Substitution when a Wider-Scope Technology Becomes Available: Evidence from Robotic Hysterectomy

Nathaniel Breg*

Veterans Health Administration, Palo Alto
Stanford University

June 17, 2024

Preliminary and Incomplete: Please Do Not Distribute

Abstract

This paper investigates substitution that occurs when a wider-scope technology is introduced. Specifically, I study substitution to robotically assisted surgery – a form of minimally invasive surgery with greater scope of applications than laparoscopic surgery – in the context of total hysterectomy. Like other new technologies, robotic surgery has high fixed costs, and so only patients with the greatest potential relative potential net benefit from robotic surgery undergo it. The patients who have the most to gain from robotic surgery are those who are not good candidates for laparoscopic surgery a limited purpose minimally invasive surgery and would undergo open surgery in the absence of the robotic option. In an event study approach using Medicare inpatient data, I find that when a Hospital Referral Region adopts robotic hysterectomy, the percent of hysterectomies performed in an open fashion decreases, and the percent performed laparoscopically remains constant. Regions that adopt robotic surgery experience lesser chances of a long length of stay after a hysterectomy – a benefit of minimally invasive surgery – without increased chances of readmissions – a sign of no increases in surgical complications on average. Even though robotic and laparoscopic hysterectomy are both considered minimally invasive procedures, the data show that robotic and open hysterectomy are in fact closer substitutes.

JEL Classifications: I1, J0

Keywords: comparative advantage, health care productivity, medical technology, physician decision-making, surgery, technology adoption, women's health

*Email: nbreg@stanford.edu. Thank you to Katherine Hicks-Courant for crucial medical knowledge and perspective and to Shruthi Venkatesh for advice regarding the Medicare data. This work was partially supported by a Fellowship in Digital Health with the Center for Machine Learning and Health at Carnegie Mellon University, and by a fellowship with the Health Economics Resource Center at the Veterans Health Administration, Palo Alto. The views expressed in this article are those of the author and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government.

1 Introduction

Technological progress is often thought to improve productivity, but it can also widen the set of cases in which technology can be improved. Innovation that expands the possible uses of a technology are important in industries as disparate as health care and agriculture. Expansion of technological capabilities is particularly important in health care, where patients are heterogeneous in clinical presentation, affecting the appropriateness of alternative treatments.

I investigate whether the introduction of a wider-scope-of-purpose surgical technology draws more cases that would have been treated with the incumbent specific-purpose technology or from an older, less desirable surgical method. I study this in the context of total hysterectomy, the removal of the uterus and cervix. This is an important and common procedure which for many years could be performed with a minimally-invasive procedure called laparoscopic hysterectomy or with an older, more invasive approach, called open hysterectomy or abdominal hysterectomy. Ideally, all patients would be treated minimally invasively, but laparoscopic surgery, in which surgeons manipulate straight, stick-like instruments through small incisions, is not appropriate for all patients. Some patients have anatomical complexities that make "straight-stick" surgery difficult or impossible, in which case a patient might be treated with open surgery. Robotically assisted surgery is also a minimally invasive technology, but the motorized wrists on the ends of the robotic arms allow surgeons to perform in anatomically complex cases that would be more difficult in laparoscopic surgery. Thus, robotic surgery is a general-purpose form of minimally invasive surgery.

When robotic hysterectomy is introduced to an area, which types of cases does it draw? I find that patients of robotic hysterectomy likely would have been treated in an open fashion in the absence of the robotic option. I describe this setting with a Roy model of patients and their physicians choosing a surgical technology. I begin with a model from my prior work (Breg, 2024), in which agents perceive the laparoscopic technology to be advantageous in low-complex cases and abdominal technology in high-complex cases. Because the robotic surgery resembles the laparoscopic surgery but with improved dexterity in more complex cases, agent utility for robotic technology is weakly greater than for laparoscopic technology, with the gap between utility for robotic and laparoscopic surgery being greater in more complex cases than in simpler cases. Robotic surgery weakly dominates both incumbent technologies. However, robotic surgery, like other newly introduced technologies, has high fixed costs, so only the cases with the greatest relative utility of robotic surgery undergo it. A greater proportion of these cases are would-be abdominal than would-be laparoscopic, because high-complex patients who would have undergone open surgery have more to gain from the robotic approach than low-complex patients who would have undergone minimally invasive surgery even without the robotic technology. I draw on the modeling approach of Chandra and Staiger (2007, 2020), who model patients and their physicians choosing between intensive treatment and medical management of heart ailments using a Roy model. In my earlier work, I develop a Roy model of patients and physicians choosing between surgical technologies that affect patients on two dimensions of health differently and show the presence of tradeoffs can explain why the two technologies coexist. This paper considers a technological improvement of an incumbent technology that allows the technology to be used in more types of cases. The improved technology poses a tradeoff on fewer cases than the prior generation of the technology did. In such a situation where the new technology has high fixed costs, patients who would have undergone the older technology, rather than the narrow-purpose version of the technology, undergo it because they have greater relative utility for the new wider-scope technology.

I find evidence for this substitution pattern in Medicare claims for total hysterectomies. Following the Callaway and Sant'Anna (2020) method for conducting a difference-in-differences study with staggered adoption timing, I consider regions that adopted robotic hysterectomy at a particular quarter, compared them to regions that adopt later, estimate that adoption cohort's treatment effect, and then do so for each adoption cohort before averaging the treatment effects across cohorts. I find that in regions that adopt robotic surgery, the percent of hysterectomies performed robotically increases, the percent performed abdominally decreases, and the percent performed laparoscopically stays the same. The substitution pattern that I find suggests that the patients who undergo robotic hysterectomy likely would have undergone abdominal surgery in the absence of the robotic alternative, and thus the patients who perceive the greatest relative utility of robotic surgery are the would-be abdominal patients. Because robotic surgery and laparoscopic surgery are both minimally invasive, one might have guessed that those two procedures are the closest substitutes. However, robotic surgery confers greater benefit to would-be abdominal patients than to patients who would have undergone a different, narrow-purpose minimally invasive surgery. The phenomenon studied here is similar to that of Gross (2018), who studies general-purpose technology in the form of tractors that are flexible enough to be used on a wider set of crops than the incumbent tractor technology, facilitating the diffusion of tractors to parts of the United States that used horses rather than tractors previously. My paper provides additional evidence supporting his point that the extent of a technology's diffusion depends on the number of a technology's uses. My paper benefits from microdata about which technology is used in which applications, in which I observe which type of minimally invasive surgery – laparoscopic or robotic – is being used in a particular case. This allows me to see that the expansion of minimally invasive surgery is not due to robotic surgery expanding into both would-be open cases and would-be laparoscopic cases, but rather by expanding into just the cases that would not have experienced the minimally invasive technology otherwise.

This substitution pattern is unlikely to be driven by a concurrent composition change, since I find evidence consistent with regions experiencing no change in patient composition when they begin robotic hysterectomy:

there is no change in number of hysterectomies, and there is no change in the percent of hysterectomies in which the patient experiences a number of conditions.

I also find evidence that robotic surgery's introduction expands the benefits of minimally-invasive surgery. In event studies, I find that hysterectomy patients experience lesser chances of a long length of stay (a length of stay of 2 or more days) than the patients before the robot was introduced. This expansion does not seem to come with additional complications, at least on average. Event studies show no change in the chance of a readmission after robotic hysterectomy's introduction. Robotic surgery, a wider-scope technology, is expanding the use of minimally invasive surgery among patients who would have undergone open surgery otherwise.

Additionally, I conduct a hospital-level difference-in-differences study. In order to compare similar hospitals, I restrict my analysis to hospitals that are the first in their respective regions to adopt robotic surgery, as first-adopters may have different characteristics than their rivals who follow.[1] I find that the abdominal rates of hysterectomy decline when first-adopting hospitals being robotic surgery, but unlike patterns at the region level, I also find that first-adopter hospitals decrease their use of laparoscopic surgery. The region-level lack of change in laparoscopic rates might be the result of first-adopting hospitals doing fewer laparoscopic procedures and rival hospitals doing more of them.

Whether robotic surgery leads to better outcomes than alternatives, particularly relative to non-robotic laparoscopic surgery, is a subject of active discussion in the medical community (Crew, 2020; Davies, 2022-October; Bakalar, 2021). My paper suggests that this might not be the relevant margin for cost-effectiveness analysis. In the hysterectomy setting, I see that many of the cases treated robotically would have been treated in an open fashion otherwise. More generally, my paper suggests that wider-scope technologies with high fixed costs may draw more heavily on applications that would have gone without the narrower-scope technology, focusing efforts to analyze cost-effectiveness of the newer technology.

2 Three Procedures for Total Hysterectomy

I study substitution to robotically assisted surgery in the context of total hysterectomy, the removal of the uterus and cervix. This procedure could be performed abdominally, laparoscopically, or, eventually, robotically. This is an ideal procedure for studying the choice of surgical mode. First, hysterectomy is a common and important procedure. 93,000 commercially insured hysterectomies (Morgan et al., 2018) and 39,000 Medicare-covered hysterectomies (author's calculations) were carried out in the United States in 2012.

[1]An extensive literature explores how the characteristics of firms affect the timing of investments, under different varieties of competition. (See Vogt (1999) for a review of the theoretical literature.)

It is used to treat several serious conditions, including uterine fibroids, endometriosis, pelvic organ prolapse, irregular bleeding, and uterine, ovarian, or cervical cancer.

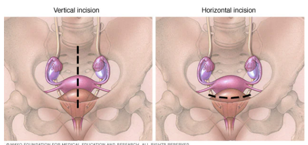
Second, hysterectomy can be performed with different technologies. It can be performed abdominally (Figure 1a), in what is called an open procedure, or it can be performed in a minimally invasive way. Laparoscopic hysterectomy was introduced in 1988. It uses long probing equipment to translate movements of the surgeon's hands into a smaller space in the patient's body (Figure 1c, 1d). It is thus minimally invasive, and as such can result in less blood loss and less scarring than abdominal surgery. Some observational clinical studies suggest that laparoscopic hysterectomy patients may have shorter lengths of stay in the hospital on average than abdominal hysterectomy patients (Aarts et al., 2015). However, laparoscopic technology has some drawbacks. For example, it features diminished dexterity and visibility for the surgeons.

Robotically assisted surgery grew among "soft tissue" (i.e., non-orthopaedic) surgeries in the 2000s. In 2000, Intuitive Surgical Systems received FDA approval for the Da Vinci robotically assistive device – under a fast track approval process because FDA deemed the device similar enough to laparoscopic devices, which were already FDA-approved. However, the Da Vinci device, which is still the only commercially available soft-tissue surgical robot, has a key difference from prior laparoscopic devices: it has motorized wrists near the end of its grasping and cutting arms that translate a surgeon's motions on the control panel to minute, dexterous movements within the patient (Figure 1f and Figure 1g). Contrast this with laparoscopic devices which are able to translate only the surgeon's in and out motion through the straight-sticks and cutting and grasping motions.

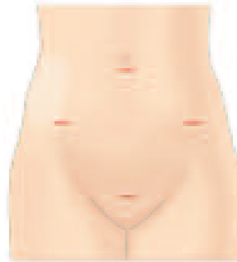
Third, different technologies for performing hysterectomy may have comparative advantages across different, heterogeneous patients. Some hysterectomy patients present with physicial complexities that make laparoscopic technology less advantageous. For example, laparoscopic hysterectomy is more difficult and less feasible on patients with large uteruses, no history of vaginal births, histories of abdominal surgery, and histories of cancer. (See American College of Obstetricians and Gynecologists (2017) and Walters and Ferrando (2021) for evidence-based guidelines.) Because the robotic device allows the surgeon more degrees of freedom, it may be capable of handing more anatomical configurations than laparoscopic equipment. [2]

Fourth, hysterectomy is an elective procedure. While it is used to treat many conditions that substantially diminish quality of life and, in some cases, threaten life, these conditions are rarely emergent. Thus, hysterectomy mode is likely to be chosen by weighing the comparative advantages of treatments in terms of the patient's clinical conditions and less likely than an emergent procedure to be chosen on some idiosyncratic

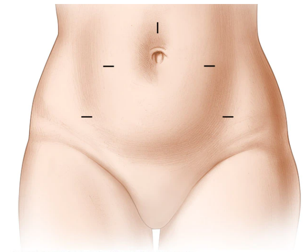
[2]There are some cases that are less appropriate for either kind of minimally invasive surgery. For both laparoscopic and robotic surgery, the procedure begins by tilting the patient slightly upside down in the Trendelenberg position and to fill the pelvic/abdominal cavity with gas. Patients with some heart and lung conditions are poor candidates for minimally invasive surgeries because these procedures are riskier for them.



