Part of

UNIVERSITY of York

CHE
Centre For Health Economics

Can financial incentives shift health care from an inpatient to an outpatient setting?

Katja Grašič, Luigi Siciliani, Jinglin Wen

CHE Research Paper 201

Can financial incentives shift health care from an inpatient to an outpatient setting?

[a] **Katja Grašič**
[b] **Luigi Siciliani**
[c] **Jinglin Wen**

[a] Centre for Health Economics, University of York
[b] Department of Economics and Related Studies, University of York
[c] Centre for Health Economics, University of York

January 2026

Background to series
CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users.

The CHE Research Paper series takes over that function and provides access to current research output via web-based publication. The RPs can take various forms and results and ideas do not necessarily represent the final position and may include work in progress not subject to peer review at the time of publication.

CHE RP 201 is an early version of a paper intended for later submission to an academic journal.

Acknowledgements
This study was funded by the National Institute for Health Research (NIHR) Doctoral Research Fellowship Program (DRF-2016-09-097) and the National Institute for Health and Care Research NIHR) Policy Research Programme (reference NIHR206116).

The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. This work uses data provided by patients and collected by the NHS as part of their care and support.

We are grateful for support by the York Data Safe Haven team and the Research Computing team.

Further copies
Only the latest electronic copy of our reports should be cited. Copies of this paper are freely available to download from the CHE website www.york.ac.uk/che/outputs/. CHE Research Papers can be reused and distributed for non-commercial purposes provided the source is properly cited. Please use https://doi.org/10.15124/yao-2x4t-xw96 in your reference.

Centre for Health Economics
Alcuin College
University of York
York,
YO10 5DD, UK

www.york.ac.uk/che

Abstract

Can financial incentives affect hospital behaviour and shift care from a high-cost (inpatient) setting to a low-cost (outpatient) setting? We exploit a unique policy that switched from a reimbursement scheme where the tariff was substantively *lower* for outpatient care relative to inpatient care, to one where the tariff was *higher* for outpatient care than for inpatient care. The policy affected three common procedures in England. After developing a theoretical model of provider incentives, we use a difference-in-difference approach to evaluate the policy and find that it significantly increased the proportion of outpatient treatments by 36 percentage points for cystoscopy, 16.3 percentage points for hysteroscopy, and 3.8 percentage points for sterilisation. This substantial increase also improved quality for cystoscopy by reducing repeated procedures and emergency readmissions, had mixed effects on quality for hysteroscopy and did not affect emergency readmissions for sterilisation. We show that the policy did not increase total volume (the extensive margin) but had positive spillover effects on related unincentivised procedures. We also explore the distribution implications for the public insurer (the funder) and hospitals (the providers).

JEL: I11, I18.
Keywords: Pay for Performance, healthcare, financial incentives, cost containment, quality.

1. Introduction

Driven by the rapid growth of healthcare spending, policymakers across OECD countries are under renewed pressure to develop policies that contain costs while preserving quality of care (OECD 2010; Cashin et al. 2014; OECD 2025). One policy lever to induce healthcare providers to reduce costs and increase efficiency is the use of financial incentives (Ellis and McGuire 1993; Hollingsworth 2008; Hussey et al. 2009; Gaughan et al. 2019).

Hospital care accounts for about 40% of all health spending across OECD countries (OECD 2021). One area that has been targeted to reduce costs in hospitals is the substitution of the more expensive inpatient care with the less expensive outpatient (ambulatory) care[1] (Davis and Russell 1972; Elnicki 1976; Vitikainen et al. 2010). In the 1970s, ambulatory care was initially limited to a few selected treatments (Vitikainen et al. 2010), but has since gained momentum due to advances in medical technology. Many other medical conditions are increasingly treated in ambulatory (office-based) settings, such as diagnostic procedures, chemotherapy, and dialysis (Dobson et al. 2013; Esparza et al. 1989; Wellenstein et al. 2017). This shift has brought about several benefits, including lower costs, shorter hospital stays, and faster patient recovery times (Ancona-Berk and Chalmers 1986; Castells et al. 2001; Robinson and Beyer 2010; Hayes et al. 2015; Gordan et al. 2019).

Despite the availability of the relevant technology and the recognised benefits, the move towards the outpatient setting for suitable procedures has been slow. In England, less than a third of these procedures were performed on an outpatient basis in 2010 (Department of Health 2012). This can be explained by (i) the financial incentives, with higher revenues and profit margins for the inpatient relative to the outpatient setting (Higgins et al. 2016; Bach and Jain 2017; Fisher et al. 2017), as well as (ii) provider reluctance to change established practices, which requires providers to invest time and resources in the reorganisation of their services (Gaughan et al. 2019).

Strategic pricing or pay for performance in the form of a financial reward, which explicitly incentivises outpatient care, can potentially encourage hospitals to change their modus operandi and shift the provision of care towards the incentivised setting (Ellis and McGuire 1986; Ma 1994; Hodgkin and McGuire 1994; Hafsteinsdottir and Siciliani 2010; Lisi et al. 2020). There have been several policy initiatives aimed at changing providers' behaviour through financial incentives, though generally with a focus on increasing quality rather than reducing costs. Existing literature reviews suggest that these initiatives had limited impact (Milstein and Schreyoegg 2016; Eijkenaar 2012; Zaresani and Scott 2021). However, with few exceptions, the size of the financial incentive was generally small or triggered by a wider overhaul of hospital reimbursement (Bach and Jain 2017; Milstein and Schreyoegg 2016; Eijkenaar 2012; OECD 2010; He and Mellor 2013; Zaresani and Scott 2021). This study instead focuses on a targeted intervention that explicitly aimed at boosting outpatient care, and therefore reducing costs as opposed to increasing quality, and where the power of the financial incentive was very high (Hoof et al. 2019).

Can financial incentives affect hospital behaviour and shift care from a high-cost (inpatient) setting to a low-cost (outpatient) setting? To answer this question, we exploit a unique policy that substantively raised the relative tariff for outpatient care. The policy switched from a regime where the tariff was substantively *lower* for the outpatient care relative to inpatient care, to one where the tariff was *higher* for

[1] The definition of outpatient care varies across countries. We refer to outpatient care as the ambulatory/office based care. Day procedures that are theatre based are considered to be performed in the inpatient setting.

the outpatient care than for inpatient care. After providing a theoretical model, we evaluate the effect of the policy in shifting care to the outpatient setting and whether it affected quality of care, volume, and had spillover effects to other procedures. We also take a regulatory perspective and explore the distribution implications of the scheme between the payer and the provider.

More precisely, we study the effect of the introduction of a Best Practice Tariff (BPT) for outpatient care that rewards providers for treating patients in an office-based ambulatory setting rather than in an inpatient setting. This BPT was one of several others introduced in the English National Health Service to either incentivise quality for specific conditions, such as stroke (Kristensen et al. 2016) or hip fracture (Grasic et al. 2025), or reduce costs while maintaining quality, such as incentivising the discharge of patients within the same day (Gaughan et al. 2019). The scheme for outpatient care was introduced across all hospitals in England in April 2012 and affected three procedures. Two are high volume diagnostic procedures (diagnostic cystoscopy, diagnostic hysteroscopy) and the third is a form of sterilisation for women (hysteroscopic sterilisation). The scheme operates by simultaneously increasing the price reimbursed for the office-based outpatient procedure by a significant amount (up to 370% increase), and, in the case of the two diagnostic procedures, by also lowering the price paid for the procedure when it is performed in the inpatient setting (up to 64% reduction).

We employ difference-in-difference methods to assess whether the implementation of the incentive scheme was successful in increasing the probability of being treated in the outpatient setting (the intensive margin). We also test if the incentive scheme affected volume (the extensive margin), quality of care, as measured by emergency readmissions and repeated procedures, and whether it had positive or negative spillover effects for closely-related but unincentivised procedures, which could be driven by cost synergies or effort diversion, respectively. Our control groups consist of procedures that were selected based on clinical relevance and suitability for treatment in both inpatient and outpatient settings (these are described in detail in Section 4).

The results show that a targeted financial stimulus can result in a swift and substantial change in the choice of the treatment setting. We find a positive and significant effect of the policy on the probability of having the procedure performed in the outpatient setting for all three incentivised procedures, with the largest effect being observed for cystoscopy and hysteroscopy (36.0 percentage points (pp) and 16.3 pp, respectively). The effect is smaller for sterilisation (3.8 pp), possibly due to the inpatient price remaining the same for this procedure in the post-policy period (the inpatient price was instead lowered for cystoscopy and hysteroscopy), and the very low pre-policy proportion of patients treated in the outpatient setting. Heterogeneity analysis also shows that for cystoscopy and hysteroscopy, the response was stronger for hospitals starting with a lower outpatient share. We do not find marked differences in the effect of the policy amongst patients with different sex, age, income deprivation and complexity (proxied by past admissions).

The effect on quality were mixed and depend on the procedure and the measure of quality. For cystoscopy, the results suggest that the policy generally improved quality. It reduced the probability of a repeated procedure within 60 days by 0.4 pp and within 365 days by 5.9 pp, and the probability of a 30-day emergency readmission by 1.0 pp. The results for hysteroscopy are mixed. The policy had no effect on a repeated procedure within 60 days (though the coefficient was positive) while it increased the probability of a repeated procedure within 365 days by 3.7 pp. Instead, it reduced the probability of a 30-day readmission rate by 0.2 pp. For sterilisation, we find no effect on 30-day readmissions. The negative effect on repeated procedures for hysteroscopy can be explained by outpatient hysteroscopies being performed without anesthesia, which can be associated with severe pain, leading to incomplete clinical investigations and the need for repeated procedures (Iaco et al. 2000). Instead, cystoscopy is generally less painful, although discomfort and anxiety may still occurr.

We find no evidence that the policy affected the overall volume (the extensive margin): the effect on patient volume was quantitatively small and statistically insignificant. In contrast, we find evidence of positive spillover effects, with the BPT increasing the probability of being treated in the outpatient setting for the non-incentivised procedures by 7.1 to 12.5 pp.

We also find different distribution implications between the National Health Service (the public insurer or funder) and the hospitals (the providers) depending on the procedure. We show that for cystoscopy and hysteroscopy, the National Health Service (NHS) achieved significant savings despite the larger tariff for the outpatient setting. The savings are sensitive to the relative volume of inpatient and outpatient care. The NHS achieved much larger savings for cystoscopy (about 49 per cent) where volume for inpatient care was relatively higher, and smaller savings for hysteroscopy (about 26 per cent). The savings for the NHS correspond to a reduction in revenues for the hospitals. But given that hospital costs are lower in the outpatient setting, whether hospital profits are lower as a result of the lower revenues is in principle indeterminate. Under some assumptions about reported hospital costs, we show that hospital profits reduced for cystoscopy and hysteroscopy but the effect was quantitatively more important for cystoscopy. (For sterilisation, both the inpatient and the outpatient tariffs increased; therefore, the NHS had larger reimbursement while the provider had higher revenues and profits.)

This research contributes to the limited literature on the effect of financial incentives specifically targeting behavioural changes across hospital settings. This is in contrast to previous studies (reviewed in more detail below) that look at the hospital response across settings following a major pricing overhaul, typically driven by the introduction of the DRG reimbursement system. In addition, the magnitude of the incentive schemes resulting from these policy changes has typically been modest, and the associated changes in behavior have been minimal.

Our paper relates to four strands of literature on the effect of financial incentives in the hospital setting. The first strand focuses on the effect of changes in the DRG prices on volume and quality of care (Gruber and Owings 1996; Yip 1998; Dafny 2005; Gruber et al. 1999; Grant 2009; Clemens and Gottlieb 2014; Papanicolas and McGuire 2015; Januleviciute et al. 2016; Verzulli et al. 2017; Batty and Ippolito 2017; Shin 2019; Milstein and Schreyögg 2024). Exploiting price shocks in reimbursement, these studies generally find that hospitals supply more treatment when the reimbursed price increases or the price of alternative (competing) treatments decreases.

A second strand of the literature studies the effect of the relative price change in the presence of multiple treatments for a given health condition (Foo et al. 2017; Papanicolas and McGuire 2015). These studies find that hospitals increase the proportion of the more profitable treatment, even when not in accordance with clinical guidance (Papanicolas and McGuire 2015). These findings are consistent with our study. However, in most of the existing research the price shock is relatively small (around 3-5%). In our study, the price increase is more dramatic (more than 50% for all procedures).

The third area of the literature pertains specifically to the impact of financial incentives on the substitution between different healthcare settings, with a particular emphasis on outpatient (ambulatory) and inpatient care, which is highly relevant to our study. Leader and Moon (1989) found that the move towards prospective payment in the US led to a large decline in inpatient visits, coinciding with the surge in outpatient appointments[2]. Similar findings were reported by Hadley and Swartz (1989) and Hadley et al. (1989). He and Mellor (2013) examined whether a change in Medicare outpatient payment rates under the Outpatient Prospective Payment System (OPPS) caused outpatient/day case care to shift towards the inpatient setting[3]. The study found a reduction in the number of hernia repair procedures performed in an outpatient setting, while the number of inpatients procedures remained the same. In these studies the shift between healthcare settings resulted from a by-product of a broader change in the payment system, such as the implementation of a DRG prospective payment. In contrast, our study differs in that the price increase was focused specifically on selected procedures with the sole aim of incentivising care in the outpatient setting.

Fourth, our paper contributes to the broader literature on the effects of Pay for Performance (P4P) schemes that are typically aimed at improving quality of care by financially rewarding the implementation of specific care processes, generally linked to best practice, or improvements in health outcomes. Only few P4P schemes incentivise measures of performance related to efficiency (Gaughan et al. 2019). P4P schemes are heterogeneous across countries and conditions and evidence regarding their effectiveness remains limited (Milstein and Schreyoegg 2016; Mendelson et al. 2017; Eijkenaar 2012; Ogundeji et al. 2016; Vlaanderen et al. 2019; Mullen et al. 2010; Zaresani and Scott 2021), with many programs lacking proper evaluation (Milstein and Schreyoegg 2016). Our study provides a comprehensive evaluation of one of the few schemes focusing on efficiency. We also provide a novel theoretical model that allows for patient heterogeneity on both benefits and costs. It differs from previous studies that model heterogeneity along one dimension, typically severity (Ellis 1998; Kifmann and Siciliani 2017). Our model allows for heterogeneity in two dimensions (patient benefits and costs), and provides a flexible framework to model hospital incentives to provide care to different patients.

The study is organised as follows. Section 2 provides the institutional background and Section 3 a theoretical model. Section 4 presents the empirical approach. Section 5 describes the data. Section 6 presents the results. Section 7 discusses distributional implications between the hospitals (the providers) and the public insurer (the funder of healthcare services). Section 8 concludes.

[2] Note that the classification of outpatient activities differs between the USA and UK, with the US definition being broader. Day case procedures that are performed in a theatre setting are classified as inpatient in the UK, while in the USA they are classified as outpatient procedures.

[3] In this case the shift is between day-case and overnight stay, rather than between theatre and non-theatre based settings.

2. Institutional background

The English National Health Service (NHS) is funded by general taxation and it is free at the point of consumption. Patients are registered with a general practitioner (GP). GPs act as gatekeeper and patients require a GP referral to access a hospital specialist. Patients have a legal right to select the hospital, nonetheless choice remains limited (NHS England and Monitor 2015). When recommending a hospital treatment for a given medical condition, providers have to inform the patient about the risks associated with the treatment and offer them possible alternatives[4] (Citizen Advice 2020).

Hospitals provide elective care for patients in both inpatient and outpatient settings. Inpatient care involves patients who are admitted for either an overnight stay or a day case procedure, which requires a hospital bed for tests or surgery but does not require an overnight stay. Outpatient care refers to ambulatory care provided by specialists in an office-based setting. It represents the largest share of NHS contacts in the hospital setting (Royal College of Physicians 2018). In 2018/19, there were over 17 million inpatient admissions (NHS Digital 2019a), and over 123 million outpatient attendances (NHS Digital 2019b).

Hospitals are reimbursed for inpatient care on a per-patient basis, with the tariff based on national average reported costs that are adjusted for local input prices (NHS England and NHS Improvement 2019). The tariff varies by Healthcare Resource Groups (HRGs), the English version of Diagnosis Related Groups (DRGs). Patients are grouped into HRG based on the recorded diagnosis and procedures codes, clinical setting as well as demographic characteristics (e.g. age).

