Medical Technologies with Comparative Advantages on Different Dimensions: Evidence from Hysterectomy

Nathaniel Breg*

Stanford University
Veterans Health Administration, Palo Alto

January 7, 2024

Please click here for the latest version

Abstract

Understanding the extent of technological diffusion is important to economics broadly and in the context of health care specifically. I show that new technologies may pose tradeoffs between different dimensions or quality or productivity. In a Roy model, I show that these tradeoffs can explain why two technologies coexist. The model also serves as a theoretical basis for using an instrumental variable to uncover evidence of tradeoffs. These local average treatment effects can be used in a benefit-cost analysis to assess whether the technology has diffused to an efficient extent. I use a patient's distance to hospitals performing laparoscopic (minimally invasive) surgery, relative to her distance to hospitals performing any surgery at all, as an instrument for whether she undergoes laparoscopic, as opposed to abdominal (open), hysterectomy. In Medicare inpatient claims, I find that laparoscopic surgery causes a shorter length of stay but a greater readmission rate, relative to abdominal hysterectomy, among patients on the margin between the alternatives with respect to this quasi-experiment. This demonstrates laparoscopic surgery's tradeoff, at least among some patient subpopulations. In a back-of-the-envelope benefit-cost analysis, I estimate that laparoscopic surgery may pose a net loss among these marginal cases, suggesting there may be too much laparoscopic surgery in this setting.

JEL Classifications: I1, J0

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

*Email: nbreg@stanford.edu. Special thank you to Lowell Taylor for his important insights that contributed to the intuition and model in this paper. I am also grateful to the other members of my dissertation committee – Martin Gaynor and David Chan – for their advice and support. Thank you to Katherine Hicks-Courant for crucial medical knowledge and perspective and to Shruthi Venkatesh for advice regarding the Medicare data. I also thank Elena Ashtari Tafti, Akshaya Jha, Emaad Manzoor, Petra Persson, Maria Polyakova, Elena Prager, Max Rubinstein, Natalia Serna, Edson Severnini, Becky Staiger, Ben Vatter, and seminar participants at the Boston University Technology & Policy Research Initiative, APPAM, ASHEcon, the Electronic Health Economics Colloquium, and the Stanford Health Economics Problem Solvers seminar for helpful advice and comments. 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. Any errors in this paper are my own.

1 Introduction

The speed and extent of technological diffusion is a broadly important subject in economics. In health care, new technology can drive health improvements but also expenditure increases. Due to asymmetric information in health care markets, it is important for policymakers to know why old and new technologies coexist and to assess their relative effectiveness. Understanding the welfare impacts of technological diffusion is tricky when innovation occurs on multiple dimensions of quality, when a technology's effectiveness differs across applications or subpopulations, and when there is selection into technology adoption on the basis of potential gains.

My paper studies how a newer technology's quality or productivity can explain its coexistence with an older technology. I construct a Roy model which shows that old and new technologies may coexist if the new technology presents tradeoffs between multiple dimensions of quality in at least some cases. It also shows that a technology's tradeoffs are apparent among marginal cases, and so evidence of those tradeoffs can be estimated using well-understood instrumental variable and marginal treatment effect methods. In turn, these estimates of the magnitudes of the tradeoffs among marginal patients can be used in benefit-cost analysis to assess whether a technology has diffused to an efficient extent. I study the choice between two alternative methods of total hysterectomy, the removal of the uterus and cervix: abdominal surgery, which entails making large incisions in the patient's abdomen, and laparoscopic surgery, in which long, straight devices are inserted through small incisions in the abdomen to detach the specimens. Despite laparoscopic surgery's promises of less blood loss and less trauma, it is only used in six percent of Medicare-covered hysterectomies. I show that laparoscopic hysterectomy poses a tradeoff between two key dimensions of quality among marginal cases.

Evaluating the extent of diffusion of technologies that are effective for some but ineffective for others is important in assessing health care productivity (Chandra and Skinner, 2012). Randomized controlled trials of medical treatments are costly to conduct, especially to estimate heterogeneous treatment effects across different subpopulations, and the selection of types of patients and providers into choosing different procedures on the basis of comparative advantage invalidates the comparison of average outcomes between procedures as an effectiveness assessment method. This paper both shows a new explanation for the coexistence of technologies and presents a way to assess the effectiveness and the efficiency of the use of new technologies using observational data, leveraging our understanding of the selection process underlying patients' observed choices and outcomes.

My main conceptual contribution is to show that old and new technologies may coexist if a technology poses tradeoffs between different dimensions of quality or productivity. The prior literature finds that products evolve along multiple dimensions of features and that consumers value these innovations, for example, in the markets for computed tomography (CT) scanners and for cars (Trajtenberg, 1989; Grieco, Murry and Yurukoglu, 2023). Different features could affect different dimensions of a technology's productivity. I demonstrate with a Roy model that two technologies may coexist because one technology offers relative improvements on one dimension but also pose setbacks on another dimension, at least in some applications. In my setting, laparoscopic surgery causes a shorter length of stay than the alternative, open procedure in all cases, but not all patients choose it. Therefore, it must cause greater readmission risk for patients near-indifferent between the two technologies. Prior work has found other factors in the speed or incompleteness of diffusion of new technologies, such as financial incentives (Finkelstein, 2007; Acemoglu and Finkelstein, 2008; Clemens and Gottlieb, 2014), information frictions Skinner and Staiger (2015), and administrative hurdles to billing for the use of new procedures (Dranove, Garthwaite, Heard and Wu, 2021). In other industries, coexistence of technologies has been attributed to the costs and benefits of different coinventions (Bresnahan and Greenstein, 1996), limitations imposed by product features (Gross, 2018), firm size (Karshenas and Stoneman, 1993), and lack of presence of complementary capital (Goldfarb, 2005). I show that technologies may coexist because old technologies may still have an advantage among some patients in terms that affect patients' physical health.

My first methodological contribution is to show how to uncover evidence of a technology's tradeoff by estimating the relative effectiveness of the technology among marginal patients using instrumental variable methods. I build on the intuitive and common approach of estimating treatment effects among patients on the margin between two alternatives by using a patient's relative distance to one alternative over the other as an instrumental variable (McClellan, McNeil and Newhouse, 1994). Similarly, I estimate the effects of laparoscopic, as opposed to abdominal, hysterectomy on two key adverse outcomes by comparing patients who live closer to hospitals that perform laparoscopic hysterectomy, relative to their distance to hospitals that perform any hysterectomies. I ground this approach with a microeconomic model of cases sorting between treatments on the basis of comparative advantage. Patients who are near indifferent between alternatives face a tradeoff between improvement on one dimension and detriment on another. They could also be induced into one or other by an instrumental variable. Marginal treatment effect methods from the labor econometrics literature identify the treatment effects of these marginal cases, and the local average treatment effect identified by two-stage least squares regression is a positively weighted combination of these marginal treatment effects (Heckman and Vytlacil, 1999, 2001; Heckman, Urzua and Vytlacil, 2006). [1]

Second, this quantification of the tradeoff can be used to show how to assess the efficiency of a technology's

[1] Other prior papers have used regression discontinuity and other evidence around policy thresholds to estimate the marginal value of care, for example, work by Almond, Doyle, Kowalski and Williams (2010).

diffusion using these estimates of marginal effects. Once you have estimates of effects among marginal uses, you can combine these estimates with valuations for the improvements and detriments on the different dimensions of quality in a benefit-cost analysis to assess the efficiency of the margin.

To illustrate the paper's central point, I build a Roy (1951) model in which patients and physicians choose a technology on the basis of how the alternatives affect two dimensions of productivity, rather than just one as is typical. This allows me to consider the role that differently signed changes on different dimensions of quality that is posed by a technological innovation. In this scenario, laparoscopic surgery must cause greater readmission risk than abdominal surgery, at least among marginal patients and inframarginal abdominal patients. The model I present is similar to that of Chandra and Staiger (2007) and (2020). In those papers, the comparative advantage of one treatment alternative versus another differs across patients, but the authors are agnostic as to what drives differences in comparative advantages across cases. In my model, I build out the utility functions so that they depend on two different outcomes, marginal treatment effects on which can be estimated empirically.

To estimate laparoscopic surgery's relative effectiveness among marginal cases, I use a patient's distance to her nearest hospital that performs laparoscopic surgery, relative to her nearest hospital performing any hysterectomy method, as an instrumental variable for undergoing laparoscopic, as opposed to abdominal, hysterectomy. I estimate the local average treatment effect in Medicare Part B insurance claims. This identification strategy, following McClellan, McNeil and Newhouse (1994), uses patients' preference for health care providers who are closer to their residence.[2] To assuage concerns raised by Hadley and Cunningham (2004) that the effect of distance on care choices may be confounded by socioeconomic conditions related to health, I control for a host of characteristics of the patient's neighborhood, some hospital characteristics, and Hospital Referral Region fixed effects. My work builds on this literature by grounding the approach in a microeconomic model that shows how the selection process allows the researcher to find evidence of tradeoffs by simply using instrumental variable regression to estimate effects among compliers who are on the margin between the two alternatives. Some prior work estimated patient preferences over improvements in overall health and avoidance of side effects using dynamic discrete choice modeling on data of patients updating their pharmaceutical choices periodically (Papageorge, 2016). In this paper, I present an approach that allows us to estimate evidence of tradeoffs using well-understood, simple-to-implement instrumental variable methods.

An important characteristic of my setting is that different alternatives have the comparative advantage for different subpopulations. I seek preliminary evidence that these differences in comparative advantage are

[2]See Burns and Wholey (1992) and Garnick et al. (1990) for evidence and reviews of literature on distance's role in patient choice of hospital, and see Card, Fenizia and Silver (2019) for a clarification of the relative distance identification strategy.

driven by heterogeneous treatment effects, by estimating marginal treatment effects according to methods from the labor econometrics literature (Björklund and Moffitt, 1987; Heckman and Vytlacil, 1999, 2001; Heckman, Urzua and Vytlacil, 2006; Brinch, Mogstad and Wiswall, 2017). My paper demonstrates the use of methods from labor economics to better understand heterogeneity of treatment effects across patients with unobserved (to the economist) characteristics that affect their treatment decisions, or, "essential heterogeneity" (Heckman et al., 2006).[3]

I find evidence that laparoscopic surgery poses a tradeoff between reducing a patient's length of stay in the hospital and increasing her readmission risk, at least for patients on the margin between the alternative hysterectomy methods. I estimate that patients who comply with the relative distance instrument experience about a 55 percentage point lesser chance of a length of stay of 2 or more days under laparoscopic surgery than under abdominal surgery, but they also experience a 23 to 36 percentage point increase in the chance of a 10-day all-cause readmission.[4] I am unaware of any other literature that uses instrumental variables to seek evidence of a tradeoff between different quality dimensions among marginal patients. Much of the health economics literature on patients' tradeoffs study their preferences for quality against cost or quality against distance in choosing among hospitals (e.g., Capps, Dranove and Satterthwaite, 2003; Ho and Pakes, 2014; Chandra, Finkelstein, Sacarny and Syverson, 2016), choosing whether to seek medical care (e.g., Manning et al., 1987; Finkelstein et al., 2012) or in choosing their use of pharmaceutical treatment (e.g., Duggan and Scott Morton, 2010).[5] My paper demonstrates that medical technologies may cause tradeoffs not just between health and costs but between one health dimension and another.

I use these point estimates to conduct a preliminary benefit-cost analysis of laparoscopic hysterectomy relative to abdominal hysterectomy among these compliers of the relative distance quasi-experiment, to demonstrate how to assess the efficiency of the extent of diffusion of a technology like laparoscopic hysterectomy. If an extra day in the hospital costs $2,490 (Foundation, 2021) and a readmission costs $15,200 (Weiss and Jiang, 2006), then my point estimates suggest that laparoscopic surgery poses a net loss of $2,054 in expectation among patients on the margin. This is likely an underestimate, since this excludes non-pecuniary costs, which are likely higher for a readmission than for an extra day in the hospital. Therefore, there may be too much laparoscopic surgery among these Medicare-covered hysterectomy patients, from the perspective of an individual patient's utility.

[3]See Basu, Heckman, Navarro-Lozano and Urzua (2007) for another application.

[4]I find that patients who live 1 mile farther from a laparoscopic-performing hospital, holding distance to any hospital constant, are 0.04 percentage points less likely to undergo laparoscopic, as opposed to abdominal, hysterectomy (off a 7 percent base rate). By something of a comparison, Chandra et al. (2016) find through conditional logit regression that patients are willing to travel 1.8 miles farther for a hospital with a 1 percentage point increase in quality.

[5]In the medical literature, Stewart, Lenert, Bhatnagar and Kaplan (2005) use vignettes to estimate patients' relative utilities over complications and quality of life under different prostate cancer treatment regimes, and Barry, Fowler, Mulley, Henderson and Wennberg (1995) conduct an experiment to see if an educational program on prostate cancer treatment alternatives affects patient decision-making and satisfaction.

The ratio of the estimates of the local average treatment effects of laparoscopic surgery on the two adverse outcomes imply that the marginal rate of substitution of a percentage point increase in the chance of a long length of stay for a percentage point reduction in readmission risk could be between -0.23 and -0.66, depending on model specification. However, because the choice of procedure could conceivably be influenced on the margin by actors like hospitals that could have different preferences over adverse outcomes than patients, this ratio may reflect a combination of different actors' preferences and objectives, rather than just a deep parameter of patient preferences.

Finally, I also find suggestive evidence that different cases perceive a different technology to have the comparative advantage because patients with the greatest unobserved resistance to (i.e., least propensity for) laparoscopic surgery would experience the greatest increases in readmission risk, even though they would experience the greatest potential reductions in length of stay due to that procedure, although these marginal treatment effect estimates are imprecise. These point estimates tell a story similar to that of Suri (2011), who finds that farmers who would experience the greatest benefit from adopting new hybrid maize technology also face the highest costs of adoption and thus do not use it.

My paper proceeds as follows. Section 2 describes the decision between laparoscopic and abdominal hysterectomy. In Section 3, I present the Roy model of patient and physician choices of surgical method and my finding that this model implies that marginal patients face a tradeoff between two health outcomes. I also present the empirical hypotheses for marginal and average patients and this implies, and I demonstrate how the ratio of the effects on marginal patients identify the marginal rate of substitution of a longer length of stay for a lesser readmission rate under certain conditions. In Section 4, I describe data, including most importantly the Medicare claims. In Section 5, I present the instrumental variable I use to identify marginal treatment effects and the local average treatment effect, the relative distance instrument, and justify its validity for these purposes. Section 6.2 presents the two-stage least squares method for estimating the local average treatment effect of laparoscopic surgery. In Section 6.3, I perform benefit-cost analysis to assess the efficiency of the extent of laparoscopic surgery's diffusion in this setting. Section 6.4 presents estimates of the marginal rate of substitution. Result from the local instrumental variable method and the separate method for estimating marginal treatment effects are shown in Section 6.5. Section 7 discusses my theoretical and empirical findings and concludes.

2 Total Hysterectomy

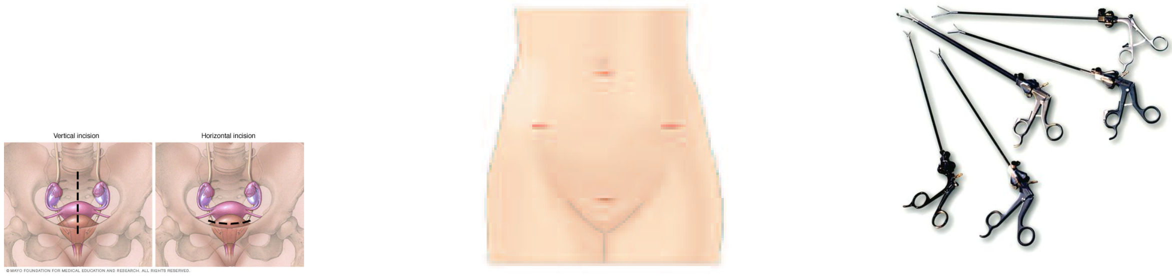
To evaluate a model of treatment decisions and the approach to estimating the marginal rate of substitution, I focus on total hysterectomy – the removal of the uterus and cervix – and the decision of whether to perform the surgery abdominally or laparoscopically. This is an ideal procedure for studying the choice of surgical mode. First, hysterectomy, the removal of the uterus, 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. It was the third most common operating room procedure among Medicaid claims, the forth most common such procedure among privately insured claims, the fifth most common such procedure among uninsured cases, and the eigth most common operating room procedure overall (Fingar, Stocks, Weiss and Steiner, n.d.). 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 1b, 1c). It thus is 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. Visibility and dexterity are important in order to, among other things, identify and track the ureter, so as not to injure it during surgery, which is a common cause of adverse outcomes after hysterectomy (Rassier, 2022).

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.)

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 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 versus abdominal surgery. Physicians reimbursements are based on a fee



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

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

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

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.

schedule with respect to CPT codes. In 2018, Medicare payments for abdominal hysterectomies was $1042. Payment for laparoscopic surgery depends on uterus size and whether tubes are removed. The laparoscopic reimbursement was $ 1048 for uterus greater than 250 grams without tube removal, and $1249 with tub removal, and it was $797 for uteruses less than 250 grams without tube removal, and $920 with removal.

3 Theory of Surgical Treatment Choice

Here I present a Roy (1951)-style model of patients and physicians jointly making treatment decisions. In this setup, patients and physicians together decide which type of surgery for the patient to undergo, laparoscopic (subscript L) or abdominal (subscript A) hysterectomy. They make this decision in order to maximize the patient's utility[6], which is primarily 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. However, in my paper, I consider treatments that affect two clinical outcomes and that might have comparative advantages for different outcomes. 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.

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

[6]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.

treatments and practice patterns, and they are of interest to health care policy makers, currently subject to regulatory scrutiny under health care finance policy.

The model also incorporates the patient's disutility of travel time to the facility for the procedure. A patient's distance to different hospitals is an important determinant of her choice of hospital. (Gaynor and Vogt (2000) review some of the prior evidence.) Different hospitals have equipment and staffs with different capabilities, so some hospitals perform laparoscopic surgery while other perform only abdominal surgery. Thus, distance of a patient to hospitals with laparoscopic technology relative to hospitals performing just open surgery affects her utility for laparoscopic surgery. This model feature will be used in the empirical strategy (section 5) for identifying effects among marginal patients.

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 $j \in L, A$, 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.

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

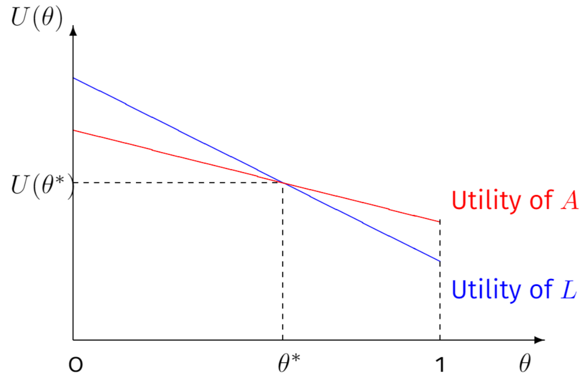


Figure 2: Utility of minimally invasive surgery and abdominal, or open, surgery as functions of patient type, θ. Types of lower θ are less appropriate for laparoscopic surgery, perhaps because of patient physical complexity, for example.

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 range 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. 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.[7]

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 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.

[7]The assumptions of the production functions here – namely, that outcomes under the two alternatives are linear with the same-signed slopes but with the one production function's slope steeper than the other – lead to similar predictions about outcomes for marginal agents as Roy (1951)'s assumptions that the log Normal-distributed random variables representing productivity in his two labor sectors are positively correlated with the fishing sector's productivity over potential workers having greater variance than the other. If the patient's utility were over just one outcome, the change in utility of the marginal patient when the nearest laparoscopic hospital is moved closer to her has the same sign as the change in the earnings of Roy's marginal worker when the (exogenous) price of fish increases.

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 by Different Patient Types

This section shows how the utility functions under laparoscopic surgery and under abdominal surgery and the adverse outcome production functions affect choices among patients with, alternatively, low and high θ types. Derivations of the findings are in Appendix A.

Consider the indifference curve of patient type $\theta = 0$ for fixed Z (Figure 4a) in terms of S and R, conditional on X, and the random shocks $W_{S,A}$, $W_{S,L}$, $W_{R,A}$ and $W_{R,L}$. Note that the slope of the indifference curve with respect to S is $m = -\frac{\omega_S}{\omega_R}$. Bliss utility, u_B, travel time, T_j, and preference weight on travel time, ω_T, are encoded in the indifference curve's R-intercept:

$$R(\theta) = \frac{u_B - \omega_T T_j}{\omega_R} - \frac{\omega_S}{\omega_R} S(\theta) \tag{5}$$

Each point represents a bundle of adverse clinical outcomes, and points L^0 and A^0 represent the bundles that type $\theta = 0$ can achieve under the two production technologies available L and A, respectively. Highest utility is achieved at the origin, and utility declines as S or R increases, conditional on (T_L, T_A). From the assumption that low complexity cases choose L, Appendix A shows that the production possibilities must lie on a line that is shallower than the indifference curve, and so type $\theta = 0$ patients experience shorter lengths of stay but greater readmission risk under laparoscopic surgery than under abdominal surgery (depicted in Figure 4a). High-complexity, type $\theta = 1$ patients choose abdominal surgery, under which they experience a lesser readmission risk but longer length of stay (Figure 4b). Appendix A.3 shows that, for a given value of the difference in distances, $T_L - T_A$, there exists a θ^* such that patients are indifferent between laparoscopic and abdominal surgery.

3.3 Predictions about Outcomes among Patients on the Treatment Margin

There is one θ for a given $Z = T_L - T_A$ such that the patient is indifferent between procedures. For a given value of Z, call this $\theta^{LA}(Z) = \theta^*$ to simplify notation.

For a patient indifferent between laparoscopic and abdominal surgery, it is true that

$$\omega_S \cdot (S_A(\theta^*) - S_L(\theta^*)) - \omega_T \cdot Z = \omega_R \cdot (R_L(\theta^*) - R_A(\theta^*)) \tag{6}$$

where $\theta^{LA}(Z) \equiv \theta^*$ is the value of θ for which a given value of $Z = T_L - T_A \geq 0$ makes the patient indifferent.[8]

Appendices A.2 and A.3 show that it follows from the comparative advantage assumption (that low-θ types choose laparoscopic surgery and high-θ types choose abdominal surgery) that the complexity type of the patient who is indifferent, θ^*, decreases when Z increases. Let's refer to the component of utility that is affected by complexity type θ but excludes the disutility of travel time, $u_B - \omega_S S(\theta) - \omega_R R(\theta)$, as clinical utility. Patients who are indifferent between the treatment methods when $Z = 0$ have less relative clinical utility from laparoscopic surgery than patients who are indifferent for a large Z – i.e., for patients who are indifferent when the laparoscopic hospital is much farther from their residence than the hospital without laparoscopic surgery.