(a) Possible incisions for abdominal – or, open – hysterectomy. Source: Mayo Clinic.

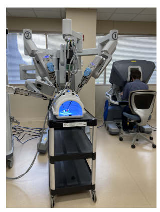


(b) Possible incisions for laparoscopic – or, straight-stick – hysterectomy. Source: Kaiser Permanente.



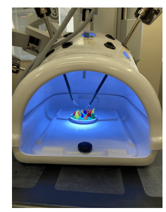
(c) Possible incisions for robotic hysterectomy. Source: Mayo Clinic.

(d) Examples of laparoscopic equipment. Source: Stryker.



(e) A Da Vinci robotically assistive device, with the author seated at the control panel.

(f) The controls for the surgeon.



(g) The DaVinci graspers, in a training exercise.

Figure 1: The long, slender nature of laparoscopic instruments allow hysterectomy to be performed with smaller incisions, but it also limit's the surgeon's dexterity. The robotic device also can be used through small incisions, but its motorized wrists facilitate greater dexterity.

provider-side basis like which doctor with which preferences or experiences was on-call on a particular night.

Finally, relative price of laparoscopic surgery likely plays a minimal role in the choice over hysterectomy methods. Hospital payments are made for Diagnosis-Related Groups (DRGs), and there are not separate Medicare DRGs for laparoscopic, abdominal, and robotic surgery. Physicians reimbursements are based on a fee schedule with respect to CPT codes. In 2018, Medicare payments for abdominal hysterectomies was \$1,042. Payment for laparoscopic surgery depends on uterus size and whether tubes are removed. The laparoscopic reimbursement was \$1,048 for uterus greater than 250 grams without tube removal, and \$1,249 with tub removal, and it was \$797 for uteruses less than 250 grams without tube removal, and \$920 with removal. Physicians are reimbursed for laparoscopic hysterectomies at the same rates as for robotic ones.

3 Model of Choice over Treatment Alternatives including a General-Purpose Technology

In this section, I analyze the decision that patients and their physicians make over surgical technologies. I modify a Roy (1951)-style model from my prior work (Breg, 2024), in which patients and their physicians choose between laparoscopic and abdominal surgery, on the basis of how these technologies affect patient potential health outcomes. Here, I extend that model to introduce a third alternative. The choice set then includes a general-purpose technology (robotic surgery), a specific-purpose technology (laparoscopic surgery), and an outside option (abdominal, or open, surgery). I use this model to clarify which case types will substitute to the robotic alternative when it is introduced.

Under the old choice set, patients and physicians together decide which type of surgery for the patient to undergo, laparoscopic (subscript L) or abdominal (subscript A) hysterectomy. Under the new choice, they may also consider robot surgery (subscript Ro). They make this decision in order to maximize the patient's utility[3]. The patient's utility is a weighted function of two adverse clinical outcomes, length of stay, S and readmission rate, R, and the distance a patient would need to travel to undergo the surgical procedure, T_L or T_A. This is in keeping with the models of Chandra and Staiger (2007, 2020), who consider treatment decisions made to maximize patient survival. If treatments have different comparative advantages over the two outcomes, then the choice will be affected by patients' (and physicians') relative marginal disutilities for the two adverse clinical outcomes.

Under the new choice set, the robotic alternative weakly dominates the laparoscopic alternative, and it may dominate the abdominal alternative or be superior to it in all but the most abdominal-favoring cases. However, the robotic technology incurs a fixed cost that is spread over the cases treated robotically, so only the cases that experience the utility from robotic surgery relative to their next-best alternative undergo it.

As I argue previously, length of stay and readmission rates are very plausible prominent features in the patient–physician indifference curve. A longer length of stay in the hospital is undesirable to the patient and exposes the patient to hospital-born infection. It is also likely correlated with the necessity for greater recuperation. The readmission rate is plausibly related to the onset of complications of the surgery. These clinical care outcomes are commonly studied in the medical and health services research literature comparing efficacy of treatments and practice patterns, and they are of interest to health care policy makers, currently subject to regulatory scrutiny under health care finance policy.

[3] One could consider the physician in Ellis and McGuire (1986)'s model, with the parameter governing the weight the physician places on patient health relative to hospital profits set so that the physician only cares about patient health.

3.1 Model

Let there be patients whose heterogeneity in clinical conditions can be characterized as a random variable θ that realizes values from zero to one. This might describe the physical complexity of a patient's case, with one representing more complex cases. Let the production of patient outcomes length of stay, S, and readmission rate, R, under each treatment method $j \in L, A, Ro$, for a given value of complexity θ be:

$$S_j(\theta, X, W_{S,j}) = \alpha_j + \beta_j\theta + \kappa_{S,j}X + W_{S,j} \tag{1}$$

$$R_j(\theta, X, W_{R,j}) = \gamma_j + \delta_j\theta + \kappa_{R,j}X + W_{R,j} \tag{2}$$

where all parameters are positive, X is a random vector of patient characteristics affecting the clinical outcomes, and $W_{S,j}$ and $W_{R,j}$ are random variables of mean zero representing idiosyncratic factors determining a patient's adverse outcomes. Condition on X and the idiosyncratic terms.

The patient–physician pair's joint indirect utility function depends on two adverse clinical outcomes – S and R – and the patient's distance or travel time to the hospital where procedure j is performed, T_j:

$$U_j(\theta, T_j) = u^B - \omega_S S_j(\theta) - \omega_R R_j(\theta) - \omega_T T_j \tag{3}$$

where u^B is "bliss utility," a maximum level of utility that could be achieved from the surgery but that is generally unattainable.

Consider the decision under the old choice set, L, A. Either one procedure type is performed for all patient types (i.e., all values of θ) or one procedure is performed for only some values of θ. Let us assume that no procedure is performed for all patient types. This is consistent with observations that both laparoscopic and open hysterectomies are performed within surgical services markets. For a given value of $Z \equiv T_L - T_A$, the laparoscopic procedure yields higher utility on one range of values of θ, and on the complementary interval, abdominal surgery yields higher utility. In this model and those of Chandra and Staiger (2007, 2020), the partition of the type interval into two intervals on which each procedure dominates follows from the linear production functions, but a "single crossing" of the utility functions with respect to θ does not require such functional form assumptions. [4]

If θ represents case complexity, I argue it is more plausible that low-θ patients experience higher utility under laparoscopic surgery than under abdominal surgery and that abdominal surgery has a comparative

[4] Indeed, Roy (1951) describes what is essentially a single-crossing without assuming functional forms of agents' utility, merely by assuming that the variance of outcomes of agents who made one choice is different from the variance of outcomes among agents who made the other choice.

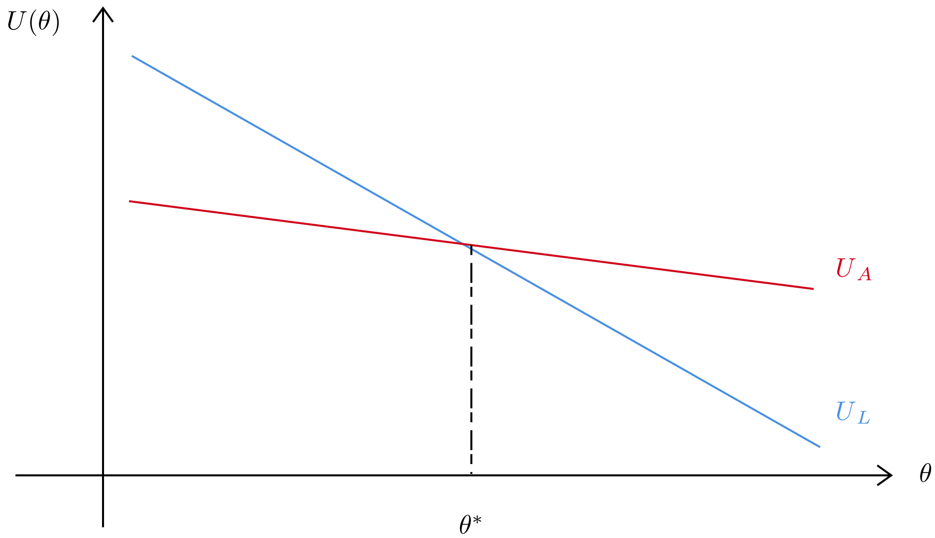


Figure 2: Utility of minimally invasive surgery and abdominal, or open, surgery as functions of patient type θ. Patients of type θ^* are indifferent.

advantage among patients with high θ, conditional on Z (Figure 2). Laparoscopic equipment has less dexterity and more limited visibility than abdominal surgery. Thus it is more difficult for surgeons to suture, make incisions, or see the anatomy of patients with trickier physical presentations and is incapable of performing some procedures like biopsies that accompany complex cases. For example, hysterectomy patients with large uteruses, patients who did not deliver any births vaginally, patients with histories of abdominal surgery, patients with history of cancer, and patients in other situations in which a specimen to be removed is near another internal organ like the colon present the surgeon with anatomical complexities for which surgery might benefit from more dexterity.

Additionally, assume that for all levels of θ,

$$S_L(\theta) < S_A(\theta) \tag{4}$$

which is consistent which the observation that laparoscopic equipment's smaller incisions are less invasive than open surgery and thus should result is less blood loss, less scarring, and shorter recovery times.

3.2 Choices under the Old Choice Set

For every relative distance to laparoscopic surgery $Z_L = T_L - T_A$, there is one patient type $\theta^{LA}(Z) \equiv \theta^*$ that is indifferent between laparoscopic and abdominal surgery. Patients with lesser θ choose laparoscopic surgery, and patients with greater θ choose abdominal surgery. Breg (2024) goes into more detailed explanation.

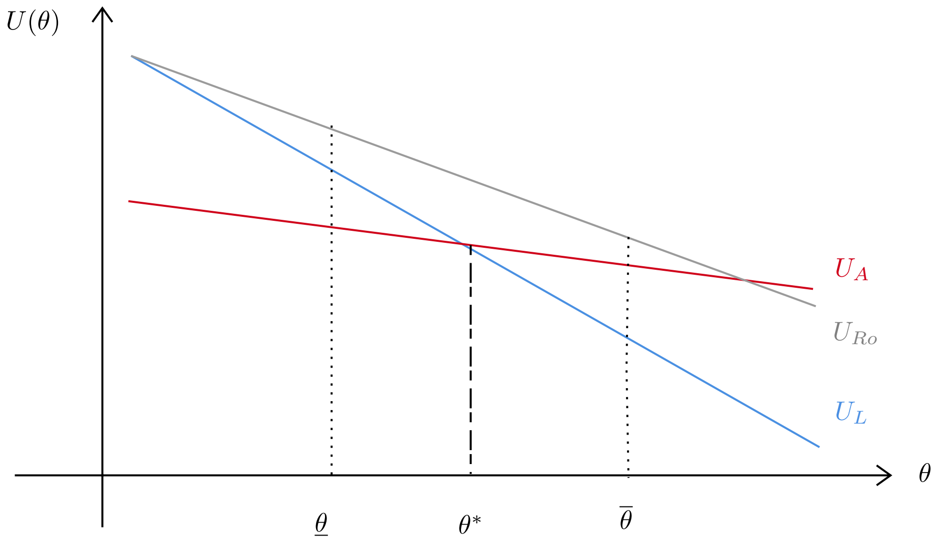
In my previous paper, I show that if laparoscopic surgery is better for length of stay in all cases, but nto everyone chooses it, it must be worse on the other dimension, the readmission dimension, at least in the marginal cases. This tradeoff can explain why laparoscopic surgery in not used in all cases despite its promises of minimal invasiveness.

3.3 Introduce a Wider-Scope-of-Purpose Minimally Invasive Alternative

Now consider the new choice set which additionally includes robotic surgery, Ro. It is a minimally invasive technology, like laparoscopic surgery, and so its potential length of stay is the same under laparoscopic surgery for each case type θ, i.e., $\beta_L = \beta_{Ro}$ and $\alpha_L = \alpha_{Ro}$. However, its improved dexterity makes it less likely to cause a complication, conditional on θ, and this advantage is increasing in θ, because dexterity is more important in more complex cases, i.e., $\delta_L > \delta_{Ro}$ and $\delta_{Ro} = \delta_A$ or is close to δ_A. Thus the utility for robotic surgery is weakly greater than the utility for laparoscopic surgery, for each θ. Robotic surgery also yields higher utility than abdominal surgery does for many if not all θ, which means it yields higher utility than abdominal surgery for at least some cases that choose abdominal surgery under the old choice set.