Outpatient care is typically reimbursed for on a per-attendance basis, with the tariff based on the clinical specialty and attendance type, such as first or follow-up attendance. Some procedures performed in the outpatient setting are based on HRGs, using the same HRG codes used for inpatients. However, in most cases, outpatient tariffs are lower than those for inpatient care (NHS England and NHS Improvement 2019).

The *BPT Outpatient scheme* is a pay-for-performance scheme that was introduced in 2012/13 to shift activity from the theatre based inpatient setting to the outpatient office based setting. The BPT includes three procedures: diagnostic cystoscopy, diagnostic hysteroscopy and hysteroscopic sterilisation. These procedures were selected by the Department of Health based on expert clinical advice, further supported by high outpatient rates achieved by a small number of providers before the start of the scheme (Department of Health 2012). The BPT for outpatient care was one of several other BPTs that were introduced in the English National Health Service. Some BPTs aimed at incentivising quality of care for specific conditions, such as stroke or hip fracture. For example, the BPT for stroke gave an additional payment if the patient was admitted in a dedicated stroke unit or given a CT scan (Kristensen et al. 2016). The BPT for hip fracture gave a bundled payment if the hospital provided a care bundle of nine process measures that need to be jointly achieved, such surgery within 36 hours, shared care by surgeon and geriatrician, bone health assessment and others (Grasic et al. 2025). Other BPTs aimed at reducing costs while maintaining quality, such as incentivising the discharge of patients within the same day (Gaughan et al. 2019).

For cystoscopy and hysteroscopy the BPT consisted of two tariffs, one for the outpatient setting and one for inpatient setting. As shown in Table 1, before 2012/13 the tariffs reflected the expected cost, with the

[4] The patient is only entitled to treatment deemed appropriate by the hospital specialist (Citizen Advice 2020).

inpatient tariff being about three times the outpatient tariff. Instead, since 2012/13 the outpatient tariff was at least 60% higher than the inpatient tariff, and therefore set at level which was significantly higher than the expected cost of the outpatient procedure while the inpatient setting was priced below the expected cost. Only providers who sustained a high proportion of outpatient care could break even under this arrangement. For hysteroscopic sterilisation, the BPT substantially increased the tariff for the outpatient setting by about four times in 2012/13, with the inpatient tariff remaining part of the conventional national price setting increasing over the years at a slower rate, by up to 26% in a given year (Department of Health 2012).

To qualify for the outpatient BPT tariff, the procedure must be coded to the outpatient department and performed in a non-theatre based setting with local or no anaesthetic. Procedures in the inpatient department are instead performed in a theatre-based setting, typically under general anaesthetic.

Table 1: National tariff, representing the price [in £] reimbursed to providers for performing the BPT procedures over time.

	Cystoscopy		Hysteroscopy		Sterilisation	
	Inpatient	Outpatient	Inpatient	Outpatient	Inpatient	Outpatient
2010/11	687	274	771	231	771	231
2011/12	714	257	733	242	733	242
2012/13	260	403	260	457	928	1,137
2013/14	251	444	268	472	1,034	1,174
2014/15	246	436	264	465	1,018	1,156
2015/16	248	438	250	498	1,123	1,238

Notes: National tariff represents the price [in £] paid to providers for performing the BPT procedures over time.

Source: NHS England Tariff Workbooks, for years 2010/11-2015/16; available online `https://improvement.nhs.uk/resources/national-tariff/`. Tariff is not readily available for financial year 2009/10, although the tariff structure is analogous to the other years in our sample.

Hysteroscopy and cystoscopy are both established diagnostic tests that are in widespread use across the UK. Cystoscopy is a diagnostic endoscopic procedure involving a telescopic examination of the bladder and urethra using a cystoscope. It is used to check for common problems such as frequent urinary tract infections, long lasting pelvic pain as well as to remove tissue for biopsy and help with the diagnosis and follow up of urogenital cancers (NHS Direct 2020). The procedure is routinely used in both male and female patients. While in many countries cystoscopy is performed in an office based setting (Casteleijn et al. 2017), it was mainly performed as an inpatient procedure in the beginning of our time series with only 12% of all cystoscopies performed in the outpatient setting in 2009/10.

hysteroscopyscopy is also classed as a endoscopic procedure and involves the use of miniaturised endoscopic equipment to directly visualise and examine the uterine cavity. It is primarily used for assessment of abnormal uterine bleeding and investigation of reproductive problems (NHS Direct 2018). While it was historically performed in the inpatient setting, advances in endoscopic technology and ancillary instrumentation have facilitated the move to an outpatient setting with or without the use of local anaesthesia (Yen et al. 2019).

Unlike cystoscopy and hysteroscopy, hysteroscopic sterilisation is a treatment rather than a diagnostic tool. It is one of two main forms of sterilisation procedures for women and it is primarily performed in the outpatient setting. The technique was first introduced in 2001. It involves insertion of titanium (nitinol) metal device into the fallopian tube (Murthy et al. 2017). The alternative method is the inpatient laparoscopic form of sterilisation, historically regarded as the gold standard by which female sterilisation techniques are measured (Greenberg 2008).

All three incentivised procedures are deemed safe when performed in the outpatient setting, albeit not without risk of complications. While generally successful, safe and well-tolerated, outpatient hysteroscopy can be associated with significant pain in up to 35% of women (Iaco et al. 2000; Ahmad et al. 2017), which is also the primary reason for early abandonment of procedure (Ahmad et al. 2017). In the UK, there were several patient groups actively advocating for better pain control and more choice in the selection of suitable hysteroscopy technique. The largest one is the Campaign Against Painful Hysteroscopies (CAPH), which started in 2012/13. Their campaign included notable presence in media; in response to their action, the pain management issue was discussed three times in the national Parliament (CAPH 2018).

While cystoscopy is also associated with discomfort and anxiety, patients generally do not experience extreme pain during the procedure (Greenstein et al. 2014; Falavolti et al. 2017). In a prospective UK study the success rate of outpatient cystoscopy was more than 96%, accompanied by high levels of tolerability and patient satisfaction (Lee et al. 2009).

Hysteroscopic sterilisation is generally associated with less pain and shorter recovery time compared to the inpatient laparoscopic procedure and it is suitable for patients with increased anaesthetic risks associated with laparoscopic technique (Royal College of Obstetricians and Gynaecologists 2016). However there is evidence of higher rates of post-operative complications. While patients undergoing hysteroscopic sterilisation have a similar risk of unintended pregnancy, they have 10-fold higher risk of undergoing the operation a second time compared with patients undergoing laparoscopic sterilisation (Mao et al. 2015). In addition, compared to the laparoscopic technique, hysteroscopic sterilisation is irreversible (Royal College of Obstetricians and Gynaecologists 2016). According to Royal College of Obstetricians and Gynaecologists (RCOG), women should be presented with a comprehensive description of benefits and risks of both techniques (Royal College of Obstetricians and Gynaecologists 2016).

3. Theoretical framework

In this section, we provide a theoretical model of provider incentives when the hospital is reimbursed through multiple prices and patients differ in both benefits and costs. It differs from previous studies that model heterogeneity along one dimension, typically severity (Ellis 1998; Kifmann and Siciliani 2017). Our model allows for heterogeneity in two dimensions (both benefits and costs), and provides a flexible framework to model provider incentives in the presence of patient heterogeneity.

Define p_I and p_O as the tariff reimbursed to a hospital to treat a patient in an inpatient and outpatient setting, respectively. Similarly, define c_I and c_O as the provider costs of treating a patient in an inpatient and outpatient setting. In line with previous literature (Chalkley and Malcomson 1998; Ellis and McGuire 1986), we assume that healthcare providers are altruistic and care about the patient health benefit that arises from treatment. Define α as the level of altruism of the provider, and b_I and b_O as patient benefit from treatment in the inpatient and outpatient setting. We assume that hospital utility U is linear and additive in profit and the altruistic component:

$$U_i = p_i + \alpha b_i - c_i, \quad i = I, O. \tag{1}$$

The provider has to decide whether to treat the patient in an inpatient or an outpatient setting. The provider has an incentive to treat the patient in an inpatient setting if $U_I > U_O$, or more explicitly if:

$$p_I + \alpha b_I - c_I > p_O + \alpha b_O - c_O, \tag{2}$$

which can be written more compactly as:

$$P + \alpha B - C > 0, \tag{3}$$

where $P = p_I - p_O$, $B = b_I - b_O$ and $C = c_I - c_O$.

In line with our institutional setting, we assume that the DRG-type tariffs are based solely on whether the treatment is provided in an inpatient or an outpatient setting and do not vary with other patient characteristics. To model heterogeneity in patients, we assume that each patient differs (i) in the differential health benefit from being treated in an inpatient rather than an outpatient setting, and (ii) in the differential treatment cost between the two settings. Formally, we assume that B and C are distributed according to the joint density function $f(B, C)$ over the support $B \in [\underline{B}, \overline{B}]$; $C \in [\underline{C}, \overline{C}]$.

Figure 1 illustrates the support of this density function, and shows which patients receive the inpatient and the outpatient treatment. In line with the empirical analysis, we assume that pre-policy intervention, the tariff for an inpatient treatment is higher than for an outpatient treatment, $P > 0$. In Figure 1, the provider is indifferent between treating a patient in an inpatient versus an outpatient setting if patients are located on the thick line where the following equality holds, $P + \alpha B = C$. Patients are treated in an inpatient setting if $C < P + \alpha B$, and in an outpatient setting if $C > P + B$. The volume of patients treated in an inpatient (V_I) and outpatient (V_O) setting is given by:

$$V_I = \int_{\underline{B}}^{\overline{B}} \int_{\underline{C}}^{\alpha B + P} f(B, C)\, \partial C\, \partial B, \tag{4}$$

and

$$V_O = \int_{\underline{B}}^{\overline{B}} \int_{\alpha B + P}^{\overline{C}} f(B, C)\, \partial C\, \partial B. \tag{5}$$

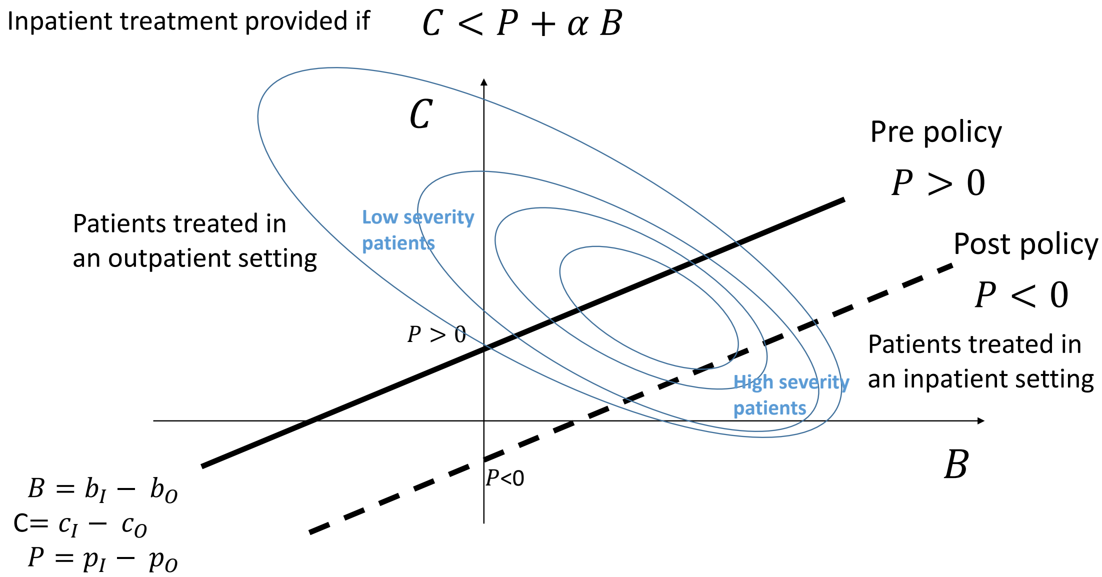


Figure 1: Supply of treatment in the inpatient and outpatient setting

Notes: The figure illustrates which patients receive the inpatient and outpatient patient before and after the policy.

Patients with a relatively high health benefit in an inpatient setting or low difference in treatment costs between the inpatient and outpatient setting tend to be treated in the inpatient setting. For example, patients with certain medical conditions (for example, neurological or cognitive impairments) may have a higher health benefit from being treated in an inpatient setting (e.g. patients may not be able to tolerate the procedure with local anesthesia alone). Treating such patients in an outpatient setting may not be practical or cost-effective, which can limit the potential cost savings associated with outpatient treatment. These patients are presented in the bottom right of the diagram.

In contrast, patients with low severity may not experience significant health losses if they receive treatment in an outpatient setting. They may even prefer the informal environment and the lower risk of infections. In addition, outpatient treatment tends to be less costly. These patients are located in the top left of the diagram. On the other hand, some patients may require inpatient care to obtain high health benefits, but the cost of such care may be significantly higher (top right). Still, others may have low health benefits and low treatment costs, placing them close to the origin.

The policy assessed in the empirical analysis increased the outpatient tariff and reduced the inpatient tariff to a level that the outpatient tariff was higher than the inpatient one after the policy, $P_{Pre} > 0$. Figure 1 also illustrates the difference in the type of patients treated under the pre-policy period, and the post-policy period where the tariff for the treatment in the outpatient setting is higher than in the inpatient setting, $P_{Post} < 0$. The difference in volume in the outpatient setting between the pre- and post-policy period is equal to:[5]

$$V_O^{Post} - V_O^{Pre} = \int_{\underline{B}}^{\overline{B}} \int_{\alpha B + P_{Post}}^{\alpha B + P_{Pre}} f(B, C)\, \partial C\, \partial B = -\left(V_I^{Post} - V_I^{Pre}\right), \quad (6)$$

[5] We can also compute the effect of a *marginal* increase in the price difference between the two settings, which is qualitatively similar. Using Leibniz rule, $dV_O/dP = \int_{\underline{B}}^{\overline{B}} f(\alpha B + P, C)\, \partial C = -dV_I/dP > 0$. Therefore, a higher price differential increases the volume of patients treated in the outpatient setting.

which is what we test in our empirical analysis. Similarly, the effect of the policy on overall patient health and treatment costs is given by

$$\int_{\underline{B}}^{\overline{B}} \int_{\alpha B+P_{Post}}^{\alpha B+P_{Pre}} Bf(B,C)\partial C\partial B, \tag{7}$$

and

$$\int_{\underline{B}}^{\overline{B}} \int_{\alpha B+P_{Post}}^{\alpha B+P_{Pre}} Cf(B,C)\partial C\partial B. \tag{8}$$

If we define welfare as the difference in patients' benefit and provider cost (weighted by the opportunity cost of public funds), the policy is welfare improving if

$$B^{Post} - B^{Pre} - (1+\lambda)(C^{Post} - C^{Pre}) > 0, \tag{9}$$

where λ is the opportunity cost of public funds.

4. Empirical strategy

We employ a difference-in-difference approach to estimate the causal effect of the introduction of the *BPT Outpatient scheme* on providers' behaviour. Our main objective is to assess whether providers increased the proportion of patients treated in an outpatient setting in response to the financial incentive scheme. Additionally, we investigate whether this change in behaviour affected the quality of care received by patients, the total volume of patients treated in either setting, and whether it resulted in a diversion of effort from non-incentivised procedures.