Now let's analyze the difference in potential readmission rates for patients who are indifferent, i.e., for whom Equation (6) holds. Recall the earlier assumption that $S_L(\theta) < S_A(\theta)$ for all values of θ, because laparoscopic surgery is always less invasive than abdominal surgery. The patient who is indifferent at $Z = 0$ must have a greater readmission rate under laparoscopic surgery than under abdominal surgery, i.e.:

$$R_L\left(\theta^{LA}(Z=0)\right) - R_A\left(\theta^{LA}(Z=0)\right) < 0 \tag{7}$$

The changes in the differences in potential outcomes among indifferent patients when Z increases, or equivalently, when considering patients with greater relative clinical utility from laparoscopic surgery, are:

$$\frac{d\left[S_L(\theta^*) - S_A(\theta^*)\right]}{dZ} = \omega_T \cdot \left(\frac{\beta_L - \beta_A}{\omega_S(\beta_A - \beta_L) - \omega_R(\delta_L - \delta_A)}\right) \tag{8}$$

$$\frac{d\left[R_L(\theta^*) - R_A(\theta^*)\right]}{dZ} = \omega_T \cdot \left(\frac{\delta_L - \delta_A}{\omega_S(\beta_A - \beta_L) - \omega_R(\delta_L - \delta_A)}\right) \tag{9}$$

It follows from the comparative advantage assumption that the two derivatives cannot both be positive (see Appendix A.4).

In summary, the model says that among marginal patients, the relative readmission rate under laparo-

[8] Recall that Z, the difference between a patient's distance to her nearest laparoscopic-performing and hysterectomy-performing hospital, T_L, and the distance to her nearest hysterectomy-performing hospital, T_A, is weakly positive. All hysterectomy-performing hospitals perform abdominal surgery, but not all hysterectomy-performing hospitals perform laparoscopic surgery.

scopic surgery, $R_L(\theta^*) - R_A(\theta^*)$, will be positive for marginal patients with the least relative clinical utility for the laparoscopic method, if not for marginal patients of all types. The differences in potential lengths of stay, $S_L(\theta^*) - S_A(\theta^*)$, (which is taken to always be negative) and the differences in potential readmission rates will each be either increasing or decreasing with respect to patient's relative clinical utility of laparoscopic surgery, but they cannot both be decreasing. If they were, then laparoscopic-resistant patients would experience both greater improvements in length of stay and lesser deteriments from readmission risk than laparoscopic-prone patients.

3.3.1 Estimands: Empirical Implications of the Model

What empirical questions does this theory lead to? This subsection shows that the predictions about outcomes among indifferent patients with a given level of unobserved resistance to laparoscopic surgery (i.e., a given level of relative health "costs" to laparoscopic surgery unobserved by the analyst) leads to predictions about *marginal treatment effects* and, in turn, *local average treatment effects*. Consider the relative utility under laparoscopic surgery, L, versus under abdominal surgery, A, rearranging terms:

$$
\begin{aligned}
&U_L(\theta, T_L, X, W_{S,L}, W_{R,L}) - U_A(\theta, T_A, X, W_{S,A}, W_{R,A}) \\
=&\underbrace{\omega_S(\alpha_L - \alpha_A + W_{S,L} - W_{S,A}) + \omega_R(\gamma_L - \gamma_A + W_{R,L} - W_{R,A}) + [\omega_S(\beta_A - \beta_L) + \omega_R(\delta_L - \delta_A)]\theta}_{\equiv V,\text{ unobserved}} \\
&+ \underbrace{\omega_T(T_L - T_A) + [\omega_S(\kappa_{S,L} - \kappa_{S,A}) + \omega_R(\kappa_{R,L} - \kappa_{R,A})]\, X}_{\equiv\, \mu(Z,X),\text{ a function of observables}} \qquad (10)
\end{aligned}
$$

The indirect utility determining whether a patient with covariates X and excluded instrument value $Z = T_L - T_A$ undergoes laparoscopic surgery can be represented as a sum of a function of observed case characteristics, $\mu(Z, X)$, and an additively separate unobserved term represented by random variable V. The indicator function for whether patients with (X, Z, V) undergo laparoscopic surgery (as opposed to abdominal surgery) is

$$
D_L(X, Z, V) = \mathbb{1}\left[\mu(Z, X) - V \geq 0\right] \qquad (11)
$$

where V has some distribution and arbitrarily depends on θ and the idiosyncratic outcome shocks, $W_{S,L}$, $W_{R,L}$, $W_{S,A}$, and $W_{R,A}$. Equivalently, it depends on all factors affecting outcomes that aren't included in X. In my empirical setting, X includes a number of comorbidities and gynecological conditions recorded in

Medicare claims (as I will detail in the data section, Section 4). Therefore, V represents determinants of the outcomes and, in turn, of the choices that I do not observe in the Medicare claims: uterus weight, history of vaginal births, history of abdominal surgery, and other anatomical conditions that I do not observe but that the physician and patient do observe and that affect the efficacy of laparoscopic surgery. V can be thought of as the unobserved (to the analyst) net "health cost" or "resistance" to choosing laparoscopic surgery. Following the literature on selection on unobservable heterogeneity (for example, Carneiro, Heckman and Vytlacil, 2011), let U_D denote the cumulative distribution function of V, $F_V(V)$, so it represents a case's percentile of unobserved "resistance" to the laparoscopic choice. Now we may consider a causal parameter of interest called the *marginal treatment effect* on outcome Y – first proposed by Björklund and Moffitt (1987) and further developed by Heckman and Vytlacil (1999, 2000, 2005, 2007). The marginal treatment effect of treatment L, relative to treatment A, on outcome Y is defined as

$$MTE_Y(x, u_D) \equiv \mathbb{E}[Y_L - Y_A | X = x, U_D = u_D] \tag{12}$$

and it is evaluated at a vector of particular covariate values, x, and at a particular percentile of unobserved "cost" of or resistance to treatment, u_D, or is commonly called, "resistance" to the laparoscopic treatment.

Different instrument values identify marginal treatment effects among patients with different levels of θ. Recall that θ is a key aspect of the theory which represents patient complexity, which makes a patient more resistant to laparoscopic surgery. Patients with lower θ have lesser V and thus a lesser U_D. Consider the propensity score for choosing laparoscopic surgery as a function of covariates and an excluded instrument, $P(z, x) \equiv Pr(D_L = 1|Z = z, X = x)$[9]. Note that U_D and $P(X, Z)$ are monotonic transformations of V and $\mu(Z, X)$, respectively. A patient who is at a lower percentile of unobserved resistance to the laparoscopic procedure, U_D, requires a lower percentile of observed net benefit, $P(Z, X)$ – induced by a greater relative distance to the laparoscopic surgery-performing hospital, Z – in order to be indifferent between laparoscopic surgery and abdominal surgery. Therefore, patients with lower θ have lesser V and lesser U_D, and thus their marginal treatment effects are identified by lesser values of P induced by greater relative distances, Z. This also demonstrates why the marginal treatment effect is sometimes described as the difference in potential outcomes among patients who have values of (Z, X) such that their unobserved resistance to laparoscopic treatment, u_D, is equal to their observed net benefit of laparoscopic surgery, $P(Z = z, X = x) = p$, so $MTE_Y(x, u_D) = MTE_Y(x, p)$.

With causal quantities defined and identification explained, let us now turn to the empirical implications of the model. Let the notation implicitly condition on X. The assumption made that $S_L(\theta) < S_A(\theta)$ for all

[9] This is sometimes characterized as the patient's mean scale utility value.

θ implies that empirically the marginal treatment effect on length of stay is

$$MTE_S(u_D) < 0 \tag{13}$$

for any given u_D. Equation (7) predicts that the marginal treatment effect on readmission risk among patients with the greatest resistance to laparoscopic surgery is positive, i.e.

$$MTE_R(P(Z = 0)) > 0 \tag{14}$$

Recall that patients at the highest percentile of resistance (i.e., greatest U_D, 1 by definition) are identified and made indifferent between surgery alternatives by the lowest instrument value, $Z = 0$. Since marginal treatment effects on readmissions is positive for the highest resistance patients, if $MTE_R(u_D)$ is continuous, then

$$MTE_R(u_D) > 0 \tag{15}$$

for an interval $u_D \in \left[u_D^0, P(Z = 0)\right]$ if $\frac{\partial MTE_R(u_D)}{\partial u_D} > 0$, where u_D^0 is some value less than one, or for all u_D otherwise. In other words, the marginal treatment effects on readmissions should be positive for the patients with the greatest resistance to treatment, if not all patients.

The analysis resulting in Equation (8) and Equation (9) also predicts that

$$\frac{\partial MTE_R(u_D)}{\partial u_D} \gtrless 0 \tag{16}$$

$$\frac{\partial MTE_S(u_D)}{\partial u_D} \gtrless 0 \tag{17}$$

as long as they are not both positive, for any u_D. Both derivatives being positive would contradict the meaning of U_D as the percentile of unobserved resistance to laparoscopic surgery. It would mean that patients with greater unobserved resistance to laparoscopic surgery, U_D, experienced both greater detriment from laparoscopic surgery (more positive MTE_R) and lesser benefit from laparoscopic surgery (less negative MTE_S).

The marginal treatment effects are related to the local average treatment effect on outcome Y,

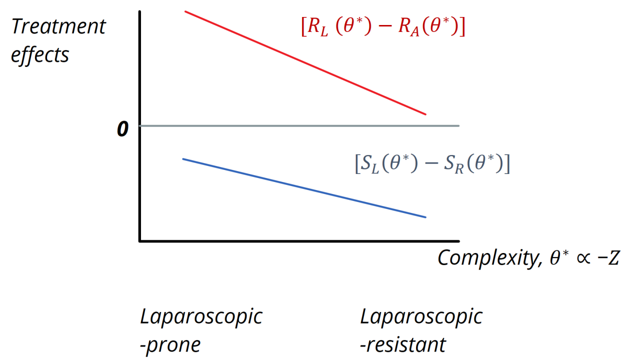


Figure 3: A hypothetical pattern of marginal treatment effects on length of stay (S) and on readmissions (R). Cases on the left are more prone to laparoscopic surgery: they have lesser values of θ and are made indifferent between the two procedures with greater values of the excluded instrument, relative distance, Z. Patients on the right are laparoscopic resistant: they have greater θ and are made indifferent with greater values of Z. The vertical axis is the magnitude of marginal treatment effects. The red line plots the marginal treatment effects on readmission with respect to resistance, $R_L(\theta^*) - R_A(\theta^*)$, and the blue line plots the marginal treatment effects on length of stay with respect to resistance, $S_L(\theta^*) - S_A(\theta^*)$. The model assumes that $S_L(\theta^*) - S_A(\theta^*)$ is negative for all θ, and it implies that $R_L(\theta^*) - R_A(\theta^*)$ is positive for at least the most resistant patients. Under the model, the marginal treatment effects trends may be both upward-sloping, or one may be upward-sloping and the other downward-sloping, but they cannot both be downward-sloping (as depicted here). If that were so, the laparoscopic-resistant cases would have lesser adverse effects from laparoscopic surgery than the laparoscopic-prone patients.

$$LATE_Y\left(p_0, p_1\right) = \frac{\text{Cov}(Y, Z)}{\text{Cov}(Z, D_L)}$$

for continuous instrument Z whose values induce a range of propensity scores from p_0 to p_1.[10] As discussed in the context of marginal treatment effects, a particular realized value of Z induces a propensity score p and identifies marginal treatment effects among patients with a u_D equal to p. Heckman and Vytlacil (1999, 2005) and Heckman, Urzua and Vytlacil (2006) show that the local average treatment effect on outcome Y for an instrument whose values induce a range of propensity scores from p_0 to p_1, is a weighted combination of marginal treatment effects:

$$LATE_Y\left(p_0, p_1\right) = \int_{p_0}^{p_1} MTE_Y\left(p\right) \varphi_{\text{IV}}^{Z}\left(u_{D_L}\right) dp \tag{18}$$

where $\varphi_{IV}^{Z}\left(u_D\right)$ are the weights for each level of u_D.[11]

[10] It is interpreted as the difference in potential outcomes among the compliers of the instrument, that is, among patients who would choose $D_L = 1$ for some values of Z and would choose $D_L = 0$ for other values. The denominator is the first stage, and the numerator is the reduced form or the intent-to-treat effect.

[11] The weights relating the MTEs to the LATE are:

$$\varphi_{IV}^{Z}\left(u_D\right) = \frac{\mathbb{E}\left[Z - \mathbb{E}[Z] \mid P(Z) > u_D\right] Pr\left(P(Z) > u_D\right)}{\text{Cov}\left(Z, D\right)}$$

This leads to the predictions

$$LATE_R\,(p_0, p_1) > 0 \qquad\qquad LATE_S\,(p_0, p_1) < 0 \tag{19}$$

for some instrument that induces changes in treatment decisions among patients with propensity scores in the range of p_0 to p_1.

To test the theory's predictions about marginal patients in my empirical setting, I will first estimate the local average treatment effect as an approximation of the marginal treatment effects, and then I will estimate the marginal treatment effects for different levels of U_D under several model specifications.

3.4 Revelation of Preferences and Objectives

Assuming that the choice of hysterectomy mode is made jointly by a patient and physician who are trying to maximize patient utility over the two clinical outcomes and travel time to surgery, the model shows how to identify patients' and physicians' joint marginal rate of substitution. Since the slope of the indifference curve for a given tuple (θ, Z) is equal to the marginal rate of substitution, $MRS_{S,R} = (\partial U/\partial S)/(\partial U/\partial R) = -\frac{\omega_S}{\omega_R}$, the ratio of the marginal treatment effects of the two outcomes equals the marginal rate of substitution:

$$MRS_{S,R}(\theta) = m = \frac{R_A(\theta) - R_L(\theta)}{S_A(\theta) - S_L(\theta)} \tag{20}$$

for each θ. Thus, in the population, the marginal rate of substitution for patients with unobserved resistance to laparoscopic surgery u_D is identified by

$$MRS_{S,R}(u_D) = \frac{MTE_R(u_D)}{MTE_S(u_D)} \tag{21}$$

the ratio of the marginal treatment effects on readmissions and on length of stay. Because the local average treatment effect is a weighted combination of the marginal treatment effects identified by the instrument, I also approximate the marginal rate of substitution across case complexity types using the local average treatment effect. If the marginal rate of substitution is the same across patient types, which would happen if indifferences were linear, then

Certain observations are weighted more heavily if their treatment covaries with particular ranges of the instrument more. The weights integrate to one, can be negative if the instrument does not satisfy monotonicity, and can be consistently estimated from the sample.

$$MRS_{S,R} = \frac{LATE_R}{LATE_S} \tag{22}$$

the ratio of local average treatment effects is the marginal rate of substitution for any patient type. Otherwise, it is a weighted combination of marginal rates of substitution across patients types.

I should make an important caveat here. The finding that the ratio of the marginal effects identifies the marginal rate of substitution for patients depends on providers fully and accurately incorporating patient preferences into their own utility function. Sepucha and Mulley (2009) review some potential reasons why physicians might not understand or implement a given patient's preferences. Additionally, this identification requires there to be no other provider-side factors influencing the choice of hysterectomy mode. For example, hospitals' profit functions, or in the case of not-for-profit hospitals, their utility functions (Pauly and Redisch, 1973), could incorporate patient length of stay or readmission risk. Patient length of stay could affect hospital profit margins on episode-based or capitated payment, and readmission risk could affect patient's quality measures which could in turn affect hospitals' bargaining leverage with insurers. (However, at the time of my observations, Medicare did not have financial penalties for readmissions.) If hospitals are able to influence surgical mode through allocation of operating room equipment, staff, and time or through other tacit ways, the "marginal rate of substitution" identified by the ratio in Equation (20) does not merely identify the patient marginal rate of substitution – i.e. some function of patients' relative elasticities of demand with respect to clinical outcomes – but rather the ratio would reflect preferences and incentives throughout the health care system.[12]

3.5 Predictions about Difference in Mean Outcomes between Treatment Groups

This section shows that the model makes predictions about differences in mean outcomes between patients making different surgical mode choices. These constitute an additional side hypothesis whose confirmation would further support the model. Taken together with the predictions for marginal patients, it also makes an interesting point that behavior under this model is consistent with a local average treatment effect on readmission being a different sign than the estimate of the difference in means. In other words, the estimated sign of the effect of laparoscopic surgery on readmissions could be different in a two-staged least squares regression that in an ordinary least squares regression. Conventionally, a difference in signs estimated form

[12] Differences in physician reimbursement or physician ergonomics between the modes, for example, could affect the decision, and these factors would be incorporated in the intercepts of the linear indifference curves considered here, rather than the slope, unless these technology-specific factors in the physicians' utility were correlated with length of stay or readmission rate. (See Newhouse (1996) for a literature review on provider response to reimbursement contract design, and see McDonald et al. (2014) for a small survey of gynecologic oncologists on ergonomics of different surgery types.)

these two types of regressions is considered a cause for concern for the estimation's validity, but under this model it is plausibly expected with a valid instrument.

The difference in means of length of stay between laparoscopic patients and abdominal patients is:

$$\bar{S}_L - \bar{S}_A < 0 \tag{23}$$

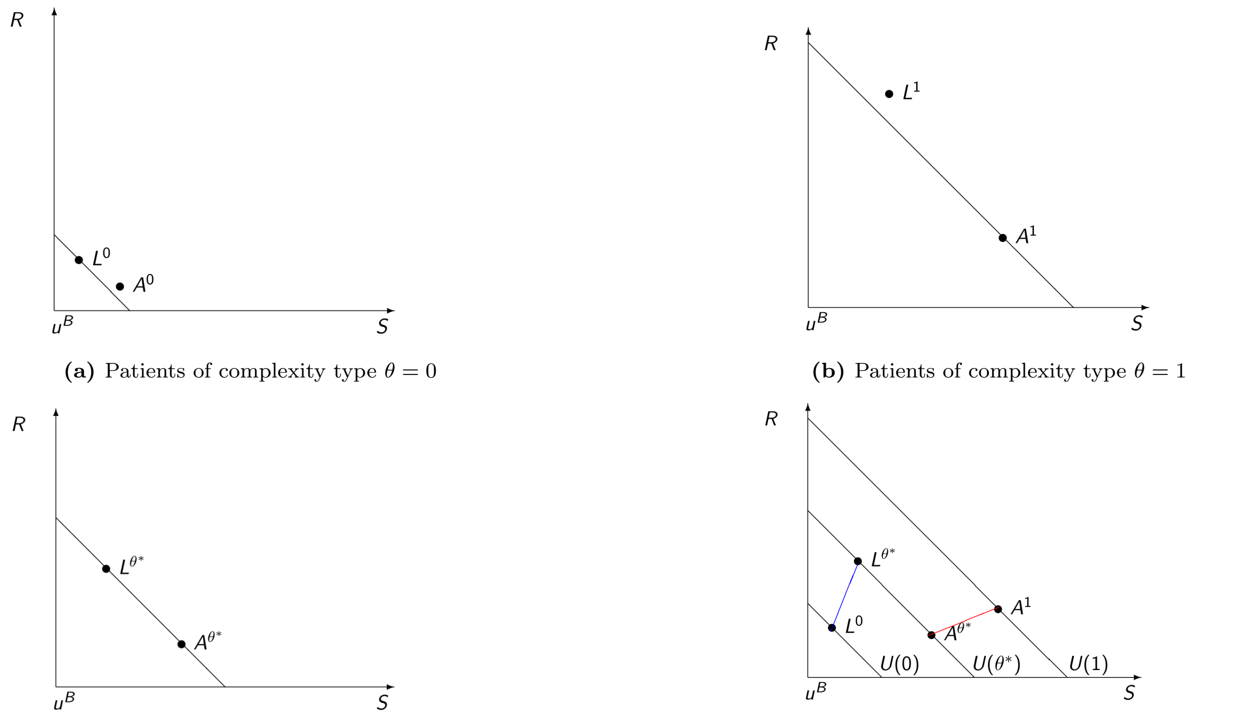
So the ordinary least squares estimate of the effect of laparoscopic surgery, relative to abdominal surgery, on length of stay, among patients who undergo either laparoscopic or abdominal surgery will be positive.

The sign of the difference between the mean readmission rate among laparoscopic patients and the mean readmission rate among abdominal patients is ambiguous under the presented assumptions. It is dependent on an interaction of the differences between technologies in readmission rates among patients without complications, in the degrees to which readmission rates increase with respect to θ, and the shares of patients of each technology choice who are of different values of θ. Appendix A.5 goes into more detail and analyzes the possible cases. The upshot is, the sign of $\bar{R}_L - \bar{R}_A$ is ambiguous. Whether the difference in means is positive or negative is not dependent on the sign of the treatment effect among the marginal patients, whose treatment effect would be approximated by the local average treatment effect. In other words, in this selection setting, theory allows for the sign of the local average treatment effect to be different from the sign of the ordinary least squares estimate of the treatment effect. This departs from the conventional notion that a contradiction between the sign of the estimated local average treatment effect and the sign of the ordinary least squares estimate of the average treatment effect is a cause for concern about the instrumental variable regression's validity. Theory predicts that the signs will be different under certain reasonable parameter assumptions and distributional assumptions.

3.5.1 Estimands

Finally, to test the model's side prediction about the difference in the average length of stay among the laparoscopic patients and among the abdominal patients in Equation (23), I estimate the difference in expected lengths of stay among patients with the same covariates, and I predict it to be

$$\mathbb{E}[S|D_L = 1, X] - \mathbb{E}[S|D_L = 0, X] < 0 \tag{24}$$



(a) Patients of complexity type $\theta = 0$

(b) Patients of complexity type $\theta = 1$

(c) Patients of complexity type $\theta = \theta^*$ such that the patient is indifferent between options.

(d) Indifference curves and production possibilities sets for patient types 0, θ^*, and 1

Figure 4: Indifference curves and production sets for patients of three different complexity types, $\theta = 0$, $\theta = \theta^*$ for the $\theta^* \in (0, 1)$ such that patients are indifferent between laparoscopic and abdominal surgery, and $\theta = 1$. The production set for patient type θ is composed of two bundles of patient outcomes labeled L^θ if laparoscopic surgery is chosen and A^θ if abdominal surgery is chosen. Bundles are composed of a readmission rate R and a length of stay in the hospital, S. Utility is highest at the origin point, u^B, and decreases outward, i.e., up and to the right. Bundles chosen by laparoscopic patients will fall in the area around the blue line connecting L^0 and L^{θ^*} in Panel D, and bundles chosen by abdominal patients will fall around the red line connecting A^{θ^*} and A^1.

