Journal Pre-proof

Information, Relative Skill, and Technology Abandonment

Bingxiao Wu Conceptualization , Guy David Conceptualization

PII: S0167-6296(22)00016-9
DOI: https://doi.org/10.1016/j.jhealeco.2022.102596
Reference: JHE 102596

To appear in: *Journal of Health Economics*

Received date: 3 September 2020
Revised date: 16 January 2022
Accepted date: 27 January 2022

Please cite this article as: Bingxiao Wu Conceptualization , Guy David Conceptualization , Information, Relative Skill, and Technology Abandonment, *Journal of Health Economics* (2022), doi: https://doi.org/10.1016/j.jhealeco.2022.102596

Information, Relative Skill, and Technology Abandonment

Bingxiao Wu[1,*] bwu@econ.rutgers.edu, Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Formal Analysis, Data curation, Visualization, Guy David[2,] Conceptualization, Methodology, Writing - original draft, Writing - review & editing

[1]Department of Economics, Rutgers University, 75 Hamilton St, New Brunswick, NJ, 08901

[2]Department of Health Care Management, The Wharton School, Department of Medical Ethics and Health Policy, Perelman School of Medicine, University of Pennsylvania, 305 Colonial Penn Center, 3641 Locust Walk, Philadelphia, PA, 19104

[*]Corresponding author: Professor Bingxiao Wu

Abstract

We study the role of relative task-specific skill in explaining the heterogeneity in physicians' technology abandonment decisions in response to negative information shocks. We show that after an unexpected FDA safety warning on the use of minimally invasive hysterectomies, physicians alter their procedural mix towards open procedures and away from the minimally invasive procedures. This effect is less pronounced for physicians more skilled in performing minimally invasive procedures relative to open procedures, highlighting relative skill as an explanation for differential technology abandonment. Since physicians with higher relative skill are more likely to use minimally invasive procedures before the FDA safety communication, we find that the FDA intervention led to a substantial increase in practice variation across physicians with different relative skill levels. These findings are consistent with a theoretical model that predicts physicians' response to new information regarding the effectiveness of medical technology.

Key Words

Physician Skill, Technology Abandonment, FDA Safety Communication

1. INTRODUCTION

Considerable variation in physician practice patterns has been documented both across-markets and within-markets (Epstein & Nicholson 2009). This variation has been attributed to productivity spillovers across physicians in high-use areas (Chandra & Staiger 2007; Epstein & Nicholson 2009; Dranove *et al.* 2011), local information spillovers driven by opinion leaders (Agha & Molitor 2018), heterogeneous beliefs across physicians about treatment effectiveness (Cutler *et al.* 2019), and physicians exhibiting heterogeneous diagnostic skills (Currie & MacLeod 2017; Chan *et al.* 2019; Currie & MacLeod 2020).

This paper proposes a new mechanism - heterogeneity in task-specific skill across physicians - as a key contributor to the variation in practice patterns across physicians. The role of individuals' task-specific skill in determining choice has long been documented in the economics literature (Roy 1950; Miller 1984; Heckman & Honoré 1990). In medical care, the command of multiple skills by a single physician may result in a set of competencies, where substitutable tasks represent alternative ways (e.g., medical procedures) to treat patients. However, empirical evidence on how such heterogeneity in physicians' procedure-specific skills affects their choices of treatments remains limited. In this paper, we fill this gap in the literature and show how a negative information shock to treatment effectiveness affects physicians' technology abandonment decisions differentially depending on their relative skill in performing two alternative surgical interventions (one subject of the negative information shock and the other not).

In the past three decades, many surgical procedures transitioned from a standard, open surgical approach to a minimally invasive one (Barbash & Glied 2010). This technological transition requires the acquisition of very different task-specific skills compared with open procedures (Noar 1991) and had an impact on physician productivity (Epstein *et al.* 2013) as well as contributed to geographical variation in physician behavior (Baicker *et al.* 2006; Kuo *et al.* 2017). In this paper, we examine physician choice in hysterectomies, a specific clinical setting where surgeons commonly perform both minimally invasive (MI) and open surgery for women diagnosed with uterine fibroids and related diseases. While the two alternative surgical techniques require very different clinical skills, they are close substitutes as treatment options. By contrast, previous literature on physician choice of technology focused on C-section versus normal delivery for childbirth, intensive therapy versus medical management for treating heart patients (Chandra & Staiger 2007; Currie *et al.* 2016), and (Currie & MacLeod 2017), and alternative drug treatments (Currie & MacLeod 2020).

The goal of this paper is to understand the role of physician relative skill (i.e., the skill of performing a procedure relative to the alternative procedure) in governing surgical procedural choice. Specifically, we ask how physicians' response to negative information shock on the appropriateness of tasks depends on their relative skill levels. Negative news regarding treatment effectiveness should alter the choice of treatment and may lead to gradual or swift abandonment of the technology in question. Our empirical analysis uses patient-level administrative data from Florida. The data allows us to measure both patient appropriateness and surgeons' relative skill in performing the two alternative procedures. Using a leave-out instrumental variable for individual physicians' skill levels, we show that after the Food and Drug Administration (FDA) unexpectedly issued a safety communication that lowered the attractiveness of MI hysterectomies, surgeons with relatively higher skill in open versus MI procedures were more likely to alter their procedural mix towards open hysterectomies and away from MI hysterectomies. Since surgeons with relatively higher open skills were more likely to use open procedures even before the FDA announcement, we find that practice variation across physicians with different relative skill levels increased substantially after the information shock. This finding highlights the importance of human capital in explaining practice variation. To better understand and provide insights into the empirical results, we also develop a simple theoretical model on physicians' technology abandonment decision in response to negative information shock, where physicians' technology-choice decision is a function of their relative skill, their patients' appropriateness of using either procedure, and the perceived effectiveness of each procedure, which can be affected by the information shock.

The literature on technology abandonment is fairly thin and heavily focused on measuring the speed of abandonment. Recent literature on technology abandonment is focused on uncovering potential mechanisms for abandonment beyond the informational shocks. These mechanisms help explain the heterogeneity in abandonment across seemingly similar providers, and include variation in organizational environments and reimbursement levels (Howard *et al.* 2017), peer effects (Berez *et al.* 2018), overconfidence (Comin *et al.* 2017), as well as different behavioral explanations (Roman & Asch 2014; Ubel & Asch 2015; Staats *et al.* 2018). This paper, as we will detail later on, argues that variation in relative skill across alternative procedures is linked to the extent of technology abandonment across surgeons. Bekelis *et al.* (2017) makes the important distinction between the terms “de-adoption,” which refers to the total relinquishment of an activity, and “exnovation,” which refers to scaling back that activity. In this paper, we chose the term “abandonment” to capture elements of de-adoption, as in the case of the technology targeted by the FDA, as well as exnovation, as in the case of minimally invasive hysterectomies.

The paper is organized as follows: Section 2 provides the clinical and institutional background. Section 3 details the theoretical framework. Section 4 describes the data and the empirical measure of physician skill. Section 5 highlights the identification strategy. Section 6 shows the results. Section 7 offers conclusions.