First, we estimate the effect of the BPT policy on the probability of patients being treated in the outpatient setting using the following regression specification:

$$Y_{iht} = \alpha + \beta \ BPT_i + \theta(D_t \ \ BPT_i) + \boldsymbol{X_i'}\boldsymbol{\delta}_1 + (\boldsymbol{X_i'} \ \ BPT_i)\boldsymbol{\delta}_2 + \boldsymbol{v_t} + \boldsymbol{v_h} + \epsilon_{iht} \tag{10}$$

where Y_{iht} is a binary variable taking value of 1 if the patient i in hospital h in month t (ranging from 1, indicating April 2009, to 84, indicating March 2016) is treated in an outpatient setting and 0 if the patient is treated in an inpatient setting. D_t is a dummy variable equal to 1 in the post-policy period (from April 2012 to March 2016), and equal to 0 in the pre-policy period (from April 2009 to March 2012). BPT_i is a dummy variable equal to 1 if the patient receives an incentivised procedure and equal to 0 if the patient receives a non-incentivised (control) procedure. $\boldsymbol{v_t}$ is a vector of year-by-month time fixed effects, which allows to control flexibly for the time trend (adding 83 dummies, one for each month in our sample period excluding the reference month). $\boldsymbol{v_h}$ is a vector of hospital fixed effects to control for time-invariant hospital factors. $\boldsymbol{X_i}$ is a vector of patient characteristics, described in more detail in section 5. We allow the effect of patient characteristics to differ across the incentivised and the non-incentivised procedure (the treatment and the control group). α is the intercept and ϵ_{iht} is the error term. We estimate (10) as a linear probability model with standard errors clustered at the hospital level. The key coefficient of interest is θ, which gives the average treatment effect on the treated over the post-policy period. In economic terms, it provides an estimate of the effect of the financial incentive scheme on the probability of patients being treated in the outpatient rather than in the inpatient setting.

Since the BPT scheme was implemented simultaneously across all hospitals in the English NHS, we identify control groups amongst non-incentivised procedures that can be performed in both inpatient and outpatient settings. To select control (non-incentivised) procedures for each of the incentivised procedures, we proceed as follows. First, we identify procedures (HRGs) that can be performed in both the inpatient and outpatient setting and hence have a separate HRG tariff in 2015/16 (242 conditions meet this criterion). Second, we include either procedures that can be performed in the urology or gynecology department, as these are the departments where the BPT conditions are performed, or are diagnostic endoscopies performed in other departments, such as colonoscopy, since they share many similarities with cystoscopy and hysteroscopy, including the use of endoscopes and similar pain relief techniques. Applying the two restrictions already leaves us with only 32 possible control groups. Third, we test for the parallel trend assumption between clinically-relevant treatment and control procedures (presented in section 6), which reduces the possible control groups to seven procedures. Last, we further evaluate their clinical relevance. Specifically, we seek medical advice regarding any adverse effects of performing the procedure in the outpatient setting, including severe pain. Additionally, we assess whether the choice between the inpatient and outpatient setting could have been affected by new technologies or new guidelines over our study period. Our final controls groups are: sigmoidoscopy (for cystoscopy), lower genital procedures (for hysteroscopy) and vacuum aspiration with cannula (for sterilisation). We provide more details about these procedures in section 5.

Equation (10) above estimates the average effect of the policy across the entire post-policy period. To test if the effect differs across years, we also use the following flexible specification:

$$Y_{iht} = \alpha + \beta\ BPT_i + \boldsymbol{\theta}\ (\boldsymbol{v_t'}\ BPT_i) + \boldsymbol{X_i'}\boldsymbol{\delta}_1 + (\boldsymbol{X_i'}\ BPT_i)\boldsymbol{\delta}_2 + \boldsymbol{v_t} + \boldsymbol{v_h} + \epsilon_{iht} \tag{11}$$

where $\boldsymbol{v_t}$ is a vector of binary year indicators for each financial year across both the pre- and post-policy period (2010/11 to 2015/16, with 2009/10 as the reference year). $\boldsymbol{\theta}$ is a vector of estimates for changes in the probability of being treated in the outpatient setting in each financial year in the study. We expect these not to be statistically significantly different from zero in the pre-policy years. As before, we estimate (11) as a linear probability model with standard errors clustered at the hospital level.

While the main objective of the *BPT Outpatient scheme* is to shift procedures from the outpatient to the inpatient setting, incentive programs may influence other aspects of patient care. To gain a better understanding of the overall impact of the BPT scheme, we explore three additional outcome measures.[6]

First, we analyse the impact of the scheme on quality of care. In particular, diagnostic endoscopic procedures (including hysteroscopy and cystoscopy) are frequently associated with severe pain when performed without anaesthesia (as it is generally the case in the outpatient setting), resulting in a failure to complete the clinical investigation. In this case the procedure has to be repeated at a later date[7], causing additional stress and inconvenience for the patient and extra demand on the healthcare system. The rate of failed procedures is a common health outcome measure for endoscopic procedures (Relph et al. 2016; Genovese et al. 2020). Since we cannot observe this metric directly in our data, we measure instead whether the procedure was repeated at a later date. This relies on the assumption that the per patient rate of other factors for failed operation (anatomical factors and structural abnormalities) remains constant throughout the period. We estimate the effect of the policy on the probability to have a repeated procedure within the following 60 or 365 days of the initial procedure[8]. In this case the dependent variable Y_{iht} takes the value 1 if the patient had the repeated procedure within the next 60/365 days and 0 otherwise. As pain is likely to be more prominent for hysteroscopy than cystoscopy, a priori we expect larger effect of the policy on repeated procedures for hysteroscopy. We also investigate the effect of the policy on emergency readmissions within 30 days from discharge for the three procedures. We estimate the effects of the BPT on patient outcomes using the same regression specification as in (10), with the same independent regressors.

Second, we investigate the effect of the financial scheme on total patient volume. The BPT significantly changes the profitability of the incentivised outpatient procedures. While this motivates providers to shift patients from the inpatient to the outpatient setting and hence the *intensive* margin, it could also affect the *extensive* margin, by inducing providers to offer the procedure to patients who would not otherwise be eligible according to their medical need. These effects can thus contribute to an increase in the total volume of procedures and affect the overall NHS healthcare costs. We analyse the effect of the BPT on

[6] sterilisation differs from the two diagnostic procedures in several aspects. Unlike hysteroscopy and cystoscopy, sterilisation is considered a treatment rather than a diagnostic procedure, it is not performed in high volumes, and it is associated with greater patient involvement in decision-making. Therefore, we do not consider the selected additional outcome measures as appropriate for this particular condition.

[7] It has been shown that the pain experienced during the procedure is the leading cause of repeated treatments, accounting for over 80% of all cases (Ahmad et al. 2017). However, occasionally procedures have to be repeated when they are originally performed in the inpatient setting, for a variety of reasons including incomplete view of the tissue.

[8] We use different time points for two reasons. First, several cancer treatments specify that patients should have recurrent endoscopic investigations every three months. We therefore include procedures done within 60-days, as patient's are unlikely to need repeated investigation in this period on medical grounds. Second, we include repeated procedures within 365-days to take into account potential waiting time for the procedure.

total volume (inpatient and outpatient cases combined) using quarterly data reported at the provider level using the following specification:

$$Y_{jhq} = \alpha + \beta\ BPT_j + \theta(D_t\ \ BPT_j) + \boldsymbol{X'_{jhq}\delta_1} + (\boldsymbol{X'_{jhq}}\ \ BPT_j)\boldsymbol{\delta_2} + \boldsymbol{v_t} + \boldsymbol{v_h} + \epsilon_{jhq} \quad (12)$$

where Y_{jhq} is the total number of procedures j performed in hospital h in quarter $q = 1, \ldots, 21$, where the quarter 1 corresponds for the period April-June 2009 whereas the last quarter 21 corresponds to the period Jan-March 2016. BPT_j equals to 1 if j is an incentivised procedure and 0 if it is a control procedure. D_t is a dummy variable equal to 1 in the post-policy period (from April 2012 to March 2016), and equal to 0 in the pre-policy period (from April 2009 to March 2012). X_{jhq} is a vector of patient characteristics for condition j averaged on a quarterly basis by provider. $\boldsymbol{v_t}$ is a vector of year-by-quarter dummies (with quarters being defined as April-June, July-September, October-December, January-March). Our key coefficient θ measures whether volume increased more quickly for the incentivised procedure than for the control one. We estimate (12) using OLS.

Third, we investigate whether the BPT policy had unintended spillover effects on closely related, but unincentivised procedures, specifically examining the probability of selecting the outpatient setting for procedures that are clinically similar to the BPT procedures but were not incentivised. We are interested in understanding whether the BPT policy had both negative or positive spillover effects. On one hand, limited capacity to treat patients in the outpatient setting could reduce and crowd out the ability of the provider to perform outpatient activity for other procedures, a form of negative spillover. On the other hand, the incentivised BPT could induce providers to invest in outpatient capacity, therefore reducing the marginal cost of treating patients in an outpatient setting even for non-incentivised procedured, a form of positive spillover.

For this estimation we select two procedures that are very similar to the incentivised hysteorsopy and cystoscopy, but do not attract the bonus: we use hysteroscopy with insertion of uterine device to test spillover effects of the BPT for hysteroscopy, and we use endoscopic urethra procedures to test the spillover effects for cystoscopy. In our estimation we use the same empirical approach as in (10) and the same control groups. A negative (positive) coefficient θ here would provide evidence of negative (positive) spillover effects, implying that the BPT reduced (increased) the probability of providing the procedure in an outpatient setting for a closely-related procedure that was not incentivised by the BPT.

All the analyses described in this section use a difference-in-difference approach and rely on the parallel trends assumption. We test the plausibility of the parallel trend assumption empirically by estimating model (11) only for the period prior to the BPT introduction (2009/10-2011/12), and testing if the coefficients associated with the year dummies interacted with the treatment group are statistically significantly different from zero. We perform this test separately for each outcome and procedure.

5. Data

We employ data from the Hospital Episode Statistics (*HES*), which separately collects information on inpatient and outpatient care. *HES Outpatients* is a dataset comprised of all office-based consultations and procedures, detailing information about the patient's hospital visit, including their socio-demographic data (ie. age, gender, income deprivation), whether the visit was patient's first attendance, and OPCS[9] codes of any procedures[10] carried out during the appointment. *HES Inpatients* includes all inpatient and day-case admissions, detailing information on patient's care pathway, including admission and discharge date, type of admission (elective or non-elective), patient's diagnosis (ICD) and procedure (OPCS) codes, and socio-demographic data.

Our study period is from April 2009 to March 2016, with the pre-policy period running from April 2009 to March 2012. Our sample consists of all patients aged 19 or older who, during this period, had either a BPT procedure (our treatment group) or any of the procedures that we use to construct the control groups. Our sample across the three BPTs and the corresponding control groups consists of 5,723,343 observations[11]. Additionally, we have 235,227 observations for the two procedures used to test for possible spillover effect[12]. Including those, the full sample consists of 5,958,570 observations, of which 3,747,670 (62.9%) are in the inpatient setting and 2,210,900 (37.1%) in the outpatient setting.

Patients are placed into the treatment and control group based on their assigned HRG codes (English equivalent to the DRG codes), mirroring the method used by the BPT scheme for payment purposes. The English HRG system is frequently updated, with groups added and removed on a yearly basis. To create consistent series throughout the study period, we use the NHS Digital Grouper software[13] to assign coherent set of HRG groups across all years. The grouper software is developed by the NHS for payment and benchmarking purposes and uses patient's clinical and demographic characteristics to assign the HRG group.

5.1. Outcome measures

We estimate the effect of the BPT scheme on several outcome variables. First, our main dependent variable is a categorical variable that equals one if the patient was treated in the outpatient setting and zero if treated in the inpatient setting. Second, we construct a measure of *hospital quality*, which is proxied by the probability of the patient receiving the same procedure again within the next 60 days and 365 days (referred to as the "repeated procedure" measure). We create a categorical variable that equals one if the patient received the same procedure (at least once) within 60 or 365 days from the original procedure. We assign a value of 1 if the number of repeated procedures is one or more, and 0 otherwise. We also measure quality as 30-day readmission rates for the three procedures. We create a

[9] OPCS Classification of Interventions and Procedures, first published by Office of Population Censuses and Surveys, is the classification of procedures used by clinical coders within the NHS. It provides codes for operations, procedures and interventions performed during inpatient stays, day case surgery and outpatient treatments in the NHS hospitals. While the codes themselves are different, the code set is comparable to the American Medical Association's Current Procedural Terminology (NHS Digital, 2020).

[10] Diagnosis code fields are included in the dataset, however they are only available for less than 1% of all consultation visits and are hence not used in our analysis.

[11] Out of 5,723,343 observations, just over 3 million correspond to the BPT incentivised procedures: 2,359,964 are for cystoscopy, 558,618 for hysteroscopy and 116,216 for sterilisation. The rest of the observations make up the control groups: 1,499,406 are for sigmoidoscopy, 1,002,097 for lower genital procedures and 187,042 for vacuum aspiration with cannula.

[12] The two control groups for spillover effect are urethra procedure (121,286 observations) and hysteroscopy with insertion of uterine device (113,941 observations).

[13] We use Payment Grouper version 2016, freely available to download from the following link `https://webarchive.nationalarchives.gov.uk/20171011074955/http://content.digital.nhs.uk/article/2063/Archive-payment`

categorical variable equal to one if the patient was readmitted as an emergency patient within 30 days from hospital discharge. Third, we measure *volume* as the total number of procedures performed across both the inpatient and outpatient setting. This variable is measured at provider level for each quarter and we only include providers present in all quarters of the study period[14] (131 out 167 of for cystoscopy and 132 out of 167 for hysteroscopy). The provider exclusion only applies to the estimation of volume; in all other estimations we use the full sample.

5.2. Control variables

We control for patients' clinical and socio-demographic characteristics, including age (measured as a categorical variable with 5-year bands and two separate categories for 19 to 24 and 90+), sex (male=1) and the number of past emergency hospital visits in the year prior to the procedure (measured as a categorical variable with values from 0 to 4 and 5+) as a proxy of patient severity[15]. As a proxy of socio-economic status, we use the income deprivation score of the English Indices of Deprivation 2010 for patient local area of residence (coded in quintiles).

5.3. Control group procedures

Procedures in the control group were selected based on their comparability to procedures in the treatment group in both inpatient and outpatient settings, their clinical relevance and whether they exhibited a parallel trend with the treatment group in the pre-policy period. The three control groups for the main outcomes are flexible sigmoidoscopy, lower genital procedure and vacuum aspiration with cannula.

Flexible Sigmoidoscopy is used as a control procedure for the BPT procedure diagnostic cystoscopy and is used to evaluate the lower part of the large intestine. Like cystoscopy, it is an endoscopic procedure (with a sigmoidoscope) used as a screening tool to detect polyps and cancerous cell. It can be safely performed in an office based setting with or without the use of pain relief. Similar to cystoscopy, pain is considered one of the main reasons for a failed procedure (Doria-Rose et al. 2005). It is mainly performed in the gastroenterology department, while the diagnostic cystoscopy is performed in the urology department. Thus the chance of a spillover effect across the two procedures is minimal (Kelly et al. 2008).

Lower genital procedure is used as a control procedure for hysteroscopy. It includes an array of procedures grouped to a common HRG group, including procedures of the Bartholian gland[16] (drainage/balloon catheter insertion) and procedures of the vulva. These procedures can typically be performed in the office based setting using local anaesthetic. While all of procedures in this group are performed in the gynaecology department, they tackle unrelated gynaecological problems.

Vacuum aspiration with cannula is used as a control procedure for hysteroscopic sterilisation. This is a safe and effective alternative method for surgical management of miscarriage. It can be performed in the

[14] There are two main reasons why hospital do not report in all quarters: (i) hospital mergers; (ii) very small numbers of cases, with frequent zeros. We are more concerned with the former. Several providers merged during the period, which affects the total volume; to avoid this issue, we exclude these providers from the analysis.

[15] Because the diagnosis information is not available for outpatient attendances, we cannot construct the Elixhauser/Charlson index to control for severity. As an alternative we use the past number of emergency admissions (in the year prior) as a proxy for patient severity.

[16] Bartholin's cyst usually appears as a lump in the genital area; it can become painful and infected, in which case it needs treatment (drainage in the first instance).

outpatient setting under local anaesthesia. While vacuum aspiration and sterilisation are very distinct procedures, they share several similarities: (i) both procedures were typically performed in the inpatient setting with effective and safe alternatives for surgical management emerging relatively recently (Sharma 2015); (ii) both procedures require strong patient involvement in the clinical process, including giving the relevant information and support[17]; (iii) patient demographic is similar across two procedures (women of childbearing age).