The prediction for the difference in the expected readmission rates

$$\mathbb{E}[R|D_L = 1, X] - \mathbb{E}[R|D_L = 0, X] \tag{25}$$

is ambiguous.

4 Data Description

Testing this paper's theory requires estimating the impacts of the choice of surgical technology on patient outcomes. For this, I require data on patients who underwent total hysterectomy performed by different physicians, in different hospitals, in different procedure markets, and I require detailed information about patient characteristics, the physicians and hospitals that provided the surgery, and patient diagnostic information recorded during the stay of the surgery and in subsequent encounters with health care providers.

Medicare claims data is well suited for this investigation. I analyze all Medicare inpatient claims throughout the United States from 2007 to 2008. This is to say that I observe virtually all inpatient stays among Americans age 65 and older, of all different demographics and clinical characteristics, in all various geographical settings and hospital market structures, treated by physicians with all different experiences and training. I use data from 2007 to 2008 because at this time, almost all Medicare-covered total hysterectomies were performed either laparoscopically or abdominally.

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.

I observe 60,889 claims for total hysterectomies from 2007 to 2008, six percent of which are for laparoscopic hysterectomies. Each claim includes a unique identifier for patients, allowing me to see information from multiple health care encounters for a given patient, such as whether a patient was readmitted to a hospital after a hysterectomy. The patient identifier also allows linking a claim to Medicare' beneficiary summary file, which contains demographic information and the Zip code of the patient's resident. It also includes ICD-9 procedure codes and diagnosis codes, providing detailed, standardized information about the clinical characteristics of the patients as well as the care provided. ICD-9 procedure codes include detailed description of the type of surgery performed, including whether a hysterectomy was open/abdominal, laparoscopic, or robotically-assisted.[13] The claim also indicates the dates of admission and discharge, allowing for calculation of the patient's length of stay in the hospital. Finally, the claims also detail the Zip codes of the hospitals and of the patients, facilitating my identification strategy that relies on comparing a patient's distance to her nearest hospital with minimally invasive surgery to her nearest hospital that does not perform minimally invasive surgery.[14]

From the claims, I derive my outcomes of interest. For each total hysterectomy, I build an indicator variable for whether the patient's length of stay in the hospital was two or more days and an indicator variable whether the patient had an inpatient claim in the 10 days since the hysterectomy. I choose use a dichotomous measure of the length of stay because the distribution of length of stay has much of its

[13] The ICD-9 procedure code for robotically-assisted surgery was not introduced until the fourth quarter of 2008; prior to that point, robotically-assisted procedures were coded as laparoscopic procedures. However, according to these claims, only 5.2 % of hysterectomy claims in that quarter were performed in a robotically-assisted manner, 5.7% in the first quarter of 2009, 7.8% in the second quarter, and 10.5% in the first quarter of 2010 (Figure 10 in the appendix). Even though I do not directly observe whether a given hysterectomy in 2007 or the beginning of 2008 was performed robotically, I infer from this trend that it is probably true that very few of the hysterectomies coded as laparoscopic from 2007 to 2008 were performed robotically.

[14] I calculate the distances between the centroids of the Census Bureau's Zip code tabulation areas, the latitude and longitudes of which are calculated and made publicly available by UDS Mapper (Bureau of Primary Health Care at the U.S. Health Resources and Services Administration; John Snow, Inc.; and the American Academy of Family Physicians; available at https://udsmapper.org), using the distHaversine function for R.

probability mass around one or two days and a long right tail (see Appendix Figure 11). Thus, much of the possible potential lengths of stay are around two days. Additionally, some unusual cases with long length of stay could have outsized influence on the treatment-specific means, so making inferences about mean length of stay may not be as informative as making inference about the frequency with which length of stay is above a common realization.

In order to condition my estimates of course of treatment on outcomes on possibly confounding factors, I augment this data with information from a few sources. To control for characteristics of the patient's neighborhood which may be correlated with their own socioeconomic characteristics, I collect Zip Code Tabulation Area-level data on race, income, rates of participation in public assistance and public insurance programs, and household income from the U.S. Census Bureau's American Community Survey's 5-Year Estimates from 2008 – 2012. I also use hospital quality measures from Medicare's Hospital Compare program. Finally, I observe some hospital characteristics through Medicare's Provider of Service (POS) file.[15] The specific covariates I control for are detailed in the section on empirical strategy.

I have several groupings of covariates, which I add sequentially to the regression specification to see how robust the estimate is to potential confounding factors. Demographic controls include indicators for whether the patient is Black, a race other than white or Black, under 65 years of age, or over 74 years of age. Clinical controls include the Charlson comorbidity index as well as indicators for whether the patient had diabetes, had a malignant neoplasm, had a non-malignant neoplasm, had a body mass index of 30 or over (considered obese), had a history of cancer indicated on the hysterectomy claim, had uterine fibroids, had endometriosis, had pelvic organ prolapse, had female genital bleeding, had post-menopausal bleeding, had an ovarian cyst, had female genital pain, or had peripheral adhesions. Variables describing the Zip code of the patient's residence include the white percent of residents, the college-educated percent of residents, the percent of residents with public assistance (including cash or nutritional assistance), the median household income, and the percent of residents on Medicaid. The hospital quality variables include how many hysterectomies the hospital performed that year, a quality measure on the appropriate use of antibiotics, a quality measure on the prevention of blood clots in heart patients, and the overall Consumer Assessment of Healthcare Providers & Systems (CAHPS) score.

Table 1 show how patient diagnoses, comorbidities, and demographics vary by hysterectomy mode.

[15] I use a version of the file cleaned and made publicly available by Adam Sacarny, http://sacarny.com/data/.

Table 1: Means and Standard Deviations of Case Characteristics, by Total Hysterectomy Approach

	Laparoscopic Mean	Std. Dev.	Abdominal Mean	Std Dev.
Percent with Any 10-Day Readmission	0.0402	0.196	0.0589	0.236
Percent with Length of Stay $\geq$ 2 Days	0.522	0.500	0.982	0.135
White	0.867	0.339	0.817	0.387
Black	0.0925	0.290	0.137	0.344
Not Black or white	0.0402	0.196	0.0461	0.210
Any Months on HMO	0.0289	0.167	0.0409	0.198
Diabetes	0.170	0.375	0.175	0.380
Malignant Neoplasm	0.505	0.500	0.463	0.499
Non-Malignant Neoplasm	0.242	0.428	0.327	0.469
BMI30+	0.0387	0.193	0.0300	0.171
History of Cancer	0.104	0.305	0.0762	0.265
Uterine Fibroid	0.239	0.427	0.287	0.452
Endometriosis	0.103	0.304	0.113	0.317
Pelvic Organ Prolapse	0.106	0.308	0.0744	0.263
Female Genital Bleeding	0.108	0.311	0.129	0.335
Postmenopausal Bleeding	0.113	0.317	0.0996	0.299
Other Ovarian Cyst	0.0753	0.264	0.0849	0.279
Female Genital Pain	0.135	0.342	0.128	0.334
Pelvic peritoneal adhesions	0.0990	0.299	0.100	0.300
Zip Percent White	0.798	0.203	0.792	0.221
Zip Percent College	0.382	0.173	0.336	0.154
Zip Percent Public Cash or Nutrition Assistance	0.114	0.0866	0.132	0.0884
Zip Median Household Income	59297.1	25353.3	53306.8	21589.3
Zip Percent Medicaid	0.106	0.0667	0.116	0.0686
Hospital Num. Hyst.s	62.63	44.07	52.29	44.40
Hospital Quality Measure: Proper Clot Prevention	0.878	0.0889	0.861	0.106
Hospital Quality Measure: Proper Antibiotic Use	0.913	0.0764	0.907	0.0872
Hospital Patient Satisfaction Score	2.549	0.115	2.534	0.117

Means for continuous variables and prevalence rates for indicator variables across hysterectomy patients, by type of hysterectomy. These procedure-level statistics describe hysterectomy outcomes, the demographic and clinical characteristics of the patients, the Zip codes of the patients' residences, and the hospitals where the procedures were performed. LOS is length of stay, MSA is Metropolitan Statistical Area, HMO is Medicare Advantage, and BMI 30+ is an indicator for Body Mass Index equalling or exceeding 30 (indicating obesity).

5 Empirical Strategy

In this section, I describe the excluded instrument that I use to identify the local average treatment effect and the marginal treatment effects. Next, I present the estimation methods and models. I consider the estimates of the local average treatment effects to be my main estimation strategy. Marginal treatment effect estimation requires estimating from smaller cells of data, so it is less precise than the two-stage least squares estimation. I consider the marginal treatment effects to constitute confirmatory, suggestive evidence. This section also shows I use the local and marginal effects to estimate the marginal rate of substitution.

5.1 Instrumental Variable Definition and Validity

In order to identify the marginal treatment effects or the local average treatment effect, I need an instrumental variable that affects the choice of hysterectomy approach but is excluded from the outcome models. The excluded instrument I use, Z, is

$$Z = T_L - T_A \tag{26}$$

the difference between the distance to a patient's nearest hysterectomy-performing hospital that performs laparoscopic surgery and the distance to her nearest hospital performing hysterectomy. Its distribution is presented in Append Table 7. This instrument meets the three criteria for the two-stage least squares estimator to identify the local average treatment effect among the compliers: relevance, exclusivity, and monotonicity (Imbens and Angrist, 1994; Angrist, Imbens and Rubin, 1996; Imbens and Rubin, 1997). Statistical inference of the results also requires that the instrument is not weak. To estimate marginal treatment effects, as well, instruments must satisfy relevance (or, the rank condition), exclusivity (or, independence), and monotonicity (or, uniformity) (Heckman, Urzua and Vytlacil, 2006).

First, I show evidence from the first stage that the instrument is relevant and not weak. I estimate the conditional correlation of Z and D_L, the indicator for whether the hysterectomy was performed laparoscopically, on all total hysterectomies in 2007 and 2008, when few robotically assisted hysterectomies were performed.

Appendix Table 8 presents the first stage results. Across all specifications, the instrument is very stable and suggests that reducing the difference between the distance to the nearest laparoscopic hospital and the distance to the nearest hospital without laparoscopic surgery by 10 miles – i.e., making the nearest laparoscopic hospital closer relative to the nearest hospital without – increases the compliers' likelihood to undergo laparoscopic rather than abdominal hysterectomy by 0.5 percentage points. In each specification, the effective F statistic far exceeds the critical values. [16] The negative relationship between relative distance and choice of hysterectomy procedure is also shown graphically in the binned scatterplots of Figure 12.

Second, the instrument arguably satisfies the exclusion restriction. A patient's relative distance to a

[16] Following the advice of Andrews, Stock and Sun (2019), I conduct a weak instrument test that is robust to heteroskedasticity proposed by Montiel Olea and Pflueger (2013) and implemented by Pflueger and Wang (2015). Their test statistic is compared against a two-stage least squares/limited information maximum likelihood critical value either for 5% bias, which is 37.418 in my sample, or the value for 10% bias, which is 23.109. When there is just one endogenous variable, as in my case, the Olea-Montiel Pflueger test statistic is equivalent to the Kleibergen and Paap (2006) statistic. This latter test is packaged with the common Stata commands `ivreg2` and Correia (2018)'s `ivreghdfe`. Evidence strongly suggests that my instrument is not weak. However, note that Andrews, Stock and Sun (2019) advise that even if a set of instruments should fail the appropriate test, that the instrument should not be discarded due to its weakness. Instead, they write that analysis with the instrument should proceed with weak instrument-robust inference methods.

hospital performing laparoscopic surgery arguably affects hysterectomy outcomes only through its effect on the patient's choice of hospital and whether that hospitals perform laparoscopic surgery. Hadley and Cunningham (2004) raise concerns that the effect of distance to care on a patient's choice of care may be confounded by socioeconomic patient characteristics correlated with distance and health. If relative distance were associated with demographic and clinical characteristics, the validity of this assumption would be in doubt. However, Table 9 shows that patients in the top half of the relative distance instrument's distribution have similar demographic and clinical characteristics as patients in the bottom half of the distribution. For further evidence for or against the independence assumption, I look for associations between the instrument and clinical- and socioeconomic-based predictions of five adverse outcomes: 10-day readmissions, 90-day readmissions, lengths of stay of two or more days, and lengths of stay of three or more days, and I show that controlling for other observed patient characteristics and Hospital Referral Region fixed effects, the associations are weak and less in magnitude than the reduced form effects of the instruments on adverse outcomes. First, I predict adverse outcomes using demographic, clinical, and neighborhood characteristics. Then I inspect binned scatterplots of the instrument against the fitted values of the adverse outcome rates, in Appendix Figure 14 and Figure 15. The plotted associations are conditional on patients' distance to any hospital and on four patient Zip-code characteristics, the percent of persons in the Zip code who are white, the percent of persons with a college degree, the percent of persons on Medicaid, and the median household income. The point estimates are small in comparison to the corresponding reduced form correlations between the adverse outcomes and the instrument, which are presented in Appendix Table 11 through Table 14, for all outcomes except for any 90-day readmission (in which case the reduced form effect has a similar magnitude to the association between the instrument and the predicted outcome). For example, predicted 10-day readmissions has a conditional correlation with relative distance of -0.00003, which is one third of the reduced form effect of relative distance on 10-day readmissions, -0.00009. Prediction of a length of stay of two or more days has a conditional correlation with the instrument of 0.00003, while the corresponding reduced form effect is 0.00022. This suggests that even if the instrument were associated with adverse outcomes of interest through some channel besides the procedure choice, such a confounding association is likely much smaller than the causal effects of interest and would not likely affect the qualitative estimates of the local average treatment effects.

Third, the instrument likely satisfies monotonicity and uniformity. Increasing the relative distance to a laparoscopic hospital arguably weakly decreases the patient's propensity to undergo laparoscopic surgery, as opposed to abdominal surgery, and in no case would not increase the propensity. This is demonstrated in the Appendix in Table 15. I estimate the first stage on several cells of patients by demographics and by diagnoses, following an approach used in the "judge IV" literature (e.g., Arnold, Dobbie and Yang, 2018;

Bhuller, Dahl, Løken and Mogstad, 2020) and in Chan, Card and Taylor (2022). In each case, the estimated effect of the distance instrument on the choice of laparoscopic hysterectomy is qualitatively the same and quantitatively similar, strongly suggesting that there are no defiers of the instrument, and the local average treatment effect identifies the treatment effect among the compliers only.

5.2 Estimation

This subsection presents the estimation methods used to estimate the local average treatment effects and the marginal treatment effects.

5.2.1 Estimating Local Average Treatment Effects

I estimate the local average treatment effect using a two-stage least squares estimator, where the first and second stages are

$$Y = \rho_{Y,0} + \rho_{Y,1} D_L + \rho_{Y,2} X + \epsilon_Y \tag{27}$$

$$D_L = \pi_0 + \pi_1 Z + \pi_2 X + \nu \tag{28}$$

where D_L is a random indicator for whether a hysterectomy was performed laparoscopically (rather than abdominally), Z is the excluded instrument described above that characterizes how much farther the nearest laparoscopic hospital is to a patient than the nearest hospital, X is a random vector of covariates, and Y is random variable representing a clinical outcome of the hysterectomy. In alternative regression specifications, the outcome is an indicator for whether the surgery resulted in any 10-day all-cause readmission, and an indicator for whether the hysterectomy inpatient stay was 2 or more days. The random variables ν, ϵ_R, and ϵ_S represent idiosyncratic shocks. I list the demographic, clinical Zip-level, and hospital covariates in Section 4. I model the standard errors of two-stage least squares estimators assuming that there is clustering of outcomes at the hospital level.

5.2.2 Estimating Marginal Treatment Effects across Heterogeneous Patients

I estimate the marginal treatment effects using the two common estimation methods, the local instrumental variables method and the separate method. The marginal treatment effect can be re-written as:

$$\mathbb{E}[Y_L - Y_A | X = x, U_D = u_D] = \kappa_{L,Y} x - \kappa_{A,Y} x + \mathbb{E}[W_{Y,L} - W_{Y,A} | U_D = u_D] \tag{29}$$

for outcome Y, which alternatingly is the long length of stay indicator or the readmission indicator. $W_{Y,L}$ and $W_{Y,A}$ are the idiosyncratic shocks to potential outcomes Y_L, under L, and Y_A, under A, respectively. Each method involves estimating an outcome model that includes an additively separable component that represents unobserved heterogeneity:

$$\mathbb{E}[Y|X = x, U_D = u_D] = \kappa_{A,Y} x + px(\kappa_{L,Y} - \kappa_{A,Y}) + K_Y(p) \tag{30}$$

where $K_Y(p) = p\mathbb{E}[W_{Y,L} - W_{Y,A}|U_D \leq p]$, the unobserved "essential heterogeneity" in the outcome that is correlated with the potential utilities under each alternative.[17] The true distribution of $K_Y(p)$ is unknown, and the function $K_Y(p)$ could be nonlinear. Thus, the outcome is alternatively modeled parametrically in terms of the unobserved term and semiparametrically (partially linear), in keeping with practices in the literature. The four parametric specifications are (1) modeling $K_Y(p)$ as Normal, (2 – 4) modeling $k_Y(p) = K'(p)$ as a first-, second-, and then third-degree polynomial in p.

The semiparametric specifications model Y as an additively separable model of two components, (1) a nonlinear function of p, $K_Y(p)$, and (2) the linear combination $\kappa_{A,Y} X + pX(\kappa_{L,Y} - \kappa_{A,Y})$. Estimation of $K_Y(p)$ proceeds as follows. The residuals $\hat{e}_Y$, $\hat{e}_X$, and $\hat{e}_{Xp}$ are acquired by regressing Y, X, and Xp each on p by local linear regression with the Epanechnikov kernel and alternative bandwidths of 0.01, 0.02, 0.03, and 0.05. The double residual regression is due to Robinson (1988) and modified by Heckman, Ichimura and Todd (1997). Next, the $\kappa_{A,Y}$ and $\kappa_{L,Y}$ are estimated by regressing $\hat{e}_Y$ on $\hat{e}_X$ and $\hat{e}_{Xp}$. $Y - X\hat{\kappa}_A - X(\hat{\kappa}_L - \hat{\kappa}_A)p$ is in turn regressed on p by second-degree local polynomial regression with the Epanechnikov kernel and the bandwidth chosen by a plug-in estimator for a rule by Fan and Gijbels (1995). This yields $\hat{K}_Y(p)$, whose derivative is taken to construct the marginal treatment effect with the estimates for the kappas. A detailed description of this is in the appendix of Heckman, Urzua and Vytlacil (2006).

Because the unobserved heterogeneity is a function of the propensity score, each method entails estimating a propensity score for undergoing laparoscopic surgery, as opposed to undergoing abdominal surgery, as a probit function of covariates and the excluded instrument. The marginal treatment effects are only identified where there is overlap of the instrument-induced propensity scores. I model the propensity score as a probit of almost the entire set of covariates used in the ordinary least squares and two-stage least squares regressions.[18]

The local instrumental variable method due to Heckman and Vytlacil (1999) and Heckman and Vytlacil

[17]Recall that $p = P(Z, X)$ is the propensity score induced by relative distance instrument Z and U_D is the patient's percentile of unobserved resistance to the laparoscopic alternative.

[18]Hospital quality measures were not available for all hospitals in my dataset. In the interest of maintaining the sample size for the information-intensive marginal treatment effect estimation, I omit these variables from the set of covariates in this section of analysis.

(2007) is to estimate $\mathbb{E}[Y|x,p]$ using one of the parametric or semiparametric models described above and take the derivative with respect to the propensity score, p.

In the so-called separate approach developed by Heckman and Vytlacil (2007) and Brinch, Mogstad and Wiswall (2017), the terms reflecting unobserved heterogeneity and the coefficients from the two separate potential outcome models

$$\mathbb{E}[Y_L|X=x,U_D=u_D] = \kappa_{L,Y}x + K_L(p) \tag{31}$$

$$\mathbb{E}[Y_A|X=x,U_D=u_D] = \kappa_{A,Y}x + K_A(p) \tag{32}$$

are estimated separately among laparoscopic patients and among abdominal patients, respectively. Then the marginal treatment effect at p is calculated by subtracting the two estimated potential outcomes at mean x.

I implement both the local instrumental variable method and the separate method using software by Andresen (2018). I estimate cluster-robust standard errors through 100 bootstrap repetitions with resampling over the hospitals (Cameron and Trivedi, 2005).

5.3 Marginal Rate of Substitution: Estimation and Inference

The model in Section 3 shows that the marginal rate of substitution among patients with resistance to laparoscopic surgery u_D is identified by Equation 20, the ratio of the marginal treatment effect on readmissions to the effect on length of stay. Thus, I estimate the marginal rate of substitution by estimating the marginal treatment effects and plug in:

$$\widehat{MRS}(\bar{x},u_D) = \frac{\widehat{MTE}_R(\bar{x},u_D)}{\widehat{MTE}_S(\bar{x},u_D)} \tag{33}$$

If the indifference curves are linear, as postulated, or if they are convex but patients under a particular alternative of surgical technology are each located on the same relative point on their respective indifference curves, then these rates will be the same across all percentiles of resistance, $u_D \in U_D$.

I calculate the standard errors of this marginal rate of substitution with 100 bootstrap iterations.

I also estimate an approximation of the marginal rate of substitution by estimating the local average treatment effects on readmission and on length of stay and plugging in:

$$\widehat{MRS} \approx \frac{\frac{\hat{\psi}_1^R}{\hat{\pi}_1}}{\frac{\hat{\psi}_1^S}{\hat{\pi}_1}} \tag{34}$$

where $\hat{\pi}_1$ is the estimate of the first-stage coefficient representing the effect of Z on D_L, and $\hat{\psi}_1^R$ and $\hat{\psi}_1^S$ are the intent to treat effects from the reduced form estimating equations for readmission and for length of stay, respectively.

I calculate standard errors on this estimate of the marginal rate of substitution in Equation 34 using the Delta method.

6 Empirical Results

6.1 Testing the Model Predictions for Average Patients

The model predicts in Equation 23 that the average length of stay among laparoscopic patients will be less than that among abdominal patients, conditional on covariates. I estimate the difference in conditional expectations of the chance of a long length of stay using ordinary least squares regression.