Regions that adopt robotic surgery incur a fixed cost that is spread over all cases that use the robotic device. All surgeries require an operating room, a trained surgical staff, and an intensive care unit on standby in case something goes wrong, but the robotic alternative stands apart from the other technologies because of the high expensive of the robotic device itself. Each Da Vinci device costs around \$ 2 million. (The device also poses a high variable costs, as parts of its robotic arms must be replaced after every several surgeries.) Even though robotic surgery dominates the other technologies in terms of how it affects length of stay and readmission risk, because of its high fixed cost, only the patients with the greatest relative net benefit from it undergo it.

For illustrative purposes, assume that patients are distributed uniformly across θ. This implies that a greater share of robotic patients are patients who would have chosen abdominal under the old choice set, $J^{old} = A$, than patients who would have chosen laparoscopic, $J^{old} = L$.



One can see in the plot that the patient type that experiences the greatest relative utility of robotic surgery is type θ^*, the type that is indifferent under the two-alternative scenario. The cases with the next-greatest relative utility are in the neighborhood around θ^*. Let the lower- and upper-bounds of the range of θ for whom the relative utility of robotic surgery exceeds the threshold imposed by the robotic technology's fixed cost be $\underline{\theta}$ and $\bar{\theta}$.

On average, cases with $\theta < \theta^*$ have less to gain from switching to robotic surgery than patients with $\theta > \theta^*$, because the former cases are already benefiting from minimally invasive surgery in the form of laparoscopic surgery. Therefore, a greater proportion of cases for whom robotic surgery has a greater relative utility, including the portion of the fixed cost, will be cases with $\theta > \theta^*$ who woudl have udnergone abdominal surgery in the absence of the robotic alternative. More details supporting this are in Appendix Appendix B.

Additionally, patients who do not switch to robotic surgery and remain abdominal patients should experience worse potential outcomes under abdominal surgery than the patients who substitute to robotic surgery from abdominal. That is because for a given procedure, patients with greater θ have worse potential outcomes, and the more inframarginal patients with greater θ will not switch to the robotic method.

In sum, if robotic surgery has the same advantages on length of stay that laparoscopic surgery has and if robotic surgery has the same complication-avoiding features as abdominal surgery, then most of the cases that would switch to robotic surgery are would-be abdominal cases. These strict relations can be loosened somewhat, and this prediction will still hold. However, the prediction becomes more ambiguous as robotic surgery becomes less like laparoscopic surgery on the length of stay dimension and less like abdominal surgery on the readmission dimension.

3.4 Estimands: Empirical Implications of the Model

3.4.1 Event Study Hypotheses

The model suggests that among hospitals that adopt robotic surgery, the percent of hysterectomies performed robotically will increase, the percent performed abdominal will decrease, and the percent performed laparoscopically will decrease by less than the percent performed abdominally. These have implications for the average adoption effect on the adopting hospitals. This quantity analogous to the well-studied average treatment effect on the treated (ATTs), so I will frame these quantities as ATTs from now on. The ATTs on the percent robotic, on the percent abdominal, and on the percent laparoscopic will be

$$ATT_{Ro} = \mathbb{E}[Ro|AdoptRobot, AfterAdoption] - \mathbb{E}[Ro|AdoptRobot, BeforeAdoption] > 0 \quad (5)$$

$$ATT_{A} = \mathbb{E}[A|AdoptRobot, AfterAdoption] - \mathbb{E}[A|AdoptRobot, BeforeAdoption] < 0 \quad (6)$$

$$ATT_{L} = \mathbb{E}[L|AdoptRobot, AfterAdoption] - \mathbb{E}[L|AdoptRobot, BeforeAdoption] \in [0, ATT_A) \quad (7)$$

These behavioral responses to the hospital adopting robotic surgery in turn imply changes in average patient adverse outcomes length of stay, S, and readmission, R:

$$ATT_{S} = \mathbb{E}[S|AdoptRobot, AfterAdoption] - \mathbb{E}[S|AdoptRobot, BeforeAdoption] < 0 \quad (8)$$

$$ATT_{R} = \mathbb{E}[R|AdoptRobot, AfterAdoption] - \mathbb{E}[R|AdoptRobot, BeforeAdoption] = 0 \quad (9)$$

These parameters can be estimated using event study approaches (see Callaway and Sant'Anna (2020), Sun and Abraham (2020), Borusyak et al. (2021), and de Chaisemartin and D'Haultfœuille (2020)).

Furthermore, because of sorting with respect to θ and the robot abdominal margin moving from the intermediate θ^*, after robotic surgery's introduction, the remaining abdominal cases will have a greater θ on average and so they should have worse adverse outcomes.

Note that these predictions assume that the potential outcomes of the hospital's cases do not change. To test the plausibility of this, I check whether the number of hysterectomies or the case mix of hysterectomies at adopting hospitals changes.

4 Data

To study substitution patterns across surgical procedures when a new technology is introduced, I require hospital-level data on the volumes of procedures performed, and to probe how much these observed patterns may be affects by patients switching hospitals and thus affecting hospital case mixes, I require demographic and clinical characteristics about the patients. To study how patient adverse outcomes may or may not change after technology adoption, I also require information about lengths of stays for surgical cases and information about patients' hospital utilization after the target surgeries of interest. In order to estimate the counterfactual utilization patterns and adverse outcomes if hospitals did not adopt the new technology, I need to observe similar hospitals before and after they adopt robotic surgery.

Medicare inpatient claims are well suited to this investigation. In them, I observe the procedures involved and clinical characteristics present on admission of all Medicare-covered total hysterectomies. I also observe which hospital the procedures were performed at and can link the claims to basic demographic information about each patient.[5]

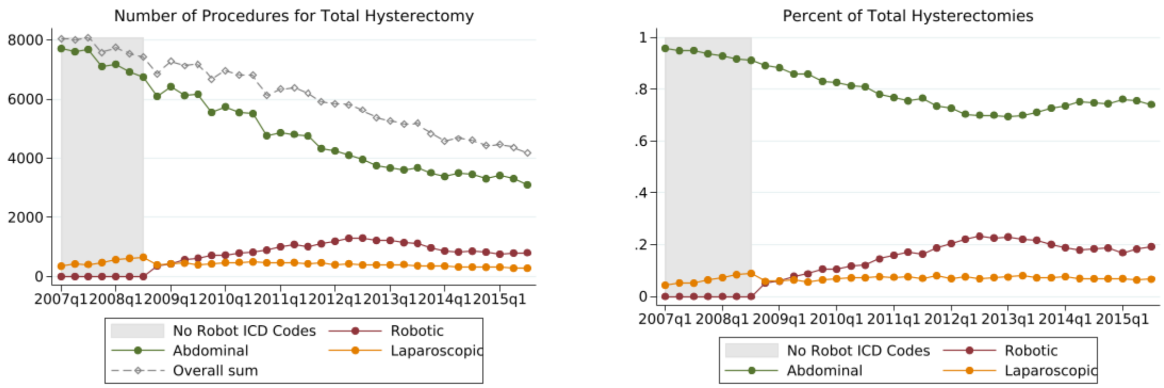
The national aggregate numbers preview the story of this paper. In Figure 10, the left panel plots the number of Medicare-covered total hysterectomies, in sum and by type. The number of abdominal hysterectomies has been declining over time, and the number of robotic hysterectomies increased through 2013. The number of laparoscopic hysterectomies has stayed the same. The right panel plots the percent of hysterectomies performed in the three modes. Again, one sees that abdominal procedures and robotic procedures make up decreasing and increasing shares of total hysterectomy, respectively, while the laparoscopic procedure's share holds constant.

As will be described in the methods section, my empirical strategy will compare utilization and outcomes in regions where robotic surgery is adopted to regions that have not yet (but will) adopt the procedure. I identify which Hospital Referral Regions (HRRs) each hospital is in, using the hospital Zip code in Medicare provider information and geographic crosswalks from the Dartmouth Atlas. Region-level adoption of robotic surgery is staggered across time.

My next empirical strategy will compare utilization and outcomes in hospitals that are the first to adopt robotic surgery in their respective regions to hospitals that are the first to adopt robotic surgery in their respective regions, which adopt later. I identify which Hospital Referral Regions (HRRs) each hospital is in, using the hospital Zip code in Medicare provider information and geographic crosswalks from the Dartmouth Atlas. Region-level adoption of robotic surgery is staggered across time.

[5] There were very few Medicare outpatient claims for hysterectomy in this period (141 hysterectomies in 2007, including total, subtotal, and radical hysterectomies). The few that I observe may be part of a different data generating process than the inpatient hysterectomies and are a very small segment of the hysterectomies in the population, so I do not include them in my analysis here.

Figure 3: Trends in Types of Hysterectomies as Seen in Medicare Inpatient Claims, over Calendar Quarters



As of the mid to late 2000s, abdominal and minimally invasive surgery had been coexisting. Most procedures were performed with the oldest technology, abdominal surgery. As the robotic procedure has been used increasingly for hysterectomies, the abdominal procedure has been used decreasingly. The share of hysterectomies performed laparoscopically – the minimally invasive technology that is newer than abdominal surgery but older than robotic – has been roughly constant over time.

Table 3 shows that Hospital Referral Regions adopted robotic surgery at different points in time from 2008 to 2012. (Robotic surgery was adopted before then, as documented in Horn, Sacarny and Zhou (2022), but as I explain in my description of the Medicare claims, I do not observe adoption before the third quarter of 2008.) Most of the regions that adopt robotic surgery do so in the earlier period of my data set. The rightmost column presents, for a given calendar quarter, the number of Medicare-covered total hysterectomies performed in Hospital Referral Regions that first adopted robotic surgery in that quarter. Earlier adopting regions had greater hysterectomy volumes at adoption time than later adopting regions.[6]

I will also compare utilization and outcomes in hospitals that are the first to adopt robotic surgery in their respective regions to hospitals that are the first to adopt robotic surgery in their respective regions, which adopt later. Table 4 show that first-adopting hospitals

My analysis compares changes in hospital utilization rates and outcomes to changes of other hospitals, so it is important to track observations of the same hospital over time. For part of that analysis, I use hospital IDs that are stable over time that Cooper et al. (2018) constructed and mapped to actual American Hospital Association hospital identifiers and to their own identifier of hospital systems sometimes comprised of multiple hospitals.[7]

[6] A coding issue requires me to drop cohorts that adopted robotic surgery in the first two calendar quarters during which the ICD-9 procedure code for robotic surgery existed. This code was introduced in the fourth quarter of 2008. Prior to that time, robotically-assisted procedures were coded as laparoscopic surgery (and, since up to 20 procedure codes can be indicated on a claim, this is additionally done after that time). In my data set, I consider procedures to be robotic if the ICD-9 code for robotic surgery is present, and I consider it to be (non-robotic) laparoscopic if the laparoscopic code is present but not the robotic code. If I were to attempt to estimate the effect of robotic adoption of utilization rates among hospitals that adopt robotic surgery during the first two quarters when the robotic ICD-9 code was in use, I would overestimate the degree to which the use of (non-robotic) laparoscopic surgery decreases after the introduction of robotic surgery. (Exploratory results, available upon request, confirm that event studies on groups treated in those two quarters have similar effect patterns as groups treated later except that they also have notable associations between adoption and decreasing use of laparoscopic surgery, clearly an artifact of this coding situation.)

[7] There is no official mapping from Medicare provider numbers to American Hospital Association hospitals. I use a machine

5 Methods

This paper seeks to estimate the average adoption effect among adopting regions and among adopting hospitals, on utilization rates and on adverse patient outcomes, as described in subsection 3.4. If adoption of robotic surgery is considered a "treatment" of regions and hospitals, then these causal quantities of interest could be recast as average treatment effects on the treated (ATTs). Different variations on the difference-in-differences approach has been shown to estimate the ATT. In short, I will compare changes in utilization rates in early adopting regions to regions that have not yet adopted robotic surgery (but will eventually, in my sample), and in early adopting hospitals to hospitals that have not yet adopted robotic surgery (but will eventually, in my sample) to estimate substitution patterns when robotic technology becomes available. I do this to estimate an event study for each treatment timing cohort or group – groups of hospitals that adopt robotic surgery in the same calendar quarter – and then aggregate the group-specific event studies together. I will also compare utilization rates in regions that have adopted robotic surgery early to regions that have not yet but will adopt it.

5.1 Identification

Causal Parameter of Interest. In my first line of empriical ivnestigation,I target the effect of adoption among regions where robotic surgery is adopted, and I also aim to estimate the effect of adoption among hospitals that are the first in their region to adopt robotic surgery. In the second line of inquiry where I estimate effects among adopting hospitals, I narrow the scope of my causal quantity of interest in order to avoid making unlike comparisons. An extensive literature in industrial organization establishes that the order in which firms in a market adopt technology may be a function of firm characteristics, depending on the form of competition. See Vogt (1999) for a review of the theoretical literature. Therefore, a clean causal parameter will represent the effect of adoption among hospitals that adopt robotic surgery at the same point in their respective markets' order. Thus, my goal is to estimate the effect of adoption among hospitals that are first adopters.