5.4. Spillover effects

We also test whether the BPT policy had spillover effects on the probability of being treated in the outpatient setting for clinically similar procedures that were not part of the BPT scheme. One concern is that incentivising an increase in the outpatient setting for one procedure may crowd out and reduce the probability of being treated in the outpatient setting for non-incentivised procedures, resulting in a negative spillover effect.[18]

To test for spillover effects for cystoscopy, we examine the *Endoscopic urethra procedure*. This collection of procedures is typically performed alongside cystoscopy and is grouped under a single HRG. Examples include retrograde pyelogram, which involves introducing contrast dye into the urinary system during cystoscopy, and endoscopic urine sampling, which involves taking a urine sample during cystoscopy to check for tumors and infections.

To test for spillover effects for hysteroscopy, we consider the *Hysteroscopy with insertion of intrauterine device*. In this case, hysteroscopy is followed by the insertion of an intrauterine device, which is generally a straightforward, office-based procedure.

For this analysis, the dependent variable is again a binary variable that equals one if the patient was treated in an outpatient setting rather than an inpatient setting. Due to the similarities between the incentivised conditions and the non-incentivised conditions where spillover effects may occur (in terms of trends and clinical similarity), we use the same control groups as in the main analysis. Specifically, we use sigmoidoscopy as the control group for the endoscopic urethra procedure, and lower genital procedures as the control group for hysteroscopy with insertion of intrauterine device.

5.5. Descriptive statistics

Table 2 presents the sample mean and standard deviation (SD) for each of the outcomes in the pre- and post- policy period (2009/10-2011/12 and 2012/13-2015/16, respectively). For the three incentivised conditions, we observe sizeable difference across the two periods in the proportion of patients treated in the outpatient setting. The outpatient rates for cystoscopy increased from 13% in the pre-policy period to 52%. Similarly, for hysteroscopy it increased from 41% to 62%.[19] The outpatient rate for sterilisation increased from 0% to 5%. Instead, for the control groups the increases in outpatient rates are much smaller. For sigmoidoscopy, the control group of cystoscopy, it increased from 13% to 15%. For

[17] NICE guidance requires providers to provide all the necessary information and give support to women who experience miscarriage: `https://tinyurl.com/yxpc5ed3`

[18] It is also possible that the effect could go in the opposite direction, leading to a positive spillover effect and a faster transition from an inpatient to an outpatient setting.

[19] Department of Health/NHS estimated maximum achievable rate is considerably lower for cystoscopy (50%) than for hysteroscopy (80%) (Department of Health 2012).

lower genital procedure, the control group of hysteroscopy, it increased from 74% to 82%. For vacuum aspiration, the control group for sterilisation, the outpatient rate increased from 1% to 2%.

Table 2: Descriptive statistics. Outcome variables

(a) Cystoscopy, corresponding control group & procedure to test for spillover effects

	Cystoscopy				Sigmoidoscopy				Urethra procedure			
	Pre-policy		Post-policy		Pre-policy		Post-policy		Pre-policy		Post-policy	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Proportion treated as outpatient	0.13	0.33	0.52	0.50	0.13	0.34	0.15	0.36	0.03	0.18	0.13	0.34
Volume (on quarterly basis)	77,292	2,973	89,529	4,526	46,622	3,864	58,746	4,268	4,029	225	4,558	213
Reoperation within 60 days	0.05	0.21	0.06	0.23	0.03	0.17	0.04	0.20	0.03	0.17	0.03	0.17
Repeated procedure within 365 days	0.36	0.48	0.31	0.46	0.09	0.29	0.10	0.30	0.20	0.40	0.20	0.40
Emergency readmission within 30 days	0.03	0.16	0.02	0.12	0.02	0.15	0.02	0.14	0.06	0.24	0.06	0.23

(a) Hysteroscopy, corresponding control group & procedure to test for spillover effects

	Hysteroscopy				Lower genital procedures				Hysteroscopy with ID			
	Pre-policy		Post-policy		Pre-policy		Post-policy		Pre-policy		Post-policy	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Proportion treated as outpatient	0.41	0.49	0.62	0.48	0.74	0.44	0.82	0.38	0.20	0.40	0.38	0.49
Volume (on quarterly basis)	18,781	435	20,828	1,211	35,196	1,040	36,234	1,355	4,340	186	3,866	113
Repeated procedure within 60 days	0.04	0.19	0.05	0.21	0.07	0.25	0.07	0.26	0.00	0.05	0.00	0.05
Repeated procedure within 365 days	0.07	0.26	0.08	0.27	0.35	0.48	0.36	0.48	0.01	0.10	0.01	0.10
Emergency readmission within 30 days	0.01	0.10	0.01	0.08	0.01	0.08	0.00	0.07	0.02	0.12	0.01	0.11

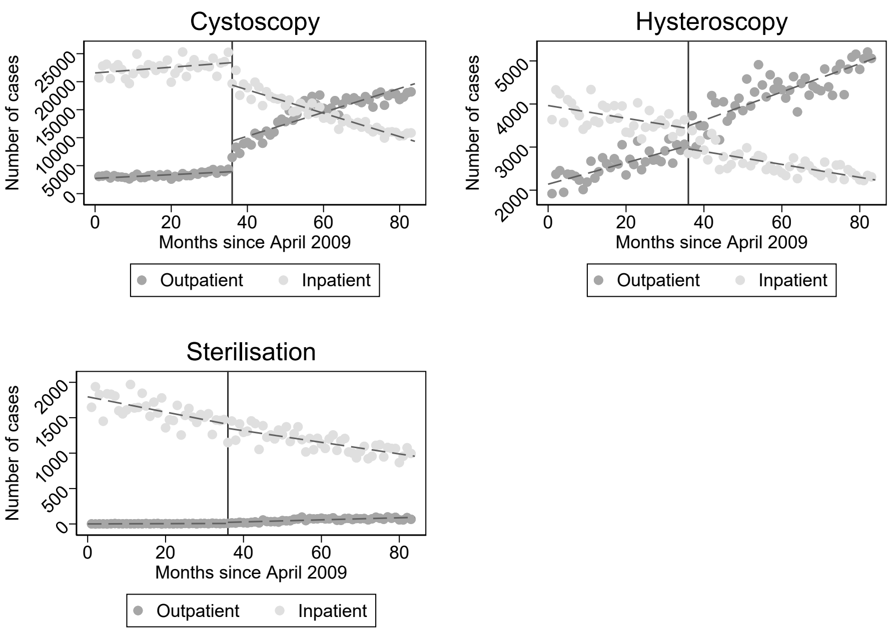
(a) Sterilisation & the corresponding control group

	Sterilisation				Vacuum aspiration			
	Pre-policy		Post-policy		Pre-policy		Post-policy	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Proportion treated as outpatient	0.00	0.05	0.05	0.21	0.01	0.09	0.02	0.13
Volume (on quarterly basis)	4,822	355	3,647	301	6,934	238	6,489	303
Emergency readmission within 30 days	0.03	0.16	0.03	0.16	0.03	0.16	0.03	0.16

Notes: This table shows the sample mean and standard deviation (SD) for the outcomes across treatment/control procedures in the pre-and post-policy period (2009/10-2011/12 and 2012/13-2015/16, respectively). The table shows the proportion of patients treated in the outpatient setting and of having a repeated procedure within 60 or 365 days and emergency readmission within 30 days. For volume, the table presents the average number of patients each quarter in the pre/post-policy period.

Figure 2 displays the monthly volume of the BPT procedures across the inpatient and outpatient setting. For cystoscopy, there is a rapid decline in inpatient activity after the start of the policy in April 2012 and concurrent increase in the number of outpatient cases. This is also the case for hysteroscopy, though it is less pronounced, with the volume in the inpatient setting already trending downwards prior to the start of the policy. The volume of outpatient hysteroscopic sterilisation increases slowly after the start of the policy and remains low throughout.

Figure 2: Volume of BPT procedures over time in an outpatient and inpatient setting



Notes: Number of inpatient and outpatient procedures measured on a monthly basis for the three incentivised BPT procedures: diagnostic cystoscopy, diagnostic hysteroscopy and sterilisation. Time period is from April 2009 to March 2016. The vertical line indicates the start of the BPT policy in April 2012. The scale (y-axis) differs between the three graphs, hence the figures are not directly comparable.

Table 2 also provides descriptive statistics for the quality measures. In the pre-policy period, the proportion of patients with a 30-day emergency readmission is 3% for cystoscopy and sterilisation, and 1% for hysteroscopy. These are similar for the post-policy period. About 5% (36%) of patients require a repeated cystoscopy within 60 days (365 days) in the pre-policy period, and 6% (31%) in the post-policy period. For hysteroscopy, 4% (7%) of patients require a repeated procedure within 60 days (365 days) in the pre-policy period, and 5% (8%) in the post-policy period. The total volume, across the inpatient and outpatient setting, averaged across quarters and providers increases over time for both cystoscopy and hysteroscopy (and their control groups) but reduces for sterilisation (and its control group).

Table 3 presents descriptive statistics for the patients characteristics in our final sample. While we observe variations across procedures, the patients composition is relatively similar between the treatment and control groups. We briefly comment on the patient characteristics in the treatment group. Patients treated for cystoscopy are on average 66 years old, 64% are male, and 83% did not have any emergency admissions in the previous year. 28% of patients are in the most deprived quintile. Patients treated for hysteroscpy are on average 53 years old, and 92% had no emergency admissions in the past year. 27% of patients are in the most deprived quintile. Patients treated for sterilisation were on average 35 years old, and 92% had no emergency admissions in the past year. Differently from hysteroscopy and sterilisation, 28% of patients are in the least deprived group.

Table 3: Descriptive statistics. Patient characteristics

(a) Cystoscopy, the corresponding control group & the procedure to test the spill-over effect

	Cystoscopy		Sigmoidoscopy		Urethra procedures	
	Mean	SD	Mean	SD	Mean	SD
Age	66.17	15.25	58.43	16.99	58.53	16.88
Male	0.64	0.48	0.48	0.50	0.55	0.50
Deprivation quintiles						
Most deprived	0.28	0.45	0.28	0.45	0.25	0.43
2nd quintile	0.18	0.39	0.18	0.38	0.17	0.38
3rd quintile	0.19	0.39	0.19	0.39	0.19	0.39
4th quintile	0.18	0.39	0.19	0.39	0.20	0.40
Least deprived	0.17	0.38	0.17	0.38	0.20	0.40
Past emergencies						
0	0.83	0.37	0.86	0.35	0.57	0.49
1	0.04	0.19	0.04	0.20	0.05	0.22
2	0.04	0.19	0.04	0.20	0.13	0.33
3	0.02	0.14	0.02	0.14	0.07	0.26
4	0.02	0.12	0.01	0.11	0.05	0.22
5+	0.05	0.23	0.03	0.16	0.12	0.33
Observations	2,359,964		1,499,406		121,286	

(c) Hysteroscopy, the corresponding control group & the procedure to test the spill over-effect

	Hysteroscopy		Lower genital procedures		Hysteroscopy with ID	
	Mean	SD	Mean	SD	Mean	SD
Age	52.71	13.01	55.20	20.87	44.17	7.69
Male	-		-		-	
Deprivation quintiles						
Most deprived	0.27	0.44	0.23	0.42	0.25	0.43
2nd quintile	0.17	0.38	0.17	0.38	0.17	0.37
3rd quintile	0.18	0.38	0.19	0.39	0.18	0.39
4th quintile	0.19	0.39	0.20	0.40	0.20	0.40
Least deprived	0.19	0.40	0.21	0.41	0.21	0.40
Past emergencies						
0	0.92	0.27	0.90	0.30	0.93	0.25
1	0.03	0.16	0.02	0.14	0.03	0.16
2	0.02	0.15	0.02	0.13	0.02	0.14
3	0.01	0.10	0.01	0.10	0.01	0.10
4	0.01	0.08	0.01	0.08	0.01	0.07
5+	0.01	0.10	0.04	0.20	0.01	0.08
Observations	558,618		1,002,097		113,941	

(c) Sterilisation & the corresponding control group

	Sterilisation		Vacuum aspiration with cannula	
	Mean	SD	Mean	SD
Age	34.66	8.37	29.75	6.71
Male	-		-	
Deprivation quintiles				
Most deprived	0.18	0.39	0.19	0.39
2nd quintile	0.14	0.35	0.14	0.35
3rd quintile	0.17	0.38	0.17	0.38
4th quintile	0.22	0.42	0.22	0.41
Least deprived	0.28	0.45	0.28	0.45
Past emergencies				
0	0.92	0.28	0.86	0.35
1	0.03	0.17	0.09	0.29
2	0.03	0.17	0.03	0.17
3	0.01	0.10	0.01	0.10
4	0.01	0.08	0.00	0.07
5+	0.01	0.09	0.01	0.07
Observations	116,216		187,042	

Notes: The table shows the descriptive statistics (mean and standard deviation, SD) for the patients' characteristics in the regression sample across all study years (2009/10-2015/16). Age is patient's age at the time of admission. Deprivation quintiles are based on the continuous IMD index of income deprivation that takes value from 0 (least deprived) to 1 (most deprived). "Past emergencies" measure the number of emergency admissions in the year prior to the procedure.

6. Results

6.1. Main effects

In this section, we present our findings on whether the introduction of the BPT policy increased the probability of being treated in an outpatient setting. The results of our difference-in-difference analysis are presented in Table 4 (see Table A1 for full regression results). The DiD estimates show a sizeable, positive and statistically significant effect (at 1% level) of the *BPT Outpatients Scheme* on the probability of being treated in the outpatient setting for all three BPT procedures. The effect is larger for cystoscopy, and equal to 36.0 percentage points (pp), while it is 16.3 pp for hysteroscopy. The effect is smaller and equal to 3.8 pp for hysteroscopic sterilisation.

Table 4: Difference-in-difference estimates of the impact of the BPT outpatient policy on patient probability of being treated in the outpatient setting

Treatment group	Cystoscopy (1)	Hysteroscopy (2)	Sterilisation (3)
DiD coefficient	0.360***	0.163***	0.038***
	(0.031)	(0.022)	(0.010)
Adjusted R^2	0.386	0.295	0.093
Number of hospitals	168	167	158
Observations	3,859,365	1,560,714	303,256

Notes: The dependent variable is the probability to be treated in the outpatient setting. Time period is from April 2009 to March 2016. Models are estimated by OLS with standard errors (presented in parenthesis under the coefficients) clustered at hospital level. Models are run separately for each treatment-control procedure pair (cystoscopy-sigmoidoscopy; hysteroscopy-lower genital procedures; sterilisation-vacuum aspiration). All models control for casemix and a set of year-by-month and hospital fixed effects.

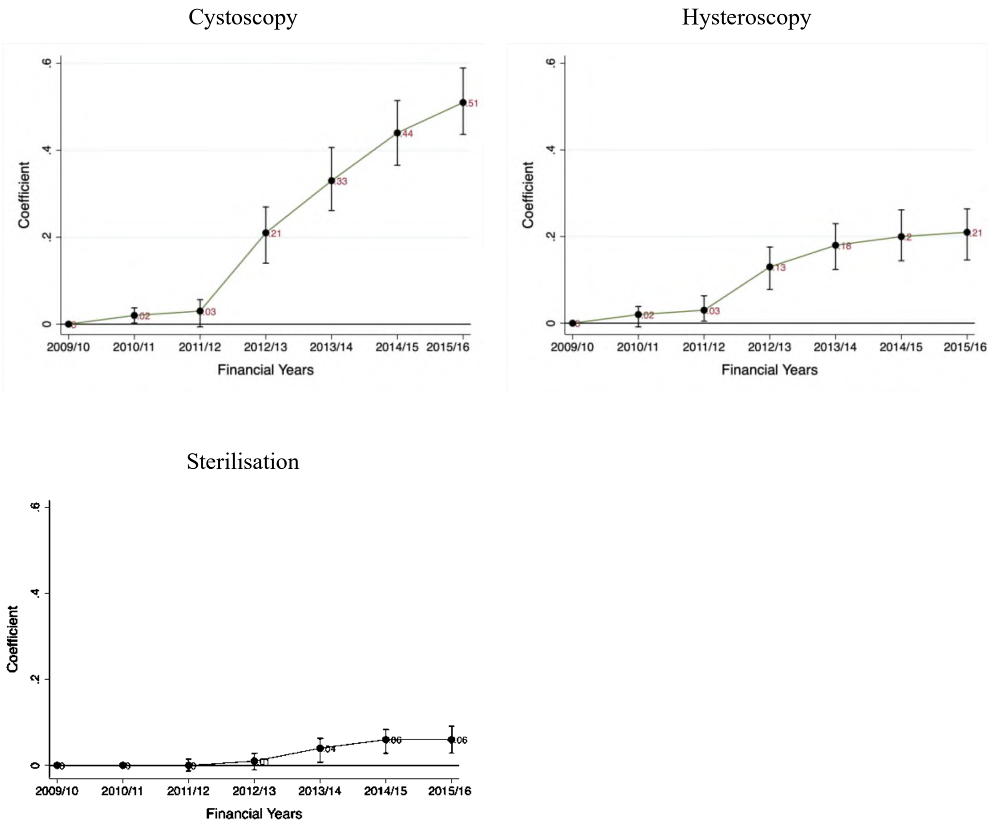
*** p<0.01, ** p<0.05, * p<0.1

Figure 3 shows the effect of the BPT scheme on being treated in the outpatient setting for each financial year both pre and post policy (using the first year of the pre-policy period as reference group) with 95% confidence intervals. The effect is close to zero in the pre-policy years (2010/11 and 2011/12) and is increasing over time for all three conditions in the post-policy years. For hysteroscopy the effect is concentrated in the first year of the BPT. For cystoscopy and sterilisation the increase in the uptake in the outpatient setting is more gradual over time.