Table 2 shows the OLS estimates of the effect of laparoscopic surgery (relative to abdominal surgery) on length of stay of 2 days or more under several specifications that each add additional covariates to control for potential confounding from patient and provider characteristics, as well as a fourth, fixed effect specification. The first specification has no covariates. The second controls for demographic covariates, the third adds comorbidities and gynecological conditions, the fourth adds characteristics of the residents in the patient's Zip code, and the fifth adds hospital characteristics. Section 4 lists the specific covariates in each category. In all specifications, standard errors assume clustering at the hospital level.

Laparoscopic hysterectomy patients have between a 41 percentage point and a 46 percentage point lesser chance of a length of stay that is 2 days or longer. This is in keeping with the model's prediction of shorter mean lengths of stay among laparoscopic patients. The point estimate is fairly stable across the different specifications.

The model predicts that whether abdominal patients or laparoscopic patients have lower or higher mean readmission rates is ambiguous. Table 3 shows that OLS and FE estimates of the association between laparoscopic surgery and any 10-day all-cause readmission is a reduction of around two percentage points percentage points. The estimate is also very stable across specifications.

Table 2: Association between Laparoscopic Procedure and Probability of Length of Stay of 2 or More Days: OLS and FE Regression

	(1)	(2)	(3)	(4)	(5)	(6)
Laparoscopic	-0.460***	-0.460***	-0.459***	-0.461***	-0.468***	-0.467***
	(0.0137)	(0.0137)	(0.0138)	(0.0137)	(0.0141)	(0.0136)
Observations	60832	60832	60832	59634	52349	52347
Dependent variable mean	0.952	0.952	0.952	0.952	0.951	0.951
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	1119.0	234.9	86.50	69.29	57.73	59.67
Adj. R^2	0.277	0.281	0.285	0.286	0.297	0.302

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Ordinary least squares and fixed effects regression estimates of the difference between laparoscopic and abdominal hysterectomies in prevalence of a length of stay being two or more days. Demographic controls: whether the patient is Black, a race other than white or Black, under 65 years of age, or over 74 years of age. Clinical controls: the Charlson comorbidity index and indicators for whether the patient had diabetes, malignant neoplasm, non-malignant neoplasm, body mass index of 30 or over (considered obese), history of cancer indicated on the hysterectomy claim, uterine fibroids, endometriosis, pelvic organ prolapse, female genital bleeding, post-menopausal bleeding, an ovarian cyst, female genital pain, or peripheral adhesions. Zip code controls of the patient's residence: white percent of residents, the college-educated percent of residents, the percent of residents with public assistance (including cash or nutritional assistance), the median household income, and the percent of residents on Medicaid. Hospital controls: number of hysterectomies the hospital performed that year, a quality measure on the appropriate use of antibiotics, a quality measure on the prevention of blood clots in heart patients, and the overall Consumer Assessment of Healthcare Providers & Systems (CAHPS) score. HRR= Hospital Referral Region. Standard errors assume clustering at the hospital level.

Table 3: Association between Laparoscopic Procedure and Probability of All-Cause 10-Day Readmission: OLS and FE Regression

	(1)	(2)	(3)	(4)	(5)	(6)
Laparoscopic	-0.0187***	-0.0179***	-0.0180***	-0.0186***	-0.0203***	-0.0208***
	(0.00335)	(0.00336)	(0.00344)	(0.00350)	(0.00376)	(0.00384)
Observations	60832	60832	60832	59634	52349	52347
Dependent variable mean	0.0577	0.0577	0.0577	0.0577	0.0572	0.0572
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	31.23	17.60	21.22	17.94	15.53	15.12
Adj. R^2	0.000369	0.00176	0.00721	0.00736	0.00874	0.00938

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Ordinary least squares and fixed effects regression estimates of the difference between laparoscopic and abdominal hysterectomies in prevalence of a 10-day all-cause readmission. Demographic controls: whether the patient is Black, a race other than white or Black, under 65 years of age, or over 74 years of age. Clinical controls: the Charlson comorbidity index and indicators for whether the patient had diabetes, malignant neoplasm, non-malignant neoplasm, body mass index of 30 or over (considered obese), history of cancer indicated on the hysterectomy claim, uterine fibroids, endometriosis, pelvic organ prolapse, female genital bleeding, post-menopausal bleeding, an ovarian cyst, female genital pain, or peripheral adhesions. Zip code controls of the patient's residence: white percent of residents, the college-educated percent of residents, the percent of residents with public assistance (including cash or nutritional assistance), the median household income, and the percent of residents on Medicaid. Hospital controls: number of hysterectomies the hospital performed that year, a quality measure on the appropriate use of antibiotics, a quality measure on the prevention of blood clots in heart patients, and the overall Consumer Assessment of Healthcare Providers & Systems (CAHPS) score. HRR= Hospital Referral Region. Standard errors assume clustering at the hospital level.

6.2 Testing the Model Predictions for Marginal Patients

Next, I test the model's assumptions and predictions about marginal patients. The theoretical model in Section 3 assumes that laparoscopic procedures have shorter lengths of stay than abdominal procedures among marginal patients, and it predicts in Equation 49 that laparoscopic procedures have greater readmission rates than abdominal procedures among marginal patients.

Section 3.3.1 explains that the predictions about marginal patients imply predictions about instrument compliers. Intuitively, patient near-indifferent are more likely to be induced into switching their choice on the basis of relative distance. In more technical detail, the model is condition on a relative distance, so there is a set of marginal patients for each level of relative distance. Each of these sets of marginal patients' treatment effects are marginal treatment effects identifiable with the use of the relative distance instrument, and the local average treatment effect is a positively weighted combination of the marginal treatment effects.

Table 4 presents the two-stage least squares estimates of the local average treatment effects on whether a hysterectomy patient has a length of stay of two days or more. Across specifications, the estimated effect is negative and statistically significant. The magnitude of the effect is greater as more factors are controlled for. Column 5 shows that controlling for all covariates, laparoscopic hysterectomy causes a 57 percentage point decline in the chance of a length of stay of two or more days, relative to abdominal hysterectomy, among patients who are induced into the laparoscopic mode by the relative distance instrument's variation. I also estimate that the local effect of laparoscopic surgery on the probability of a length of stay of 3 or more days is to lower it by 55 percentage points, though the effect is noisily estimated and not statistically significant (Table 16 in Appendix D).

Table 4: Local Effect of Laparoscopic Procedure on the Probability of Length of Stay is 2 or More Days: 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)
Laparoscopic	-0.332***	-0.397***	-0.443***	-0.542***	-0.567***	-0.504***
	(0.0681)	(0.0644)	(0.0634)	(0.0950)	(0.109)	(0.129)
Observations	54992	54992	54992	54972	48553	48553
Dependent Variable Mean	0.950	0.950	0.950	0.950	0.949	0.949
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	23.82	56.07	27.99	24.05	21.90	21.40
Adj. R^2	0.260	0.281	0.290	0.283	0.289	0.289
First-Stage Effective F	94.41	102.1	104.4	60.45	43.97	29.09

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Two-staged least squares estimates, using relative distance to hospital with laparoscopic surgery as instrument. Across specifications, the effective F statistic (due to Montiel Olea and Pflueger (2013) and Kleibergen and Paap (2006)) should be compared against either the two-stage least squares/limited information maximum likelihood critical value for 5% bias, which is 37.418, or for 10% bias, which is 23.109. Demographic controls: whether the patient is Black, a race other than white or Black, under 65 years of age, or over 74 years of age. Clinical controls: the Charlson comorbidity index and indicators for whether the patient had diabetes, malignant neoplasm, non-malignant neoplasm, body mass index of 30 or over (considered obese), history of cancer indicated on the hysterectomy claim, uterine fibroids, endometriosis, pelvic organ prolapse, female genital bleeding, post-menopausal bleeding, an ovarian cyst, female genital pain, or peripheral adhesions. Zip code controls of the patient's residence: white percent of residents, the college-educated percent of residents, the percent of residents with public assistance (including cash or nutritional assistance), the median household income, and the percent of residents on Medicaid. Hospital controls: number of hysterectomies the hospital performed that year, a quality measure on the appropriate use of antibiotics, a quality measure on the prevention of blood clots in heart patients, and the overall Consumer Assessment of Healthcare Providers & Systems (CAHPS) score. HRR= Hospital Referral Region. Standard errors assume clustering at the hospital level. Standard errors assume clustering at the hospital level.

Estimates of the local treatment effects on the chance of a 10-day readmission are shown in Table 5. Across specifications, the estimated effect on readmissions is positive and economically significant. It is statistically significant controlling for demographic, clinical and Zip-code level socioeconomic factors. When hospital factors – including some Hospital Compare quality measures which are not available for all hospitals – are additionally controlled for, the point estimate is a statistically significant increase in the readmission rate of 23 percentage points. I conclude from this evidence that there is good reason to believe that compliers experience greater readmission risk under laparoscopic hysterectomy than under abdominal hysterectomy. As a robustness check, I also estimate that the local effect of laparoscopic surgery on the chance of a 90-day readmission is a 17 percentage point increase, under the specification with all covariates (Appendix 17).

My study in its current form cannot explain why marginal laparoscopic hysterectomy patient experience greater readmission rates than marginal abdominal patients. One possibility is that marginal laparoscopic patients experience greater injury rates than inframarginal laproscopic patients and marginal abdominal patients. One metastudy suggests that laparoscopic patients have greater rates of bladder and ureter injuries than abdominal patients (Teeluckdharry et al., 2015). Indeed, I find that evidence that marginal laparoscopic hysterectomy patients experience greater rates of readmissions in which it was indicated they had urogenital infections (Table 18 in Appendix D), which are associated with such injuries.

In sum, these results are consistent with the model's assumptions and predictions for marginal patients: the two-stage least squares procedures estimate that the chance of a hysterectomy having a long length of stay is greater among marginal abdominal patients than among marginal laparoscopic patients, and the chance of a readmission is greater among marginal laparoscopic patients than among marginal abdominal patients.

Table 5: Local Effect of Laparoscopic Procedure on the Probability of Any 10-day Readmission: 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)
Laparoscopic	0.362***	0.312***	0.261***	0.326***	0.233*	0.228*
	(0.0763)	(0.0709)	(0.0687)	(0.102)	(0.120)	(0.136)
Observations	54992	54992	54992	54972	48553	48553
Dependent Variable Mean	0.0583	0.0583	0.0583	0.0583	0.0576	0.0576
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	22.53	13.49	17.06	12.88	12.22	11.96
Adj. R^2	-0.166	-0.123	-0.0816	-0.127	-0.0660	-0.0685
First-Stage Effective F	94.41	102.1	104.4	60.45	43.97	29.09

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Two-staged least squares estimates, using relative distance to hospital with laparoscopic surgery as instrument. Across specifications, the effective F statistic (due to Montiel Olea and Pflueger (2013) and Kleibergen and Paap (2006)) should be compared against either the two-stage least squares/limited information maximum likelihood critical value for 5% bias, which is 37.418, or for 10% bias, which is 23.109. Demographic controls: whether the patient is Black, a race other than white or Black, under 65 years of age, or over 74 years of age. Clinical controls: the Charlson comorbidity index and indicators for whether the patient had diabetes, malignant neoplasm, non-malignant neoplasm, body mass index of 30 or over (considered obese), history of cancer indicated on the hysterectomy claim, uterine fibroids, endometriosis, pelvic organ prolapse, female genital bleeding, post-menopausal bleeding, an ovarian cyst, female genital pain, or peripheral adhesions. Zip code controls of the patient's residence: white percent of residents, the college-educated percent of residents, the percent of residents with public assistance (including cash or nutritional assistance), the median household income, and the percent of residents on Medicaid. Hospital controls: number of hysterectomies the hospital performed that year, a quality measure on the appropriate use of antibiotics, a quality measure on the prevention of blood clots in heart patients, and the overall Consumer Assessment of Healthcare Providers & Systems (CAHPS) score. HRR= Hospital Referral Region. Standard errors assume clustering at the hospital level. Standard errors assume clustering at the hospital level.

6.3 Efficiency of the Extent of Diffusion: Benefit-Cost Analysis on the Margin

There are a number of reasons that medical technologies may not diffuse to all cases for which they would be efficient or that they might diffuse to too many cases, including cases for which they are not effective. These include asymmetric information, regulation, and underinsurance. If a technology poses tradeoffs between different dimensions of quality, patients who are indifferent between the alternatives should face roughly similar expected benefits from choosing one option as they would under the other option. This section presents a back-of-the-envelope benefit-cost analysis using estimates of the relative effectiveness of a technology on the margin of an instrumental variable quasi-experiment. From this analysis, one can infer whether the technology has diffused to an efficient extent.

The expected differential benefit of laparoscopic hysterectomy could be estimated as the estimated benefit of a reduction in the length of stay in the hospital, relative to the length of stay under abdominal hysterectomy. According to descriptive analysis from the American Hospital Association's Annual Survey, the cost of a day in the hospital in Washington state, the U.S. state with the highest daily hospital cost, was $2,490 in 2008 (Foundation, 2021). Combined with the estimate of laparoscopic surgery's effect among marginal patients on the chance of having a length of stay of two or more days (a 56.7 percentage point increase), and I estimate that the differential benefit of laparoscopic surgery is roughly $ 1,411.83. To estimate the differential cost laparoscopic surgery poses by increasing the patient's readmission risk, I use an estimate from hospital discharge reports that the average cost of a readmission in the U.S. is $15,200 in 2010 (Weiss and Jiang, 2006). This implies that the expected differential cost of laparoscopic surgery is $3,465.60, so laparoscopic surgery poses an expected $2,054 loss among marginal patients, relative to abdominal surgery. Since suffering an acute surgical complication and being readmitted to a hospital on an inpatient basis arguably imposes greater non-pecunicary costs than discharge from a planned inpatient stay being delayed by a day, this net loss estimate is likely an underestimate. This suggests that laparoscopic surgery may have diffused beyond the efficient extent in this setting, from the perspective of the individual patient considering the adverse outcomes under alternative hysterectomy procedures.

Why there may be too much laparoscopic surgery is beyond the scope of this paper. However, I will briefly speculate some potential causes of a wedge that would cause this. Reimbursement incentives could favor one treatment over the other. As I described in Section 2, the hospitals are reimbursed the exact same rate for laparoscopic hysterectomy as they are for abdominal hysterectomy, and the physician reimbursement rates across procedures are similar.

Alternatively, a wedge could be introduced by another actor in the health care system who has different preferences from the patient over adverse outcomes and is able to influence treatment decisions on the margin. Say that hospitals maximize either profits or population health, and thus have incentives to perform additional surgeries as long as the marginal surgeries yield positive utility. Hospitals that are near full have an incentive to switch indifferent patients from abdominal surgery to a technology that results in shorter lengths of stay, in order to increase surgical volume. This would result in more laparoscopic surgery than is efficient from the individual patient's perspective. I investigate this in work outside this paper.

6.4 Estimation of the Marginal Rate of Substitution from Two-Stage Least Squares

Here I estimate the marginal rate of substitution of a greater chance of a long length of stay for a lesser chance of a readmission, by taking the ratio of the local effect on readmissions to the local effect on length of stay (Equation (34)). The estimates under different outcome model specifications are listed in Table 6. In the specifications controlling demographic and clinical characteristics as well as the specification additionally controlling for characteristics of the patients' neighborhoods, I estimate the marginal rate of substitution to be around -0.60. In the fifth specification, where the effect is estimated to be less and with greater uncertainty, the estimate of the marginal rate of substitution is -0.41.

Table 6: Estimates of the marginal rate of substitution from two-stage least squares

	(1)	(2)	(3)	(4)	(5)
MRS	-1.090***	-0.786***	-0.590***	-0.601***	-0.411*
	(0.328)	(0.229)	(0.187)	(0.226)	(0.242)
Observations	54992	54992	54992	54972	48553
Demographic Controls		✓	✓	✓	✓
Clinical Controls			✓	✓	✓
Zip Code Controls				✓	✓
Hospital Controls					✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Estimates of the marginal rate of substitution of a greater chance of a long length of stay for a lesser chance of a readmission. They are calculated by dividing the two-stage least squares estimate of the local average treatment effect on the probability of a patient's length of stay being 2 or more days (relative to abdominal surgery) by the two-stage least squares estimate of the local effect on the probability of an all-cause 10-day readmission. Standard errors were calculated by the Delta method. The model in Column 5 includes quality measures from Hospital Compare which are not available for all hospitals.

The results from the fifth specification with all covariates implies that patients are willing to trade off a 55 percentage point increase in the chance of long length of stay for a 23 percentage point decrease in the probability of a readmission. The standard errors of the marginal rate of substitution estimate are calculated by the Delta method and are presented in the parentheses.

6.5 Estimation of Marginal Treatment Effects across Heterogeneous Patients

In this section, I test the model's predictions about outcomes among marginal patients through estimation of marginal treatment effects of laparoscopic surgery, as opposed to abdominal surgery, on readmission and on length of stay among patients with a given level of unobserved "cost" or "resistance" to the laparoscopic option, which the theory section showed is partially dependent on the patient complexity characteristic, θ. I estimate these effects for different levels of unobserved resistance to laparoscopic surgery.

The propensity score as a function of observable covariates and excluded instruments is integral to the estimation of marginal treatment effects. The marginal treatment effects are identified only for propensity scores that are induced by the variation in the available instrument and that are observed under both surgical options. Figure 5 presents the distributions of propensity scores, generated from a probit regression, among laparoscopic patients and among abdominal patients. Much of the probability mass of the overlap of propensity score distributions under the two surgical alternatives is among propensity scores between five and ten percent, so the marginal treatment effects among patients with propensity scores outside that range are not identified.

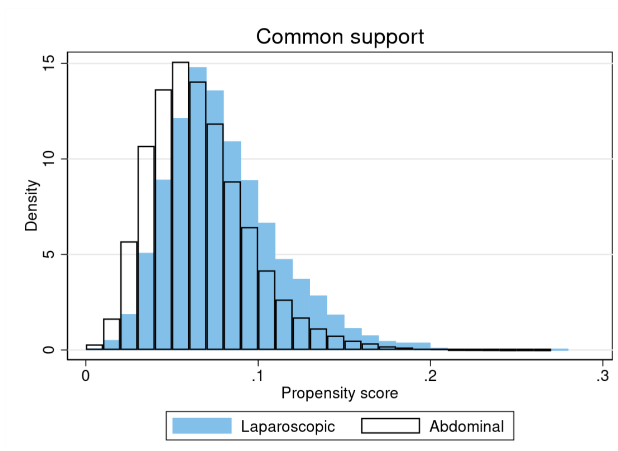


Figure 5: Overlapping Distributions of Propensity Scores

Distributions of propensity scores among laparoscopic cases and abdominal cases. The local instrumental variable method and the separate method of estimating marginal treatment effects can identify where there is overlap of the propensity scores of the two groups.

The covariates X that I control for are the demographic controls, the clinical controls, the Zip code-level controls, and one of the hospital controls, the number of hysterectomies that the hospital performed that year. I omit the remaining the hospital controls, the quality measures from Medicare's Hospital Compare program, because they are not available for all hospitals and thus their inclusion would reduce my sample of hysterectomies for this data-intensive estimation approach considerably.

The first two rows of Figure 7 graphically present estimates of marginal treatment effects on the probability of a length of stay being 2 or more days, with respect to the patient's percentile of unobserved resistance to (or, net unobserved "cost" of) laparoscopic surgery as opposed to abdominal surgery. The effects are estimated by local instrumental variable method (row one) and by the separate method (row two). The first plot in the first row summarize the estimates from each of the parametric and semiparametric models estimated. The estimates are very similar across models within estimation method. In the local instrumental variable method, estimates of the marginal effects across model specifications are about -0.5 among the patients most likely to undergo laparoscopic surgery. The effects among patients at the 15th percentile range from -1 to -2.5, with estimates from the parametric models sloping down more steeply than the estimates from the semiparametric models.

These point estimates are quite large, but keep in mind the wide 90 percent confidence intervals, shaded in gray, particularly for percentiles of resistance with less support from the data. To give a sense of the variation in the estimations, plots in the second column present point estimates and 90% confidence intervals from the most restrictive model, the parametric model assuming that the unobserved heterogeneous component of the outcome, $K_Y(p)$, is Normal, and plots in the third column present results from the most flexible model, the semiparametric models with the narrowest bandwidth, 0.01. Estimates of the effects on length of stay are mostly significant at the 90% level.

As shown in the second row, the separate method estimates that the effects on length of stay among the patients at the 5th percentile range from about -0.4 to 0.6. Among patients at the 15th percentile, the estimates are between -0.7 and -0.9.

Across all models, the estimates suggest that the effects of laparoscopic surgery on length of stay are greater for patients with greater resistance to the laparoscopic alternative. However, the width of the confidence intervals relative to the downward slope of the point estimates make this finding merely suggestive. The full set of results on length of stay from the instrumental variable method and the separate method are in Appendix Figure 16 and Figure 18, respectively.

The third and fourth rows of Figure 7 presents estimates of the marginal treatment effects on the chance of a readmission from the local instrumental variable method and the separate method, respectively. Across most models, the local instrumental variable estimators in row 3 suggest that the effects of readmission are positive and increasing with respect to unobserved resistance to laparoscopic surgery. An exceptional set of results are from estimating the model assuming the unobserved heterogeneity is distributed Normal. Estimates of that model suggest the effect on readmission is decreasing. At the 5th percentile of resistance, the estimates are clustered around an effect size of 0.2, and at the 15th percentile, they range from 0.15 to

0.65. Economically, this fits with the prior finding that patients with greater resistance experience differential lengths of stay of greater magnitude under laparoscopic surgery than patients with lesser resistance: if patients who have greater resistance to laparoscopic surgery experience more beneficial laparoscopic treatment effects on lengths of stay than patients who are more willing to choose laparoscopic surgery, then those higher-resistance patients must experience worse outcomes under laparoscopic surgery on some other dimension than the lower-resistance patients. Statistically, these findings must be taken with caution as the 90% confidence intervals almost always include zero and are wide, particularly so for higher-resistance patients.

Estimates from the separate method tell a somewhat different story. While most of the point estimates for the range of percentiles of resistance that are most supported by the data, from the 5th to the 10th percentiles, are positive, the series of estimates from each model are flat or downward sloping. The effects are much smaller in magnitude than those estimated from the local instrumental variables. Evidence from the most restrictive models, the Normal model and the polynomial of degree one model, suggest that the effect on readmissions may be constant over percentiles of resistance, whereas point estimates from the other models suggest that the effects may be decreasing with respect to resistance. In all cases, the confidence sets for the estimates are quite wide. The full set of results on readmissions from the instrumental variable method and the separate method are in Appendix Figure 17 and Figure 19, respectively.