Following the new literature on difference-in-differences with staggered treatment timing (Borusyak, Jaravel and Spiess, 2021; Callaway and Sant'Anna, 2020; Callaway, Goodman-Bacon and Sant'Anna, 2021; Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2020; Sun and Abraham, 2020), I consider the

learning method that matches entities in different datasets based on the similarity of string descriptions as implemented in a publicly available Python program called Dedupe (https://dedupe.io), which is based on the greedy matching algorithm proposed by Chvatal (1979). I use the algorithm to match hospitals using the official names and Zip codes listed in the Medicare provider file and the American Hospital Association's Annual Survey. The algorithm learns from the author's manually matched training set of Medicare provider file–AHA Annual Survey entry pairs. I am unable to map some hospitals in Medicare claims to the AHA set. As a result, of the 204,284 Medicare-covered total hysterectomies that I observe from 2007 to 2015, 13,121 are not able to be included in the sample of hospitals tracked across time.

treatment-timing "groups," or cohorts, of regions or of hospitals that were the first in their Hospital Referral Regions to adopt robotic surgery and set out to estimate the dynamic average treatment effect on the treated (ATT) for each treatment group at each time period:

$$ATT(g,t) = \mathbb{E}[Y_t(1) - Y_t(0)|G_g = 1] \tag{10}$$

where t is a time period and g denotes the time period in which the region or hospital first adopted robotic surgery. Y alternatingly refers to the utilization rates or patient adverse outcomes about which I detail hypotheses in subsection 3.4. For simplicity of this methods section, let's call these all outcomes. $Y_t(1)$ is the hospital's potential outcome if they did adopt robotic surgery at time g, $Y_t(0)$ is their potential outcome if they did not, and G_g is an indicator for whether the hospital first adopted robotic surgery at time g.

I follow the identification argument and estimation procedure of group-specific ATTs (GATTs) laid out by Callaway and Sant'Anna (2020), which require three assumptions: parallel trends, irreversibility of treatment, and overlap of propensities to be treated or to be not treated. The remainder of these sections on identification and estimation largely follows their paper.

Irreversibility. I assume irreversibility of treatment, technically that $D_{t-1} = 1$ implies D_t. In a policy context such as this, this merely means that once a hospital or region adopts robotic surgery, it is considered to be "treated" for the remainder of the setting. I also assume that the panel data are independent and identically distributed, which merely allows one to view the potential outcomes as random and does not impose restrictions between potential outcomes and treatment timing.[8]

Parallel trends. Next, I assume parallel trends between the group treated at time g and a group treated at a later period:

$$\mathbb{E}\left[Y_t(0) - Y_{t-1}(0)|G_g = 1\right] = \mathbb{E}[Y_t(0) - Y_{t-1}(0)|G_g = 0] \text{ a.s.} \tag{11}$$

Treated and not-yet-treated regions or hospitals are assumed to have had the same year-on-year change in untreated potential outcomes at a particular year t, so the two groups would have had parallel outcome trends if neither had received treatment.

Loosely speaking, the parallel trends assumption states that the treated and untreated hospital did not differently experience anything before "treatment" that would affect their outcomes – that is, their utilization patterns or adverse outcomes. As an example, the difference-in-differences approach to estimating the ATT

[8] de Chaisemartin and D'Haultfœuille (2020) provide a method of identifying average treatment effects on the treated without requiring the irreversibility assumption.

would be invalid if hospitals that adopted robotic surgery had different trends in practice patterns or patient case mix that affected utilization or case potential outcomes from hospitals that were the first to adopt robotic surgery *later*. Differences in mortality *levels* between these two groups does not invalidate the approach.

One cannot directly empirically verify the parallel trends assumption, but one can look to different sources of suggestive evidence that it holds.

Applied economists conventionally estimate the difference-in-differences in periods before treatment. If the null cannot be rejected and estimates seem reasonably precise, researchers often take this as evidence that treated and not-yet-treated trends were parallel before treatment and further interpret this as suggestive evidence that their potential untreated outcomes after treatment would have been parallel. I follow this approach.

Overlap. Finally, I also assume overlap: for all time periods after the first one, $P(G_g = 1) > 0$ and for some $\epsilon > 0$, $p_g < 1 - \epsilon$ almost surely. In other words, a nonzero fraction of the population start to be treated at time g and there is a nonzero probability that an individual is not treated.

Identification of Group-specific Average Treatment Effect on the Treated. Callaway and Sant'Anna prove that under these assumptions, the group-time average treatment effect $ATT(g,t)$ is non-parametrically identified by the outcome regression-based estimand[9]:

$$ATT^{y}_{or}(g,t) = \mathbb{E}\left[\frac{G_g}{\mathbb{E}[G_g]}\left(\underbrace{Y_t - Y_{g-1}}_{\text{long diff among treated}} - \underbrace{m^{ny}_{g,t}}_{\text{estimated diff among not-yet-treateds}}\right)\right] \tag{12}$$

where $m^{ny}_{g,t} = E[Y_t - Y_{g-1} | D_t = 0, G_g = 0]$.[10] This estimand represents the difference between hospitals treated at time g and hospitals treated in the sample after time g in their differences in outcomes over time, that is, their difference-in-differences.

I estimate $ATT^{ny}_{or}(g,t)$ for each g to estimate a sequence of dynamic treatment effects for hospitals in cohort g. I estimate these treatment effects unconditionally.

Event Study Summary across Groups. To summarize across these groups and arrive at a single estimand $ATT(t)$ for a given event time t, I follow one of Callaway and Sant'Anna's suggestions for combining the $ATT(g,t)$ for a given t: I combine the group-specific effects by length of exposure to treatment, weighting more heavily groups treated later. Specifically, the ATTs are aggregated thusly:

[9] They also provide a doubly-robust estimator and an inverse probability weighting estimator that identify interpretable weighted combinations of GATTs.

[10] Callaway and Sant'Anna show that, when the parallel trends assumption holds unconditionally, the $ATT(g,t)$ for one particular group at a particular time t is consistently estimated by a two-way fixed effects estimator run on a sample composed only of that adoption group and the units that never adopted.

$$\theta(e) = \sum_{g \in \mathcal{G}} \sum_{t=2}^{\mathcal{T}} w(g,t) \cdot ATT(g,t) \tag{13}$$

for event time $e \equiv t - g$, where $\mathcal{T}$ is the last period in the sample, $\mathcal{G}$ is the set of treatment timing groups, and

$$w_e^{es}(g,t) = 1\{g + e \leq \mathcal{T}\} e\{t - g = e\} P\left(G = g | G + e \leq \mathcal{T}\right) \tag{14}$$

are nonnegative weights. The expression in Equation 13 that aggregates the group-specific effects estimands can be loosely interpreted as a sample-wide average treatment effect on the treated for event time e, $ATT(e)$.

Simple Aggregation. One final aggregation can be made to summarizing the treatment effect on the treated. Callaway and Sant'Anna suggest averaging across a group's dynamic treatment effects for all periods after treatment until some time – in this paper, I average dynamic effects from the first quarter until the eighth quarter after adoption – and then averaging those group-specific estimates across groups to arrive at one quantity

$$\theta_{sel}^{O} = \sum_{g \in \mathcal{G}} \theta_{sel}(g) P(G = g | G \leq \mathcal{T}) \tag{15}$$

where $\theta_{sel}(g)$ is the average of group G's dynamic treatment effects from quarter 1 to quarter 8, $P(G = g | G \leq \mathcal{T})$ is group G's share of units – either Hospital Referral Regions or hospitals – in the sample. An advantage of using this quantity rather than estimating the common two-way fixed effects estimator of the model

$$Y = \beta Post \times Treat + \alpha Post + \gamma Treat \tag{16}$$

in a setting where there are dynamic treatment effects and heterogeneous treatment effects across staggered treatment timing is that θ_{sel}^{O} does not make comparisons between treated units and already-treated units and does not weight the simple group-specific treatment effects by negative weights, which the two-way fixed effects estimator may do.

5.2 Estimation

The expectations in the estimands in Equation 12 can be estimated using conditional sample averages, so I use linear regression. Equation 12 nonparametrically identifies the treatment effect, so to estimate the average treatment effect on the treated, I plug in averages for expectations.

In the group-level regressions, I model standard errors assuming clustering at the hospital level, and in the event studies that aggregate results across groups, I estimate standard errors using a cluster-robust bootstrap procedure that accounts for multiple hypothesis testing. I implement estimation, aggregation, and inference using the R package by Callaway and Sant'Anna (2021).

6 Results

6.1 Adoption Effects on Procedure Choices

Here I present the main empirical results of the paper. I estimate how physicians and patients change their treatment decisions when robotic surgery becomes available in a region. The top row of Figure 4 presents the aggregation of the timing group-specific point estimates of the average adoption effect on the adopted regions, which are made by comparing utilization changes over time between regions that have adopted robotic surgery and regions that have not yet adopted it, calculated according to Equation 13. The left chart plots the difference-in-differences of the number of robotic hysterectomies performed, the middle plots the difference-in-differences in the number of abdominal hysterectomies performed, and the right chart plots the difference-in-differences in the number of laparoscopic hysterectomies. There is no evidence of the number of laparoscopic hysterectomies changing after robotic surgery's introduction.

The point estimates suggest there are no differential pretrends in utilization numbers of any of these procedures. (Of course, for robotic surgery, this is obviously so.) I estimate that eight quarters after adoption, the number of robotic hysterectomies increases by 2, the number of abdominal hysterectomies decreases by 0.75, and the number of laparoscopic hysterectomies stays the same. The point estimates suggest that utilization of robotic surgery is increases and utilization of abdominal surgery is decrease over time among adopting regions, relative to not-yet-adopting regions.

I further aggregate these estimates of dynamic effects into simple two-by-two difference-in-differences estimates in Table 1. Since these are convex combinations of the event study estimates by design[11], the point estimates are not of great surprise here, but the significance tests have more power to detect differences

[11] The algorithm by Callaway and Sant'Anna to construct these avoid the problems plagued by two-way fixed effects estimators in the presence of dynamic effects and heterogeneous effects across staggered timing groups, namely that those two-way fixed effects estimates are non-convex combinations of group-specific effects.

from zero. Because laparoscopic procedures make up a small shares of total hysterectomies at baseline, it is important to probe whether any effect of adoption on laparoscopic hysterectomy is truly zero. Again, the effects on numbers of robotic and abdominal procedures are significant and the estimated effect on laparoscopic procedures is too small for me to reject the hypothesis of no effect.

Results from compare hospitals that adopt robotic surgery to those that have not yet are presented in the second row. Again, numbers of robotic surgeries increase and the number of abdominal surgeries decrease. However, the point estimates for laparoscopic surgery suggest that numbers of those procedure relatively decrease in adopting hospitals, but these effects are both small and statistically insignificant. Table 1 shows that the simple difference-in-differences estimates for the effect on the number of laparoscopic hysterectomies is also statistically insignificant.

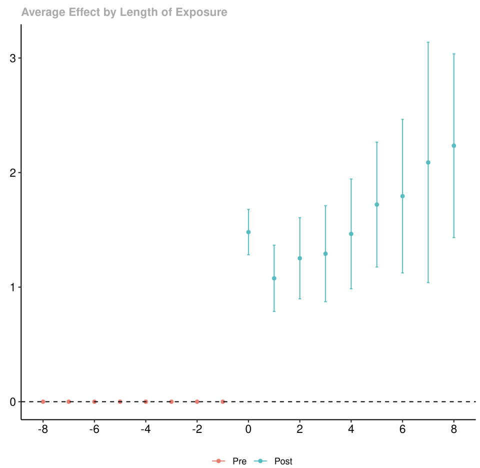
Hospital Referral Regions are large geographic boundaries that approximate markets for tertiary care, so they are likely larger than most patients' choice sets for hospitals for hysterectomies. Thus, comparing utilization rates between such large regions may underestimate a population's response to the availability of a new surgical alternative. That being said, the point estimates of the effects on numbers of robotic hysterectomies and abdominal hysterectomies are larger in magnitude in the hosptial-level studies than in the region-level studies.

Theses analyses sought changes in levels of procedures. Next, I conduct event studies on the percentages of hysterectomies perform robotically, abdominally, and laparoscopically. Figure 5 show that at both the region (top row) and hospital levels of analysis (bottom row) that any effects on utilization are discrete shifts, rather than dynamic effects. The percent of hysterectomies performed robotically increases by ten percentage points at the region level or 20 percentage points at the hospital level, the percent performed abdominal decreases by about ten percentage points at the region level or 20 points at the hospital level, and the percent performed laparoscopically doesn't change at the region level and perhaps drops by a statistically insignificant five percentage points at the hospital level. Again, Table 1 presents simple difference-in-differences estimates.