The effect differs across the three conditions. While cystoscopy and hysteroscopy both had similar estimated achievable rates (Department of Health 2012), the starting point is different across the two conditions: cystoscopy had a much lower probability of outpatient treatment in the pre-policy period compared to cystoscopy (13% vs 41%). Furthermore, the outpatient treatment for hysteroscopy received negative publicity (CAPH 2018), which might have negatively affected the uptake rates. The effect is small for sterilisation. Compared to the other two procedures, for which the inpatient prices decreased after the introduction of the BPT, they inpatient prices increased for sterilisation. The small increase might be due to the relative novelty of the outpatient sterilisation technique (Murthy et al. 2017) and increased involvement of the patient in the decision making process (Royal College of Obstetricians and Gynaecologists 2016).

For cystoscopy and hysteroscopy, we can also compute elasticities. The elasticity of outpatient volume (as a proportion of total volume) to a relative price increase is 0.71. This is computed as $\frac{d[V_O/(V_0+V_1)]/[V_O/(V_0+V_1)]}{d(p_O/p_1)/(p_O/p_1)} = \frac{2.77}{3.92} = 0.71$. The elasticity for hysteroscopy was instead much lower and

Figure 3: Probability of being treated in the outpatient setting by year



Notes: Effect of the BPT across financial years relative to 2009/10 with 95% confidence interval.

equal to 0.08 ($\frac{0.40}{4.77}$). We do not compute the elasticity for sterilisation because the price of outpatient care was zero before the policy was introduced.

6.1.1. Parallel trends assumption

Table 5: Test for the parallel trends assumption for the probability of being treated in the outpatient setting

Treatment group	Cystoscopy (1)	Hysteroscopy (2)	Sterilisation (3)
DiD coefficient for 2010/11	0.018**	0.016	-0.000
	(0.009)	(0.011)	(0.002)
DiD coefficient for 2011/12	0.021	0.035**	0.001
	(0.016)	(0.015)	(0.006)
Adjusted R^2	0.466	0.374	0.122
Number of hospitals	167	159	155
Observations	1,486,961	647,729	141,069

Notes: Dependent variable is the probability to be treated in the outpatient setting. Time period is from April 2009 to March 2012. Standard errors (presented in parenthesis under the coefficients) are clustered at hospital level. Models are run separately for each treatment-control procedure pair (cystoscopy-sigmoidoscopy; hysteroscopy-lower genital procedures; sterilisation-vacuum aspiration). All models include a constant, case mix variables and a full set of year-by-month and hospital dummies. The null hypothesis for the parallel trends assumption is that the DiD coefficients are jointly zero.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

The estimates from the difference-in-difference analyses rely on the parallel trends assumption. We test for this assumption in Table 5. We restrict the sample to the pre-policy years, and test whether the probability of being treated in an outpatient setting differs between the treatment and the control group in 2010/11 and 2011/12 relative to the first year of the policy in 2009/10. The coefficients are very small, especially compared to the effects of the policy (presented in Table 4) and mostly statistically insignificant. These results are also consistent with Figure 3 that covers the whole sample period, and again show very little differences in the probability of being treated in the outpatient setting in the pre-policy years.[20] Overall, these findings suggest that the parallel trend assumption is likely to hold.

6.1.2. Sensitivity analyses

We perform additional sensitivity analyses to confirm the robustness of our estimates. We first adjust for the transition period, by excluding from the analysis the period within 6 months on either side of the start of the policy in April 2012. This controls for possible anticipatory effect or delayed transition after the start of the policy. Using the same specification as in the main model, we find that the estimates are similar. Table A2(a) in the Appendix shows that the effect of the policy increases from 36.0 pp to 37.9 pp for cystoscopy, from 16.3 pp to 18.0 pp for hysteroscopy and from 3.8 pp to 4.3 pp for sterilisation. All of the coefficients are statistically significant.

We further control for changes in the number of hospitals in the sample by excluding those hospitals that did not report data in all quarters of our study period (therefore using a balanced panel). By doing so, we remove any hospitals that merged over the sample period, as this could affect their medical practice. In this case the estimates are slightly smaller: 35.7 pp for cystoscopy, 15.2 pp for hysteroscopy and 4.0 pp for sterilisation (see Table A2(b) in the Appendix). All coefficients are statistically significant.

One possible concern is that the policy itself may affect the patient casemix. We therefore estimate a model where we exclude all patients characteristics. Again, the results are similar to those reported in Table 4: 36.5 pp for cystoscopy, 13.8 pp for hysteroscopy and 3.8 pp for sterilisation (see Table A2(c) in the Appendix). All coefficients are statistically significant.

6.2. Effect on quality and volume

Table 6 reports the effect of the policy on hospital quality. For cystoscopy, the results suggest that the BPT policy generally improved quality. It reduced the probability of a repeated procedure within 60 days by 0.4 pp and within 365 days by 5.9 pp, and the probability of a 30-day emergency readmission by 1.0 pp. The results for hysteroscopy are mixed. The BPT policy had no statistically significant effect on a repeated procedure within 60 days while it increased the probability of a repeated procedure within 365 days by 3.7 pp. Instead, it reduced the probability of a 30-day readmission rate by 0.2 pp. For sterilisation, we find no effect on 30-day readmissions. The negative effect on repeated procedures for hysteroscopy can by explained by outpatient hysteroscopies being performed without anesthesia, which can be associated with severe pain, leading to incomplete clinical investigations and the need for repeated procedures (Iaco et al. 2000). In contrast, cystoscopy is generally less painful, although discomfort and anxiety may still be present.

[20] Figure A1 in the Appendix displays the trends in the proportion of patients treated in the outpatient setting which appear to be parallel across all treatment-control group pairs.

Table 6: Difference-in-difference estimates of the impact of the BPT outpatients scheme on the probability of having a repeated procedure and emergency readmission

	Repeated procedure within 60 days		Repeated procedure within 365 days	
	Cystoscopy	Hysteroscopy	Cystoscopy	Hysteroscopy
	(1)	(2)	(3)	(4)
DiD coefficient	-0.004**	0.006	-0.059**	0.037***
	(0.002)	(0.004)	(0.004)	(0.008)
Adjusted R^2	0.008	0.022	0.137	0.290
Number of hospitals	168	167	168	167
Observations	3,859,365	1,560,714	3,859,365	1,560,714

	Emergency readmission within 30 days		
	Cystoscopy	Hysteroscopy	Sterilisation
	(1)	(2)	(3)
DiD coefficient	-0.010***	-0.002***	-0.000
	(0.001)	(0.001)	(0.001)
Adjusted R^2	0.020	0.007	0.011
Number of hospitals	168	167	158
Observations	3,859,365	1,560,714	303,256

Notes: Dependent variable is the probability of having a repeated procedure within 60/365 days and emergency readmission within 30 days. Time period is from April 2009 to December 2015. Models are estimated by OLS with standard errors (presented in parenthesis under the coefficients) clustered at hospital level. Models are run separately for each treatment-control procedure pair (cystoscopy-sigmoidoscopy; hysteroscopy-lower genital procedures; sterilisation-vacuum aspiration). All models control for casemix and include a set of year-by-month and hospital dummies.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7 reports the effect of the BPT policy on total volume (the sum of procedures in the inpatient and outpatient setting). Although the coefficient is positive for both cystoscopy (by 17.4 patients, less then 3%, per hospital and quarter) and hysteroscopy (by 22.9 patients, less than 5%), it is not statistically significant at 5% level. Therefore, although the policy significantly affected the intensive margin (by increasing substitutions across settings), it did not affect the extensive margin.

Table 7: Difference-in-difference estimates of the impact of the BPT scheme on volume

	Cystoscopy	Hysteroscopy
	(1)	(2)
DiD coefficient	17.436	22.942*
	(13.786)	(12.286)
Adjusted R^2	0.706	0.545
Number of hospitals	131	132
Observations	7,336	7,392

Notes: The dependent variable is the volume of procedures at the quarter level. Time period is from April 2009 to March 2016. Models are estimated by OLS with standard errors (presented in parenthesis under the coefficients) clustered at hospital level. Models are run separately for each treatment-control procedure pair (cystoscopy-sigmoidoscopy; hysteroscopy-lower genital procedures). Both models control for casemix and include a set of year-by-quarter and hospital dummies. Only hospitals who report to all quarters are included in the analysis.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

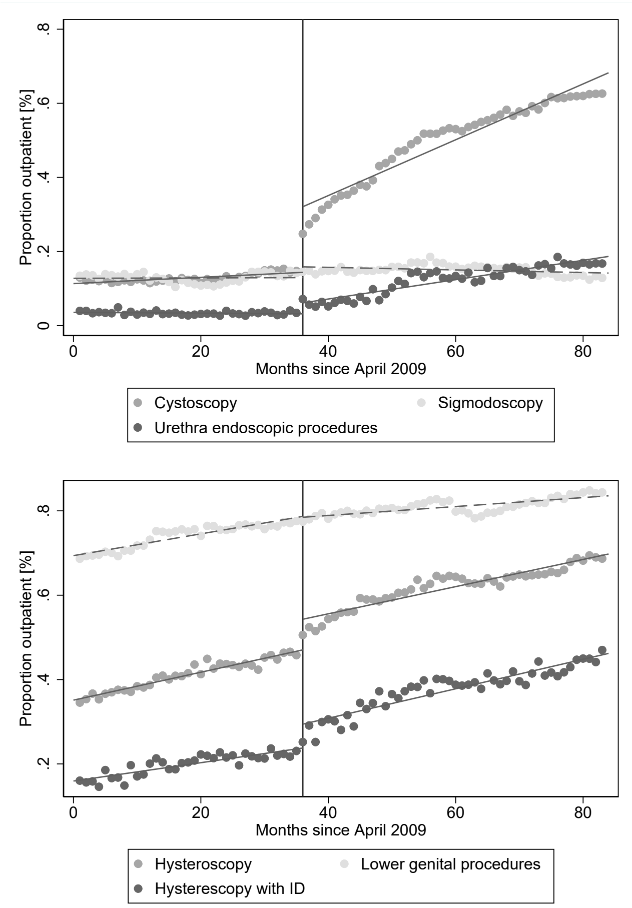
We further test for the plausibility of the parallel trends assumption for the additional analyses on quality and volume. Figures A2 and A3 in the Appendix show the trends over time for volume and for the probability of having repeated procedure within 60 and 365 days. The pre-trends appear parallel for both treatment-control pairs. The results of the empirical test are presented in Tables A3(a)-(c) in the Appendix. All coefficients are small or insignificant, confirming that the parallel trend assumption is likely to hold.

6.3. Spillover effects

Figure 4 displays trends in the proportion of outpatient cases for the treatment, control, and spillover conditions. In the case of cystoscopy, the spillover procedure is endoscopic urethra procedures. The figure shows a modest increase in the proportion of cases for endoscopic urethra procedures following the implementation of the policy. For hysteroscopy, the spillover procedure is hysteroscopy with insertion of an intrauterine device (ID). The time trends are similar, both before and after the policy, for both the treatment (hysteroscopy) and the spillover (hysteroscopy with ID) conditions.

The difference-in-difference analysis on the effect of the BPT policy on closely related procedures is provided in Table 8. The results suggest that the BPT policy had a sizeable impact on the proportion of patients treated in an outpatient setting for selected closely-related procedures that were not incentivised by the BPT. For endoscopic urethra procedures, the policy increased the probability of outpatient treatment by 7.1 pp. While this estimate is statistically significant, the effect is much lower than for cystoscopy (36.0 pp). For hysteroscopy with insertion of inuterine device, the effect of the policy on outpatient treatment is 12.5 pp and statistically significant. This is more in line with the effect observed for hysteroscopy (16.3 pp), suggesting a substantive positive spillover effect.

Figure 4: Proportion of patients treated in the outpatient setting over time for the treatment, control and spillover procedures



Notes: The figure shows the trends in the proportion of outpatient procedures for the two diagnostic BPT procedures (cystoscopy and hysteroscopy), the corresponding control groups (sigmoidoscopy and lower genital procedures) and the applicable spillover procedures (endoscopic urethra procedure and hysteroscopy with ID).

Table 8: Difference-in-difference results of the main effects for the spillover procedures

Treatment group	Endoscopic urethra procedures (1)	Hysteroscopy with ID (2)
DiD coefficient	0.071***	0.125***
	(0.020)	(0.021)
Adjusted R^2	0.302	0.368
Number of hospitals	165	167
Observations	1,620,688	1,116,036

Notes: Dependent variable is the probability to be treated in the outpatient setting. Time period is from April 2009 to March 2016. Models are estimated by OLS with standard errors (presented in parenthesis under the coefficients), clustered at hospital level. Models are run separately for each spill over-control procedure pair (endoscopic urethra procedure-sigmoidoscopy; hysteroscopy with ID-lower genital procedures). All models include case mix variables and a full set of year-by-month and hospital dummies.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

We test for the parallel trends assumption for the spillover procedures. The results of the empirical test are in Tables A3(e) in the Appendix. All coefficients are small or insignificant, confirming that the parallel trend assumption is likely to hold.

6.4. Heterogeneity analysis

In this section, we perform a heterogeneity analysis. First, we split hospitals into two groups, depending on whether the outpatient share for each hospital is above or below the median in the pre-policy period. Hospitals with a lower outpatient share may respond more to the BPT policy because there is larger scope for improvement. On the other hand, these hospitals may be slow adopters of new technologies which implies that they respond less to the policy. The results in Table 9 suggest that hospitals whose share of patients treated in an outpatient setting were below the median tend to respond more strongly to the policy for cystoscopy and hysteroscopy but the differences are not marked. The probability of being treated in an outpatient setting for cystoscopy increased by 38.7 pp for hospitals below the median and by 33.7 pp for hospitals above the median. Similarly, for hysteroscopy the probability of being treated in an outpatient setting increased by 16.7 pp for hospitals below the median and by 13.8 pp for hospitals above the median. Instead, the response was slightly larger for hospitals above the median for sterilisation. The probability of being treated in an outpatient setting increased by 4.3 pp for hospitals above the median and 3.4 pp for hospitals below the median.

The effect of the policy may also differ by the characteristics of the patient. Table 10 shows the differential effect of the policy by sex, age, income deprivation and patient complexity (proxied by past emergency admissions). Overall, we do not find marked differences between different groups. For cystoscopy the effect was 36.9 pp for male and 34.2 pp for females; 36.6 pp for patients whose age is above the median and 35.5 pp for those below the median; 35.6 pp for more deprived and 36.4 pp for less deprived patients; 38.2 pp for those with at least one past emergency admission and 35.6 pp for those with no past admissions.