2. BACKGROUND

We explore physicians’ choice between two broad surgical alternatives for hysterectomy surgery. The first option is laparotomy hysterectomy, an open procedure involving a large incision in the lower abdomen. The second option is laparoscopic hysterectomy, a minimally invasive (MI) procedure involving a number of small incisions and aided by camera and monitor. MI hysterectomy for benign indications (e.g., uterine fibroids) normally uses a device called a power morcellator to slice up the tissue, allowing surgeons to work through small slits rather than large open cuts. As a result, the patient heals faster with a lower risk of bleeding, infection, and other postoperative complications; however, MI hysterectomy is not appropriate for patients with certain contraindications.[1] MI hysterectomy requires a surgeon who is experienced with laparoscopic techniques. In addition, surgeons must command skills in both surgical options, as a potential complication of laparoscopic surgery is the need for the surgeon to switch to a laparotomy incision during the procedure.[2,3] This setting allows us to study relative performance, its impact on the choice of procedure, and its role in affecting surgeons' response to negative information shocks.

Triggered by an unexpected tragic event in which the power morcellation during a laparoscopic hysterectomy treating "benign" uterine fibroids caused the spread of a hidden type of rare cancer in a patient,[4] a petition to the FDA was filed in December 2013 calling for an end to the use of power morcellation during MI hysterectomies, as it was a potential cause of cancer.[5] Following widespread national news coverage, the FDA published a safety communication on April 17, 2014, discouraging the use of laparoscopic power morcellation during *all* MI hysterectomies and highlighting the risk of spraying malignant tissue in the event of an unsuspected sarcoma. The FDA also noted that uterine sarcoma is a hidden type of rare cancer whose signs and symptoms are similar to uterine fibroids, and there are no tests or exams to detect uterine sarcoma prior to surgery.[6] [7]

This setting provides a unique opportunity to study physicians' heterogeneous responses in reaction to negative news for a number of reasons. First, the negative information shock was unexpected, and most physicians were caught by surprise by the FDA

---

1 Major contraindications to laparoscopic hysterectomy are the anesthetic contraindications to perform the pneumoperitoneum, which is required for laparoscopic hysterectomy but not for laparotomy hysterectomy (Ferreira & Barga 2018).
2 This might occur for many reasons, including the need for better visualization of the pelvis or controlling bleeding during the procedure.
3 Due to the possibility of switching to laparotomy hysterectomy during the laparoscopic procedure, all laparoscopic hysterectomies, including both inpatient and outpatient surgeries, are performed in hospital operating rooms rather than ambulatory surgery centers. All laparotomy hysterectomies are inpatient surgeries and are also performed in hospital operating rooms.
4 In October 2013, a patient, Dr. Amy Reed, received laparoscopic surgery in a teaching hospital in Boston, Massachusetts, where she herself was working as a physician at the time. The procedure sprayed malignant cells from a hidden type of rare cancer, uterine sarcoma, around her abdomen. Shortly after the procedure, Dr. Reed was diagnosed with stage 4 uterine cancer. (Source: https://www.nytimes.com/2017/05/24/us/amy-reed-died-cancer-patient-who-fought-morcellation-procedure.html Retrieved on 3/31/2021)
5 Source: https://www.wbur.org/hereandnow/2013/12/18/fibroid-removal-cancer (Retrieved on 3/31/2021.)
6 Link to the FDA safety communication: http://wayback.archive-it.org/7993/20170722215731/https://www.fda.gov/MedicalDevices/Safety/AlertsandNotices/ucm393576.htm (Retrieved on 3/31/2021.)
7 Cited from the FDA Safety communication: "Based on an FDA analysis of currently available data, it is estimated that 1 in 350 women undergoing hysterectomy or myomectomy for the treatment of fibroids is found to have an unsuspected uterine sarcoma."

announcement because it was triggered by an event rather than the systematic accumulation of scientific knowledge; in fact, no adverse event was ever reported to the FDA before the petition was filled.[8] Second, the FDA safety communication was *not* accompanied by a product recall. Instead, it served to reduce the perceived effectiveness of using power morcellation, which up to that point was commonly used during MI hysterectomies, and thus lowered the overall attractiveness of MI hysterectomies. Third, the alternative procedure, open hysterectomy, is not risk-free. It is associated with relatively higher postoperative complications and has been shown to be sensitive to physician skill. Siedhoff *et al.* (2015) modeled outcomes using a decision tree analysis and found that the risk of death related to sarcoma after MI hysterectomy with morcellation was fully balanced by procedure-related complications (including deaths) associated with open hysterectomy; put differently, depending on surgeon competency with open hysterectomy, MI hysterectomy with morcellation may result in more quality-adjusted life years compared with open hysterectomy. Not surprisingly, a *Wall Street Journal* article documented that physicians nationwide were still using MI hysterectomy with morcellation months after the FDA warning (Levitz & Kamp 2014). These facts suggest that heterogeneity in physician compliance decisions may be linked to their relative skill. Finally, we note that our study only focuses on the short-term response (i.e., 6 calendar quarters) by physicians; during this time period, technology innovation in response to the FDA announcement, such as containing power morcellation within a bag when performing MI hysterectomy, was still quite experimental and did not happen systematically.

Consistent with studies of many different medical technologies that were abandoned as a result of a negative information shock (Shah *et al.* 2010; Howard *et al.* 2016; Howard *et al.* 2017; Staats *et al.* 2018), the decrease in the share of MI hysterectomies and the corresponding increase in the share of open hysterectomies following the FDA safety communication is well documented in the medical literature (Barron *et al.* 2015; Harris *et al.* 2016; Wright *et al.* 2016; Multinu *et al.* 2018). However, these studies do not provide an explanation for the high degree of heterogeneity in physicians' responses.

[8] Source: https://www.medicaldesignandoutsourcing.com/power-morcellation-lingering-questions/#:~:text=That%20all%20changed%20in%202013,at%20nearby%20Brigham%20%26%20Women's%20Hospital. (Retrieved on 3/31/2021.)

3. THEORETICAL FRAMEWORK

In this section, we develop a theoretical framework for understanding the role of technology-specific skills in determining procedural mix and in altering it under an information shock. We consider two alternative procedures for hysterectomy, the first is a Minimally Invasive (*MI*) procedure and the second is an Open (*O*) procedure. The two procedures may vary in their perceived effectiveness, their appropriateness for different patients, and the levels of physicians' performance (i.e., skill). We begin by introducing notations for these three elements.

Physicians vary in their procedure-specific skill levels. Specifically, the skill of a physician *j* is denoted by a tuple $\{s_{MI,j}, s_{O,j}\} \in [0,1]^2$, where s_{MI} is the skill of the physician on the *MI* procedure and s_O is the skill of the physician in the *O* procedure. $s_{ij} = 0$ corresponds to the lowest level of skill and $s_{ij} = 1$ corresponds to the highest level of skill on each procedure ($i \in \{MI, O\}$). Physicians are differentiated based on their skill levels and are aware of their skill levels.