In sum, there is strong evidence that patients experience lower lengths of stay under laparoscopic surgery than under abdominal surgery, and there is evidence that this effect could be declining in patient resistance to laparoscopic surgery. This raises the question of what relative outcome from laparoscopic surgery could be worsening as resistance increases that counterweights this declining relative length of stay. Estimates from local instrumental variable regression suggest that the risk of readmissions is greater under laparoscopic surgery and that this effect is greater among patients with greater resistance to laparoscopic surgery. This would be the countervailing consideration that makes laparoscopic surgery less attractive among patients with greater resistance. It also is consistent with the evidence from the two-stage least squares procedures, which is no surprise since the local average treatment effect is a weighted combination of the marginal treatment effects, and the weights are all positive because the instrument satisfies monotonicity, or, uniformity. (See Appendix Figure 20 for the estimated weights at each percentile of unobserved resistance.) Evidence from the separate method largely confirms the signs of the effects on length of stay and readmission and the slope of the effects on length of stay, but they are largely at odds with the local instrumental variable estimates of the sign of the slope of the effects on readmissions. It is not clear how to definitively settle this discrepancy, but it is relevant to note that, the separate method estimates all the effects twice, once among laparoscopic patients and once among abdominal patients, while the local instrumental variable method performs this once. Therefore, it's possible that the separate method is underpowered in my sample.

6.6 Estimation of the Marginal Rate of Substitution Using Marginal Treatment Effects

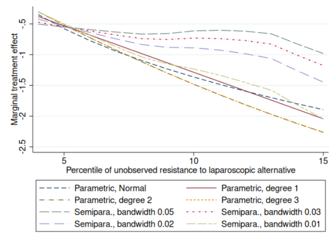
This section attempts to estimate the marginal rate of substitution at different percentiles of resistance using a ratio of marginal treatment effects, as in Equation 21. I divide the estimated marginal effect on having a readmission by the estimated marginal effect on having a length of stay of two or more days, and I plot the results over the resistance.

For brevity, I present estimates of marginal rates of substitution from the most restrictive model and the most flexible model estimated by local instrumental variable. In Figure 8, (a) plots estimates of the effect on readmission, the effect on length of stay, and the marginal rate of substitution from the model assuming that the unobserved component has a Normal distribution. The 90 percent confidence intervals on the estimates of the marginal rate of substitution are represented by the gray regions. The marginal rate of substitution is estimated to be -0.5 at the 5th percentile and to slope upward to just less than zero at the 14th percentile. Subfigure (b) plots results from the semiparametric model with a bandwidth of 0.01. The marginal rate of substitution is estimated to be -0.5 at the 5th percentile and slopes slightly upward at the 14th percentile.

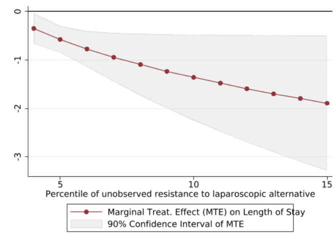
The separate estimates from the restrictive, Normal model are stable, from -0.3 to -0.2. The estimates from the most flexible model are nonmonotonic and volatile, ranging from -0.8 at the 5th percentile to 0.6 at the 14th. The full set of estimates of the marginal rate of substitution are in Appendix Figure 21. In no cases are the marginal rates of substitution statistically significantly different from zero, which follows from combining two noisy sets of estimates of the marginal treatment effects on length of stay and on readmission.

Figure 7: Summary of Marginal Treatment Effect Estimates

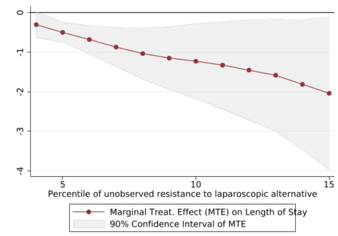
Effect on Probability of Length of Stay of 2 or More Days, Local Instrumental Variable Method



(a) Local instrumental variable estimates from all length of stay models.

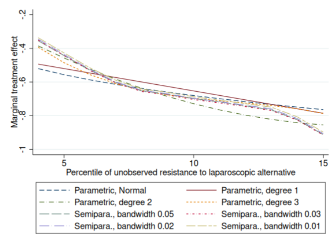


(b) Parametric: Normal.

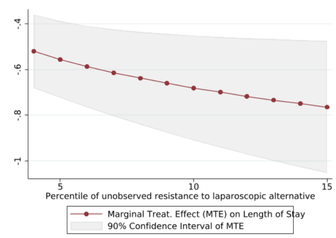


(c) Semiparametric: bandwidth 0.01.

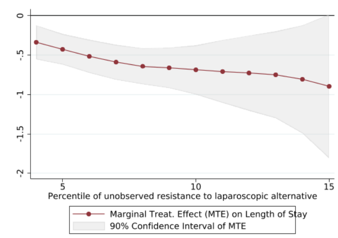
Effect on Probability of Length of Stay of 2 or More Days, Separate Method



(d) Separate method estimates from all length of stay models.

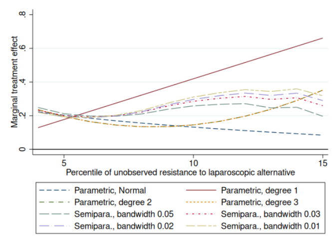


(e) Parametric: Normal.

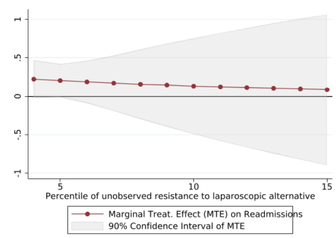


(f) Semiparametric: bandwidth 0.01.

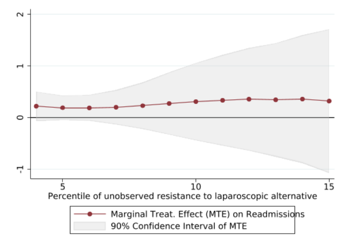
Effect on Probability of All-Cause 10-Day Readmission, Local Instrumental Variable Method



(g) Local instrumental variable method estimates from all readmission models.

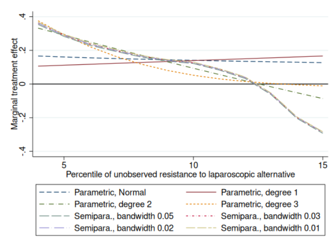


(h) Parametric: Normal.

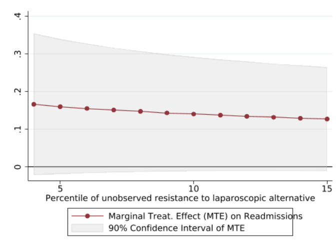


(i) Semiparametric: bandwidth 0.01.

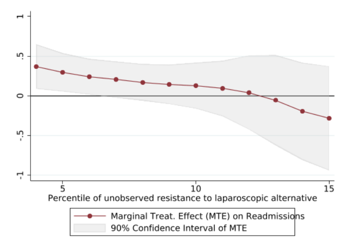
Effect on Probability of All-Cause 10-Day Readmission, Separate Method



(j) Separate method estimates from all readmission models.



(k) Parametric: Normal.

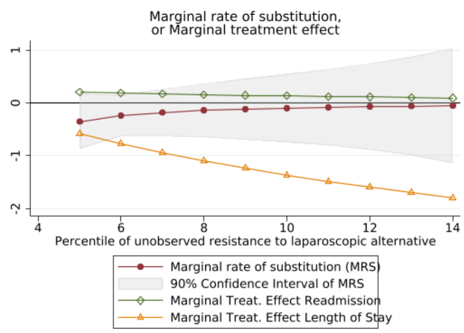


(l) Semiparametric: bandwidth 0.01.

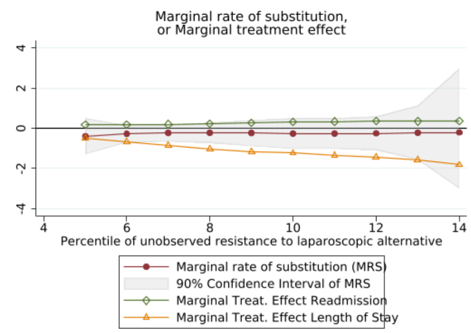
The horizontal axis in each plot is U_D, the case's percentile on the distribution of unobserved "resistance" to or "cost" of the laparoscopic choice. Gray bands are 90% confidence intervals, bootstrapped with 100 repititions. Parametric models presented in middle column model unobserved heterogeneity (functions of the propensity score) as Normal. Semiparametric models in right column model the unobserved heterogeneity with a local polynomial using the Epanechnikov kernel.

Figure 8: Marginal Rates of Substitution across Heterogeneous Patients, Calculated from Marginal Treatment Effects

Local Instrumental Variable Method

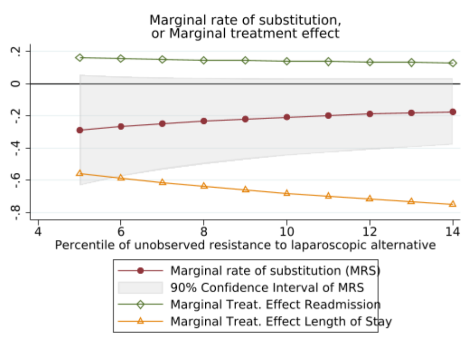


(a) Estimates from Normal Parametric Model

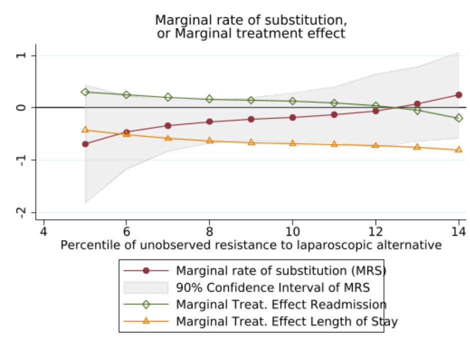


(b) Estimates from Semiparametric Model

Separate Method



(c) Estimates from Normal Parametric Model



(d) Estimates from Semiparametric Model

Estimates of the marginal rate of substitution of readmission risk for length of stay, at each percentile of unobserved "resistance" to or "cost" of the laparoscopic approach. Also plotted are the bootstrapped 90% confidence intervals of the marginal rates of substitution and the marginal treatment effects on readmission rates and on chance of a length of stay of 2 days or more. Semiparamteric results come from the Epanechnikov filter with a 0.01 bandwidth. The marginal rate of substitution is calculated by dividing the marginal tratment effect on readmission by the marginal treatment effect on length of stay.

7 Conclusion

Medical technologies may present patients with tradeoffs between improvements on different dimensions of care. I have shown that hysterectomy patients on the margin between laparoscopic and abdominal surgery face a trade-off between shorter lengths of stay and greater readmission risk. I presented a Roy model in which patients and physicians choose surgical technology based on how it affects two clinical outcomes. The model predicts that indifferent patients and their physicians face shorter lengths of stay but greater readmission rates under laparoscopic surgery than abdominal surgery. These differences in outcomes among indifferent patients are identified by marginal treatment effects, which can be estimated for patients with different levels of unobserved resistance to the laparoscopic alternative. The local average treatment effects identified by two-stage least squares regressions are positively weighted averages of the marginal treatment effects across patient types. Empirically I find that compliers of a distance-based instrument for the choice of laparoscopic procedure experienced shorter lengths of stay under laparoscopic hysterectomy than under abdominal hysterectomy but also experienced greater readmission rates. The estimation of these local average treatment effects leads to an estimation of a ratio that equals patients' marginal rate of substitution of longer length of stay for lesser readmission risk if providers had no influence over treatment but more broadly reflects preferences and objectives in the health care system otherwise. The marginal treatment effect estimates suggest that the (negative, beneficial) effect on length of stay is greater in magnitude for patients with greater unobserved resistance to laparoscopic surgery, and some model specifications suggest that the (positive, detrimental) effect on readmission rates also increases in magnitude with respect to unobserved resistance to the laparoscopic alternative, although these marginal treatment effect estimates are imprecise.

Taking a wider view, this paper emphasizes the multidimensionality of technology. Technological innovation does not proceed along a single continuum and may lead to new products and services that prompt end users to make tradeoffs between different dimensions of quality, not just between quality and cost. When end users are heterogeneous, as in the case of hysterectomy patients with different diagnoses, different comorbidities, and different complexities, diffusion of new technologies may proceed unevenly or incompletely. In health care, understanding these tradeoffs is important to understanding the welfare impacts of new medical technologies and to assessing their allocative efficiency.

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.

Acemoglu, Daron and Amy Finkelstein, "Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector," *Journal of Political Economy*, October 2008, *116* (5), 837–880.

Almond, Douglas, Joseph J. Doyle, Amanda E. Kowalski, and Heidi Williams, "Estimating Marginal Returns to Medical Care: Evidence from At-risk Newborns," *The Quarterly Journal of Economics*, May 2010, *125* (2), 591–634.

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.

Andresen, Martin Eckhoff, "Exploring Marginal Treatment Effects: Flexible Estimation Using Stata," *The Stata Journal*, March 2018, *18* (1), 118–158.

Andrews, Isaiah, James Stock, and Liyang Sun, "Weak Instruments in IV Regression: Theory and Practice," *Annual Review of Economics*, 2019, *11*, 727–753.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin, "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 1996, *91* (434), 444–455.

Arnold, David, Will Dobbie, and Crystal S Yang, "Racial Bias in Bail Decisions," *The Quarterly Journal of Economics*, November 2018, *133* (4), 1885–1932.

Barry, M. J., F. J. Fowler, A. G. Mulley, J. V. Henderson, and J. E. Wennberg, "Patient Reactions to a Program Designed to Facilitate Patient Participation in Treatment Decisions for Benign Prostatic Hyperplasia," *Medical Care*, August 1995, *33* (8), 771–782.

Basu, Anirban, James J. Heckman, Salvador Navarro-Lozano, and Sergio Urzua, "Use of Instrumental Variables in the Presence of Heterogeneity and Self-Selection: An Application to Treatments of Breast Cancer Patients," *Health Economics*, 2007, *16* (11), 1133–1157.

Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad, "Incarceration, Recidivism, and Employment," *Journal of Political Economy*, April 2020, *128* (4), 1269–1324.

Björklund, Anders and Robert Moffitt, "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models," *The Review of Economics and Statistics*, 1987, *69* (1), 42–49.

Bresnahan, Timothy F. and Shane Greenstein, "Technical Progress and Co-invention in Computing and in the Uses of Computers," January 1996.

Brinch, Christian N., Magne Mogstad, and Matthew Wiswall, "Beyond LATE with a Discrete Instrument," *Journal of Political Economy*, August 2017, *125* (4), 985–1039.

Burns, Lawton R. and Douglas R. Wholey, "The Impact of Physician Characteristics in Conditional Choice Models for Hospital Care," *Journal of Health Economics*, May 1992, *11* (1), 43–62.

C., Jr Chan David, David Card, and Lowell Taylor, "Is There a VA Advantage? Evidence from Dually Eligible Veterans," Working Paper 29765, National Bureau of Economic Research February 2022.

Cameron, A. Colin and Pravin K. Trivedi, *Microeconometrics: Methods and Applications*, Cambridge University Press, May 2005.

Capps, Cory, David Dranove, and Mark Satterthwaite, "Competition and Market Power in Option Demand Markets," *The RAND Journal of Economics*, 2003, *34* (4), 737–763.

Card, David, Alessandra Fenizia, and David Silver, "The Health Impacts of Hospital Delivery Practices," Working Paper 25986, National Bureau of Economic Research June 2019.

Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil, "Estimating Marginal Returns to Education," *American Economic Review*, October 2011, *101* (6), 2754–2781.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson, "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector," *American Economic Review*, August 2016, *106* (8), 2110–2144.

_ 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.

_ and Jonathan Skinner, "Technology Growth and Expenditure Growth in Health Care," *Journal of Economic Literature*, July 2012, *50* (3), 645–680.

Clemens, Jeffrey and Joshua D. Gottlieb, "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?," *American Economic Review*, April 2014, *104* (4), 1320–1349.

Correia, Sergio, "IVREGHDFE: Stata Module for Extended Instrumental Variable Regressions with Multiple Levels of Fixed Effects," Boston College Department of Economics September 2018.

Dranove, David, Craig Garthwaite, Christopher Heard, and Bingxiao Wu, "The Economics of Medical Procedure Innovation," October 2021.

Duggan, Mark and Fiona Scott Morton, "The Effect of Medicare Part D on Pharmaceutical Prices and Utilization," *American Economic Review*, March 2010, *100* (1), 590–607.

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.

Fan, Jianqing and Irene Gijbels, "Data-Driven Bandwidth Selection in Local Polynomial Fitting: Variable Bandwidth and Spatial Adaptation," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1995, *57* (2), 371–394.

Fingar, Katherine, Carol Stocks, Audrey Weiss, and Claudia Steiner, "Most Frequent Operating Room Procedures Performed in U.S. Hospitals, 2003-2012 #186," Technical Report 186, AHRQ HCUP.

Finkelstein, Amy, "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare," *The Quarterly Journal of Economics*, February 2007, *122* (1), 1–37.

_ , Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, and Katherine Baicker, "The Oregon Health Insurance Experiment: Evidence from the First Year," *The Quarterly Journal of Economics*, August 2012, *127* (3), 1057–1106.

Foundation, Kaiser Family, "Hospital Adjusted Expenses per Inpatient Day, 2008," https://www.kff.org/health-costs/state-indicator/expenses-per-inpatient-day/ 2021.

Garnick, Deborah W., Erik Lichtenberg, Ciaran S. Phibbs, Harold S. Luft, Deborah J. Peltzman, and Stephen J. McPhee, "The Sensitivity of Conditional Choice Models for Hospital Care to Estimation Technique," *Journal of Health Economics*, February 1990, *8* (4), 377–397.

Gaynor, Martin and William B. Vogt, "Chapter 27 Antitrust and Competition in Health Care Markets," in "Handbook of Health Economics," Vol. 1, Elsevier, January 2000, pp. 1405–1487.

Goldfarb, Brent, "Diffusion of General-Purpose Technologies: Understanding Patterns in the Electrification of US Manufacturing 1880–1930," *Industrial and Corporate Change*, October 2005, *14* (5), 745–773.

Grieco, Paul L E, Charles Murry, and Ali Yurukoglu, "The Evolution of Market Power in the U.S. Automobile Industry," *The Quarterly Journal of Economics*, September 2023, p. qjad047.

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.

Hadley, Jack and Peter Cunningham, "Availability of Safety Net Providers and Access to Care of Uninsured Persons," *Health Services Research*, 2004, *39* (5), 1527–1546.

Heckman, James J. and Edward J. Vytlacil, "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects," *Proceedings of the National Academy of Sciences*, April 1999, *96* (8), 4730–4734.

_ and _ , "The Relationship between Treatment Parameters within a Latent Variable Framework," *Economics Letters*, January 2000, *66* (1), 33–39.

_ and _ , "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments," *Handbook of Econometrics*, 2007, *6*, 4875–5143.

_ and Edward Vytlacil, "Policy-Relevant Treatment Effects," *American Economic Review*, May 2001, *91* (2), 107–111.

_ and _ , "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 2005, *73* (3), 669–738.

_ , Hidehiko Ichimura, and Petra E. Todd, "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *The Review of Economic Studies*, 1997, *64* (4), 605–654.

Heckman, James J, Sergio Urzua, and Edward Vytlacil, "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics*, August 2006, *88* (3), 389–432.

Ho, Kate and Ariel Pakes, "Hospital Choices, Hospital Prices, and Financial Incentives to Physicians," *American Economic Review*, December 2014, *104* (12), 3841–3884.

Imbens, Guido W. and Donald B. Rubin, "Estimating Outcome Distributions for Compliers in Instrumental Variables Models," *The Review of Economic Studies*, October 1997, *64* (4), 555–574.

_ and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475.

Karshenas, Massoud and Paul L. Stoneman, "Rank, Stock, Order, and Epidemic Effects in the Diffusion of New Process Technologies: An Empirical Model," *The RAND Journal of Economics*, 1993, *24* (4), 503–528.

Kleibergen, Frank and Richard Paap, "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 2006, *133* (1), 97–126.

Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz, "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *The American Economic Review*, 1987, *77* (3), 251–277.

McClellan, Mark, Barbara J. McNeil, and Joseph P. Newhouse, "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality?: Analysis Using Instrumental Variables," *JAMA*, September 1994, *272* (11), 859–866.

McDonald, Megan E., Pedro T. Ramirez, Mark F. Munsell, Marilyn Greer, William M. Burke, Wendel T. Naumann, and Michael Frumovitz, "Physician Pain and Discomfort during Minimally Invasive Gynecologic Cancer Surgery," *Gynecologic Oncology*, August 2014, *134* (2), 243–247.

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.

Newhouse, Joseph P., "Reimbursing Health Plans and Health Providers: Efficiency in Production Versus Selection," *Journal of Economic Literature*, 1996, *34* (3), 1236–1263.

Olea, José Luis Montiel Montiel and Carolin Pflueger, "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, July 2013, *31* (3), 358–369.

Papageorge, Nicholas W., "Why Medical Innovation Is Valuable: Health, Human Capital, and the Labor Market," *Quantitative Economics*, 2016, *7* (3), 671–725.

Pauly, Mark and Michael Redisch, "The Not-For-Profit Hospital as a Physicians' Cooperative," *The American Economic Review*, 1973, *63* (1), 87–99.

Pflueger, Carolin E and Su Wang, "A Robust Test for Weak Instruments in Stata," *Stata Journal*, 2015.

Rassier, Sarah L Cohen, "Hysterectomy: Laparoscopic," in "UpToDate," Wolters Kluwer, October 2022.

Robinson, P. M., "Root-N-Consistent Semiparametric Regression," *Econometrica*, 1988, *56* (4), 931–954.

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

Sepucha, Karen and Albert G. Mulley, "A Perspective on the Patient's Role in Treatment Decisions," *Medical Care Research and Review*, February 2009, *66* (1_suppl), 53S–74S.

Skinner, Jonathan and Douglas Staiger, "Technology Diffusion and Productivity Growth in Health Care," *The Review of Economics and Statistics*, August 2015, *97* (5), 951–964.

Stewart, Susan T., Leslie Lenert, Vibha Bhatnagar, and Robert M. Kaplan, "Utilities for Prostate Cancer Health States in Men Aged 60 and Older," *Medical Care*, April 2005, *43* (4), 347–355.

Suri, Tavneet, "Selection and Comparative Advantage in Technology Adoption," *Econometrica*, 2011, *79* (1), 159–209.