Similarly, for hysteroscopy the effect was 17.1 pp for patients whose age is above the median and 14.9 pp for those below the median; 16.7 pp for more deprived and 16.0 pp for less deprived patients; 16.6 pp for those with at least one past emergency admission and 16.2 pp for those with no past admissions. For sterilisation, the effect was 4.2 pp for patients whose age is above the median and 2.3 pp for those below the median; 2.6 pp for more deprived and 4.3 pp for less deprived patients; 2.2 pp for those with at least one past emergency admission and 4.0 pp for those with no past admissions.

Table 9: The impact of the BPT scheme on the probability of treatment in the outpatient setting: heterogeneity by provider

	Cystoscopy	Hysteroscopy	Sterilisation	Cystoscopy	Hysteroscopy	Sterilisation
	Hospital outpatient share *above* the median			Hospital outpatient share *below* the median		
	(1)	(2)	(3)	(4)	(5)	(6)
DiD estimates	0.337***	0.138***	0.043*	0.387***	0.167***	0.034**
	(0.042)	(0.024)	(0.019)	(0.044)	(0.039)	(0.012)
Adjusted R^2	0.352	0.152	0.098	0.401	0.305	0.091
No. hospitals	83	79	52	79	80	102
Mean value	0.349	0.813	0.023	0.186	0.486	0.014
Observations	2,120,690	1,002,192	128,348	1,689,618	541,999	168,080

Notes: The dependent variable is the probability to be treated in the outpatient setting. The sample is divided into two groups based on whether the hospital has a share of outpatient procedures that is above or below the median value in the pre-policy period. Columns 1 to 3 report the estimates for hospitals whose outpatient share is above median share, while columns 4 to 6 report the estimates for hospitals whose outpatient share is below the median. Time period is from April 2009 to March 2016. Models are estimated by OLS with standard errors (presented in parenthesis under the coefficients) clustered at hospital level. Models are run separately for each treatment-control procedure pair (cystoscopy-sigmoidoscopy; hysteroscopy-lower genital procedures; sterilisation-vacuum aspiration). All models control for casemix and a set of year-by-month and hospital fixed effects.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 10: The impact of the BPT scheme on the probability of treatment in the outpatient setting: heterogeneity by patient characteristics

	Cystoscopy	Hysteroscopy	Sterilisation	Cystoscopy	Hysteroscopy	Sterilisation
		Male			Female	
	(1)	(2)	(3)	(4)	(5)	(6)
DiD estimates	0.369***	-	-	0.342***	-	-
	(0.031)			(0.030)		
Mean value	0.296			0.252		
Observations	2,232,287			1,627,074		
		Above median age			Below median age	
	(1)	(2)	(3)	(4)	(5)	(6)
DiD estimates	0.366***	0.171***	0.042***	0.356***	0.149***	0.023***
	(0.033)	(0.020)	(0.011)	(0.030)	(0.026)	(0.009)
Mean value	0.306	0.787	0.022	0.249	0.605	0.013
Observations	1,948,032	794,649	162,692	1,911,330	766,064	140,563
		More deprived			Less deprived	
	(1)	(2)	(3)	(4)	(5)	(6)
DiD estimates	0.356***	0.167***	0.026**	0.364***	0.160***	0.043***
	(0.031)	(0.024)	(0.013)	(0.033)	(0.022)	(0.010)
Mean value	0.275	0.705	0.017	0.280	0.693	0.018
Observations	1,766,997	647,813	98,884	2,092,366	912,900	204,372
		Past emergency>0			Past emergency=0	
	(1)	(2)	(3)	(4)	(5)	(6)
DiD estimates	0.382***	0.166***	0.022**	0.356***	0.162***	0.040***
	(0.032)	(0.023)	(0.010)	(0.031)	(0.022)	(0.011)
Mean value	0.290	0.758	0.015	0.275	0.692	0.018
Observations	607,560	139,715	36,277	3,251,805	1,420,996	266,978

Notes: Dependent variable is the probability to be treated in the outpatient setting. The sample is divided into groups based on patient characteristics. Time period is from April 2009 to March 2016. Models are estimated by OLS with standard errors (presented in parenthesis under the coefficients) clustered at hospital level. Models are run separately for each treatment-control procedure pair (cystoscopy-sigmoidoscopy; hysteroscopy-lower genital procedures; sterilisation-vacuum aspiration). All models control for a set of year-by-month and hospital fixed effects.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

7. Effect on revenues and profits

We further explore the effect of the BPT on hospitals' revenues and profits. The effect on the revenues is important for the National Health Service, as any reductions in hospital revenues will translate into savings for the NHS. Our results show that the BPT increased the probability of being treated in the outpatient setting compared to the inpatient setting. The effect of the policy on revenues and profits is the result of the combined effect of changes in tariffs, the volume in the inpatient and outpatient setting, and costs.

7.1. Effect on revenues

In line with our theoretical framework, define p_O^1 and p_I^1 as the prices paid to hospitals in the post-policy period 1 for treating the patient in an outpatient and inpatient setting, respectively; and p_O^0, p_I^0 as the outpatient and inpatient prices in the pre-policy period 0 . Similarly, define V_O^1 and V_I^1 as the volumes of patients treated in an outpatient and inpatient setting in the post-policy period 1, and V_O^0 and V_I^0 as the outpatient and inpatient volumes in the pre-policy period 0. The hospital revenues in the post-policy (R_{Post}) and pre-policy period (R_{Pre}) are then given by:

$$R_{Post} = p_O^1 V_O^1 + p_I^1 V_I^1 \quad (13)$$

$$R_{Pre} = p_O^0 V_O^0 + p_I^0 V_I^0, \quad (14)$$

with the difference in revenues corresponding to:

$$\Delta R = R_{Post} - R_{Pre} \quad (15)$$

Table 11 presents the pre- and post-policy prices and volumes for the three conditions. The prices are those used to pay hospitals in 2011/12, which is the year prior to the introduction of the policy, and in 2012/13, which is the year the BPT was introduced. The pre-policy volumes are based also on 2011/12. The volumes for 2012/13 are estimated using our regression results, which show that the policy increased to a great extent the proportion of patients treated in the outpatient setting, and to a small extent the volume. We can therefore use the data from Table 11 to compute the hospital revenues for 2011/12 both before the policy and post policy, as if the policy had been in place for that year.

For cystoscopy the revenues pre-policy and post-policy are respectively equal to:

$$\begin{aligned} R_{Pre} &= p_O^0 V_O^0 + p_I^0 V_I^0 \\ &= 257 * 45,010 + 714 * 276,487 \\ &= 208.98M \end{aligned}$$

$$\begin{aligned} R_{Post} &= p_O^1 V_O^1 + p_I^1 V_I^1 \\ &= 403 * 160,749 + 260 * 160,748 \\ &= 106.58M \end{aligned}$$

The results indicate a significant reduction in hospital revenues for cystoscopy, decreasing from £208.98 million in the pre-policy period to £106.58 million in the post-policy period. This corresponds to a reduction of £102.40 million. Additionally, this implies that the National Health Service achieved substantial savings as a result of this policy, amounting to approximately 49.0% of the original hospital reimbursement. Note

Table 11: Information used to calculate the effect of the BPT on hospital's revenue and costs

(a) Prices

	Cystoscopy	Hysteroscopy	Sterilisation
Pre-policy price			
Outpatient p_O^0	257	242	242
Inpatient p_I^0	714	733	733
Post-policy price			
Outpatient p_O^1	403	457	1,137
Inpatientp_I^1	260	260	928

(b) Volumes

	Cystoscopy	Hysteroscopy	Sterilisation
Pre-policy volume			
Outpatient V_O^0	45,010	34,227	348
Inpatient V_I^0	276,487	41,834	11,268
Post-policy volume			
Outpatient V_O^1	160,749	46,625	789
Inpatient V_I^1	160,748	29,436	10,827

(c) Costs

	Cystoscopy	Hysteroscopy	Sterilisation
Pre-policy costs			
Outpatient c_O^0	159	197	269
Inpatient c_I^0	422	775	1,155
Post-policy costs			
Outpatient c_O^1	154	179	274
Inpatient c_I^1	453	733	1,181

Notes: The post-policy and pre-policy prices are based on the prices paid to the hospitals in 2012/13 (year of the BPT introduction) and 2011/12 (one year prior the introduction). The volume is based on the 2011/12 volume of inpatient and outpatient attendances.

that even in the extreme case where 100% of patients were treated in the outpatient setting, hospital revenues would reach a maximum of £129.56 million, still resulting in a 38% savings for the NHS.

A qualitatively similar trend is observed for hysteroscopy, although the reduction in hospital revenue is much smaller. The pre-policy revenues amounted to £38.95 million, which decreased to £28.96 million in the post-policy period, resulting in a decrease of £9.99 million or 25.6%. The effect on hospital revenues is thus less dramatic compared to cystoscopy. While prices (pre-policy and post-policy) and outpatient volumes for hysteroscopy are comparable to those of cystoscopy, the volumes in an inpatient setting for hysteroscopy were considerably smaller in the pre-policy period. This explains why the reduction in revenues is more contained, as it is primarily driven by the lower tariff in the inpatient setting. In this case, under the extreme scenario where 100% of patients were treated in the outpatient setting during the post-policy period (while the estimation indicates 61.3% being treated in the outpatient setting), the hospital could have increased revenues to £34.76 million.

In contrast to cystoscopy and hysteroscopy, the revenue dynamics for sterilisation follows a different pattern. Both the inpatient and outpatient tariffs increased after the policy implementation, although the outpatient tariff experienced a much larger increase. As a result, hospital revenues and subsequent reimbursement by the public insurer increased. Specifically, hospital revenues for sterilisation rose from £8.34 million in the pre-policy period to £10.94 million in the post-policy period, reflecting an increase of £2.60 million or 31.2%. The NHS reimbursement also increased by an equivalent amount.

The number of patients treated in the outpatient setting only increased from 348 to 789. Therefore, the overall increase in revenues is primarily driven by the higher tariff for the inpatient setting, where the majority of the volume is concentrated. The outpatient setting, although experiencing a four-fold increase in tariff, contributes to revenue growth to a lesser extent due to the low volume of cases.

The results for cystoscopy and hysteroscopy highlight the substantial savings obtained by the public health insurer through the introduction of an incentive scheme that generously rewards the outpatient setting while simultaneously reducing the inpatient tariff.

7.2. Profits

Even if the BPT reduced hospital revenues for cystoscopy and hysteroscopy, this does not necessarily imply that profits also reduced because the higher proportion of patients treated in an outpatient setting also reduced provider total cost of provision across the two settings. Hospital profit for the pre- and post-policy period is:

$$\pi_{Post} = \left(p_O^1 - c_O^1\right) V_O^1 + \left(p_I^1 - c_I^1\right) V_I^1 \quad (16)$$

$$\pi_{Pre} = \left(p_O^0 - c_O^0\right) V_O^0 + \left(p_I^0 - c_I^0\right) V_I^0 \quad (17)$$

The effect of the BPT on the change in profit is then as follows:

$$\Delta\pi = \pi_{Post} - \pi_{Pre} \quad (18)$$

We obtain information on costs pre and post policy, presented in Table 11, from the yearly Reference Costs publication, which provides treatment information across hospitals and HRGs.[21] Using information on prices, costs and volumes in the post-policy period from Table 11, we can calculate the post policy profit for cystoscopy from equation (16).

For cystoscopy, the pre-policy profit was £85.15 million, which decreased to £9.00 million in the post-policy period, resulting in a reduction of £76.14 million. Thus, the BPT policy led to a decrease in both revenues and profits for this procedure. Despite the positive price mark-up and the increase in the volume of patients treated in the outpatient setting, the negative price mark-up in the inpatient setting, coupled with the high volume in that setting, outweighed the positive effects.

The effect on profit for hysteroscopy was less pronounced. Hospitals made a (relatively small) loss of -£0.22 million in the pre-policy period and this loss further increased to -£0.96 million in the post-policy period. Therefore, the policy generated a loss of £0.74 million, which is much less pronounced relative to cystoscopy (equal to £76.14 million). This can be explained by the lower volume of inpatients in both the pre-policy and post-policy periods.

Regarding sterilisation, hospitals were operating at a loss in both periods. The loss was -£4.76 million in the pre-policy period, and decreased to -£2.06 million in the post-policy period, representing a reduction in losses of £2.71 million. In this case, prices and price mark-ups increased in both settings, leading to higher revenues and reduced losses.

7.3. Distribution implications between public health insurer and provider

In this section, we briefly describe the distribution implications of the scheme between the public health insurer (the payer) and the provider. For cystoscopy, the implementation of the incentive scheme resulted in significant savings for the payer, amounting to approximately £102.40 million. These savings outweighed the profit loss experienced by hospitals, which amounted to around £74.14 million. The distributional impact of the scheme favored the payer, indicating a transfer of financial burden from the payer to the hospitals.

For hysteroscopy, the payer achieved savings of approximately £9.99 million, and hospitals increased their losses by £0.74 million. The savings for payer were substantially higher than the provider losses. For sterilisation, the payer increased both the inpatient and outpatient tariffs, resulting in a £2.60 million increase in reimbursement. However, hospitals gained £2.71 million by shifting care to the low-cost outpatient setting.

In terms of welfare implications, as highlighted by the theory model in equation (9), the effect of the policy depends on difference between health benefits and costs (also weighted by the opportunity cost of public funds). For cystoscopy, costs reduced and quality also improved. We therefore conclude that welfare improved as a result. For hysteroscopy and sterilisation, costs also reduced but the effect on quality for hysteroscopy was mixed and for sterilisation we are only able to measure quality with 30-day readmissions. Therefore, whether welfare improved is less clear-cut.

[21] NHS Reference Costs are detailed financial data, providing comprehensive overview of the average treatment costs incurred by NHS providers. Reference Costs follow a standardized costing methodology, ensuring consistency in cost calculation across different NHS providers, promoting comparability and reliability of the data. They undergo a rigorous quality assurance process to ensure accuracy and reliability. This process includes data validation checks, verification against audited financial statements, and external reviews.

These examples highlight the significant variation in the implications of the scheme across different treatments, underscoring the importance of the inpatient-outpatient mix in pre-policy volume interventions, as well as the role of provider behavioral response to the scheme.

8. Conclusion

Our study shows that financial incentives can be successful in changing hospital behaviour by shifting patients from the inpatient to the outpatient setting. We document positive effects for all three incentivised conditions with the largest effects for diagnostic hysteroscopy and cystoscopy, ranging between 16 and 36 percentage points. The effect on quality were however mixed depending on the procedure. The effect on quality was positive for cystocopy, which featured the largest increase in patients in the outpatient setting, while they were mixed for hysteroscopy and no effect was found for sterilisation.

Hospitals did not respond to the scheme by increasing the overall volume of patients at the extensive margin. We also observed positive spillover effects on related conditions, suggesting that providers adapted their working patterns to accommodate further shifts across settings. The distribution implications of these financial schemes can vary. We show that the public insurer (the funder of health services) can achieve significant savings when the inpatient tariff is reduced, while the impact on providers' profit varies depending on the inpatient-outpatient mix post policy (and mitigated only if the proportion of patients in the inpatient setting is low post policy) .

These findings have significant policy implications, particularly in the context of the rapid growth in medical spending observed in many healthcare systems. Our study demonstrates that provider response can be economically important when the scheme is targeted and the bonus is substantial.

References

Ahmad, G., S. Saluja, H. O'Flynn, A. Sorrentino, D. Leach and A. Watson (2017). 'Pain relief for outpatient hysteroscopy'. *Cochrane Database of Systematic Reviews*, URL: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD007710.pub3/abstract.

Ancona-Berk, V. A. and T. Chalmers (1986). 'An analysis of the costs of ambulatory and inpatient care'. *American Journal of Public Health*, 76: (9), 1102–1104.

Bach, P. B. and R. H. Jain (2017). 'Physician's Office and Hospital Outpatient Setting in Oncology: It's About Prices, Not Use'. *Journal of Oncology Practice*, 13: (1), 4–5.

Batty, M. and B. Ippolito (2017). 'Financial Incentives, Hospital Care, and Health Outcomes: Evidence from Fair Pricing Laws'. *American Economic Journal: Economic Policy*, 9: (2), 28–56.

CAPH (2018). Last accessed on 30th of July 2020. URL: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD007710.pub3/abstract.