The perceived effectiveness of a procedure is publicly known and is subject to updating over time as new research is made available. This information is common to all physicians. Effectiveness of procedure $i \in \{MI, O\}$ is given by $e_i \in [0,1]$. That is, the effectiveness of the *MI* procedure is given by e_{MI} and the effectiveness of the *O* procedure is given by e_O. New information (shocks) may alter the value assigned to procedure effectiveness.

Each physician faces a patient population with varying degree of appropriateness for the two procedures. Assume a patient is diagnosed and is assigned a number $k \in [0,1]$. This number represents the suitability of the patient to a specific procedure. $k = 0$ corresponds to the case where the patient is perfectly suitable for the *MI* procedure, and $k = 1$ corresponds to the case where the patient is perfectly suitable for the *O* procedure. Intermediate values of *k* may reflect a degree of uncertainty about the suitability of either procedure. More specifically, a value $k \in [0,1]$ tells us that the MI procedure is more suitable than the O procedure with probability $(1 - k)$. Note that *k* is independent of the physician's skill level as well as the effectiveness of each procedure.

The procedure chosen by physician *j* depends on the diagnosis *k*, the skill tuple of the physician, $s_{ij} = \{s_{MI,j,}s_{O,j}\}$, and the known effectiveness of each procedure, $e = \{e_{MI,}e_O\}$. The physician's choice of procedure *i* is represented by $a_{MI,j} = 0$ if the *MI* procedure is chosen and by $a_{O,j} = 1$ if the *O* procedure is chosen. A physician makes the choice of a procedure by maximizing her utility function given by:

$$U_j(k, s_{ij}, e_i, a_{ij}) = -\frac{f(a_{ij}, k)}{g(s_{ij}, e_i)} \ \ for\ i \in \{MI, O\}$$

Note that the diagnosis *k* can take any number in the interval [0,1] while the action of the physician is dichotomous (i.e., it can be either $a_{MI,j} = 0$ or $a_{O,j} = 1$). Note that the highest level of utility is zero. The numerator, $f(a_{ij}, k)$, is a penalty function that measures the magnitude of mismatch between the procedure chosen and the appropriateness of that procedure to the patient. The denominator, $g(s_{ij}, e_i)$, captures the scaling effect of the mismatch. That is, the more effective the chosen treatment is in general and the more skilled the physician is in performing the chosen procedure, the lower is the negative effect of mismatch on the physician's utility. More specifically, for simplicity, consider the following functional forms for *f* and *g*: $f(a_{ij}, k) = (|a_{ij} - k|)^p$ and $(s_{ij}, e_i) = (s_{ji} + e_i)^q, p, q \in \mathbb{R}_+$[9] . Higher values of *p* correspond to higher levels of loss from the mismatch, while higher values of *q* correspond to a lower impact of the mismatch. This formulation assumes perfect substitutability between the physician's procedure-specific skill and the overall effectiveness of the procedure. Using the specific functional form, the physician's utility function is given by

$$U_j(k, s_{ij}, e_i, a_{ij}) = -\frac{(|a_{ij} - k|)^p}{(s_{ij} + e_i)^q} \ \ for\ i \in \{MI, O\}$$

[9] It is common in literature to use p = 2 which is referred as the quadratic loss function; however, we need not restrict ourselves to this specific case.

Note that higher levels of procedure-specific skill (s_{ij}) and procedure overall effectiveness (e_i) lower the penalty at any level of diagnosis k. Physicians select the procedure that maximizes their utility, or in our formulation, minimizes the loss from the mismatch. This is represented using the following value function, V_j:

$$V_j(k, s_{ji}, e_i) = \max_{i \in \{MI,O\}} -\frac{(|a_{ij} - k|)^p}{(s_{ij} + e_i)^q}$$

Note that for a given set of skill s and effectiveness e, the physician is indifferent between the two procedures at diagnosis level $\bar{k}$ if

$$-\frac{(|a_{MI,j} - \bar{k}|)^p}{(s_{MI,j} + e_{MI})^q} = -\frac{(|a_{O,j} - \bar{k}|)^p}{(s_{O,j} + e_O)^q}$$

Since $a_{MI,j} = 0$ and $a_{O,j} = 1$ we have,

$$-\frac{(|0 - \bar{k}|)^p}{(s_{MI,j} + e_{MI})^q} = -\frac{(|1 - \bar{k}|)^p}{(s_{O,j} + e_O)^q}$$

which gives us

$$\left(\frac{\bar{k}}{1 - \bar{k}}\right) = \left(\frac{s_{MI,j} + e_{MI}}{s_{O,j} + e_O}\right)^{q/p}$$

Defining $q/p = n$ and solving for $\bar{k}$ in terms of other parameters we get

$\bar{k} = \frac{(s_{MI,j}+e_{MI})^n}{(s_{MI,j}+e_{MI})^n+(s_{O,j}+e_O)^n}$ (1)

Note that $\bar{k}$ is the clinical appropriateness of the marginal patient. It also represents the proportion of *MI* procedures conditional on effectiveness levels e_{MI} and e_O and physician *j*'s procedure-specific skill levels $s_{MI,j}$ and $s_{O,j}$. If the effectiveness of both the procedures is identical $(e_{MI} = e_O)$ and the procedure-specific skill of physician *j* is identical for both procedures $\left(s_{MI,j} = s_{O,j}\right)$ then $\bar{k} = \frac{1}{2}$. Put differently, if the physician's skill and the procedure effectiveness are identical across procedures, there should be no bias in procedure choice. That is, *MI* is chosen for $k \leq \bar{k} = \frac{1}{2}$ and *O* is chosen for $k \geq \bar{k} = \frac{1}{2}$.

Note that the cutoff $\bar{k}$ is decreasing in s_O and e_O, and increasing in $s_{MI,j}$ and e_{MI} for any positive value of *n*. This has a number of implications. First, physicians with higher relative skill (favoring *MI* over *O*), will have a higher cutoff $\bar{k}$. In other words, these physicians will choose *MI* for patients appropriate for open surgery (i.e. high *k* patients). Secondly, the relative effectiveness of *MI* versus *O* will affect the cutoff in a similar way. Note that relative effectiveness is not physician-specific, yet it is affected by external information shocks. Negative information shocks related to the effectiveness of *MI* procedures, the central focus of our research, would reduce e_{MI}/e_O and hence lower the cutoff $\bar{k}$.

The notion of minimally invasive technology abandonment in this model is measured by reductions in $\bar{k}$, the proportion of physician *j*'s patients treated using a MI procedure. Put differently, technology abandonment need not be a binary choice for a physician over their entire patient population, but rather measures how the use of a specific procedure decreases in response to the negative information. We turn now to an analysis of the relationship between procedure-specific skill levels and the intensity of *MI* abandonment.