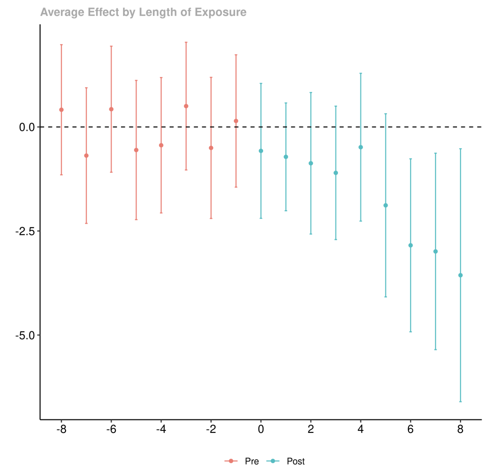
Note here that the effect on laparoscopic surgery at the region level is estimated to be very small and statistically insignificant, but the effect at the hospital level is estimated to be a seven percentage point decrease, statistically significant. Also note that the estimated effects on the shares of procedures are smaller in magnitude at the region level than at the hosptial level, which one would expect simply on the basis that more patients' choice sets should be affected at the hospital level of analysis, though this result on the shares of hysterectomies is the opposite of what is found in estimates the effects on the numbers of procedures.

Figure 4: Levels: Event Study Estimates of the Effect of Robotic Surgery Adoption on Number of Robotic/Abdominal/Laparoscopic Total Hysterectomies among Adopting Regions and Hospitals

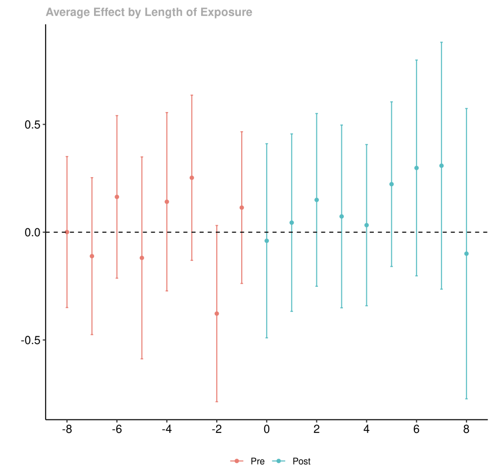
Hospital Referral Region-level Event Studies



(a) Robotic total hysterectomies

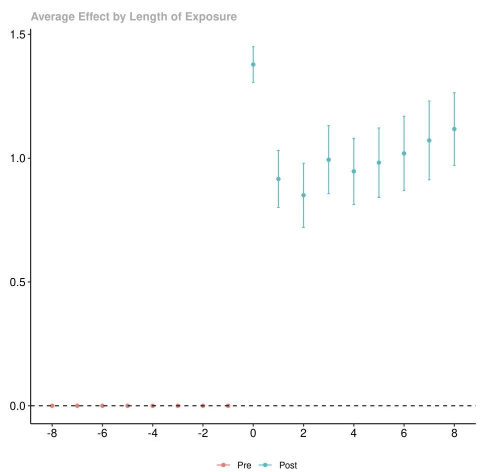


(b) Abdominal total hysterectomies

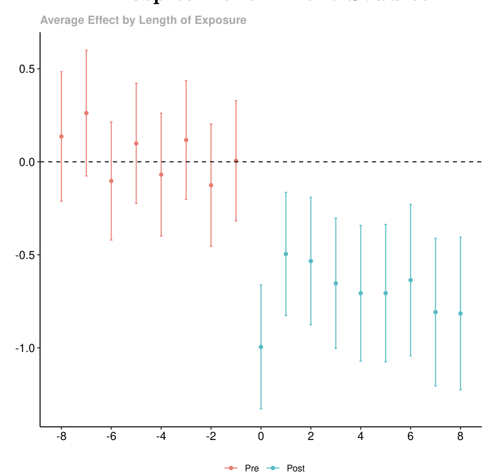


(c) Laparoscopic total hysterectomies

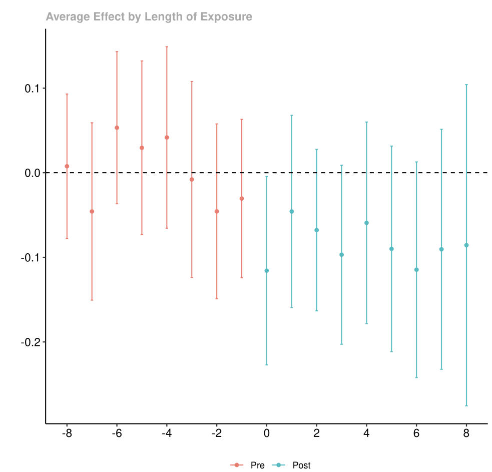
Hospital-level Event Studies



(d) Robotic total hysterectomies



(e) Abdominal total hysterectomies

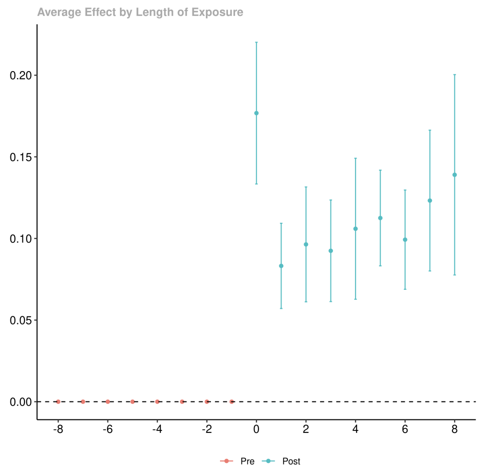


(f) Laparoscopic total hysterectomies

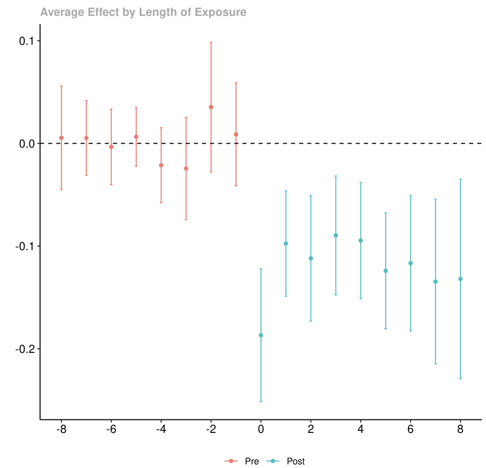
Point estimates at the calendar-quarter level. Average treatment effects on the treated are estimated for each treatment-timing group.

Figure 5: Shares: Event Study Estimates of the Effect of Robotic Surgery Adoption on Robotic/Abdominal/Laparoscopic Shares of Total Hysterectomies among Adopting Regions and Hospitals

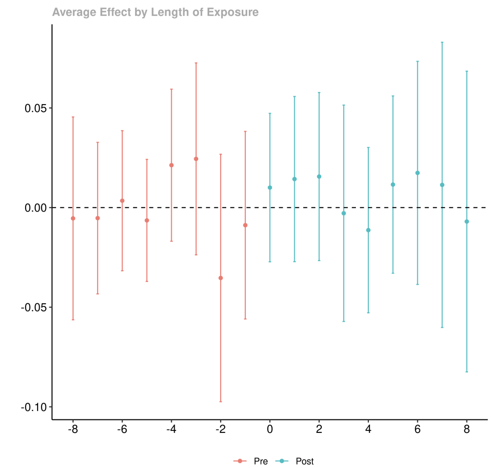
Hospital Referral Region-level Event Studies



(a) Robotic percent of total hysterectomies

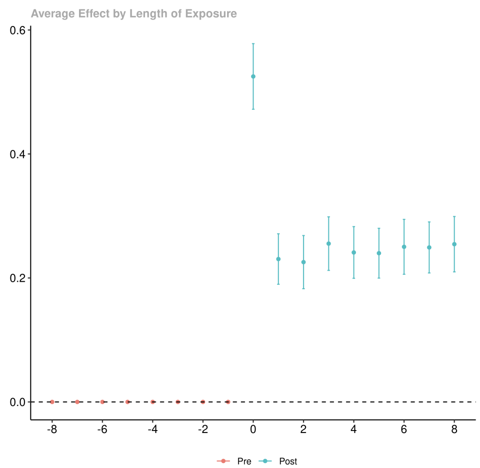


(b) Abdominal percent of total hysterectomies

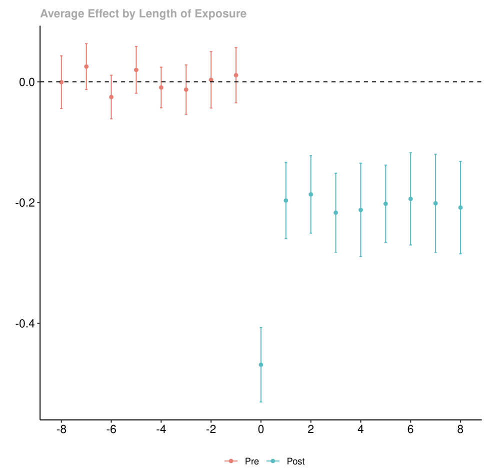


(c) Laparoscopic percent of total hysterectomies

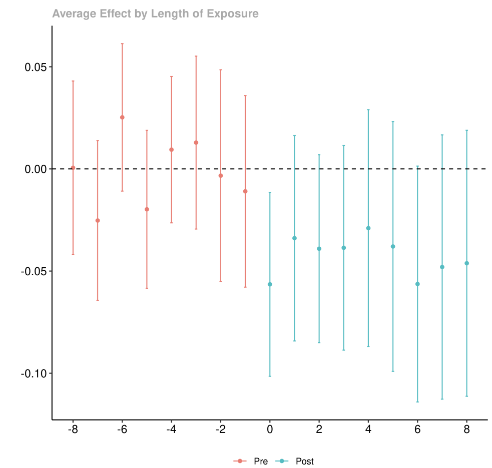
Hospital-level Event Studies



(d) Robotic percent of total hysterectomies



(e) Abdominal percent of total hysterectomies



(f) Laparoscopic percent of total hysterectomies

Point estimates at the calendar-quarter level. Average treatment effects on the treated are estimated for each treatment-timing group

6.2 Adoption Effects on Adverse Clinical Outcomes

Because different procedures have comparative advantages on different dimensions of care, substitutions from incumbent procedures towards new procedures should result in different rates of adverse patient outcomes. Figure 6 presents event studies at the region level (top) and hospital level (bottom) on the effects on having a length of stay of 2 or more days (left), on having a length of stay of three or more days (center), and on having an all-cause ten-day readmission. Point estimates suggest that incidence of a length of stay of two or more days differentially increases after the adoption of robotic surgery and that the incidence of a ten-day readmission stays the same, at both the region and hospital level. Effects on length of stay of three or more less conclusive and less consistent. At the region level, incidence of this adverse outcome may be small and somewhat grow, though these estimates magnitudes are statistically insignificant, and at the hospital level, the effect may be comparable to that of the effect on a length of stay of two or more days, though this effect is imprecisely estimated for later quarters.

Table 1 shows that the simple difference-in-differences estimates are similar across levels for each outcomes. Effects on length of stay are estimates to be a statistically significant relative reduction of 5.7 to 6.7 percentage points, the effect on length of stay of three or more days is a reduction of six (statistically insignificant) to seven (statistically significant) percentage points, and the effect on the probability of having a readmission are economically very small, statistically insignificant increases of 0.5 to 0.8 percentage points.

In sum, it appears that the introduction of robotic surgery and the substitution patterns that ensue lead to a reduction in average length of stay – a benefit of minimally invasive surgery – without any change in readmission rates – a sign of no additional serious surgical complications that would require inpatient care. Robotic surgery appears to have expanded the benefits of minimally invasive surgery to patients who would have undergone open surgery, without an increase in complications which work in my prior paper suggests would occur if minimally invasive surgery expanded by increasing the use of pre-robotic laparoscopic surgery (Breg, 2024). The new technology is increasing the scope of the minimally invasive technology, and the fact that readmission rates do not increase suggest that it is able to increase the scope because it is appropriate for more kinds of cases than laparoscopic surgery would be.

If cases are being sorted according to comparative advantage and if the margin between robotic and abdominal surgery starts from the original laparoscopic–abdominal margin and moves into more complex cases, then the remaining abdominal cases after robotic surgery has diffused should have worse potential outcomes under abdominal surgery that the cases choosing abdominal surgery before robotic surgery's introduction. In Table 1 and Figure 7, we see that after robotic surgery was used in a region, abdominal cases had greater risk of readmission than before.