Teeluckdharry, Brahmananda, Donna Gilmour, and Gordon Flowerdew, "Urinary Tract Injury at Benign Gynecologic Surgery and the Role of Cystoscopy: A Systematic Review and Meta-analysis," *Obstetrics and Gynecology*, December 2015, *126* (6), 1161–1169.

Trajtenberg, Manuel, "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners," *Journal of Political Economy*, 1989.

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

Weiss, Audrey J. and H. Joanna Jiang, "Overview of Clinical Conditions With Frequent and Costly Hospital Readmissions by Payer, 2018," in "Healthcare Cost and Utilization Project (HCUP) Statistical Briefs," Rockville (MD): Agency for Healthcare Research and Quality (US), 2006.

A Theory: Derivations of Findings

A.1 Choices by Patients with Different Levels of Complexity

This section shows how the utility functions under laparoscopic surgery and under abdominal surgery and the adverse outcome production functions affect choices among patients with, alternatively, low and high θ types.

Consider the indifference curve of patient type $\theta = 0$ for fixed Z (Figure 4a) in terms of S and R, conditional on X and T_j. Note that the slope of the indifference curve is $m = -\frac{\omega_s}{\omega_R}$, and bliss utility, T_j, and ω_T are encoded in the indifference curve's R-intercept:

$$R(\theta) = \frac{u_B - \omega_T T_j}{\omega_R} - \frac{\omega_S}{\omega_R} S(\theta) \tag{35}$$

Each point represents a bundle of adverse clinical outcomes, and points L^0 and A^0 represent the bundles that type $\theta = 0$ can achieve under the two production technologies available L and A, respectively. Highest utility is achieved at the origin, and utility declines as S or R increases. From the assumption that low complexity cases choose L, it follows that

$$U_L(0) < U_A(0) \tag{36}$$

$$-m = \frac{\omega_s}{\omega_R} < \frac{\delta_A - \delta_L}{\alpha_L - \alpha_A} \tag{37}$$

$$< \frac{\gamma_L - \gamma_A}{\alpha_L - \alpha_A} \tag{38}$$

The last line follows from

$$U_A(1) > U_L(1) \tag{39}$$

$$\gamma_A - \gamma_L < \delta_L - \delta_A \tag{40}$$

Inequality 38 implies that the bundles L^0 and A^0 be oriented relative to type 0's indifference curve as depicted in Figure 4a. Patients with the lowest complexity choose L, which yields lower S but higher R than A.

By analogous reasoning, patients with the highest complexity (type $\theta = 1$) choose abdominal surgery bundle A^1, which yields a lesser readmission risk but longer length of stay than L (Figure 4b).

A.2 Comparative Advantage

Assuming that low-θ types choose laparoscopic surgery and high-θ types choose abdominal surgery, i.e., $U_L(\theta = 0) > U_A(\theta = 0)$ and $U_A(\theta = 1) > U_L(\theta = 1)$ implies:

$$\omega_S(\alpha_L - \alpha_A) < \omega_R(\gamma_A - \gamma_L) \tag{41}$$

$$\omega_S(\underbrace{\alpha_A - \alpha_L}_{>0 \text{ by } S_A(\theta) > S_L(\theta) \forall \theta}) > \omega_R(\gamma_A - \gamma_L) \tag{42}$$

and

$$\omega_S(\beta_A - \beta_L) - \omega_R(\delta_L - \delta_A) < \omega(\alpha_L - \alpha_A) - \omega_R(\gamma_A - \gamma_L) < 0 \tag{43}$$

A.3 Existence of A Type of Patient Who is Indifferent Conditional on Z

Consider the indifferent patient, conditioning on X and the idiosyncratic shocks. Setting the utility of laparoscopic surgery, as a function of θ equal to the utility of abdominal surgery, substituting into the utility functions, , set equal to each other and solving for θ yields:

$$\theta^* = \frac{\omega_R\left(\gamma_L - \gamma_A\right) + \omega_T Z - \omega_S\left(\alpha_A - \alpha_L\right)}{\omega_S\left(\beta_A - \beta_L\right) - \omega_R\left(\delta_L - \delta_A\right)} \tag{44}$$

where θ^* is the $\theta^{LA}(Z)$ that makes a patient indifferent for a particular value of Z. The type of patients θ who are indifferent at value Z is a linear function of Z, and as Z increases, θ^* decreases:

$$\frac{\partial \theta^*}{\partial Z} = \frac{\omega_T}{\omega_S(\beta_A - \beta_L) - \omega_R(\delta_L - \delta_A)} < 0 \tag{45}$$

where the denominator is negative due to the findings derived from comparative advantage.

A.4 Outcome Predictions on the Margin

Consider a particular combination of values of θ, T_L, and T_A such that a patient with those values is indifferent. Conditioning on X, there is one θ for a given $Z = T_L - T_A$ such that the patient is indifferent between procedures. Call this $\theta^{LA}(Z) = \theta^*$.

$$U_L(\theta^*, T_L) = U_A(\theta^*, T_A) \tag{46}$$

$$\frac{\omega_S}{\omega_R}\left(\alpha_L + \beta_L\theta^* - \alpha_A - \beta_A\theta^*\right) + \frac{\omega_T}{\omega_R}\left(Z\right) = \left(\gamma_A + \delta_A\theta^*\right) - \left(\gamma_L + \delta_L\theta^*\right) \tag{47}$$

If we assume that $\alpha_L + \beta_L\theta < \alpha_A + \beta_A\theta$ for all values of θ, because laparoscopic surgery is always less invasive than abdominal surgery, then the left hand side must be negative. So the right-hand side must be negative when $Z = 0$: the indifferent patient experiences a higher readmission rate under laparoscopic surgery than under abdominal surgery:

$$(\gamma_A + \delta_A\theta_{LA}) - (\gamma_L + \delta_L\theta_{LA}) < 0 \tag{48}$$

i.e.,

$$R_A(\theta_{LA}) - R_L(\theta_{LA}) < 0 \tag{49}$$

Substituting Equation (44) for θ^* in Equation (47) and differentiating with respect to Z yields:

$$\frac{d[S_L(\theta^*) - S_A(\theta^*)]}{dZ} = \omega_T\left(\frac{\beta_L - \beta_A}{\omega_S(\beta_A - \beta_L) - \omega_R(\delta_L - \delta_A)}\right) \tag{50}$$

$$\frac{d[R_L(\theta^*) - R_A(\theta^*)]}{dZ} = \omega_T\left(\frac{\delta_L - \delta_A}{\omega_S(\beta_A - \beta_L) - \omega_R(\delta_L - \delta_A)}\right) \tag{51}$$

Substituting these into Equation (43) yields:

$$-\omega_S\frac{d[S_L(\theta^*) - S_A(\theta^*)]}{dZ} - \omega_R\frac{d[R_L(\theta^*) - R_A(\theta^*)]}{dZ} > 0 \tag{52}$$

One can see that both derivatives cannot be simultaneously positive.

A.5 Predicted Difference in Mean Readmission Rates

Restating the difference in means between readmission rate among laparoscopic patients and among abdominal patients:

$$\bar{R}_L - \bar{R}_A \tag{53}$$

$$= \underbrace{\frac{1}{N_L}\left[\sum_{\{i:\theta_i<\theta^*\}} R_i + \sum_{\{i:\theta_i=\theta_{LA}\&D_A=1\}} R_i\right]}_{\text{Average } R \text{ over inframarginal and marginal } L \text{ patients}} - \underbrace{\frac{1}{N_A}\left[\sum_{\{i:\theta_i>\theta_{LA}} R_i + \sum_{\{i:\theta_i=\theta_{LA}\&D_A=1} R_i\right]}_{\text{Average } R \text{ among inframarginal and marginal } A \text{ patients}} \tag{54}$$

$$\tag{55}$$

$$= \frac{1}{N_L}\left[\sum_{\theta<\theta^*}(\gamma_L + \delta_L\theta_i) + N_{LA\&A}(\gamma_L + \delta_L\theta^*)\right] - \frac{1}{N_A}\left[\sum_{\theta>\theta^*}(\gamma_A + \delta_A\theta_i) + N_{LA\&A}(\gamma_A + \delta_A\theta^*)\right] \tag{56}$$

It follows that $\bar{R}_L - \bar{R}_A < 0$ if:

$$\frac{1-N_{\theta^*\&L}}{N_L}\gamma_L - \frac{1-N_{\theta^*\&A}}{N_A}\gamma_A + \frac{1}{N_L}\sum_{\theta<\theta^*}\delta_L\theta_i - \frac{1}{N_A}\sum_{\theta>\theta^*}\delta_A\theta_i < \frac{N_{\theta^*\&A}}{N_A}(\gamma_A + \delta_A\theta^*) - \frac{N_{\theta^*}\&L}{N_L}(\gamma_L + \delta_L\theta^*) \tag{57}$$

One case see from Equation 57 that the sign of the difference in means is dependent on an interaction of the differences between technologies in readmission rates among patients without complications, in the degrees to which readmission rates increase with respect to θ, and the shares of patients of each technology choice who are of different values of θ.

The right-hand side is the difference in weighted readmissions rates among $\theta_L A$-type patients and among abdominal patients, where the weights are the indifferent shares of patients of a particular choice. The left-hand side is the difference in weighted readmission rates among $\theta = 0$ types, where the weights are the inframarginal shares of patients of the respective technology choice, added to the difference in weighted average "complexity-sensitive" components of the readmission rates among inframarginal laparoscopic patients and among inframarginal abdominal patients, where the weights for a given patient type is that patient type's share of patients undergoing the respective type of surgery, and where "extra" readmission component is δ_j, the degree to which readmission rates increase under technology j with θ.

Here one can see that whether the difference in means is positive or negative is not dependent on the sign of the treatment effect among the marginal patients, whose treatment effect would be approximated by the local average treatment effect. In other words, in this selection setting, theory allows for the sign of the local average treatment effect to be different from the sign of the ordinary least squares estimate of the treatment effect. This suggests departing from the conventional notion that a contradiction between the sign of the estimated local average treatment effect and the sign of the ordinary least squares estimate of the average treatment effect is a cause for concern about the instrumental variable's validity. Theory predicts that the signs will be different under certain reasonable parameter assumptions and distributional assumptions.

Note that the finding about relative readmission rates among patients on the margin in Equation 48 implies

$$\gamma_A - \gamma_L < (\delta_L - \delta_A)\theta^* \tag{58}$$

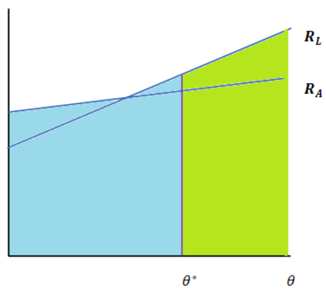
Consider three cases:

Case 1: $\gamma_A - \gamma_L > 0$. Then, $\delta_L - \delta_A > 0$. I.e., if readmission is worse for A than for L at $\theta = 0$, then readmissions must worsen faster, w.r.t. θ, under L than under A in order for readmissions to be higher under L than under A for the θ^* types.

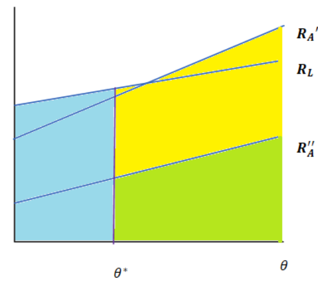
This is illustrated in Figure 9a. The average readmission rate among laparoscopic patients is the integral of R_L times the patient population density w.r.t. θ. The blue hatched area represents the average readmission rate if $\theta \sim Uniform$. The average readmission rate among abdominal patients under that distributional assumption is the green area. One can see that the average among laparoscopic patients relative to the average among abdominal patients rises if (1) the number of patients between the θ such that $R_L(\theta) = R_A(\theta)$ and θ^* rises, (2) the difference in slopes $\delta_L - \delta_A$ rises, and/or (3) $(\gamma_L - \gamma_A)$ rises.

Case 2a: $\gamma_A - \gamma_L < 0$**, and** $\delta_L - \delta_A > 0$. This case is represented by R_L and R'_A in Figure 9b. The average readmission rate under laparoscopic assuming uniformly distributed patients with respect to θ is represented in blue in Figure 9b. The average among abdominal patients is the yellow area plus the green area. Now readmissions mean under laparoscopic rises relative to the mean under abdominal patients if (1) the number of patients between the θ such that $R_L(\theta) = R_A(\theta)$ and θ^* declines (2) the relative slopes decline, and/or (3) $R_L(\theta) - R_A(\theta)$ is lesser.

Case 2b: $\gamma_A - \gamma_L < 0$**, and** $\delta_L - \delta_A < 0$. In this case, the mean readmission rate under laparoscopic surgery is always greater than the rate under abdominal surgery, $R''_A(\theta)$. If $\theta \sim Uniform$, the blue area in Figure 9b is the average among laparoscopic patients and the green area is the average among abdominal patients.



(a) Case 1. Readmissions with respect to θ, under each treatment alternative.

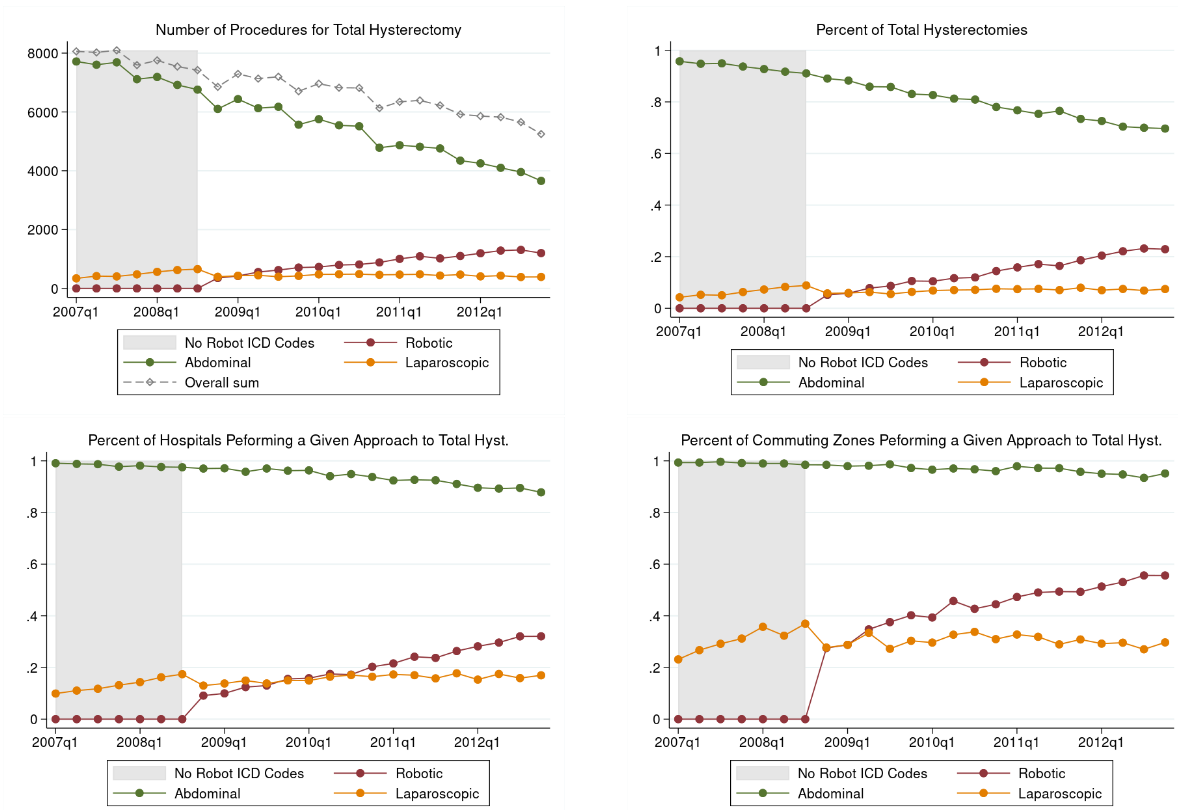


(b) Case 2a and Case 2b. Readmissions with respect to θ, under laparoscopic treatment ($R_L(\theta)$), under abdominal treatment in Case 2a ($R'_A(\theta)$), and under abdominal treatment in Case 2b ($R''_A(\theta)$).

B Further Data Description

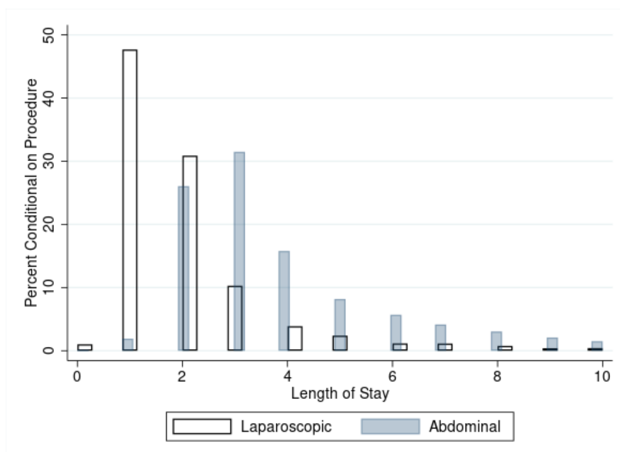
Figure 10 shows the trends in hysterectomy from 2007 to 2012. Note that the ICD-9 procedure code for robotically assisted procedure was introduced in the last quarter of 2008, and robotically assisted hysterectomies up to the point were coded as laparoscopic.

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.

Figure 11: Distribution of Inpatient Length of Stay across Hysterectomies, by Procedure Type



C Instrument Validity

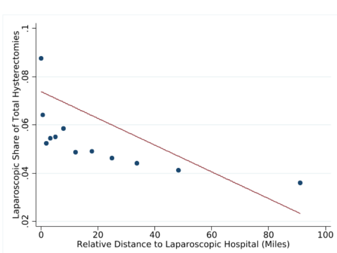
This section presents evidence supporting the validity of the relative distance instrument, which is equal to the difference between a patient's distance to her nearest hysterectomy-performing hospital that performs laparoscopic surgery and the distance to her nearest hysterectomy-performing hospital. Appendix Table 8 presents the first stage results, showing instrument relevance. Appendix Figure 12 graphically shows the negative relationship between relative distance and probability of choosing laparoscopic rather than abdominal hysterectomy. Appendix Table 9 presents evidence of the instrument's exclusion from the outcome function by comparing the characteristics of patients whose relative distance is greater than the median to those whose relative distance is less than the median. Appendix Table 15 tests for instrument monotonicity by estimating the first stage in demographic- and diagnostic-based subsamples.

Table 7: Distribution of the Relative Distance Instrument

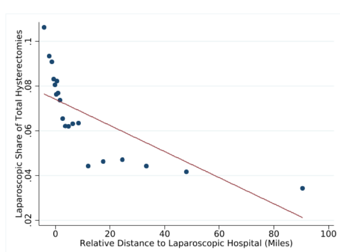
	p10	p25	p50	p75	p90
Relative Distance to Lap. Hospital	0	0	1	15	39

Distribution of values of the instrumental variable, a patient's distance to her nearest hospital that performs laparoscopic surgery and hysterectomy, relative to her nearest hospital that performs hysterectomy, in miles.

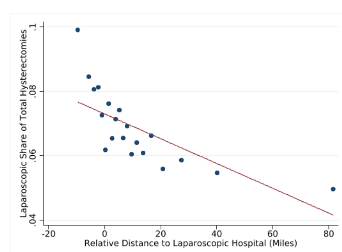
C.1 Relevance



(a) Without controls



(b) Controlling for patient demographics and clinical characteristics



(c) Additionally controlling for patient neighborhood and hospital characteristics

Figure 12: Binned scatter plots showing the association between a patient's relative distance to a hospital with laparoscopic surgery (in miles) and her likelihood of undergoing laparoscopic (as opposed to abdominal) hysterectomy. The second panel controls for patient demographic and clinical characteristics, and the third panel additionally controls for patient neighborhood and hospital characteristics. Lines of best fit are in red.

Table 8: First Stage Results: Linear Regression of Relative Hospital Distance Predicting Laparoscopic Choice

	(1)	(2)	(3)	(4)	(5)	(6)
Relative Distance	-0.000555***	-0.000581***	-0.000585***	-0.000414***	-0.000385***	-0.000392***
	(0.0000571)	(0.0000575)	(0.0000572)	(0.0000532)	(0.0000581)	(0.0000727)
Observations	54992	54992	54992	54972	48553	48553
Laparoscopic Rate	0.0670	0.0670	0.0670	0.0670	0.0686	.
Instrument Mean	12.32	12.32	12.32	12.31	11.35	.
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
Effective F	94.41	102.1	104.4	60.45	43.97	29.09

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

First stage with continuous instrumental variable for the estimation strategy to test the hypothesis about marginal patients by two-stage least squares. Across specifications, the instrumental variable is the difference between the patient's distance to her nearest hysterectomy-performing hospital with laparoscopic surgery and the distance to her nearest hysterectomy-performing hospital. The endogenous variable is an indicator for whether the hysterectomy was performed laparoscopically, rather than abdominally. Relative distance is measured in miles. Across specifications, the effective F statistic (due to Montiel Olea and Pflueger (2013) and Kleibergen and Paap (2006)) should be compared against either the two-stage least squares/limited information maximum likelihood critical value for 5% bias, which is 37.418, or for 10% bias, which is 23.109. Demographic controls: whether the patient is Black, a race other than white or Black, under 65 years of age, or over 74 years of age. Clinical controls: the Charlson comorbidity index and indicators for whether the patient had diabetes, malignant neoplasm, non-malignant neoplasm, body mass index of 30 or over (considered obese), history of cancer indicated on the hysterectomy claim, uterine fibroids, endometriosis, pelvic organ prolapse, female genital bleeding, post-menopausal bleeding, an ovarian cyst, female genital pain, or peripheral adhesions. Zip code controls of the patient's residence: white percent of residents, the college-educated percent of residents, the percent of residents with public assistance (including cash or nutritional assistance), the median household income, and the percent of residents on Medicaid. Hospital controls: number of hysterectomies the hospital performed that year, a quality measure on the appropriate use of antibiotics, a quality measure on the prevention of blood clots in heart patients, and the overall Consumer Assessment of Healthcare Providers & Systems (CAHPS) score. Standard errors assume clustering at the hospital level.