The rate of change of the threshold with respect to a small change in the effectiveness of the minimally invasive procedure e_{MI} is given by:

$$\frac{d\bar{k}}{de_{MI}} = \frac{n(s_{MI}+e_{MI})^{n-1}(s_O+e_O)^n}{[(s_{MI}+e_{MI})^n+(s_O+e_O)^n]^2} \quad (2)$$

The elasticity of the threshold $\bar{k}$ with respect to *MI* effectiveness is given by

$$\xi_{\bar{k},MI} = \frac{e_{MI}}{\bar{k}} \cdot \frac{d\bar{k}}{de_{MI}} = n \cdot \left(\frac{e_{MI}}{s_{MI}+e_{MI}}\right) \cdot \frac{(s_O+e_O)^n}{(s_{MI}+e_{MI})^n+(s_O+e_O)^n} \quad (3)$$

For ease of interpretation, this expression can be written as follows

$$\xi_{\bar{k},MI} = \frac{q}{p} \cdot \left(\frac{e_{MI}}{s_{MI}+e_{MI}}\right) \cdot \left(1-\bar{k}\right) \quad (4)$$

The elasticity is a product of three expressions. The first term, $\frac{q}{p}$ or *n*, relates the elasticity of abandonment to the magnitude of the loss function. *n* can be thought of as measuring the risk tolerance of the physician, since lower values of *n* correspond to higher levels of risk aversion. The second term, $\frac{e_{MI}}{s_{MI}+e_{MI}}$, measures the relative role of clinical effectiveness set against the physician's skill level. If procedural skill is low or does not matter, this term will approach 1.[10] Note that as minimally invasive procedure skill increases (holding open procedure skill constant), the elasticity of abandonment becomes smaller in magnitude, suggesting that physicians who perform relatively better on *MI* compared with *O* procedures, exhibit a weaker propensity to shift a larger proportion of their patients away from *MI* procedures when faced with a negative information shock. Finally, the third term, $(1-\bar{k})$, represents the proportion of patients who are treated using open surgery. The greater this proportion is, the higher is the elasticity of abandonment with respect to new information on reduced *MI* effectiveness.

Our model has a number of empirical implications. First, Equation (1) suggests that identical patients are more likely to receive MI if treated by a physician with higher relative skill. This is a natural result from the model: patients are sorted into different types of treatments based on the returns to each type.[11] It also implies that the marginal patient treated by a physician with higher relative skill is more appropriate for O, and less appropriate for MI. We provide empirical support of this implication in Section 5.1.

10 For example, taking prescription drugs does not require procedural skill.
11 Chandra & Staiger (2007) show a similar result: Surgical intensive areas have better quality of intensive care, and worse quality of medical management.

Second, Equation (2) shows that $\frac{d\bar{k}}{de_{MI}} > 0$, which suggests that a negative information shock on the effectiveness of MI will cause physicians to decrease the cutoff $\bar{k}$, i.e., the proportion of patients receiving MI. This means that on average, physicians will abandon MI and switch to O in response to the information shock. Since $\bar{k}$ also represents the clinical appropriateness of using the open procedure for the marginal patient, this finding suggests that physicians are expected to cut back on MI procedures for patients who are less appropriate for MI (or more appropriate for O) in response to the information shock. We test this prediction in Section 6.4.

Third, Equation (3) shows our main hypothesis: physicians with higher relative skills are less sensitive to the information shock, and are less likely to abandon MI and switch to O. The results from testing this main hypothesis are shown in Section 6.1.

4. DATA AND MEASURES

Our empirical setting focuses on the choice that physicians made between MI hysterectomy and open hysterectomy. There are a number of reasons why this is an ideal setting for testing our hypotheses. First, hysterectomy is the most common major gynecologic surgery—approximately 600,000 women undergo hysterectomies in the United States each year. Second, minimally invasive and open procedures require different skill sets.[12] Expertise and experience from performing one procedure are not directly transferable to the other (Rogers *et al.* 2001). Third, the vast majority of physicians acquire the expertise needed to perform both procedures. The need to command both skills is driven by heterogeneity in patient clinical appropriateness for a particular procedure. This allows us to observe and measure each physician's procedure-specific skill, which subsequently allows for the construction of a relative skill measure. Fourth, the unanticipated announcement of the FDA warning exogenously lowers the perceived attractiveness of MI hysterectomies for all physicians, allowing us to evaluate physician response to information shock.

[12] For example, minimally invasive hysterectomies are guided remotely through videos, and therefore require an extra level of spatial ability and perceptual motor skills (Silvennoinen *et al.* 2009).

Our data include hospital inpatient discharges and outpatient visits for all patients who received either minimally invasive or open hysterectomies between January 2012 and September 2015 in Florida.[13] The data contains the license number of each operating physician, a hospital identifier, a rich set of procedure codes and diagnostic codes for conditions present on admission, indicators for postoperative complications, as well as patient demographic characteristics such as age, race, and payer type.

4.1 *Measuring Patient MI Appropriateness*

To construct a measure of patient appropriateness for a MI hysterectomy (*MI appropriateness*), we estimate a logistic regression model of the probability of receiving MI hysterectomy for patient i in year-quarter *t* in the years *prior to* the information shock (Equation 5).[14]

$$\Pr(\mathrm{MI}_{it}) = G(X_{it}\Phi + \varepsilon_{it}) \ (5).$$

In Equation (5), X_{it} is a vector of patient characteristics. Following Harris *et al.* (2016), we control for a set of medical risks recorded for conditions present on admission, including the Charlson Comorbidity Index (CCI) (Charlson *et al.* 1987),[15] severe pelvic adhesion, morbid obesity, cervical dysplasia, and a set of indicator variables for top principal diagnoses. Other patient characteristics include age and its square, race, the total number of other diagnoses for conditions on admission, whether the patient is admitted under an emergent situation, and the type of insurance. The estimates are presented in Appendix Table A1.[16] Patient *MI appropriateness* is then measured using fitted values from the logit regression: $MI\ appropriateness_{it} = \hat{G}(X_{it}\Phi)$.

---

[13] The receipt of MI or open hysterectomy is defined using the principal procedure ICD-9 code. The data switched from reporting ICD-9 to ICD-10 starting from the fourth quarter of 2015. To minimize measurement error by construction, we exclude 2015 Q4 from the sample.

[14] Results are quantitatively similar when we use a machine learning technique, the penalized logistic regression, to more flexibly estimate Equation (5) in order to capture potential patient sorting based on observable characteristics.

[15] The Charlson Comorbidity Index is often used to describe the clinical severity of cardiopulmonary diseases. Cardiopulmonary disease is a major contraindication for MI hysterectomy, because MI procedures require pneumoperitoneum and/or trendelenburg positioning, which cardiopulmonary patients cannot tolerate.

[16] The model fits the data well, with a pseudo R-squared of 0.29.