Table 1: Estimates of Effect of Adoption on Procedure Choices among Adopting Regions and Hospitals

(a) Effects on Procedure Choices

	% abdominal	% laparoscopic	% robotic	N. abdominal	N. laparoscopic	N. robotic
HRR level	-0.132*	-0.0004	0.132*	-2.795*	0.0620	1.950*
	(0.018)	(0.015)	(0.008)	(0.654)	(0.123)	(0.191)
Hospital level	-0.192*	-0.069*	0.261*	-0.937*	-0.138	1.133*
	(0.034)	(0.030)	(0.006)	(0.141)	(0.071)	(0.041)

(b) Effects on Incidence of Adverse Outcomes

	LOS 2+	LOS 3+	Any 10-Day Readmission
HRR level	-0.067*	-0.073*	0.008
	(0.015)	(0.029)	(0.010)
Hospital level	-0.057*	-0.060	0.005
	(0.027)	(0.043)	(0.012)

(c) Conditional on Abdominal Cases, Effects on Incidence of Adverse Outcomes

	LOS 2+	LOS 3+	Any 10-Day Readmission
HRR level	0.002	0.061	0.043*
	(0.011)	(0.043)	(0.010)

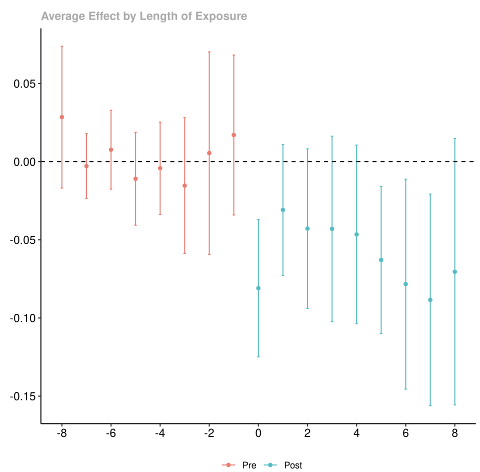
(d) Effects on Case Volume and Mix

	Num. hysterectomies	Age 75+	Malignant neoplasm	Pelvic organ prolapse	White	Black	Diabetes	Heart failure
HRR level	-0.782	-0.023	-0.036	0.002	-0.045*	0.048*	0.023	0.024*
	(0.625)	(0.034)	(0.031)	(0.019)	(0.022)	(0.016)	(0.026)	(0.011)
Hospital level	0.059	0.001	-0.019	-0.009	0.001	0.012	0.044*	0.006
	(0.143)	(0.030)	(0.048)	(0.033)	(0.028)	(0.024)	(0.023)	(0.014)

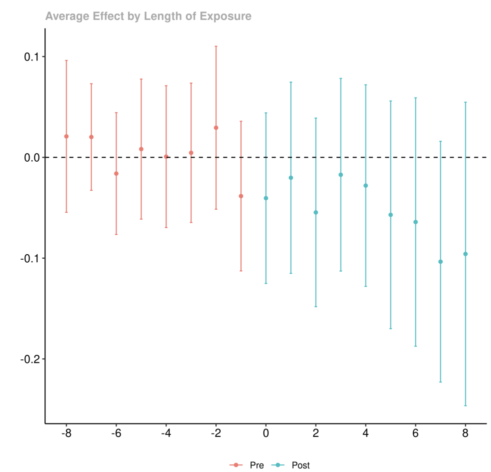
Simple difference-in-difference estimates calculated by aggregating cohort-specific estimates according to Callaway and Sant'Anna's procedure. Here are effects of adoption effect on share of hysterectomies performed a certain way and on the number of hysterectomies performed by type.

Figure 6: Event Study Estimates of the Effect of Robotic Surgery Adoption on Adverse Clinical Outcomes

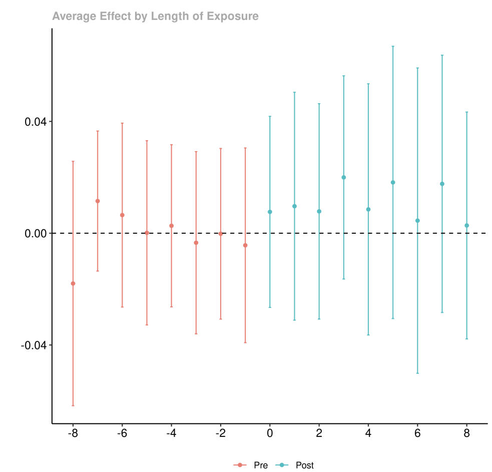
Hospital Referral Region-level Event Studies



(a) Probability that Length of Stay is 2 or More Days

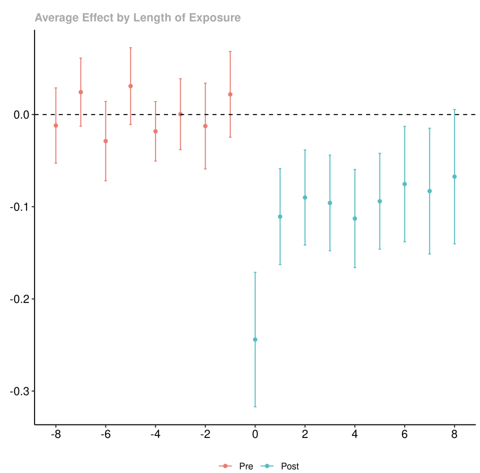


(b) Probability that Length of Stay is 3 or More Days

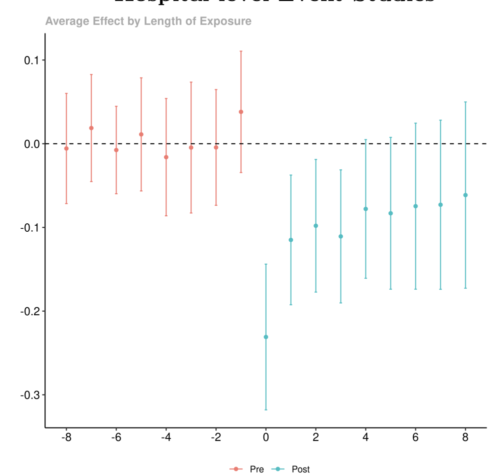


(c) Probability of Any 10-Day Readmission

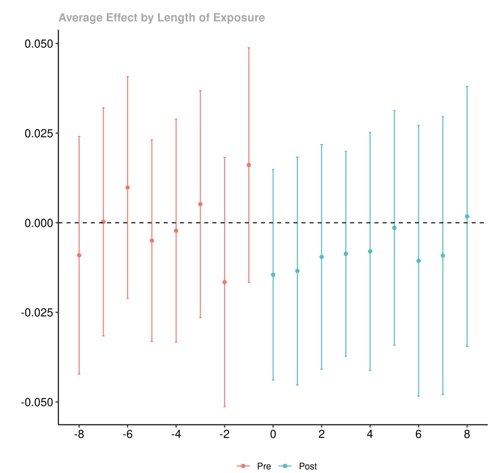
Hospital-level Event Studies



(d) Probability that Length of Stay is 2 or More Days



(e) Probability that Length of Stay is 3 or More Days

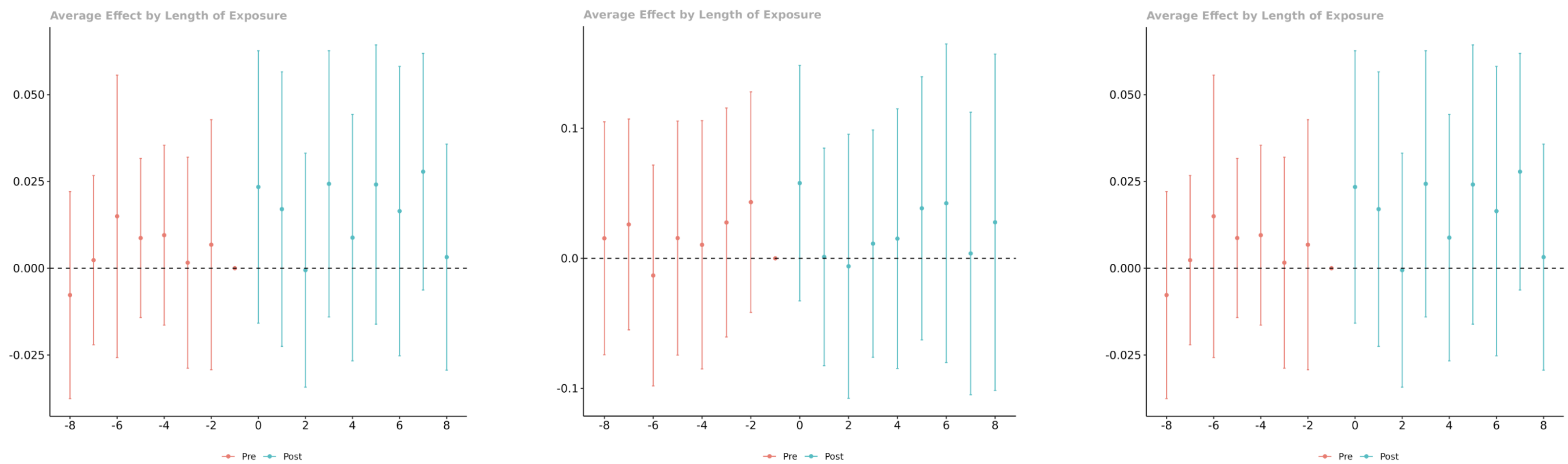


(f) Probability of Any 10-Day Readmission

Point estimates at the calendar-quarter level. Average treatment effects on the treated are estimated for each treatment-timing group.

Figure 7: Evidence from Cases That Choose Abdominal

Conditional on Abdominal Cases, Hospital Referral Region-level Event Studies



(a) Probability that Length of Stay is 2 or More Days **(b)** Probability that Length of Stay is 3 or More Days **(c)** Probability of Any 10-Day Readmission

Point estimates at the calendar-quarter level. Average treatment effects on the treated are estimated for each treatment-timing group.

6.3 Potential for Adoption Effects on Case Volume and Mix

At the population level, the introduction of a new treatment alternative could induce an increase in the number of hysterectomies among patients who prefer an outside option to the two incumbent technologies but prefer the new treatment to the outside option. Conceivably, these hypothetical patients would have different potential outcomes under the different treatment options than the existing population of hysterectomy patients, and this could be driven by them having different clinical characteristics. This story is also possible at the hospital level. Different kinds of patients may be newly interested in going to a hospital that newly adopts an additional technology. Such a shift could also be driven by non-clinical factors, such as different socioeconomic groups of patients being made aware of and having a different willingness to travel for the new technology.

Such possible effects are both interesting on their own, and they are necessary to investigate in order to more fully understand the effects on utilization patterns and on adverse outcomes that the previous results subsections present. If different types of patients are attending a hospital for surgery or going to a region for treatment, then estimated effects of adoption on adverse outcomes could be at least partially explained by these case mix changes.

I do not see clear evidence of change in hysterectomy volumes or case mix changes among hysterectomies at the region or hospital level. Figure 8 presents event study estimates of adoption's effect on the number of total hysterectomies, the percent of them performed on patients aged 75 and older, the percent of patients with a malignant neoplasm, and the percent with pelvic organ prolapase, at the region (top row) and hospital (bottom row) levels. Figure 9 presents analogous estimates on the effect on the percent of hysterectomies performed on patients who are white, who are black, have diabetes, and have heart failure.

The point estimates suggest that there were no differential pretrends in case volume or case mix in the lead up to hospitals adopting robotic surgery.

Point estimates suggest that the number of hysterectomies performed at the hospital level may increase, though those estimated dynamic effects are statistically insignificant, but the estimate of the simple difference-in-differences estimate in Table 1 is also statistically insignificant. This finding is somewhat noisy. This stands in contrast to the finding in Horn, Sacarny and Zhou (2022), who find that the adoption of robotic radical prostatectomy increases the numbers of prostatectomies at the hospital level and that it does so more than it increases the number at the hospital market level. One important difference between the hospital-level analysis in this study and that in theirs is that I am comparing hospitals that are the first in their region to adopt to hospitals that will eventually be the first in their region to adopt, whereas the previous work compares utilization in all prostatectomy-performing hospitals. A consequence of my approach is that I am likely estimating treatment effects among less heterogeneous hospitals since hospitals with different characteristics may adopt new technologies in different places in the order (Vogt, 1999) and while I am not comparing utilization in hospitals that adopt against utilization in their rivals. However, the effect among first-adopters may be different than the effect among second- or third-adopters, if first-adopters are already technology and/or quality leaders. It is conceivable that second- or third-adopters with less sterling reputations might experience a greater increase in demand after adopting a new technology than a market leader. Of course, it is also possible that the nature of demand for prostatectomies is different than that of the demand for hysterectomies.

