(1.16)</td>
<td>(1.02)</td>
<td>(0.73)</td>
<td>(1.20)</td>
<td>(4.24)</td>
<td>(1.63)</td>
<td>(2.19)</td>
<td>(0.47)</td>
<td>(0.30)</td>
</tr>
<tr>
<td>2013/14</td>
<td>3.15*</td>
<td>1.30</td>
<td>1.36</td>
<td>0.74</td>
<td>1.38</td>
<td>2.39</td>
<td>0.25</td>
<td>−1.42**</td>
<td>−0.40</td>
</tr>
<tr>
<td></td>
<td>(1.36)</td>
<td>(1.28)</td>
<td>(0.86)</td>
<td>(1.43)</td>
<td>(4.15)</td>
<td>(1.84)</td>
<td>(2.71)</td>
<td>(0.54)</td>
<td>(0.47)</td>
</tr>
<tr>
<td>2014/15</td>
<td>3.18*</td>
<td>1.92</td>
<td>1.34</td>
<td>0.02</td>
<td>2.97</td>
<td>1.37</td>
<td>1.83</td>
<td>−1.57*</td>
<td>−0.46</td>
</tr>
<tr>
<td></td>
<td>(1.55)</td>
<td>(1.55)</td>
<td>(0.92)</td>
<td>(1.69)</td>
<td>(4.52)</td>
<td>(2.03)</td>
<td>(3.21)</td>
<td>(0.65)</td>
<td>(0.38)</td>
</tr>
<tr>
<td>2015/16</td>
<td>3.22</td>
<td>2.02</td>
<td>1.06</td>
<td>0.01</td>
<td>2.42</td>
<td>0.90</td>
<td>3.13</td>
<td>−1.79**</td>
<td>−0.43</td>
</tr>
<tr>
<td></td>
<td>(1.68)</td>
<td>(1.79)</td>
<td>(1.03)</td>
<td>(1.99)</td>
<td>(5.27)</td>
<td>(2.17)</td>
<td>(3.52)</td>
<td>(0.67)</td>
<td>(0.36)</td>
</tr>
</tbody>
</table>

Providers 126 123 127 122 38 118 106 122 107
Procedures 14 8 12 20 2 15 4 11 6
Patients 619,167 450,396 2,076,701 794,472 38,707 463,787 358,211 891,381 2,666,771

Estimates are the average marginal effect of the BPT incentives on the probability that a non-targeted patient was treated as a daycase at the midpoint of each year, measured in percentage points. The shaded cells represent when a speciality became incentivised by the scheme (i.e. included a targeted procedure). Parentheses include cluster robust standard errors at the hospital trust level. * p < 0.05, ** p < 0.01, *** p < 0.001.
The findings of the study have multiple contributions. On the one hand, the study addresses a lack of research on the extent to which spillover effects may differ between non-targeted measures of performance based on their proximity to the targeted area of care (Mullen et al., 2010; Kristensen et al., 2014). In doing so, we were able to investigate the appropriateness of using different non-targeted patients as a control group when evaluating the targeted effect of a reform. Secondly, the study contributes to existing evidence on the effectiveness of the BPT scheme by expanding the focus of previous evaluations to consider the impact of the policy on non-targeted procedures and testing the assumptions of the previous evaluations that estimated the intended effect of the scheme using patients undergoing a non-targeted procedure as a control group (Allen et al., 2016; Gaughan et al., 2019).

However, there are limitations of the study that should be noted. First, whilst the use of an ITS model allowed us to estimate the impact of the BPT incentives without imposing assumptions on the spillover effect of the reform, the model still relied on the alternative assumption that the trend throughout the before period would have continued without interruption had the BPT incentives not been introduced (Wagner et al., 2002). To test the sensitivity of our analysis to this assumption, we estimated the ITS model using both linear and logit regression, estimated changes in response to the announcement and introduction of the BPT incentives, and adjusted for a range of time-varying patient level controls.

Second, the study only focused on the spillover effect of the incentives on non-targeted patients that were also recommended for treatment as a daycase. This allowed us to meaningfully evaluate the spillover effect of incentives using the same outcome that was targeted by the scheme. However, there may have also been wider spillover effects on other non-targeted patient groups. For example, treating more patients as a daycase may have freed up overnight capacity, allowing hospitals to admit emergency patients overnight. Future research should consider the potential impact of the BPT scheme on these patients.

Nonetheless, the findings still have important implications on the design of future incentive schemes and on the methods used to evaluate the intended effect of these initiatives. First, we find that positive spillover effects of the scheme were almost exclusively concentrated on patients undergoing a non-targeted procedure in an incentivised speciality, where patients were more likely to have shared commonalities in production and to have been treated by a physician that also treated a targeted patient. This finding implies that future incentives could maximise the potential positive spillover effects of a scheme by targeting areas of care from a range of different specialities. In particular, the findings suggests that the expansion of the BPT scheme in April 2017 to include head and neck, ophthalmology and vascular surgery procedures from the three non-incentivised specialities considered in this study may have led to further positive spillover effects in these areas. This finding is closely linked to theoretical evidence on the optimal design of incentive schemes, which suggests that incentivised and non-incentivised tasks should be shared between agents when the potential spillover effects of the incentives are positive (Holmstrom and Milgrom, 1991; Dewatripont et al., 2000).

Furthermore, the findings also highlight the importance of considering spillover effects when evaluating the intended impact of a scheme in order to rule out potential unintended effects of the policy. In particular, the lack of any initial negative short-term spillover effect supports previous observations that hospitals made real improvements to their daycase capacity instead of reallocating effort and resources away from other patients that were suitable for treatment as a daycase (Allen et al., 2016; Gaughan et al., 2019).

Finally, the analysis allowed us to investigate the underlying assumptions of previous evaluations that estimated the intended effect of the scheme using patients undergoing a non-targeted procedure as a control group (Allen et al., 2016; Gaughan et al., 2019). The results imply that previous studies may have both underestimated and overestimated the intended effect of the policy depending on whether patients in the control group were predominantly from an incentivised or non-incentivised speciality, respectively. This finding highlights an important issue with using
non-targeted patients within an incentivised organisation as a control group when analysing the intended effect of an incentive scheme, even if the non-targeted patients were further removed from the targeted area of care. When possible, we recommend that future studies should consider the use of patients treated in non-incentivised organisations as a control group in order to avoid bias from potential within-organisation spillover effects.

In conclusion, the study provides important evidence on how spillover effects of financial incentives may differ between different non-targeted patient groups based on their proximity to the targeted area of care. The findings highlight the need for future evaluations to consider the spillover effects of a policy in order to understand the full impact of its effects and to avoid potential bias, when using non-targeted patients as a control group in the analysis. Understanding the common causes of spillover effects across different quality incentive schemes is critical for both the design of future incentive schemes and the monitoring and evaluation of their impact.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1744133123000063.

Acknowledgements. PB was supported by a University of Manchester Alumni Research Impact Scholarship. MS is a NIHR Senior Investigator. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Conflict of interest. None.

References


Cattel D and Eijkenaar F (2020) Value-based provider payment initiatives combining global payments with explicit quality incentives: a systematic review. Medical Care Research and Review 77, 511–537.