Figure 1 plots the distribution of *MI appropriateness* for those who received MI hysterectomies and open hysterectomies, respectively. The figure shows that among patients who received MI hysterectomies, density concentrated at appropriateness levels between 0.73 and 0.95; among patients who received open hysterectomies, the distribution is more even, but the density is sparse for appropriateness above 0.8. We, therefore, define a patient to be *appropriate for MI* if the MI appropriateness is above 0.73 (among whom 91% received MI hysterectomies), and *appropriate for open* if the MI appropriateness is below 0.4 (among whom 78% received open hysterectomies).[17]

4.2 *Physician Skill Measures and Sample Statistics*

To construct empirical measures of physician skill, we first follow Harris *et al.* (2016) in defining a negative outcome as at least one major postoperative complication.[18] Appendix Table A2 presents the proportion of patients with no postoperative complications separately for patients who received open and MI hysterectomies, both unadjusted and adjusted by patient characteristics. As expected, open hysterectomy is associated with a lower probability of postoperative non-complications compared to MI procedures, with risk adjustment shrinking the difference in non-complication rates between the two procedures, as open hysterectomies are typically performed on higher-risk patients.

Next, we use two different methods to construct separate measures of physicians' *absolute* skill in performing MI and open hysterectomy (defined as "MI skill" and "open skill" hereafter), as well as *relative* skill in MI versus open.

Preferred Measure: The major challenge for measuring physician skill in a specific procedure is that the choice of procedure is potentially endogenous to physician skill and patient characteristics. To address this concern, we follow Currie & MacLeod (2017) and measure each physician's MI and open skill using the proportion of patients who do not have any postoperative complications

[17] Results are robust when we use alternative thresholds, including MI appropriateness above 0.7 or 0.8 for "*appropriate for MI*", and below 0.3 or 0.5 for "*appropriate for open.*"

[18] These complications include blood transfusions, vaginal cuff infection, vaginal cuff dehiscence, ureteral obstruction, vesicovaginal fistula, deep and organ space surgical site infection, acute renal failure, respiratory failure, sepsis, pulmonary embolism, deep vein thrombosis requiring therapy, cerebral vascular accident, and cardiac arrest. We also consider death in this category.

among his or her patients who are predicted to be appropriate for each procedure (defined in Section 4.1). Similar to Currie & MacLeod (2017), the relative skill of MI versus open is calculated as the difference between MI skill and open skill. This is our preferred measure of skill because the appropriateness measure only depends on the underlying risk factors, and is unlikely to suffer from the endogeneity of procedure choice.

Alternative Measure: For a robustness check, we adopt a second measure of physician skill. We define the MI and the open skill to be the risk-adjusted non-complication rate calculated among patients who actually received the procedure. Relative skill is again measured by the difference between MI and open skill. The method for constructing this measure is detailed in Appendix II.

Our sample contains patients who received hysterectomies between January 2012 and September 2015 in Florida.[19] In order to compute these two relative skill measures, we focus on patients whose physicians have treated at least one patient deemed appropriate for MI and at least one patient deemed appropriate for open (as defined in Section 4.1), and have performed at least one MI and one open hysterectomy, in both the pre- and the post-information shock periods. Next, as will be discussed in Section 5.3, we constructed a leave-out instrumental variable where each physician's performance is instrumented by all other physicians who practice in the same county and same type of hospital. Therefore, we have to restrict the sample to physicians who practice at hospitals with "same type" peer hospitals in the county. Furthermore, since physicians with few patients appropriate for the procedure (either MI or open) will have noisy skill measures for that procedure, we exclude patients treated by physicians whose number of patients appropriate for either type of procedure is in the bottom quartile (i.e., less than 7 MI-appropriate patients or 4 open-appropriate patients) in the pre-period. Our preferred main sample includes 23,461 patients treated by 215 physicians, spanning eight calendar quarters before (January 2012-December 2013) and six calendar quarters after (April 2014-September 2015) the information shock.[20] To further reduce the noisiness of the skill measure, we also use a restrictive subsample of patients treated by physicians whose number of MI-

[19] We focus on patients with benign gynecologic indications following the medical literature on the evaluation of practice patterns and outcomes in response to the FDA warning (Multinu et al. 2018).

[20] The original sample includes 30,254 patients treated by 394 physicians. We show in Appendix Table A3 the distribution of the number of open-appropriate and MI-appropriate patients among the original sample.

appropriate and open-appropriate patients in the pre-period are both greater than 10.[21] We exclude the interim period, 2014 Q1, from the sample used in the difference-in-differences analysis, because it is after the national news release (December 2013) and before the formal release of the FDA warning (April 2014).[22]

Table 1 presents the summary statistics of the final main sample.[23] Both the preferred and alternative skill measures exhibit considerable variation across physicians, with a mean of 0.25 and a standard deviation of 0.22 for the preferred measure, and a mean of 0.23 and a standard deviation of 0.22 for the alternative measure.[24] The two relative skill measures are highly correlated (correlation coefficient is 0.66). When performing the empirical analysis, we normalize both skill measures using z-scores for ease of interpretation.

4.3 *Variation in Relative Skill and Relative Performance*

Do the observed variations in relative skill represent heterogeneity in human capital, or are they haphazardly driven by some random chance? In this section, we use a simulation method to verify that our relative skill measure captures systematic differences as opposed to random ones. The underpinning of the simulation is as follows.[25] In each iteration of the simulation, we randomly assign patients to physicians and calculate the implied relative difference in MI and open skill for each physician, i.e., the relative skill. We then calculate the variance of the distribution in relative skill. We run 1,000 iterations and report the following two results. First, we show the distribution of the simulated variances of relative skill and compare it with the observed variance using both the main sample and the subsample. As shown in Figure 2 Column 1, the observed variance is greater than the simulated ones. We can reject the null

[21] This subsample contains 14,052 patients treated by 76 physicians. Restricting to this subsample allows us to measure physician skill more precisely with the trade-off of sacrificing the sample size. As shown in Section 6, the results are highly robust when using different samples. We note that the relatively small sample could limit the statistical power especially in the event-study estimations.
[22] We included 2014 Q1 when conducting the event-study analysis as discussed in Section 6.2.
[23] Appendix Table A4 shows the summary statistics of the subsample of patients whose physicians have treated more than 10 MI-appropriate patients and more than 10 open-appropriate patients in the pre-period.
[24] The mean of relative skill is positive because complications are more frequent for patients appropriate for open (or who received open procedures) than those appropriate for MI (or who received MI procedures).
[25] We thank one anonymous reviewer for this suggestion.

hypothesis that the observed variance is equal to the simulated variance with p-value less than 0.001. Next, we compute the average of the 1,000 simulated skills for each physician and compare the distribution of the average simulated skill with that of the observed skill (Figure 2, Col 2). Again, the results show that the observed variation of relative skill is much greater than that of the simulated ones. This simulation result suggests that the observed skill variation is not driven by random chance; instead, it reflects the heterogeneity in relative abilities and human capital, which allows us to identify the effect in our empirical estimation.

5. EMPIRICAL ESTIMATION

5.1 Basic Patterns from the Raw Data

We begin by providing some initial evidence, based on the raw data, of our main question—the extent to which physicians with differential relative skill levels differ in their responses to the FDA warning by abandoning MI hysterectomies. Having constructed the physician skill measures in Section 4.2, we define physicians whose relative skill is above the median (weighted by patient volume) to be the *Top MI Performers*, those whose relative skill is below the median (weighted by patient volume) to be the *Top Open Performers*. We follow this definition in the remainder of the paper.