Evidence on changes in case mix is not clear or consistent. The event studies suggest that adoption had no effect on the percent of hysterectomy patients who were age 75 or older or had pelvic organ prolapse. There is some statistically insignificant evidence of a dip in the rates of malignant neoplasms at the region level, though this seems unlikely in the simple difference-in-differences evidence, and there's no evidence for this at the hospital level. There is some evidence for a statistically significant four percentage point increase in the percent of patients with diabetes at the hospital level. The estimated effect is smaller and not significant at the region level. Diabetes patients would benefit from an additional alternative of minimally invasive surgery, since that condition makes healing more difficult.

The simple differences-in-difference suggest that the percentage of patients with heart failure increase by a statistically significant two percentage points at the region level, though this effects is much smaller and insignificant at the hospital level. Heart failure patients are bad candidates for any minimally invasive surgery, since patients undergoing those procedures need to be put in physical positions that are dangerous with heart or lung problems. This combined with the fact that the effect is evidence at the region level but not the hospital level raises questions as to whether this effect is real or just due to random variation.

There is evidence that the percent of patients who are white decreases and the percent who are black increases at the region level but not the hospital level. Again, because one would expect adoption effects to be strongest at the region level, this finding seems suspect. It is possible it is being driven by Hospital Referral Regions that adopt robotic surgery have different secular population characteristics, though it would be a strange coincidence for these differences in populations to change at the time that the first hospital in the region to adopt robotic surgery does so.

Finally, the event study point estimates suggest that the percent of cases with malignant cancer decrease right after adoption, in a way that dissipates over time and that is statistically insignificant. The simple difference-in-differences point estimates suggest any such effect is very small and statistically insignificant. From clinical and economic perspectives, there are stories in which such a decrease in cancer cases at the hospital level are conceivable or not. One the one hand, patients with greater socioeconomic status may be both more likely to know of and willing to travel for robotic surgery and less likely to have severe illness. On the other hand, the introduction of robotic surgery could be a meaningful addition to the choice set of cancer patients, since laparoscopic surgery is not well suited to various aspect of treating cancer, including taking biopsies and navigating around potentially difficult anatomical features, whereas robotic surgery is supposed more well suited to these challenges.

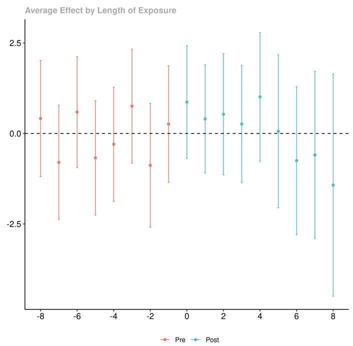
A caveat here is that many patient characteristics that affect the type of hysterectomy performed are not observable in Medicare claims, as described in **?**. A patient's full history of abdominal surgery, history of vaginal delivery, and uterus size and weight are important determinants in whether a patient is a good candidate for laparoscopic surgery, but there are no ICD-9 codes, which are the diagnosis codes on the claims, for these characteristics. Therefore, it is a possibility that the frequency of patients with unobservable characteristics that are contraindications for laparoscopic surgery increase at hospitals that adopt robotic surgery if those patients would have undergone surgery elsewhere if no hospitals in their choice sets had robotic surgery. That being said, even if that were true, it's not clear how that could affect the estimates of the adoption effect on adverse outcomes.

In sum, evidence that robotic surgery increases a hospital's volume of hysterectomies is noisy and inconclusive. Future work should investigate the effect among second- or third-adopters of robotic surgery. There's no sign of a discrete shift in case mix that follows the discrete shifts in patterns of utilization or adverse outcomes. There might be a (puzzling) change in racial composition at the regional level. Any change in volume or case mix at the region level would be surprising in this analysis, since Hospital Referral Regions are quite large.[12]

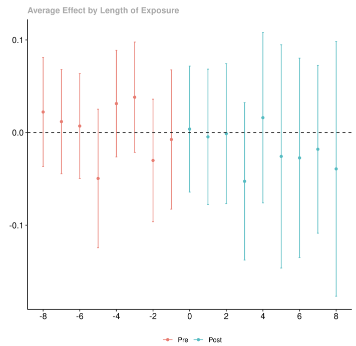
[12] One possibility is that patients in the northeastern United States, which is more densely populated and which has geographically smaller Hospital Referral Regions, are more likely to cross HRRs for surgical care.

Figure 8: Event Study Estimates of the Effect on Total Hysterectomy Volume and Case Mix among Adopting Regions and Hospitals

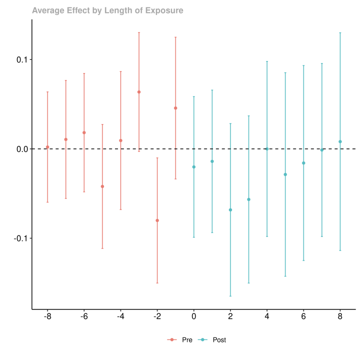
Hospital Referral Region-level Event Studies



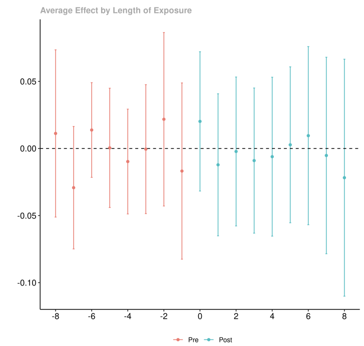
(a) Total Hysterectomies (Number of)



(b) Age 75+

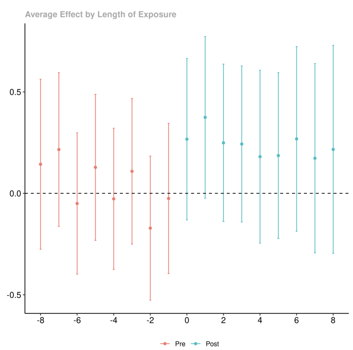


(c) Malignant Neoplasm

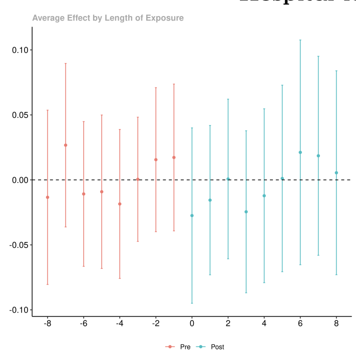


(d) Pelvic Organ Prolapse

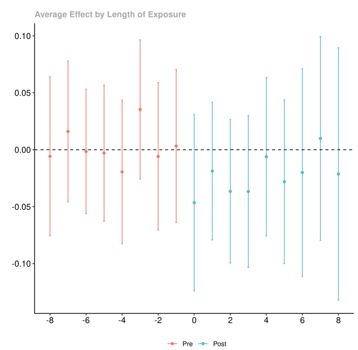
Hospital-level Event Studies



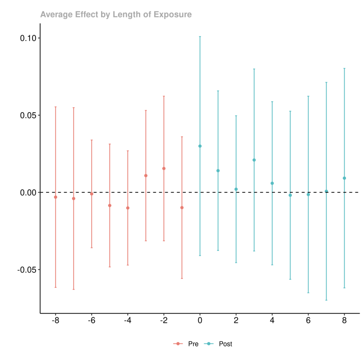
(e) Total Hysterectomies (Number of)



(f) Age 75+



(g) Malignant Neoplasm

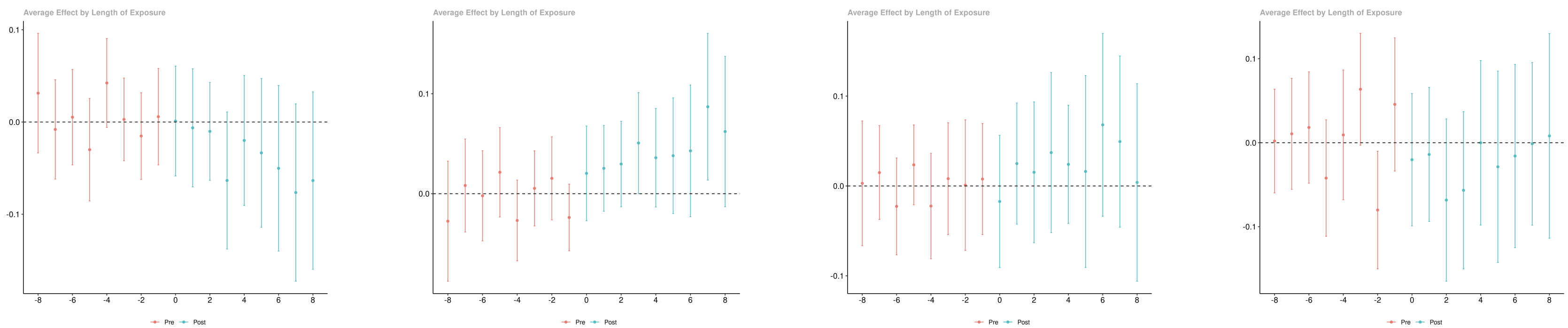


(h) Pelvic Organ Prolapse

Point estimates at the calendar-quarter level. Average treatment effects on the treated are estimated for each treatment-timing group.

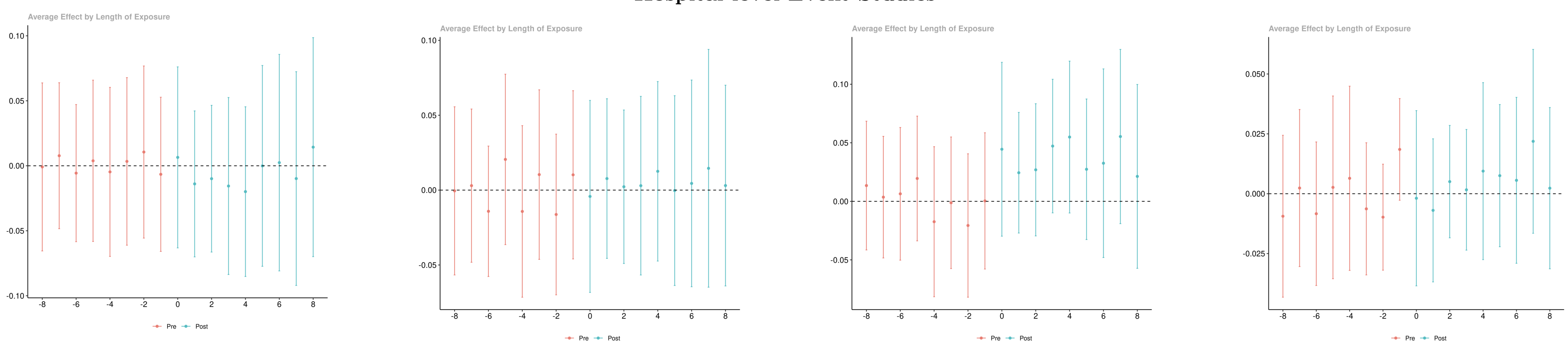
Figure 9: Event Study Estimates of the Effect on Total Hysterectomy Volume and Case Mix among Adopting Regions and Hospitals, Continued

Hospital Referral Region-level Event Studies



(a) White **(b)** Black **(c)** Diabetes **(d)** Heart Failure

Hospital-level Event Studies



(e) White DiD Estimate **(f)** Black **(g)** Diabetes **(h)** Heart Failure

Point estimates at the calendar-quarter level. Average treatment effects on the treated are estimated for each treatment-timing group.

7 Discussion

This paper explores how choices over technologies change when a new technology becomes available in the setting of choices of surgical procedures. Before robotic surgery become available, total hysterectomies were performed using either an older, more invasive approach, abdominal surgery, or a minimally invasive approach that yielded benefits over the older approach but was not able to be used in a broad scope of applications, laparoscopic surgery. Robotic surgery is a newer form of minimally invasive technology that has features that facilitate a broader scope of applications. Thus, it could be considered a wider-scope-of purpose form of minimally invasive surgery. Like other newly released technologies, robotic surgery has high fixed costs. Therefore, only the patients with the greatest utility for the new technology relative to their preferences among the incumbent technologies undergo the wider-scope technology. A Roy model of physician and patient choice over technologies suggests that the share of "switchers" to the new technology that come from the more limited minimally invasive technology rather than the older, open approach is dependent on how similar the wider-scope technology is to the limited technology on one dimensions – its potential for shorter lengths of hospital stays – and how similar it is to the older, open technology on another dimensions – its potential for surgical complications that lead to readmissions.