Cashin, C., Y. Chi, P. Smith, M. Borowith and S. Thomson (2014). 'Paying for Performance in Healthcare'. *European Observatory on Health Systems and Policies Series*,

Casteleijn, N., J. Vriesema, S. Stomps, O. van Balen and E. Cornel (2017). 'The effect of office based flexible and rigid cystoscopy on pain experience in female patients'. *Investigative and Clinical Urology*, 58: (1), 48–53.

Castells, X., J. Alonso, M. Castilla, C. Ribó, F. Cots and J. M. Antó (2001). 'Outcomes and costs of outpatient and inpatient cataract surgery: a randomised clinical trial'. *Journal of Clinical Epidemiology*, 54: (1), 23–29.

Chalkley, M. and J. Malcomson (1998). 'Contracting for health services with unmonitored quality'. *The Economic Journal*, 108: (449), 1093–1110. DOI: https://doi.org/10.1111/1468-0297.00331.

Citizen Advice (2020). 'NHS Patients Rights'. Last accessed 4th of August 2020. URL: https://www.citizensadvice.org.uk/health/nhs-healthcare/nhs-patients-rights/.

Clemens, J. and J. D. Gottlieb (2014). 'Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?' *American Economic Review*, 104: (4).

Dafny, L. S. (2005). 'How do hospitals respond to price changes?' *American Economic Review*, 95: (5), 1525–1547.

Davis, K. and L. Russell (1972). 'The Substitution of Hospital Outpatient Care for Inpatient Care'. *The Review of Economics and Statistics*, 54: (2), 109–120.

Department of Health (2012). 'Payment by Results Guidance for 2012-13'. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216212/dh_133585.pdf.

Dobson, A., A. M. El-Gamil, M. T. Shimer, J. E. DaVanzo, A. Q. Urbanes, G. A. Beathard and T. F. Litchfield (2013). 'Clinical and economic value of performing dialysis vascular access procedures in a freestanding office-based center as compared with the hospital outpatient department among Medicare ESRD beneficiaries'. *Seminars in dialysis*, 26: (5). DOI: 10.1111/sdi.12120.

Doria-Rose, V., P. A. Newcomb and T. R. Levin (2005). 'Incomplete screening flexible sigmoidoscopy associated with female sex, age, and increased risk of colorectal cancer'. *Gut*, 54: 1273–1278. DOI: 10.1136/gut.2005.064030.

Eijkenaar, F. (2012). 'Pay for performance in health care: an international overview of initiatives'. *Medical Care Research and Review*, 69: (3).

Ellis, R. and T. McGuire (1993). 'Supply-Side and Demand-Side Cost Sharing in Health Care'. *Journal of Economic Perspectives*, 7: (4), 135–151.

Ellis, R. P. and T. G. McGuire (1986). 'Provider behavior under prospective reimbursement. Cost sharing and supply'. *Journal of Health Economics*, 5: (2), 129–151.

Ellis, R. (1998). *Journal of Health Economics*, 17: 537–555. DOI: https://doi.org/10.1111/j.1756-2171.2009.00090.x.

Elnicki, R. A. (1976). 'Substitution of outpatient for inpatient hospital care: a cost analysis'. *Inquiry*, 13: (3), 245–261.

Esparza, D. M., N. Young and J. A. Luongo (1989). 'Effective planning for office and outpatient chemotherapy administration'. *Seminars in oncology nursing*, 5: (2). DOI: https://doi.org/10.1016/0749-2081(89)90077-6.

Falavolti, C., F. Sergi, T. Petitti and M. Buscarini (2017). 'Does listening to music during flexible cystoscopy and bladder biopsy decrease patient's pain?' *Journal of Clinical Urology*, DOI: 10.1177/2051415816666689.

Fisher, M. D., R. Punekar and Y. M. Yim (2017). 'Differences in health care use and costs among patients with cancer receiving intravenous chemotherapy in physician offices versus in hospital outpatient settings'. *Journal of Oncology Practice*, 13: (1), e37–e44.

Foo, P. K., R. S. Lee and K. Fong (2017). 'Physician prices, hospital prices, and treatment choice in labor and delivery'. *American Journal of Health Economics*, 3: (3), 422–453.

Gaughan, J., N. Gutacker, K. Grasic, N. Kreif, L. Siciliani and A. Street (2019). 'Paying for efficiency: Incentivising same-day discharges in the English NHS'. *Journal of Health Economics*, 68:

Genovese, F., G. D'Urso, F. D. Guardo, G. Insalaco, A. Tuscano, L. Ciotta, A. Carbonaro, V. Leanza and M. Palumbo (2020). 'Failed diagnostic hysteroscopy: Analysis of 62 cases.' *European journal of obstetrics, gynecology, and reproductive biology*, 245: 193–197. DOI: 10.1016/j.ejogrb.2019.10.031.

Gordan, L., M. Blazer, V. Saundankar, D. Kazzaz, S. Weidner and M. Eaddy (2019). 'Cost Differential of Immuno-Oncology Therapy Delivered at Community Versus Hospital Clinics'. *The American Journal of Managed Care*, 25: (3), e66–e70.

Grant, D. (2009). 'Physician financial incentives and cesarean delivery: New conclusions from the healthcare cost and utilization project'. *Journal of Health Economics*, 28: 244–250.

Grasic, K., A. Villasenor, J. Gaughan, N. Gutacker and L. Siciliani (2025). 'Incentivizing Hospital Quality Through Care Bundling'. *Health Economics*, 34: 2140–2160. DOI: https://doi.org/10.1111/j.1756-2171.2009.00090.x.

Greenberg, J. A. (2008). 'Hysteroscopic Sterilization: History and Current Methods'. *Reviews in Obstetrics and Gynecology*, 1: (3), 113–121.

Greenstein, A., I. Greenstein, S. Senderovich and N. J. Mabjeesh (2014). 'Is diagnostic cystoscopy painful? Analysis of 1,320 consecutive procedures'. *International Brazilian Journal of Urology*, 40: (4). DOI: 10.1590/S1677-5538.IBJU.2014.04.13.

Gruber, J., J. Kim and D. Mayzlin (1999). 'Physician fees and procedure intensity: The case of cesarean delivery'. *Journal of Health Economics*, 18: 473–490.

Gruber, J. and M. Owings (1996). 'Physician financial incentives and cesarean section delivery'. *RAND Journal of Economics*, 27: 99–123.

Hadley, J. and K. Swartz (1989). 'The Impacts on Hospital Costs between 1980 and 1984 of Hospital Rate Regulation, Competition, and Changes in Health Insurance'. *Inquiry*, 26: (1), 35–47.

Hadley, J., S. Zuckerman and J. Feder (1989). 'Profits and Fiscal Pressure in the Prospective Payment System: Their Impacts on Hospitals'. *Inquiry*, 26: (3), 354–365.

Hafsteinsdottir, E. J. G. and L. Siciliani (2010). 'DRG prospective payment systems: refine or not refine?' *Health economics*, 19: 1226–1239.

Hayes, J., J. R. Hoverman, M. E. Brow, D. C. Dilbeck, D. K. Verrilli, J. Garey, J. L. Espirito, J. Cardona and R. Beveridge (2015). 'Cost Differential by Site of Service for Cancer Patients Receiving Chemotherapy'. *The American Journal of Managed Care*, 21: (3), 189–196.

He, D. and J. M. Mellor (2013). 'Do Changes in Hospital Outpatient Payments Affect the Setting of Care?' *Health Services Research*, 48: (5).

Higgins, A., G. Veselovskiy and J. Schinkel (2016). ‘National estimates of price variation by site of care’. *American Journal of Managed Care*, 22: (3), e116–e121.

Hodgkin, D. and T. G. McGuire (1994). ‘Payment levels and hospital response to prospective payment’. *Journal of Health Economics*, 13: (1), 1–29.

Hollingsworth, B. (2008). ‘The measurement of efficiency and productivity of health care delivery’. *Health Economics*, 17: (10), 1107–1128.

Hoof, S. van, T. Quanjel, M. Kroese, M. Spreeuwenberg and D. Ruwaard (2019). ‘Substitution of outpatient hospital care with specialist care in the primary care setting: A systematic review on quality of care, health and costs’. *PLoS One*, 14: (8).

Hussey, P., H. de Vries, J. Romley, M. Wang, S. Chen, P. Shekelle and E. McGlynn (2009). ‘A Systematic Review of Health Care Efficiency Measures’. *Health Services Research*, 44: (3), 784–805.

Iaco, P. D., A. Marabini, M. Stefanetti, C. D. Vecchio and L. Bovicelli (2000). ‘Acceptability and pain of outpatient hysteroscopy’. *Journal of the American Association of Gynecologic Laparoscopists*, 7: (7), 71–75.

Januleviciute, J., J. E. Askildsen, O. Kaarboe, L. Siciliani and M. Sutton (2016). ‘How do Hospitals Respond to Price Changes? Evidence from Norway’. *Health Economics*, 25: (5), 620–636.

Kelly, S. B., J. Murphy, A. Smith, H. Watson, S. Gibb, C. Walker and R. Reddy (2008). ‘Nurse specialist led flexible sigmoidoscopy in an outpatient setting’. *Colorectal Disease*, 10: 390–393. DOI: 10.1111/j.1463-1318.2007.01271.x.

Kifmann, M. and L. Siciliani (2017). ‘Average-cost pricing and dynamic selection incentives in the hospital sector’. *Health Economics*, 26: 1566–1582. DOI: https://doi.org/10.1111/j.1756-2171.2009.00090.x.

Kristensen, S. R., L. Siciliani and M. Sutton (2016). ‘Optimal Price-Setting in Pay for Performance Schemes in Health Care’. *Journal of Economic Behavior Organization*, 123: 57–77. DOI: https://doi.org/10.1111/j.1756-2171.2009.00090.x.

Leader, S. and M. Moon (1989). ‘Medicare Trends In Ambulatory Surgery’. *Health Affairs*, 8: (1). DOI: https://doi.org/10.1377/hlthaff.8.1.158.

Lee, J., S. Doumouchtsis, S. Jeffery and M. Fynes (2009). ‘Evaluation of outpatient cystoscopy in urogynaecology’. *Archives of Gynecology and Obstetrics*, 279: (5). DOI: 10.1007/s00404-008-0773-6.

Lisi, D., L. Siciliani and O. R. Straume (2020). ‘Hospital competition under pay-for-performance: Quality, mortality, and readmissions’. *Journal of economics & management strategy*, 29: 289–314.

Ma, A. (1994). ‘Health Care Payment Systems: Cost and Quality Incentives’. *Journal of Economics and Management Strategy*, 3: (1), 93–112.

Mao, J., S. Pfeifer, P. Schlegel and A. Sedrakyan (2015). ‘Safety and efficacy of hysteroscopic sterilization compared with laparoscopic sterilization: an observational cohort study’. *BMJ*, DOI: https://doi.org/10.1136/bmj.h5162.

Mendelson, A., K. Kondo, C. Damberg, A. Low, M. Motúapuaka, M. Freeman, M. O’Neil, R. Relevo and D. Kansagara (2017). ‘The Effects of Pay-for-Performance Programs on Health, Health Care Use, and Processes of Care: A Systematic Review’. *Annals of Internal Medicine*, 166: (5), 341–353. DOI: 10.7326/M16-1881.

Milstein, R. and J. Schreyoegg (2016). ‘Pay for performance in the inpatient sector: A review of 34 P4P programs in 14 OECD countries’. *Health Policy*, 120: (10), 1125–1140.

Milstein, R. and J. Schreyögg (2024). ‘The end of an era? Activity-based funding based on diagnosis-related groups: A review of payment reforms in the inpatient sector in 10 high-income countries’. *Health Policy*, 141: 104990.

Mullen, K. J., R. G. Frank and M. B. Rosenthal (2010). 'Can you get what you pay for? Pay-for-performance and the quality of healthcare providers'. *The RAND Journal of Economics*, 41: 64–91. DOI: https://doi.org/10.1111/j.1756-2171.2009.00090.x.

Murthy, P., J. Edwards and M. Pathak (2017). 'Update on hysteroscopic sterilisation'. *The Obstetrician adn Gynaecologist*, 22: (3), 227–235.

NHS Digital (2019a). 'Hospital Admitted Patient Care Activity 2018-19'. Last accessed 4th of August 2020. URL: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/2018-19.

— (2019b). 'Hospital Outpatient Activity 2018-19'. Last accessed 4th of August 2020. URL: ttps://digital.nhs.uk/data-and-information/publications/statistical/hospital-outpatient-activity/2018-19.

NHS Direct (2018). Acessed on 28th of July 2020. URL: https://www.nhs.uk/conditions/hysteroscopy/.

— (2020). Acessed on 28th of July 2020. URL: https://www.nhs.uk/conditions/cystoscopy/.

NHS England and NHS Improvement (2019). '2019/20 National Tariff Payment System'. URL: https://improvement.nhs.uk/documents/4980/1920_National_Tariff_Payment_System.pdf.

NHS England and Monitor (2015). 'Populus Summary: Outpatient Appointment Referrals'. Last accessed 4th of August 2020. URL: https://www.england.nhs.uk/wp-content/uploads/2015/09/monitor-nhse-outpatient-appointments-summary.pdf.

OECD (2010). 'Health care systems: Getting more value for money'. *OECD Economics Department Policy Notes*, No. 2: 1–204.

— (2021). *Health at a Glance 2021*, 274. DOI: https://doi.org/https://doi.org/10.1787/ae3016b9-en. URL: https://www.oecd-ilibrary.org/content/publication/ae3016b9-en.

— (2025). 'PROMoting quality of care through patient reported outcome measures'. *OECD Health Working Paper*, No. 183: 1–65.

Ogundeji, Y., J. Bland and T. Sheldon (2016). 'The effectiveness of payment for performance in healthcare: a meta-analysis and exploration of variation in outcomes'. *Health Policy*, 120: (10), 1141–1150.

Papanicolas, I. and A. McGuire (2015). 'Do financial incentives trump clinical guidance? Hip Replacement in England and Scotland'. *Journal of Health Economics*, 44: 25–36.

Relph, S., T. Lawton, M. Broadbent and M. Karoshi (2016). 'Failed hysteroscopy and further management strategies'. *The Obstetrician & Gynaecologist*, 18: 65–68. DOI: 10.1111/tog.12261.

Robinson, W. R. and J. Beyer (2010). 'Impact of Shifting From Office- to Hospital-Based Treatment Facilities on the Administration of Intraperitoneal Chemotherapy for Ovarian Cancer'. *Journal of Oncology Practice*, 6: (5), 232–235.

Royal College of Obstetricians and Gynaecologists (2016). 'Femele Sterilisation; Consent advice no. 3'. *Annals of Internal Medicine*, URL: https://www.rcog.org.uk/globalassets/documents/guidelines/consent-advice/consent-advice-3-2016.pdf.

Royal College of Physicians (2018). *Outpatients: The future*. URL: https://www.rcplondon.ac.uk/file/11479/download.

Sharma, M. (2015). 'Manual vacuum aspiration: an outpatient alternative for surgical management of miscarriage'. *The Obstetrician & Gynaecologist*, 17: (3). DOI: https://doi.org/10.1111/tog.12198.

Shin, E. (2019). 'Hospital responses to price shocks under the prospective payment system'. *Health Economics*, 28: 245–260.

Verzulli, R., G. Fiorentini, M. L. Bruni and C. Ugolini (2017). 'Price Changes in Regulated Healthcare Markets: Do Public Hospitals Respond and How?' *Health Economics*, 26: (11), 1429–1446.

Vitikainen, K., M. Linna and A. Street (2010). 'Substituting inpatient for outpatient care: what is the impact on hospital costs and efficiency?' *The European Journal of Health Economics*, 11: (4), 395–404.

Vlaanderen, F., M. Tanke, B. Bloem, M. Faber, F. Eijkenaar, F. Schut and et al. (2019). 'Design and effects of outcome-based payment models in healthcare: a systematic review'. *European Journal of Health Economics*, 20: 217–232.

Wellenstein, D. J., H. W. Schutte, H. M., J. Honings, P. Belafsky, G. Postma, R. Takes and G. van den Broek (2017). 'Office-based procedures for diagnosis and treatment of esophageal pathology'. *Head & Neck*, 39: (9). DOI: 10.1002/hed.24819.