Figure 3 plots the quarterly trend of MI utilization (i.e., the percentage of MI hysterectomies) for patients treated by the *Top MI Performers* and *Top Open Performers*. There are several key observations. First, in the pre-FDA warning period, *Top MI Performers* are more likely to use MI hysterectomies than *Top Open Performers*, though the difference is small. This is consistent with the hypothesis that physicians who are more skilled in a specific procedure are more likely to use the procedure. To further establish the association between relative skill and MI utilization, we run a simple OLS model regressing whether the patient receives MI on patient characteristics and physician relative skill while controlling for county-fixed effects and hospital-fixed effects, respectively. Consistent with the graphs, the regression results shown in Appendix Table A5 confirm that conditional on patient characteristics and location or institution, physicians with higher relative skill (MI relative to open) are more likely to use the MI procedure (Equation 1).

The key observation from Figure 3 is that after the FDA warning, *Top Open Performers* are more likely to abandon MI, substantially increasing the utilization gap between the two groups of physicians. This observation is consistent with our main hypothesis that physician skill explains how they respond to negative information shock and abandon MI (Equation 3). One may wonder if the trend in Figure 3 is driven by patient selection over time, i.e., if patients who are more appropriate for MI are more likely to choose *Top MI Performers* after the FDA warning. To address this concern, we show in Table 2 the average patient MI appropriateness, the unadjusted and adjusted (by patient MI appropriateness) MI utilization by *Top MI Performers* and *Top Open Performers*, before and after the FDA warning, respectively. There are several important insights to take from Table 2. First, for both types of physicians, *Top MI Performers,* and *Top Open Performers*, we do not observe any statistically significant change in the average *MI appropriateness* after the information shock. This indicates that the extent to which patient sorting occurs (i.e., patients who are more appropriate for MI are sorted to physicians who are relatively better at performing MI) is minor, at least based on observables. This result relieves our concern that the main results are driven by patient sorting. Second, a comparison between Columns (3) and (6) in Table 2 indicates that consistent with our prediction, *Top Open Performers* are more likely to abandon MI after the FDA warning, even after we adjust for *MI appropriateness*.

5.2 Identifying the Effect of Relative Performance on Technology Abandonment

In this section, we aim to assess the extent to which relative skill affects physician technology abandonment in response to negative information shock. Equation 3 predicts that physicians with higher relative skill (MI relative to open) are less likely to abandon MI hysterectomy. To determine the magnitude of this effect, we conduct a within-physician analysis by employing the following linear probability model.

$$MI_{ijkt} = a + b_1 Skill_j \times Post_t + b_2 Z_j \times Post_t + b_3 X_{it} + b_4 Physician_j + b_5 Time_t + e_{it} \quad (7),$$

For a patient i who receives a hysterectomy performed by physician j in county k and year-quarter t, $MI_{ijkt}$is an indicator variable which is equal to one if the patient receives a MI hysterectomy, and zero if the patient receives an open hysterectomy. $Skill_j$ represents the physician's relative skill (MI relative to open). X_{it} are the observed patient characteristics. $Post_t$ indicates whether the patient receives the hysterectomy after the FDA warning. We include physician fixed effects $Physician_j$ to control for unobserved time-invariant physician-specific characteristics that may affect the utilization of MI, and the year-quarter fixed effects, $Time_t$, to control for seasonal and other longer-term trends of MI utilization. e_{it} is a random error term. The coefficient of interest is b_1, which is expected to be positive for a physician with higher relative skill.

For robustness, we also include an interaction term between physician characteristics other than skill that may also affect technology abandonment decisions and the post-time dummy, $Z_j \times Post_t$. Specifically, we use the following measures for Z_j. First, we include the physician's average MI share in the pre-period (i.e., 2012-2013) to eliminate mean reversion and control for the potentially mechanical relationship between a physician's baseline MI share and the implied impact of the information shock. We also consider a physician's MI volume (i.e., the number of MI hysterectomies performed by the physician) in the pre-period to be a potential factor influencing physician response to negative news regarding MI procedure (Staats *et al.* 2018).

The OLS estimation of Equations (7) may lead to biased estimates due to patient-physician sorting and measurement error with regards to physician skills. For example, if there is dynamic patient sorting—i.e., patients who are more appropriate for MI (in unobserved ways) are more likely to see physicians with higher relative skill in MI over open—the observed trend of MI abandonment by physicians with higher relative skill could be explained by the unobserved patient characteristics that changed over time.

Although the results shown in Table 2 suggest that the magnitude of time-varying selection bias in response to the FDA warning, if it exists, is rather small, we strive to mitigate further identification concerns using a leave-out instrumental variable (IV) strategy. Specifically, we instrument for each physician's relative skill using *the weighted average of the relative skill of all physicians who practice in the same county and type of hospital, excluding the focal hospital*. We focus on three attributes of hospitals to define “type”—an indicator variable for teaching status, an indicator variable for hospitals with more than 500 beds, and an indicator variable for rural hospitals—leading to 8 “type” bins. Hospitals are defined to be of the same type if they fall in the same bin (i.e., are identical on all three attributes). For each index physician who works in a hospital, we then focus on all the other hospitals of the same type in the same county and calculate the weighted average of performance for physicians who work in these hospitals, where the weights are given by the total number of hysterectomies performed by each *within*-county *within*-type peer physician.[26]

The aforementioned measure is a valid instrument because it is likely to explain physicians’ relative skill and is likely to be exogenous. The first assumption is justified by the abundant evidence on peer effects and observational (social) learning effects in the medical profession. For example, studies have shown practice style and outcomes to be tantamount *within* geographic areas (Chandra & Staiger 2007; Agha & Molitor 2018) and *within* hospitals (Berez *et al.* 2018; Staats *et al.* 2018). Therefore, we expect that the performance of peer physicians who practice at similar hospitals in the same county is a good predictor of the index physician’s relative skill. This statement is verified by the highly significant first-stage estimations.

The exclusion restriction of our IV requires that the instrument is uncorrelated with unobserved characteristics of patients treated by the index physician. This is a valid assumption because when constructing the IV, we exclude the index physician’s colleagues at the focal hospital and use only peer physicians who practice at the other similar hospitals in the same county. This is a crucial step because it is possible that within the focal hospital, patients who are more appropriate for a procedure are matched to physicians who are relatively better at performing the procedure, and thus the performance of the index physicians’ colleagues at the focal hospital might be correlated with unobserved characteristics of patients treated by the index physician. This is not likely to be the case across hospitals within the same region. Furthermore, there is no public information channel through which patients or referring primary care physicians can learn about providers’ relative performance. Therefore, it is unlikely that patients are sorted to physicians or hospitals based on unobserved patient characteristics in response to the FDA warning.

[26] Note that this measure is not feasible if there is only one hospital of the same type in a county. We thus exclude physicians in these hospitals from our sample.