I empirically study changes in utilization rates of different hysterectomy procedures following the introduction of the robotic alternative. In a difference-in-differences approach, I compare rates in Hospital Referral Regions where robotic surgery is adopted to rates in regions where robotic surgery is not yet adopted but will be. I find that the numbers of robotic hysterectomies increases, the number of abdominal hysterectomies decreases, and the number of laparoscopic hysterectomies stays the same. I find the same pattern in the percentages of hysterectomies performed one way or another. Then I compare hospitals that are the first in their regions to adopt robotic surgery to hospitals that will become the first in their region to adopt. Again, I see numbers and shares of robotic procedures increasing and numbers and shares of abdominal surgeries decreasing. However, among first-adopter hospitals, I find evidence the numbers and rate of laparoscopic hysterectomies may decrease, though these estimates are sufficiently small and noisy to be statistically insignificant. This suggests that the patients who undergo robotic surgery either all would have undergone abdominal surgery in the absence of the robotic alternative, or a share of patients would have switched that is too small to be detected by my design.

I also find that the adoption of robotic surgery leads to a decrease in the chance of a long length of stay without any change in the chance of a readmission to the hospital. This suggests that the introduction and take-up of the robotic alternative expands the benefits of robotic surgery without increasing the chance of serious complications. My prior work suggests than an expansion of laparoscopic surgery would lead to an increase in readmissions, so it appears that robotic surgery is a wider-scope type of minimally invasive surgery, that is, a form of minimally invasive surgery that can be used in more kinds of cases than laparoscopic surgery.

These changes in utilization patterns and adverse outcomes do not seem to be driven by changes in the types of patients undergoing hysterectomy. Evidence that robotic surgery's introduction may increase the number of cases at the hospitals that are the first in their respective regions to adopt is noisy and inconclusive. One may speculate or caution that this result may not hold if the effect on second- or third-adopters of robotic surgery in a market were estimated, since first-adopters may already be quality-leaders with reputations for high technology up-take. There is also no clear change in case mix in terms of observable patient demographics or conditions, certainly none that match the discrete shifts in utilization and adverse outcomes that I observe.

References

Aarts, Johanna WM, Theodoor E Nieboer, Neil Johnson, Emma Tavender, Ray Garry, Ben Willem J Mol, and Kirsten B Kluivers, "Surgical Approach to Hysterectomy for Benign Gynaecological Disease," *The Cochrane Database of Systematic Reviews*, August 2015, *2015* (8), CD003677.

American College of Obstetricians and Gynecologists, "Choosing the Route of Hysterectomy for Benign Disease," Technical Report 701, American College of Obstetricians and Gynecologists June 2017.

Bakalar, Nicholas, "Are Robotic Surgeries Really Better?," *The New York Times*, August 2021.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event Study Designs," *Mimeo*, May 2021.

Breg, Nathaniel, "How Much Is Too Much? Assessing the Efficiency of Medical Technology Diffusion," June 2024.

Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, December 2020.

— **and Pedro H.C. Sant'Anna**, "Did: Difference in Differences," 2021.

— **, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna**, "Difference-in-Differences with a Continuous Treatment," July 2021.

Chandra, Amitabh and Douglas O. Staiger, "Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks," *Journal of Political Economy*, 2007, *115* (1), 103–140.

— **and** — , "Identifying Sources of Inefficiency in Healthcare," *The Quarterly Journal of Economics*, May 2020, *135* (2), 785–843.

Chvatal, V., "A Greedy Heuristic for the Set-Covering Problem," *Mathematics of Operations Research*, August 1979, *4* (3), 233–235.

Cooper, Zack, Stuart V. Craig, Martin Gaynor, and John Van Reenen, "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured," *The Quarterly Journal of Economics*, 2018.

Crew, Bec, "Worth the Cost? A Closer Look at the Da Vinci Robot's Impact on Prostate Cancer Surgery," *Nature*, April 2020, *580* (7804), S5–S7.

Davies, Benjamin, "Waiting for Godot Robot," *The Journal of Urology*, 2022-October.

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, September 2020, *110* (9), 2964–2996.

Ellis, Randall P. and Thomas G. McGuire, "Provider Behavior under Prospective Reimbursement: Cost Sharing and Supply," *Journal of Health Economics*, June 1986, *5* (2), 129–151.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, June 2021.

Gross, Daniel P., "Scale versus Scope in the Diffusion of New Technology: Evidence from the Farm Tractor," *The RAND Journal of Economics*, 2018, *49* (2), 427–452.

Horn, Danea, Adam Sacarny, and Annetta Zhou, "Technology Adoption and Market Allocation: The Case of Robotic Surgery," *Journal of Health Economics*, December 2022, *86*, 102672.

Morgan, Daniel M., Neil S. Kamdar, Carolyn W. Swenson, Emily K. Kobernik, Anne G. Sammarco, and Brahmajee Nallamothu, "Nationwide Trends in the Utilization of and Payments for Hysterectomy in the United States among Commercially Insured Women," *American Journal of Obstetrics and Gynecology*, April 2018, *218* (4), 425.e1–425.e18.

Roy, A. D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

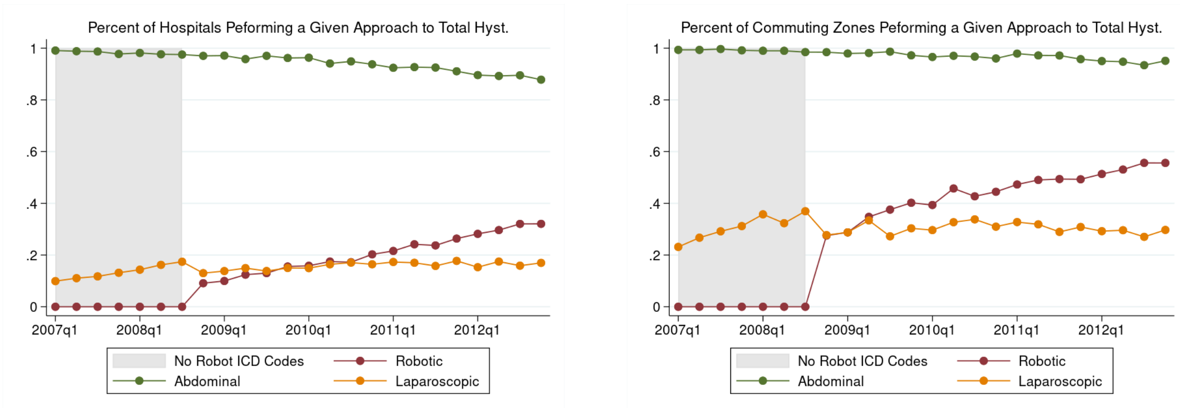
Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, December 2020.

Vogt, William B., "Existence and Welfare Implications of Strategic Behavior in Technology Adoption." PhD dissertation, Stanford University 1999.

Walters, Mark and Cecile Ferrando, "Hysterectomy: Selection of Surgical Route (Benign Indications)," in "UpToDate," Wolters Kluwer, November 2021.

A Further Data Description

Figure 10: Trends in Types of Hysterectomies as Seen in Medicare Inpatient Claims, over Calendar Quarters



As of the mid to late 2000s, abdominal and minimally invasive surgery had been coexisting. Most procedures were performed with the oldest technology, abdominal surgery. As the robotic procedure has been used increasingly for hysterectomies, the abdominal procedure has been used decreasingly. The share of hysterectomies performed laparoscopically – the minimally invasive technology that is newer than abdominal surgery but older than robotic – has been roughly constant over time.

Table 3: Number of Hospital Referral Regions in Each Adoption Cohort

Adoption	Number of Hospital Referral Regions	Number of Total Hysterectomies
Q3 2008	7	464
Q4 2008	117	4,076
Q1 2009	28	641
Q2 2009	23	499
Q3 2009	17	361
Q4 2009	11	129
Q1 2010	9	167
Q2 2010	10	83
Q3 2010	3	31
Q4 2010	7	56
Q1 2011	12	93
Q2 2011	3	10
Q3 2011	7	54
Q4 2011	5	28
Q1 2012	4	23
Q2 2012	2	12
Q3 2012	2	11
Q4 2012	1	15

This presents the number of Hospital Referral Regions from the Dartmouth Atlas in which a Medicare-covered inpatient robotically assisted surgery was first performed in a given calendar quarter. Number of hospital beds are reported in the AHA Annual Survey. The final column prints the number of Medicare-covered inpatient total hysterectomies performed in that quarter (not across all quarters in my sample) in that cohort.

Table 4: Number of Hospitals and Hospital Beds in Each Adoption Cohort

Adoption	Number of Hospitals	Cohort Total Number of Beds	Cohort Avg. Number of Beds	Number of Inpatient Total Hysterectomies
Q3 2008	*	5,465	781	107
Q4 2008	156	72,281	476	1,821
Q1 2009	82	40,734	509	727
Q2 2009	62	22,939	382	339
Q3 2009	64	28,803	450	441
Q4 2009	58	22,714	406	382
Q1 2010	39	15,612	400	225
Q2 2010	63	22,172	363	258
Q3 2010	29	8,843	328	125
Q4 2010	45	17,738	413	163
Q1 2011	48	13,609	296	166
Q2 2011	36	9,415	269	120
Q3 2011	39	10,679	281	137
Q4 2011	35	10,731	325	134
Q1 2012	45	11,989	272	136
Q2 2012	45	11,362	264	148
Q3 2012	42	12,319	316	130
Q4 2012	32	7,655	247	68
Q1 2013	31	6,485	224	78
Q2 2013	28	6,092	226	61
Q3 2013	19	4,516	251	35
Q4 2013	23	5,734	273	46
Q1 2014	17	3,754	221	28
Q2 2014	11	2,794	279	26
Q3 2014	11	2,170	197	22
Q4 2014	*	2,446	272	25
Q1 2015	20	6,143	307	48
Q2 2015	15	3,138	224	40
Q3 2015	11	2,339	234	21

This presents the number of hospitals in Medicare claims which I can map to the American Hospital Association's Annual Survey that first perform robotically assisted surgery in Medicare inpatients claims in a given calendar quarter. Number of hospital beds are reported in the AHA Annual Survey. * indicates the number is suppressed to comply with the Centers for Medicare and Medicaid Services' small cell size policy. The final column prints the number of Medicare-covered inpatient total hysterectomies performed in that first quarter of adoption (not across all quarters observed for that hospital) in such hospitals.

B Model details

The least θ that undergoes robotic surgery is the solution to

$$\max_{\underline{\theta}} \int_0^{\underline{\theta}} U_L(\theta)d\theta + \int_{\underline{\theta}}^{\underline{\theta}+N} U_{Ro}(\theta)d\theta + \int_{\underline{\theta}+N}^{1} U_A(\theta)d\theta \tag{17}$$

The first order condition implies that the relative benefit of robotic surgery will be the same for the case on the robotic–laparoscopic margin as on the robotic–abdominal margin:

$$U_{Ro}(\underline{\theta}+N) - U_A(\underline{\theta}+N) = U_{Ro}(\underline{\theta}) - U_L(\underline{\theta}) \tag{18}$$

Among the patients for whom the relative benefit of robotic surgery exceeds the average fixed cost, would more of them undergo abdominal surgery in the two-alternative scenario, or laparoscopic surgery? This restatement of the first order condition implies

$$(\theta^* - \underline{\theta}) = \frac{\frac{\partial U_A}{\partial \theta} - \frac{\partial U_{Ro}}{\partial \theta}}{\frac{\partial U_{Ro}}{\partial \theta} - \frac{\partial U_L}{\partial \theta}}(\bar{\theta} - \theta^*) \tag{19}$$

After plugging in the expressions for utility from Equation 3 and for adverse outcomes from Equation 1 and simplifying, one can see that the left-hand side of Equation 19, the laparoscopic-switching interval, is less than the abdominal-switching interval if

$$\beta_{Ro} - \frac{1}{2}(\beta_L + \beta_A) < \frac{\omega_R}{\omega_S}\left(\frac{1}{2}(\delta_L + \delta_A) - \delta_{Ro}\right) \tag{20}$$

which could be recast in more interpretable notation as

$$\frac{\partial S}{\partial \theta}|_{Ro} - \mathbb{E}\left[\frac{\partial S}{\partial \theta}|L \text{ or } A\right] < \frac{\omega_R}{\omega_S}\left(\mathbb{E}\left[\frac{\partial R}{\partial \theta}|L \text{ or } A\right] - \frac{\partial R}{\partial \theta}|_{Ro}\right) \tag{21}$$

Because laparoscopic and robotic surgery are both minimally invasive forms of surgery, the lengths of stay under both procedures would be the same and $\beta_{Ro} = \beta_L < \beta_A$, so the left-hand side is negative. Robotic surgery offers similar dexterity as abdominal surgery, which affects complication rate. If abdominal and robotic surgery have the same readmission rates, then $\beta_{Ro} = \beta_A < \beta_L$, and so the right-hand side is positive. The right-hand side thus exceeds the left-hand side, and so the laparoscopic-switching interval is exceeded by the abdominal-switching interval.