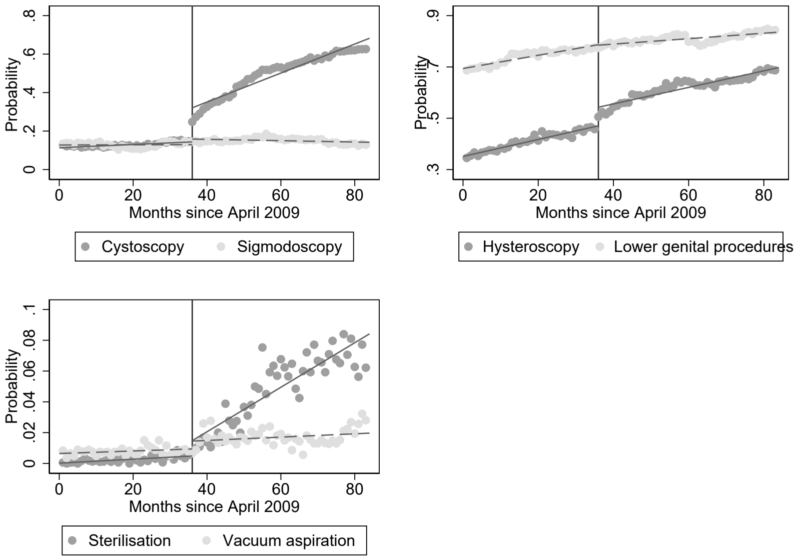
Yen, C. F., H. H. Chou, H. M. Wu, C. L. Lee and T. C. Chang (2019). 'Effectiveness and appropriateness in the application of office hysteroscopy'. *Journal of the Formosan Medical Association*, 118: (11), 1480–1487.

Yip, W. C. (1998). 'Physician response to medicare fee reductions: Changes in volume of coronary artery bypass graft (CABG) surgeries in the medicare and private sectors'. *Journal of Health Economics*, 17: 675–696.

Zaresani, A. and A. Scott (2021). 'Is the evidence on the effectiveness of pay for performance schemes in healthcare changing? Evidence from a meta-regression analysis'. *BMC health services research*, 21: 1–10.

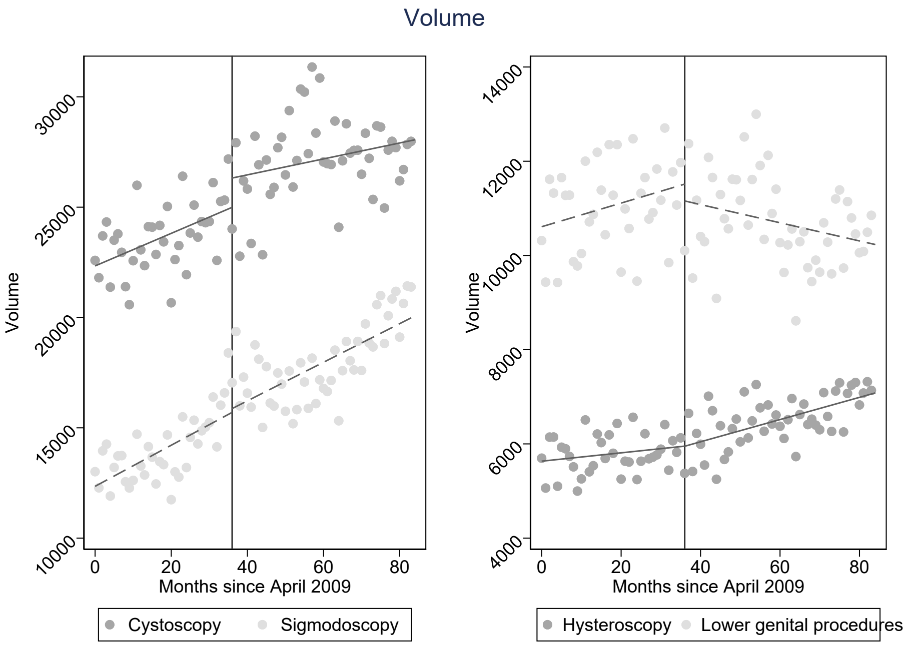
9. Appendix

Figure A1: Probability of treatment in the outpatient setting - trends over time



Notes: Figure shows pre- and post- trends of the treatment/control group pairs for the primary outcome: probability of treatment in the outpatient setting

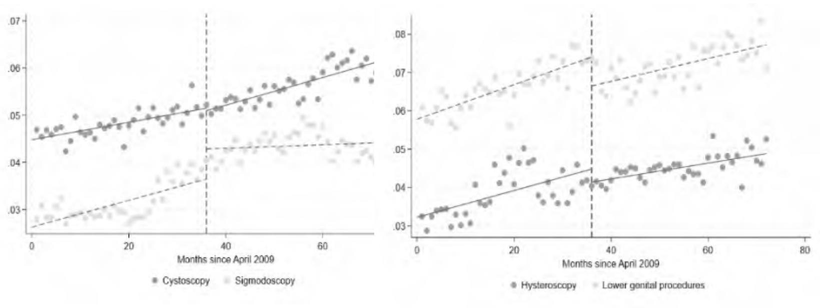
Figure A2: Volume over time



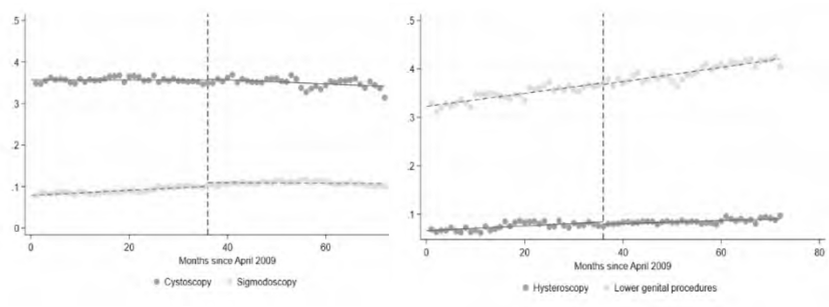
Notes: Figure shows pre- and post- trends in volume for the treatment/control group pairs.

Figure A3: Health Outcomes

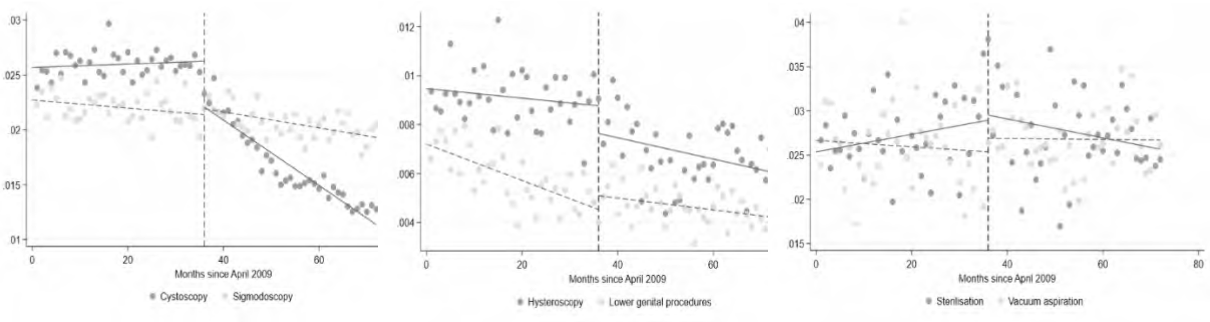
30-Days Reoperation



365-Days Reoperation



30-Days Readmission



Notes: Figure shows pre- and post- trends in the probability in having the procedure repeated within 60/365 days and emergency readmission within 30 days for the treatment/control group pairs.

Table A1: Main regression results: coefficients of patient characteristics for the main analysis

	Cystoscopy		Hysteroscopy		Sterilisation	
	Estimate	SE	Estimate	SE	Estimate	SE
Diff-in-diff estimates						
Treatment (BPT group)	-0.051*	0.027	-0.269***	0.033	-0.034***	0.010
Diff-in-diff coefficient	0.360***	0.031	0.163***	0.022	0.038***	0.010
Age group						
Reference: 19-24						
26-30	0.015***	0.003	-0.001	0.008	0.002	0.002
31-35	0.021***	0.003	0.008	0.006	0.002	0.003
36-40	0.026***	0.003	0.034***	0.008	0.003	0.004
41-45	0.031***	0.005	0.062***	0.010	0.006	0.004
46-50	0.033***	0.005	0.092***	0.013	0.032***	0.012
51-55	-0.050***	0.015	0.124***	0.015	0.211*	0.108
56-60	0.026***	0.007	0.165***	0.019	0.630***	0.168
61-65	0.027***	0.008	0.216***	0.023	0.722***	0.117
66-70	0.017**	0.008	0.257***	0.025	0.630***	0.196
71-75	0.016*	0.008	0.286***	0.026	0.708***	0.004
76-80	0.017*	0.009	0.305***	0.027	0.040**	0.018
81-85	0.014	0.009	0.319***	0.027	0.091*	0.046
86-89	0.015	0.010	0.321***	0.027	0.152**	0.071
90+	0.028**	0.012	0.326***	0.027	0.001	0.005
BPT x 26-30	0.010*	0.005	0.081***	0.013	0.012***	0.004
BPT x 31-35	0.014**	0.007	0.100***	0.015	0.020***	0.005
BPT x 36-40	0.018**	0.007	0.090***	0.016	0.021***	0.005
BPT x 41-45	0.019**	0.008	0.087***	0.021	0.017***	0.006
BPT x 46-50	0.023***	0.009	0.059***	0.022	-0.014	0.012
BPT x 51-55	0.113***	0.018	0.026	0.023	-0.197*	0.108
BPT x 56-60	0.035***	0.011	-0.016	0.026	-0.599***	0.168
BPT x 61-65	0.032***	0.012	-0.076***	0.029	-0.670***	0.117
BPT x 66-70	0.042***	0.012	-0.124***	0.031	-0.613***	0.196
BPT x 71-75	0.043***	0.013	-0.155***	0.032	-0.664***	0.018
BPT x 76-80	0.044***	0.013	-0.183***	0.032		
BPT x 81-85	0.048***	0.014	-0.191***	0.033		
BPT x 86-89	0.053***	0.015	-0.195***	0.034		
BPT x 90+	0.049***	0.017	-0.126***	0.038	0.062	0.055
Sex						
Reference: Female						
Male	0.002	0.002				
BPT x Male	0.021***	0.004				
Past emergency visits						
Reference: no visits						
1 visit	-0.020***	0.004	-0.052***	0.006	0.000	0.002
2 visits	-0.026***	0.005	-0.008	0.006	0.001	0.003
3 visits	-0.030***	0.006	0.004	0.006	0.008**	0.004
4 visits	-0.037***	0.007	0.017***	0.006	0.003	0.007
5+ visits	-0.044***	0.008	0.030***	0.004	0.003	0.003
1 visit	-0.001	0.005	0.034***	0.008	-0.002	0.004
2 visits	0.039***	0.006	0.017**	0.007	-0.004	0.004
3 visits	0.050***	0.008	0.018**	0.008	-0.014**	0.006
4 visits	0.043***	0.009	-0.006	0.010	-0.012	0.008
5+ visits	0.060***	0.010	-0.016	0.010	-0.011**	0.005
Deprivation						
Reference: least deprived						
2nd decile	0.006	0.010	0.000	0.004	-0.002	0.002
3rd decile	0.007	0.009	0.004	0.005	-0.002	0.003
4th decile	0.011	0.011	0.007	0.006	-0.002	0.003
Most deprived	0.006	0.015	0.007	0.008	-0.003	0.003
BPT x 2nd decile	-0.008	0.017	-0.014	0.010	0.003	0.004
BPT x 3rd decile	-0.012	0.015	-0.027**	0.012	0.007	0.006
BPT x 4th decile	-0.016	0.017	-0.034**	0.014	0.008	0.007
BPT x Most deprived	-0.007	0.024	-0.030	0.023	0.016**	0.007
Constant	0.126***	0.015	0.613***	0.018	0.014***	0.003
Observations	3,859,365		1,560,714		303,256	

Notes: Dependent variable is the probability to be treated in the outpatient setting. Time period is from April 2009 to March 2016. Models are estimated by OLS with standard errors (presented in parenthesis under the coefficients), clustered at hospital level. Models are run separately for each treatment-control procedure pair (cystoscopy-sigmoidoscopy; hysteroscopy-lower genital procedures; sterilisation-vacuum aspiration). All models control for casemix and a set of year-by-month and hospital dummies.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A2: Sensitivity analyses of the difference-in-difference estimates of the impact of the BPT Outpatients scheme on the probability of treatment in the outpatient setting

Treatment group	Cystoscopy (1)	Hysteroscopy (2)	Sterilisation (3)
(a) Anticipatory/adjustment period			
DiD coefficient	0.395***	0.181***	0.045***
	(0.033)	(0.024)	(0.012)
Adjusted R^2	0.397	0.294	0.100
Number of hospitals	167	166	158
Observations	3,261,528	1,324,047	255,430
(b) Balanced panel			
DiD coefficient	0.357***	0.152***	0.040***
	(0.032)	(0.023)	(0.011)
Adjusted R^2	0.374	0.270	0.092
Number of hospitals	131	132	135
Observations	3,528,821	1,417,426	279,392
(c) Excluding covariates			
DiD coefficient	0.365***	0.138***	0.038***
	(0.031)	(0.022)	(0.011)
Adjusted R^2	0.384	0.249	0.087
Number of hospitals	168	167	158
Observations	3,859,365	1,560,714	303,256

Notes: Dependent variable is the probability to be treated in the outpatient setting. Time period is from April 2009 to March 2016, with the model (a) excluding period from October 2011 to October 2012 (6 month on either side of the start of the policy in April 2012). Model (b) excludes providers that did not report in all quarters of our study period. Models are estimated by OLS with standard errors (presented in parenthesis under the coefficients), clustered at hospital level. Model (c) excludes covariates. Models are run separately for each treatment-control procedure pair (cystoscopy-sigmoidoscopy; hysteroscopy-lower genital procedures; sterilisation-vacuum aspiration). All models control for casemix and a set of year-by-month and hospital dummies.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A3: Empirical test for the parallel trends assumption for the primary outcome measure

Treatment group	Cystoscopy (1)	Hysteroscopy (2)	Sterilisation (3)
(a) Repeated operation within 60-days			
DiD coefficient for 2010/11	0.001	0.004	-
	(0.001)	(0.006)	
DiD coefficient for 2011/12	-0.002	-0.003	-
	(0.002)	(0.003)	
Observations	1,486,961	647,729	
(b) Repeated operation within 365-days			
DiD coefficient for 2010/11	0.004	0.009	-
	(0.003)	(0.008)	
DiD coefficient for 2011/12	-0.014***	0.005	-
	(0.005)	(0.006)	
Observations	1,486,961	647,729	
(c) Emergency readmission within 30 days			
DiD coefficient for 2010/11	0.000	0.001	-0.002
	(0.001)	(0.001)	(0.002)
DiD coefficient for 2011/12	0.001	0.000	0.002
	(0.001)	(0.001)	(0.002)
Observations	1,486,961	647,729	141,069
(d) Volume			
DiD coefficient for 2010/11	9.762	0.899	-
	(7.075)	(9.620)	
DiD coefficient for 2011/12	-10.986	12.474	-
	(10.881)	(13.123)	
Observations	3,144	3,168	
(e) Spill-over effect			
DiD coefficient for 2010/11	0.026**	0.008	-
	(0.013)	(0.011)	
DiD coefficient for 2011/12	0.004	0.008	-
	(0.017)	(0.015)	
Observations	607,814	474,434	

Notes: Time period is from April 2009 to March 2012. Standard errors (presented in parenthesis under the coefficients) are clustered at hospital level. Models are run separately for each treatment-control procedure pair (cystoscopy-sigmoidoscopy; hysteroscopy-lower genital procedures; sterilisation-vacuum aspiration) and each outcome. All models include a constant, case mix variables and a full set of year-by-month and hospital dummies. The null hypothesis for the parallel trends assumption is that the DiD coefficients are jointly zero.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$