To validate the use of the instrumental variable, we present the first stage results in Column (4) of Table 3. The highly significant first-stage estimations indicate that, for a given physician, other physicians' relative skills at the hospital are good predictors of their relative skill. The F-statistics of the excluded instrument in the first stage are above 10 (Table 3, Col 5-7), indicating that the instrument generates adequate variation for the identification of the model (Staiger & Stock 1997).

6. RESULTS AND ROBUSTNESS

6.1 Baseline Results—Impact of Relative Skill and Absolute Skill on Technology Abandonment

Our main results from the baseline estimation model (Equation 7) are presented in Table 3. We show how physicians respond to the negative information shock based on their relative skill using both the main sample and the restricted subsample excluding low-volume physicians (Panel 1 and 2 in Table 3). Both the OLS and two-stage least squares (2SLS) results are highly significant. The results are robust with the inclusion of the share of patients who received MI procedures and the number of patients who received MI procedures in the pre-period. The 2SLS IV regression results from the main sample suggest that a one standard deviation increase in relative skill of MI over open would reduce the probability of MI abandonment by 4.6-4.9 percentage points (a 6.2%-6.5% reduction from the pre-period level), depending on the specification (Col 5-7 in Panel 1 of Table 3). The magnitude becomes larger when we focus on the subsample excluding patients treated by low-volume physicians (Col 5-7 in Panel 2 of Table 3).[27] These findings confirm the key hypothesis derived from Equation 3, suggesting that relative skill is an important determinant of technology abandonment in response to negative information shock.[28]

Next, we explore the role of the absolute skill (i.e., MI skill and open skill) in affecting procedure choice and abandonment decision in response to the negative information shock. To do so, we substitute the absolute skill measures of performing MI or open

[27] We expect the skill measures to be less noisy when using the subsample excluding low-volume physicians.
[28] Since the dependent variable is an indicator variable, we conduct a robustness test using *Probit* regressions instead of linear probability models. The results remain highly significant and are quantitatively similar to that of the main results.

hysterectomy for the relative skill measure in Equation (7) and re-estimate the model. The estimations have the expected signs (Appendix Table A6): physicians with higher MI (open) skill are less (more) likely to abandon MI and switch to open. But the coefficients are quantitatively smaller than those on the relative skill (as shown in Table 3) and some are insignificant. These results indicate the role of relative skill is more important than absolute skill in determining technology abandonment decisions.

6.2 Validating the Difference-in-Differences Design

A crucial assumption of the difference-in-differences estimation in Equation 7 requires that the use of MI procedures would have moved in parallel in the post-period for both high- and low-relative skill physicians had the FDA announcement not occurred. We validate this assumption using the following analyses. The first analysis is an event-study estimation—we replace the "post" dummy variable in Equation 7 with a set of event-time (quarterly) indicators, omitting the quarter before the information shock (the 4th quarter of 2013). We show the estimated coefficients and 95% confidence intervals of the interaction terms (i.e., the interactions between physicians' relative skill and the quarterly indicator variable) in Figure 4 for both the main sample and the subsample using both OLS and 2SLS IV estimations. As shown in the figure, no statistically significant pre-trends are detected. The post-period results suggest a one-time change in practice pattern in response to the information shock.[29]

Next, to rule out the possibility that the difference-in-differences results could reflect changes in volume or composition of patients across physicians of different relative skill levels, we conduct two falsification tests by re-estimating the difference-in-differences specification with patient volume and patient MI-appropriateness as the outcome variables. First, we regress the patient MI-appropriateness estimated from Section 4.1 on the interaction between the physician relative skill and the post-time dummy variable, controlling for physician fixed effects and year-quarter fixed effects (Table 4, Col 1). As robustness checks, we also include

[29] We conduct a robustness test of the event-study estimation in Appendix Figure A1, where we treat the 4th quarter of 2013 as the starting point of the information shock (i.e., time period t). We conduct this robustness test because in December 2013, the FDA petition was filed and national news started to report on the incidence (link: https://www.wbur.org/hereandnow/2013/12/18/fibroid-removal-cancer).

the interaction terms between the share of MI procedure, or the number of MI patients, in the pre-period and the post-time dummy (Table 4, Col 2 and 3). The insignificant results suggest there is no significant change in patients' MI-appropriateness treated by physicians of different relative skill levels after the information shock. Second, we regress the number of patients at the physician-hospital-year-quarter level on the same interaction terms, while controlling for physician fixed effects and year-quarter fixed effects. Results shown in Col 4-6 in Table 4 suggest there is no significant change in patient volume treated by each type of physicians after the information shock.

6.3 Robustness Tests—Controlling for the Role of Hospitals, Alternative Skill Measures, and Alternative Samples

In this section, we conduct a number of analyses to show our results are robust to the specification of the model, the measure of physician skill, and the selection of the sample.

The first test adds hospital fixed effects in the main analysis in Equation 7 to capture potential differences in technology, quality, management, and organizational structure across hospitals. For example, some hospitals may be better equipped to perform MI procedures than others, and may have different rules, guidelines, or managerial decisions regarding the extent to which they comply institutionally with FDA warnings. Because hospital fixed effects are collinear with our physician-hospital level time-invariant instrumental variable (i.e., the performance of physicians at the "same type" hospitals in the same county), we only conduct OLS regressions for this analysis. The effects are now identified through within-hospital variation in relative skill across physicians. The highly significant results shown in Table 5 suggest that after controlling for potential impacts of hospitals, physicians' relative skill is still an important determinant of the heterogeneity in their procedure choice and abandonment decisions within each hospital.

The second test aims to show the robustness of our results when using alternative measures for physician relative skill. First, as explained in section 4.2, our preferred skill measure relies on defining patients for their *MI appropriateness*, predicted from patient characteristics only (Equation 5). To address the concern that patients of different characteristics might be sorted to physicians of a specific relative skill level, and thus the possibility that our empirical measure of patient *MI appropriateness* might indeed be influenced by physician skill, we add physician fixed effects in the prediction model to control for potential patient sorting (Equation 8).

$\Pr(\mathrm{MI}_{ijt}) = G(X_{it}\Phi + Physician_j + \varepsilon_{ijt})$ (8).

Patient *MI appropriateness* is then measured using the fitted values from the estimated coefficients in Equation 6 for only the patient characteristic variables: $MI\ appropriateness_{it} = \hat{G}(X_{it}\Phi)$. The distribution of this new measure is very similar to that of the original one, where density concentrated at appropriateness levels above 0.72 (Appendix Figure A2). We thus define a patient to be appropriate for MI if the new measure of *MI Appropriateness* is above 0.72 (among whom 86% received MI hysterectomy) and appropriate for open if below 0.4 (among whom 80% received MI hysterectomy). The results using this new measure of physician skill are similar to our main results, suggesting the robustness of our results (Appendix Table A7).

Our second alternative skill measure is based on the risk-adjusted non-complication rate for patients who received each type of procedure. (See Appendix II for the construction of the alternative measure.) The purpose of this robustness check is to relieve further concern that our preferred measure may capture physicians' diagnostic ability in addition to their procedural skill. Results shown in Appendix Table A8 are similar to our main findings, suggesting robustness to the measure of physician skill.