C.2 Independence

Table 9: Patient Characteristics in Top and Lower Half of Instrument's Distribution

	Lesser Relative Distance	Greater Relative Distance	Overall
Relative Distance to Lap. Hospital	0.0504	24.50	12.28
	(0.188)	(28.70)	(23.69)
Lap % of Hyst.	0.0846	0.0484	0.0665
	(0.278)	(0.215)	(0.249)
White	0.798	0.833	0.815
	(0.402)	(0.373)	(0.388)
Black	0.153	0.123	0.138
	(0.360)	(0.328)	(0.345)
Not Black or white	0.0492	0.0442	0.0467
	(0.216)	(0.206)	(0.211)
HMO	0.0432	0.0371	0.0402
	(0.203)	(0.189)	(0.196)
Charlson index	4.152	3.967	4.060
	(2.618)	(2.654)	(2.637)
Diabetes	0.171	0.179	0.175
	(0.376)	(0.383)	(0.380)
Malignant Neoplasm	0.486	0.451	0.469
	(0.500)	(0.498)	(0.499)
Non-Malignant Neoplasm	0.317	0.326	0.321
	(0.465)	(0.469)	(0.467)
BMI30+	0.0365	0.0256	0.0310
	(0.188)	(0.158)	(0.173)
History of Cancer	0.0799	0.0766	0.0783
	(0.271)	(0.266)	(0.269)
Uterine Fibroid	0.283	0.285	0.284
	(0.451)	(0.451)	(0.451)
Endometriosis	0.103	0.118	0.111
	(0.303)	(0.323)	(0.314)
Pelvic Organ Prolapse	0.0721	0.0796	0.0759
	(0.259)	(0.271)	(0.265)
Female Genital Bleeding	0.118	0.135	0.127
	(0.322)	(0.342)	(0.333)
Postmenopausal Bleeding	0.0985	0.101	0.0999
	(0.298)	(0.302)	(0.300)
Other Ovarian Cyst	0.0807	0.0850	0.0828
	(0.272)	(0.279)	(0.276)
Female Genital Pain	0.118	0.137	0.127
	(0.322)	(0.344)	(0.333)
Pelvic peritoneal adhesions	0.0980	0.0998	0.0989
	(0.297)	(0.300)	(0.299)

Characteristics among total hysterectomy patients. Lap = Laparoscopic. Hyst=Hysterectomies. HMO = Any months that year on Medicare Advantage (managed care). BMI30+ = Body mass index ≥ 30, considered obese.

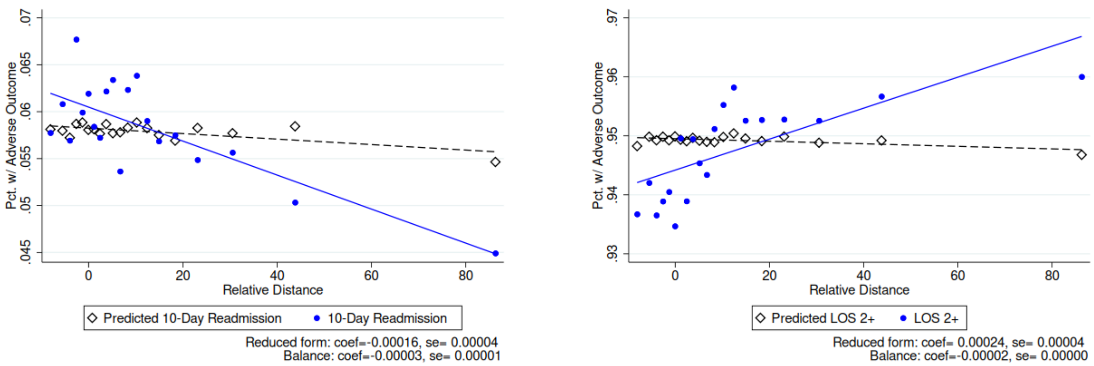
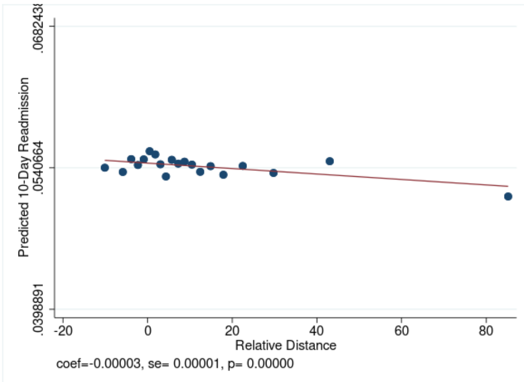
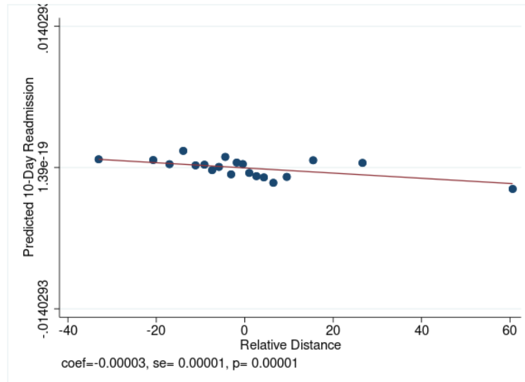


Figure 13: Main balance tests. Solid round binned scatterplots visually represent the reduced form regressions. Hollow diamonds constitute the balance test, showing the relationship between (1) the variation in adverse outcomes explained by patient and neighborhood characteristics and (2) the patient's relative distance to laparoscopic surgery. The latter correlation appears to be very small and an order of magnitude smaller than the reduced form effect, allaying concerns that the instrument's relationship with adverse outcomes of interest may be confounded by patients' geographic determinants of health.

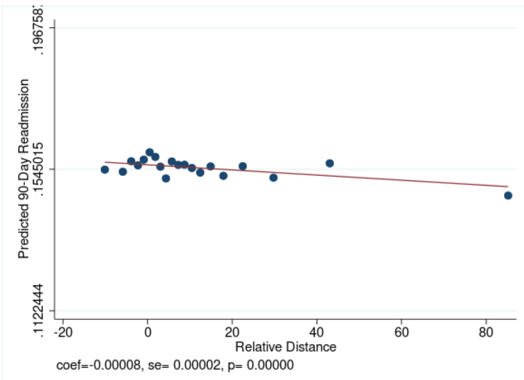
Figure 14: Binned Scatter Plots of Predicted Readmissions versus Relative Distance, controlling for minimum distance to any hospital and all Zip-level characteristics



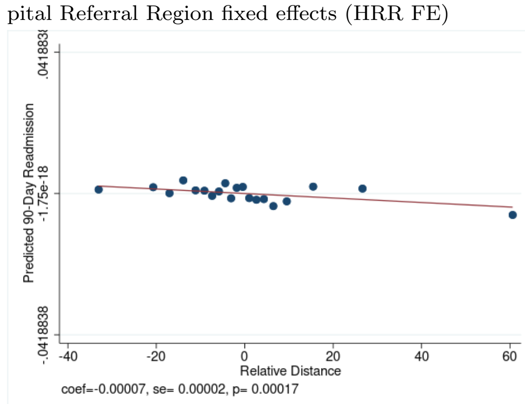
(a) Predicted Any 10-Day Readmission



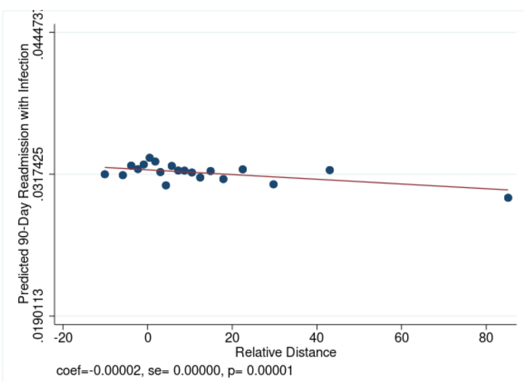
(b) Predicted Any 10-Day Readmission, controlling for Hospital Referral Region fixed effects (HRR FE)



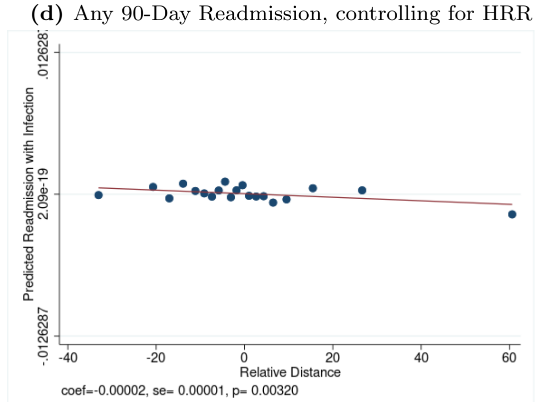
(c) Any 90-Day Readmission



(d) Any 90-Day Readmission, controlling for HRR FE

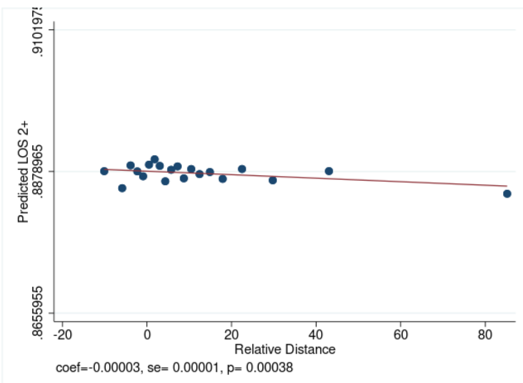


(e) Predicted 90-Day Readmission with Urogenital Infection

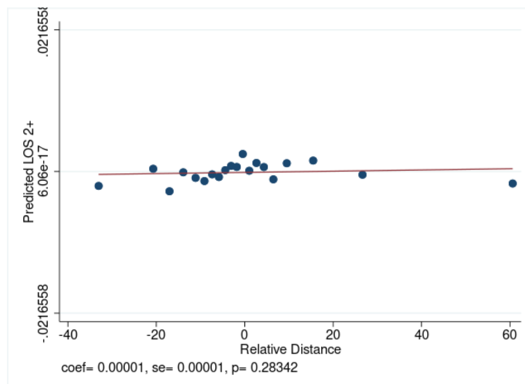


(f) Predicted 90-Day Readmission with Urogenital Infection, controlling for HRR FE

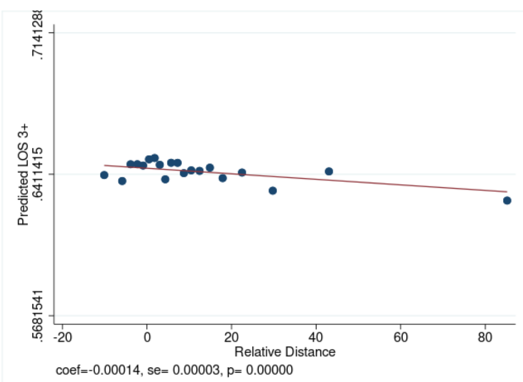
Figure 15: Binned Scatter Plots of Predicted Readmissions versus Relative Distance, controlling for minimum distance to any hospital and all Zip-level characteristics



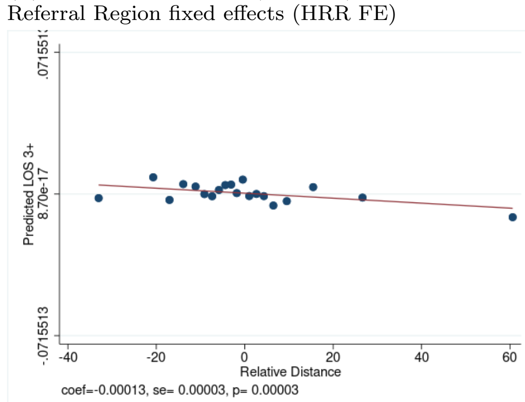
(a) Long Length of Stay (2+ Days)



(b) Long Length of Stay (2+ Days), controlling for Hospital Referral Region fixed effects (HRR FE)



(c) Long Length of Stay (3+ Days)



(d) Long Length of Stay (3+ Days), controlling for HRR FE

C.3 Reduced Form

Table 10: Reduced Form: Any 10-Day Readmission

	(1)	(2)	(3)	(4)	(5)	(6)
Relative Distance	-0.000201***	-0.000181***	-0.000153***	-0.000135***	-0.0000898**	-0.0000896*
	(0.0000387)	(0.0000385)	(0.0000381)	(0.0000393)	(0.0000441)	(0.0000498)
Observations	54992	54992	54992	54972	48553	48553
Dependent variable mean	0.0583	0.0583	0.0583	0.0583	0.0576	0.0576
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	27.02	15.50	21.08	17.48	14.87	14.14
Adj. R^2	0.000397	0.00174	0.00733	0.00739	0.00858	0.00869

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All specs control for Minimum Distance to Nearest Hospital.

Table 11: Reduced Form: Any 90-Day Readmission

	(1)	(2)	(3)	(4)	(5)	(6)
Relative Distance	-0.000292***	-0.000263***	-0.000174***	-0.000149**	-0.0000669	-0.000111
	(0.0000692)	(0.0000678)	(0.0000650)	(0.0000655)	(0.0000732)	(0.0000820)
Observations	54992	54992	54992	54972	48553	48553
Dependent variable mean	0.163	0.163	0.163	0.163	0.162	0.162
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	17.83	35.45	72.35	59.18	48.35	48.59
Adj. R^2	0.000334	0.00430	0.0291	0.0291	0.0306	0.0326

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All specs control for Minimum Distance to Nearest Hospital.

Table 12: Reduced Form: Any 90-Day Readmission with Urogenital Infection

	(1)	(2)	(3)	(4)	(5)	(6)
Relative Distance	-0.000113***	-0.0000973***	-0.0000735**	-0.000102***	-0.0000751**	-0.0000807**
	(0.0000291)	(0.0000288)	(0.0000286)	(0.0000297)	(0.0000334)	(0.0000369)
Observations	54992	54992	54992	54972	48553	48553
Dependent variable mean	0.0326	0.0326	0.0326	0.0326	0.0325	0.0325
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	15.09	31.04	27.25	22.46	17.78	19.27
Adj. R^2	0.000210	0.00424	0.0110	0.0111	0.0113	0.0111

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All specs control for Minimum Distance to Nearest Hospital.

Table 13: Reduced Form: LOS ≥ 2

	(1)	(2)	(3)	(4)	(5)	(6)
Relative Distance	0.000184***	0.000231***	0.000259***	0.000224***	0.000218***	0.000197***
	(0.0000447)	(0.0000454)	(0.0000457)	(0.0000462)	(0.0000501)	(0.0000519)
Observations	54992	54992	54992	54972	48553	48553
Dependent variable mean	0.950	0.950	0.950	0.950	0.949	0.949
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	17.01	40.68	18.52	15.40	13.50	13.66
Adj. R^2	0.000387	0.00493	0.0113	0.0118	0.0143	0.0296

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All specs control for Minimum Distance to Nearest Hospital.

Table 14: Reduced Form: LOS ≥ 3

	(1)	(2)	(3)	(4)	(5)	(6)
Relative Distance	-0.000342**	-0.000116	0.000102	0.000235*	0.000213	0.000190
	(0.000136)	(0.000127)	(0.000123)	(0.000122)	(0.000137)	(0.000133)
Observations	54992	54992	54992	54972	48553	48553
Dependent variable mean	0.707	0.707	0.707	0.707	0.707	0.707
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	6.295	322.3	173.5	149.6	126.4	125.8
Adj. R^2	0.000300	0.0490	0.109	0.111	0.115	0.138

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All specs control for Minimum Distance to Nearest Hospital.

C.4 Monotonicity

Table 15: Test for Monotonicity of Instrument

	Age < 65	Age ≥ 65	Age ≥ 75	Age < 75	64 < Age < 75	Age < 65 or Age > 74	White	Not White
Relative Distance	-0.000487***	-0.000332***	-0.000289***	-0.000416***	-0.000359***	-0.000405***	-0.000391***	-0.000286***
	(0.0000837)	(0.0000696)	(0.0000927)	(0.0000606)	(0.0000792)	(0.0000672)	(0.0000648)	(0.0000966)
Observations	14751	33808	13703	34856	20105	28454	39713	8846

	Malignant Neoplasm	No Malignant Neoplasm	Fibroids	No Fibroids	Pelvic Prolapse	No Prolapse	Genital Pain	No Genital Pa
Relative Distance	-0.000247***	-0.000482***	-0.000413***	-0.000372***	-0.000718***	-0.000354***	-0.000782***	-0.000311***
	(0.0000827)	(0.0000664)	(0.0000772)	(0.0000694)	(0.000173)	(0.0000600)	(0.000122)	(0.0000624)
Observations	22698	25861	13744	34815	3687	44872	6117	42442

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

First stage regression run on subsets of the sample, where the dependent variable is whether a hysterectomy was performed laparoscopically and the independent variable is the instrumental variable, relative distance. Headers describe patient subsample. Relative distance is the difference between a patient's distance to her nearest hysterectomy-performing hospital with laparoscopic surgery and her distance to her nearest hysterectomy-performing hospital. First stage estimates are qualitatively the same and quantitatively similar across subsamples, suggesting that different types of patients respond to the instrument in the same way and that the instrument satisfies monotonicity.

D Additional Two-Stage Least Squares Results

Table 16: Effect of Laparoscopic Procedure on the Probability of Length of Stay is 3 or More Days: 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)
Laparoscopic	0.616**	0.199	-0.174	-0.568**	-0.552	-0.484
	(0.258)	(0.223)	(0.208)	(0.285)	(0.344)	(0.323)
Observations	54992	54992	54992	54972	48553	48553
Dependent Variable Mean	0.707	0.707	0.707	0.707	0.707	0.707
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	5.694	299.9	183.6	162.3	135.6	136.7
Adj. R^2	-0.311	-0.0267	0.155	0.196	0.203	0.195
First-Stage Effective F	94.41	102.1	104.4	60.45	43.97	29.09

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Two-staged least squares estimates, using relative distance to hospital with laparoscopic surgery as instrument. Across specifications, the effective F statistic (due to Montiel Olea and Pflueger (2013) and Kleibergen and Paap (2006)) should be compared against either the two-stage least squares/limited information maximum likelihood critical value for 5% bias, which is 37.418, or for 10% bias, which is 23.109. Demographic controls: whether the patient is Black, a race other than white or Black, under 65 years of age, or over 74 years of age. Clinical controls: the Charlson comorbidity index and indicators for whether the patient had diabetes, malignant neoplasm, non-malignant neoplasm, body mass index of 30 or over (considered obese), history of cancer indicated on the hysterectomy claim, uterine fibroids, endometriosis, pelvic organ prolapse, female genital bleeding, post-menopausal bleeding, an ovarian cyst, female genital pain, or peripheral adhesions. Zip code controls of the patient's residence: white percent of residents, the college-educated percent of residents, the percent of residents with public assistance (including cash or nutritional assistance), the median household income, and the percent of residents on Medicaid. Hospital controls: number of hysterectomies the hospital performed that year, a quality measure on the appropriate use of antibiotics, a quality measure on the prevention of blood clots in heart patients, and the overall Consumer Assessment of Healthcare Providers & Systems (CAHPS) score. Standard errors assume clustering at the hospital level.

Table 17: Local Effect of Laparoscopic Procedure on the Probability of Any 90-day Readmission: 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)
Laparoscopic	0.527***	0.452***	0.298***	0.361**	0.174	0.284
	(0.128)	(0.119)	(0.112)	(0.163)	(0.191)	(0.215)
Observations	54992	54992	54992	54972	48553	48553
Dependent Variable Mean	0.163	0.163	0.163	0.163	0.162	0.162
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	16.89	31.29	66.94	52.58	47.34	44.93
Adj. R^2	-0.148	-0.107	-0.0236	-0.0449	0.00935	-0.0256
First-Stage Effective F	94.41	102.1	104.4	60.45	43.97	29.09

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Two-staged least squares estimates, using relative distance to hospital with laparoscopic surgery as instrument. Across specifications, the effective F statistic (due to Montiel Olea and Pflueger (2013) and Kleibergen and Paap (2006)) should be compared against either the two-stage least squares/limited information maximum likelihood critical value for 5% bias, which is 37.418, or for 10% bias, which is 23.109. Demographic controls: whether the patient is Black, a race other than white or Black, under 65 years of age, or over 74 years of age. Clinical controls: the Charlson comorbidity index and indicators for whether the patient had diabetes, malignant neoplasm, non-malignant neoplasm, body mass index of 30 or over (considered obese), history of cancer indicated on the hysterectomy claim, uterine fibroids, endometriosis, pelvic organ prolapse, female genital bleeding, post-menopausal bleeding, an ovarian cyst, female genital pain, or peripheral adhesions. Zip code controls of the patient's residence: white percent of residents, the college-educated percent of residents, the percent of residents with public assistance (including cash or nutritional assistance), the median household income, and the percent of residents on Medicaid. Hospital controls: number of hysterectomies the hospital performed that year, a quality measure on the appropriate use of antibiotics, a quality measure on the prevention of blood clots in heart patients, and the overall Consumer Assessment of Healthcare Providers & Systems (CAHPS) score. Standard errors assume clustering at the hospital level.

Table 18: Effect of Laparoscopic Procedure on the Probability of Any 90-day Readmission Accompanied by Urogenital Infection: 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)
Laparoscopic	0.204***	0.168***	0.126**	0.247***	0.195**	0.206**
	(0.0550)	(0.0516)	(0.0502)	(0.0778)	(0.0923)	(0.102)
Observations	54992	54992	54992	54972	48553	48553
Dependent Variable Mean	0.0326	0.0326	0.0326	0.0326	0.0325	0.0325
Demographic Controls		✓	✓	✓	✓	✓
Clinical Controls			✓	✓	✓	✓
Zip Code Controls				✓	✓	✓
Hospital Controls					✓	✓
Fixed Effects						HRR
F	13.68	28.68	24.52	16.78	14.37	15.43
Adj. R^2	-0.0937	-0.0610	-0.0274	-0.122	-0.0765	-0.0906
First-Stage Effective F	94.41	102.1	104.4	60.45	43.97	29.09

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Two-staged least squares estimates, using relative distance to hospital with laparoscopic surgery as instrument. Across specifications, the effective F statistic (due to Montiel Olea and Pflueger (2013) and Kleibergen and Paap (2006)) should be compared against either the two-stage least squares/limited information maximum likelihood critical value for 5% bias, which is 37.418, or for 10% bias, which is 23.109. Demographic controls: whether the patient is Black, a race other than white or Black, under 65 years of age, or over 74 years of age. Clinical controls: the Charlson comorbidity index and indicators for whether the patient had diabetes, malignant neoplasm, non-malignant neoplasm, body mass index of 30 or over (considered obese), history of cancer indicated on the hysterectomy claim, uterine fibroids, endometriosis, pelvic organ prolapse, female genital bleeding, post-menopausal bleeding, an ovarian cyst, female genital pain, or peripheral adhesions. Zip code controls of the patient's residence: white percent of residents, the college-educated percent of residents, the percent of residents with public assistance (including cash or nutritional assistance), the median household income, and the percent of residents on Medicaid. Hospital controls: number of hysterectomies the hospital performed that year, a quality measure on the appropriate use of antibiotics, a quality measure on the prevention of blood clots in heart patients, and the overall Consumer Assessment of Healthcare Providers & Systems (CAHPS) score.

E Additional Results from Marginal Treatment Effect Estimations

E.0.1 Results Using Local Instrumental Variable Method

This subsection provides support for the model's predictions for marginal patients using the local instrumental variable method. Figure 16 graphically presents the estimates of marginal treatment effects with respect to percentiles of unobserved resistance to laparoscopic hysterectomy, under four different parametric approaches to modeling the outcomes as functions of unobserved heterogeneity and four different semiparametric approaches. Subfigure (a) summarizes the estimates across all model specifications. Figures (b) through (i) show the estimates one at a time from each model, as well as 90 percent confidence intervals. Standard errors are analytically derived for parametric models and bootstrapped for semiparametric models. Effects are statistically significant across most percentiles in each model result. The local average treatment effect estimate from Table 4, Column 4, -0.54, is near the middle of marginal treatment effects estimated over the supported range of percentiles of resistance.

Figure 17 presents analogous results of marginal treatment effects on the chance of an all-cause 10-day readmission. Subfigure (a) summarizes results across models and shows that estimated marginal treatment effects are positive and upward sloping as functions of unobserved resistance to laparoscopic surgery, across model specifications. Point estimates from parametric models, presents with confidence intervals in subfigures (b) through (i), are not statistically significant at the 90 percent level at any levels of unobserved resistance, although for some percentiles around 0.05 and 0.1, for which there is substantial common support, much of the probability mass of the point estimates are positive. The local average treatment effect estimate from Table 17, Column 4, 0.36, is near the middle of marginal treatment effects estimated over the supported range of percentiles of resistance.

The estimates from local instrumental variable estimation of marginal treatment effects on length of stay and readmission rate are supportive of the model's prediction of a tradeoff among marginal patients between the two adverse clinical outcomes. This is not surprising, since the local average treatment effects, estimated above, are known to be weighted combinations of marginal treatment effects across the support of the instrumental variable.

E.1 Results Using Separate Method

This subsection presents estimates of marginal treatment effects on length of stay and readmission risk from the separate estimation method, presented analogously to the results from local instrumental variable estimation. The results largely resemble the results from the local instrumental variable approach. Figure 18 shows that across model specifications, marginal treatment effects on having a length of stay of two or more days are estimated to be negative and, as a function of unobserved resistance, is estimated to be downward sloping.

Figure 19 shows that the estimates of effects on any readmission from the separate method, like in the local instrumental variable approach, are positive across most of the support, but not statistically significant at the 90 percent confidence level. One difference is that estimates from the separate method suggest that the marginal treatment effects as a function of unobserved resistance is upward sloping, whereas the local instrumental variables method suggested it is downward sloping. Over the support, the separate method estimates that the marginal treatment effect varies from about 0.5 to zero. The two-stage least squares estimate of the local average treatment effect on readmission risk is 0.361, controlling for demographic, clinical, and Zip code-level controls, and it is 0.173 when additionally controlling for hospital characteristics. These estimates of the local effect fall within the range of estimated marginal effects.

E.2 Marginal Treatment Effect Weights

The local average treatment effect is a weighted combination of the marginal treatment effects across all percentiles of unobserved resistance. Figure 20 plots the weights, estimated from data, that relate the marginal treatment effect at a particular level of unobserved resistance to the local average treatment effect.

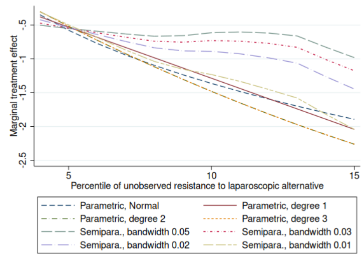
Recall from Heckman and Vytlacil (1999, 2005) and Heckman, Urzua and Vytlacil (2006) that

$$LATE_Y\left(p_0, p_1\right) = \int_{p_0}^{p_1} MTE_Y\left(p\right) \varphi_{\mathrm{IV}}^{Z}\left(u_{D_L}\right) dp \tag{59}$$

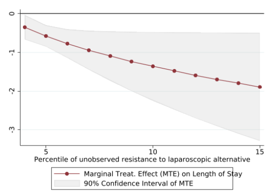
where the weights relating the MTEs to the LATE are:

$$\varphi_{IV}^{Z}\left(u_D\right) = \frac{\mathbb{E}\left[Z - \mathbb{E}[Z] \mid P(Z) > u_D\right] Pr\left(P(Z) > u_D\right)}{\mathrm{Cov}\left(Z, D\right)}$$

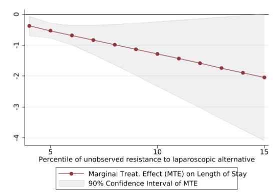
Certain observations are weighted more heavily if their treatment covaries with particular ranges of the instrument more. The weights integrate to one, can be negative if the instrument does not satisfy monotonicity, and can be consistently estimated from the sample.



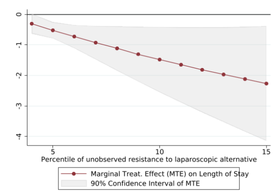
(a) All estimates of effect on readmission, from separate method. Summarizes estimates in plots (b) through (g).



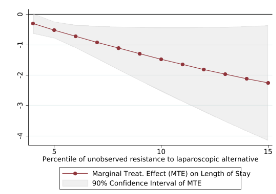
(b) Parametric: Normal.



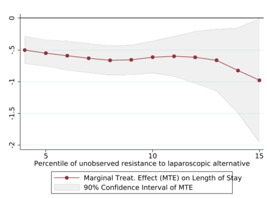
(c) Parametric: first-degree polynomial.



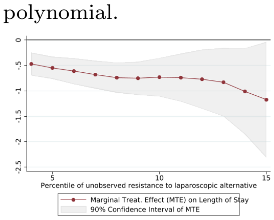
(d) Parametric: second-degree polynomial.



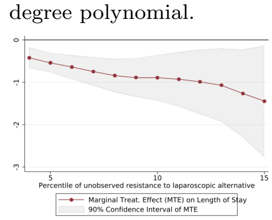
(e) Parametric: third-degree polynomial.



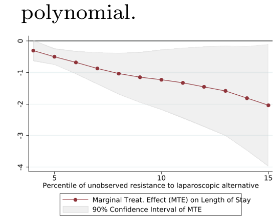
(f) Semiparametric: bandwidth 0.05.



(g) Semiparametric: bandwidth 0.03.

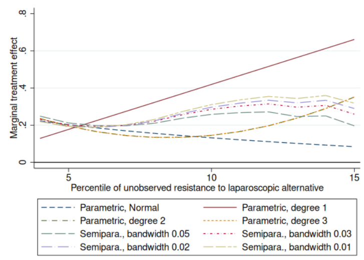


(h) Semiparametric: bandwidth 0.02.

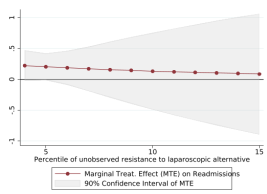


(i) Semiparametric: bandwidth 0.01.

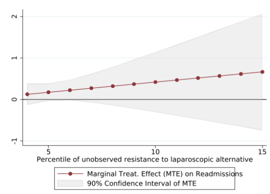
Figure 16: Local Instrumental Variable Method: Length of Stay. Estimates of marginal treatment effects of laparoscopic surgery, as opposed to abdominal surgery, on the probability of the length of stay being two or more days, using the separate approach. The horizontal axis in each plot is U_D, the case's percentile on the distribution of unobserved "resistance" to or "cost" of the laparoscopic choice. Gray bands are 90% confidence intervals. Unobserved heterogeneity, modeled as a function of the propensity score, p, is alternatively modeled parameterically (either Normal or as a polynomial of p) or semiparametrically, using the Epanechnikov kernel with alternative bandwidths. Standard errors are bootstrapped with 100 repititions. Subfigure (a) summarizes the point estimates in plots (b) through (h).



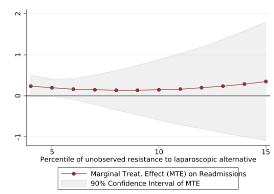
(a) All estimates of effect on readmission, from separate method. Summarizes estimates in plots (b) through (g).



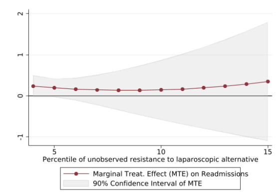
(b) Parametric: Normal.



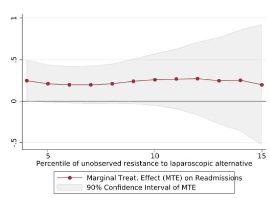
(c) Parametric: first-degree polynomial.



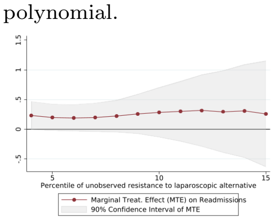
(d) Parametric: second-degree polynomial.



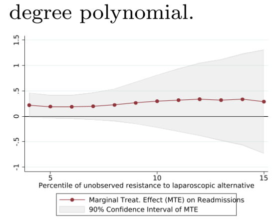
(e) Parametric: third-degree polynomial.



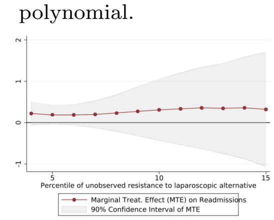
(f) Semiparametric: bandwidth 0.05.



(g) Semiparametric: bandwidth 0.03.

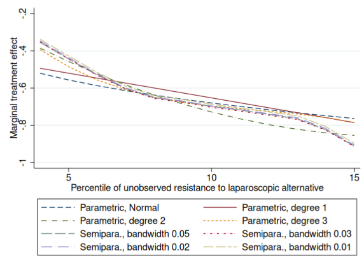


(h) Semiparametric: bandwidth 0.02.

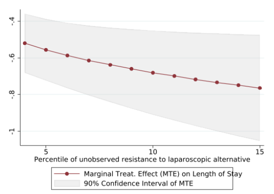


(i) Semiparametric: bandwidth 0.01.

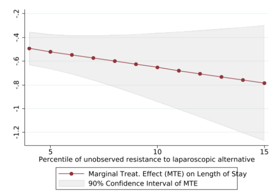
Figure 17: Local Instrumental Variable Method: Readmissions. Estimates of marginal treatment effects of laparoscopic surgery, as opposed to abdominal surgery, on the probability of an all-cause 10-day readmission, using the local instrumental variable approach. The horizontal axis in each plot is U_D, the case's percentile on the distribution of unobserved "resistance" to or "cost" of the laparoscopic choice. Gray bands are 90% confidence intervals. Unobserved heterogeneity, modeled as a function of the propensity score,p, is alternatively modeled parameterically (either Normal or as a polynomial of p) or semiparametrically, using the Epanechnikov kernel with alternative bandwidths. Standard errors are bootstrapped with 100 repititions. Subfigure (a) summarizes the point estimates in plots (b) through (h).



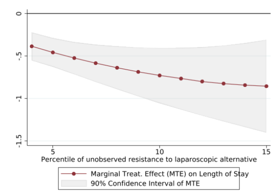
(a) All estimates of effect on readmission, from separate method. Summarizes estimates in plots (b) through (g).



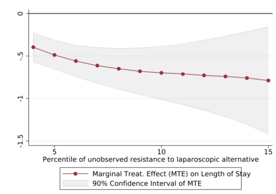
(b) Parametric: Normal.



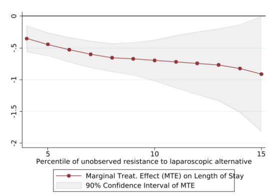
(c) Parametric: first-degree polynomial.



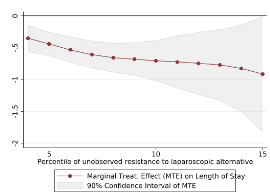
(d) Parametric: second-degree polynomial.



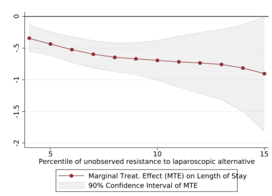
(e) Parametric: third-degree polynomial.



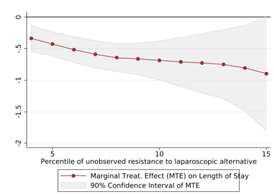
(f) Semiparametric: bandwidth 0.05.



(g) Semiparametric: bandwidth 0.03.

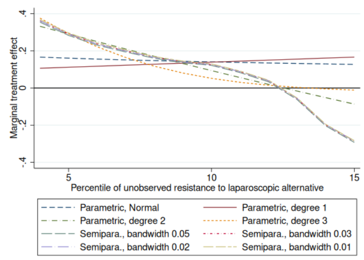


(h) Semiparametric: bandwidth 0.02.

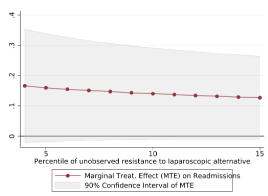


(i) Semiparametric: bandwidth 0.01.

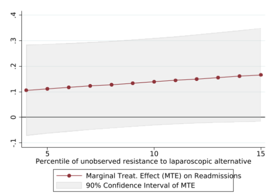
Figure 18: Separate Method: Length of Stay. Estimates of marginal treatment effects of laparoscopic surgery, as opposed to abdominal surgery, on the probability of the length of stay being two or more days, using the separate approach. The horizontal axis in each plot is U_D, the case's percentile on the distribution of unobserved "resistance" to or "cost" of the laparoscopic choice. Gray bands are 90% confidence intervals. Unobserved heterogeneity, modeled as a function of the propensity score,p, is alternatively modeled parameterically (either Normal or as a polynomial of p) or semiparametrically, using the Epanechnikov kernel with alternative bandwidths. Standard errors are bootstrapped with 100 repititions. Subfigure (a) summarizes the point estimates in plots (b) through (f).



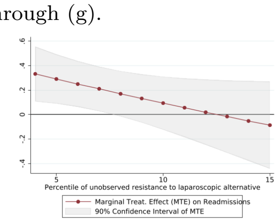
(a) All estimates of effect on readmission, from separate method. Summarizes estimates in plots (b) through (g).



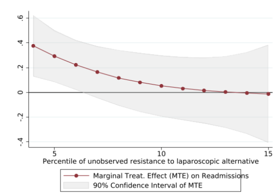
(b) Parametric: Normal.



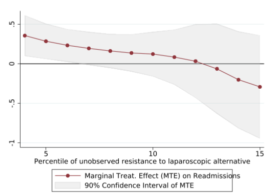
(c) Parametric: first-degree polynomial.



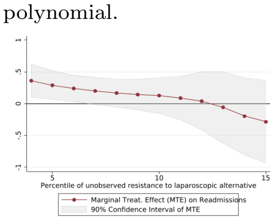
(d) Parametric: second-degree polynomial.



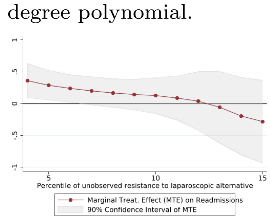
(e) Parametric: third-degree polynomial.



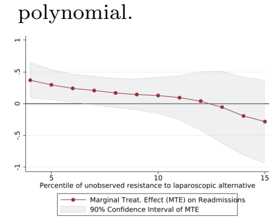
(f) Semiparametric: bandwidth 0.05.



(g) Semiparametric: bandwidth 0.03.



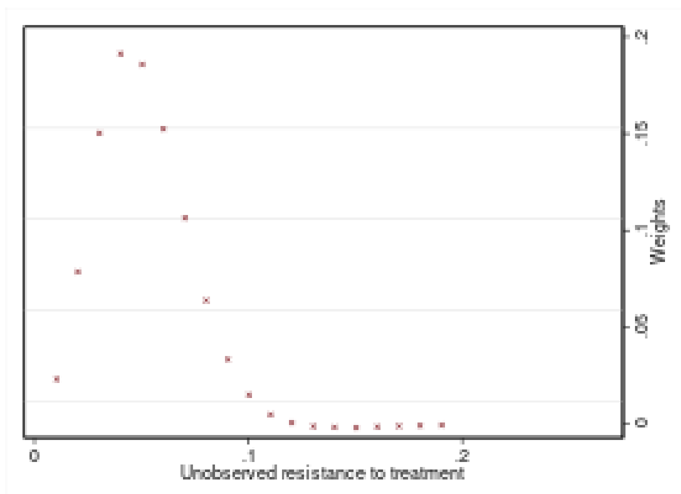
(h) Semiparametric: bandwidth 0.02.



(i) Semiparametric: bandwidth 0.01.

Figure 19: Separate Method: Readmission. Estimates of marginal treatment effects of laparoscopic surgery, as opposed to abdominal surgery, on the probability of an all-cause 10-day readmission, using the separate approach. The horizontal axis in each plot is U_D, the case's percentile on the distribution of unobserved "resistance" to or "cost" of the laparoscopic choice. Gray bands are 90% confidence intervals. Unobserved heterogeneity, modeled as a function of the propensity score,p, is alternatively modeled parameterically (either Normal or as a polynomial of p) or semiparametrically, using the Epanechnikov kernel with alternative bandwidths. Standard errors for parametric models are calculated analytically, while standard errors for semiparametric models are bootstrapped with 100 repititions. Subfigure (a) summarizes the point estimates in plots (b) through (g).

Figure 20: Weights Relating the Marginal Treatment Effects to the Local Average Treatment Effects



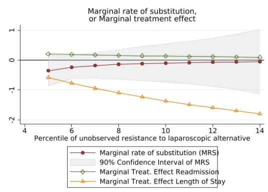
The horizontal axis in each plot is U_D, the case's percentile on the distribution of unobserved "resistance" to or "cost" of the laparoscopic choice. The Xs indicate the weight that relates the marginal treatment effect at that percentile to the local average treatment effect.

F Full Marginal Rate of Substitution Estimates

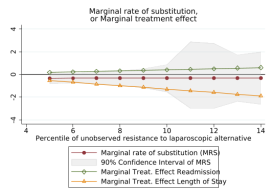
Figure 21 presents the estimates of the marginal rates of substitution, and their respective confidence intervals, calculated from the marginal treatment effects on length of stay and on readmission, which are also plotted alongside.

Figure 21: Estimates of Marginal Rate of Substitution, Their Confidence Intervals, and Marginal Treatment Effects

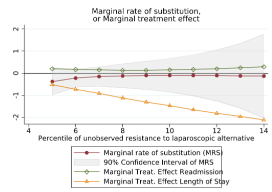
Estimates from Local Instrumental Variable Method



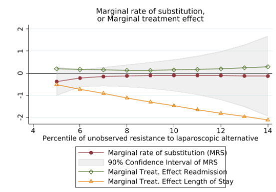
(a) Parametric: Normal.



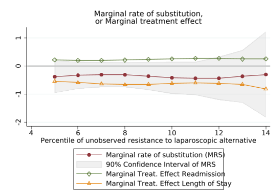
(b) Parametric: first-degree polynomial.



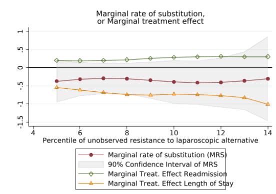
(c) Parametric: second-degree polynomial.



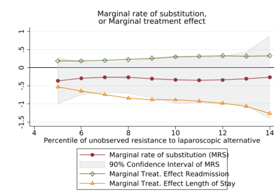
(d) Parametric: third-degree polynomial.



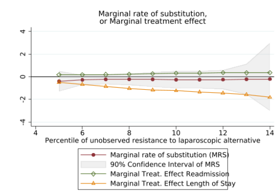
(e) Semiparametric: bandwidth 0.05.



(f) Semiparametric: bandwidth 0.03.

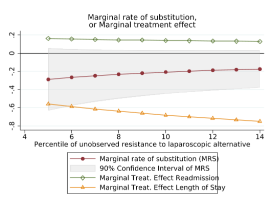


(g) Semiparametric: bandwidth 0.02.

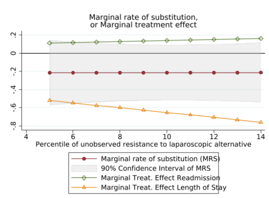


(h) Semiparametric: bandwidth 0.01.

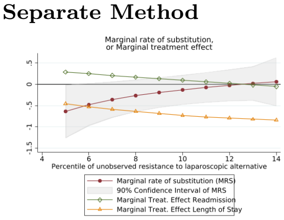
Estimates from Separate Method



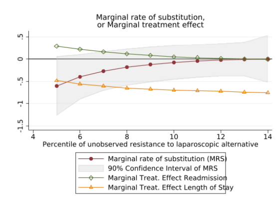
(i) Parametric: Normal.



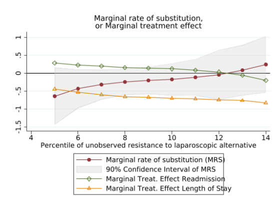
(j) Parametric: first-degree polynomial.



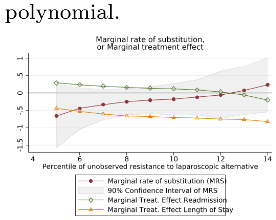
(k) Parametric: second-degree polynomial.



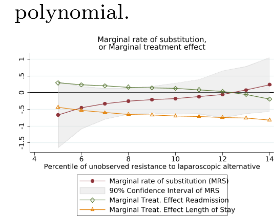
(l) Parametric: third-degree polynomial.



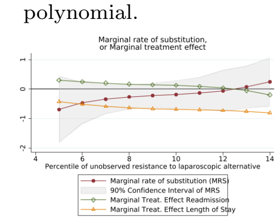
(m) Semiparametric: bandwidth 0.05.



(n) Semiparametric: bandwidth 0.03.



(o) Semiparametric: bandwidth 0.02.



(p) Semiparametric: bandwidth 0.01.

Estimates of marginal rate of subtitution (MRS) and, in gray bands, their 90% confidence intervals, as well as the marginal treatment effects of the separate method from which they are calculated. The horizontal axis in each plot is U_D, the case's percentile on the distribution of unobserved "resistance" to or "cost" of the laparoscopic choice. Unobserved heterogeneity, modeled as a function of the propensity score,p, is alternatively modeled parameterically (either Normal or as a polynomial of p) or semiparametrically, using the Epanechnikov kernel with alternative bandwidths. Standard errors for parametric models are calculated analytically, while standard errors for semiparametric models are bootstrapped with 100 repititions. In Panel (b), the plotted lower confidence bounds for the MRS at $U_D = 11, ..., 14$ are truncated. They reach a minimum of -5.