The final robustness test uses an alternative sample—the whole sample of patients without dropping those who are treated by physicians in the bottom quartile in terms of their number of MI-appropriate patients and open-appropriate patients in the pre-period. As expected, the results shown in Appendix Table A9 are smaller in magnitude compared with the main results shown in Table 3. This is because the skill measures when using the whole sample are much noisier. However, even with the noisy skill measures, results remain highly significant.

6.4 Extensions – Heterogeneity by Patient MI-Appropriateness and the Extensive Margin

In this section, we conduct two additional analyses to shed some light on the potential implications of patient welfare. The first analysis focuses on the heterogeneity in physician response by patient type (MI-appropriateness). Previously, we have shown that physicians are less likely to use MI hysterectomies after the information shock, and this is true even controlling for patient characteristics (see Figure 3 and Table 2). However, it remains unclear what type of patients experience more scaling back of MI hysterectomies.

Our theoretical model predicts that conditional on physician skill, negative information shock on the effectiveness of MI should cause physicians to modify the procedure-specific appropriateness cut-off and hence refrain from offering MI hysterectomies to patients with low MI-appropriateness (Equation 2). It is crucial to verify this empirically because understanding the characteristics of the marginal patients is important for welfare implications—the information shock would lead to higher patient welfare if physicians were more likely to decrease the use of MI hysterectomies for patients who have relatively lower appropriateness for MI procedures. To analyze the heterogeneity in physician response by patient type, we ran the following regression:

$$MI_{ijkt} = a + b_1 MIAppropriateness_i \times Post_t + b_2 Z_j \times Post_t + b_3 MIAppropriateness_i + b_4 Physician_j + b_5 Time_t + e_{it} \quad (8).$$

In Equation 8, $MIAppropriateness_i$ is the MI appropriateness measure of patient i we have estimated from Equation 5. We regress the use of MI on patient MI-appropriateness $MIAppropriateness_i$ and the interaction between patient MI Appropriateness and the post-time dummy $MIAppropriateness_i \times Post_t$, while controlling for physician fixed effects $Physician_j$ and year-quarter fixed effects $Time_t$. As robustness tests, we also control for the interaction between the physician's pre-period MI share or MI patient volume and the post-time dummy $Z_j \times Post_t$. If physicians were more likely to reduce the use of MI hysterectomies for MI-inappropriate patients after the information shock, we would expect the estimated coefficient of b_1 to be positive. Results shown in Appendix Table A10 confirm this—a one standard deviation decrease in patient MI appropriateness is associated with a 3.7-4.3 percentage point (5%-5.8% of the pre-period level) higher probability in the abandonment of MI hysterectomies in response to the information shock.

The second analysis focuses on the extensive margin of the effect of the information shock on physician practice. In our main analysis, we have examined the intensive margin effect, i.e., conditional on receiving hysterectomies, how physicians' relative skill affected the mix of procedures they have offered their patients. As an extension of the main analysis, we ask how the extensive margin—the total number of hysterectomies—was affected by the information shock. In particular, we are interested in documenting whether the total number of hysterectomies decreased after the FDA warning, and how physicians' relative skill is associated with such volume reduction. We expect this margin to be driven by both demand-side contraction (e.g., patients may stop seeking care after viewing the FDA warning in the news) and supply-side declines (e.g., physicians might be more inclined to suggest watchful waiting as an alternative to any elective, non-urgent intervention).

Appendix Table A11 presents the number of patients who receive hysterectomies before and after the information shock. Panel 1 shows that the total number of patients decreased by 14.1% on average after the FDA announcement, suggesting an extensive margin response to the information shock. Furthermore, the decrease was more pronounced among *Top MI Performers* (quarterly volume decreased by 16.4%) than the *Top Open Performers* (quarterly volume decreased by 11.9%), indicating greater scope for and comfort in switching from MI to open procedures among *Top Open Performers*. This is supported by the results presented in Panels 2-4 of Appendix Table A11, which shows how differential declines across physician types affect the overall practice variation. Both types of physicians decreased the use of MI hysterectomies after the information shock (Panel 2); yet only the *Top Open Performers* switched to performing more open hysterectomies (Panel 3), which is consistent with the prediction of our theoretical model. As a result, the decrease in the share of MI hysterectomies was concentrated among *Top Open Performers* (Panel 4).

7. CONCLUSIONS AND DISCUSSIONS

This paper highlights the importance of relative skill in governing procedural choice and in altering these choices in response to negative information shock, leading to what is commonly referred to as technology abandonment. The heterogeneity in individuals' relative performance across substitutable tasks has received ample attention in the economics literature, but little or no attention in the literature on technology adoption and abandonment. Our paper presents a step in closing that gap by using a unique empirical setting, which allows us to observe individual performance as it relates to alternative tasks, and thus assess how relative performance interacts with the choice of procedural mix as well as how it is affected by informational shocks.

The underpinning of our analysis ties back to the basic relationship between scarcity and choice. The fundamental idea is that when a physician diagnoses a medical problem, oftentimes she will face a choice between mutually exclusive procedural alternatives, each requiring a set of specific skills and conferring specific costs and benefits. The physician's relative skill level is therefore tied to the cost and benefit associated with each procedure. In this context, negative information shocks serve to reduce the expected benefit of performing the minimally invasive procedures; this reduction in expected benefit is more pronounced for surgeons with higher relative skill at performing the alternative intervention, i.e., the open procedures, compared with the minimally invasive ones.

The choice of mutually exclusive procedural alternatives, each requiring a set of specific skills and conferring specific costs and benefits, is related to the fundamental notion of the division of labor introduced by Adam Smith's (1776). Subsequent theoretical and empirical work focused primarily on the role of interdependencies in production, such as task complementarities (Rosen 1983) or the cost of coordinating tasks across workers (Becker & Murphy 1992). A second stream of the literature suggested that task interdependencies can arise on the demand side, mainly through differential payment for tasks in a variety of industries and professions (MacDonald & Marx 2001). More specifically, the effect of demand interdependencies for physicians' division of labor within a medical specialty was shown to be important (Baumgardner 1988; David & Helmchen 2011). Regardless of why the division of labor might be incomplete, the command of multiple skills by a single individual indicates that the individual could adopt substitutable tasks represent alternative ways to achieve the same goal. For example, physicians often face a choice between different courses of medical treatment for a given condition, as is the case in this paper.

While the division of labor does not require individuals to be endowed with different task-specific skills, a large related literature relies on this notion of heterogeneity to explain task and occupational choices (Roy 1950; Miller 1984; Heckman & Honoré 1990; Chandra & Staiger 2007). This literature is primarily focused on overcoming the fact that once a choice of task is made by an individual, performance of the counterfactual task by that same individual is not observable and empirical identification relies on assumptions about the joint distribution of skills across tasks and individuals. The cases described above, in which division of labor is incomplete, provide a unique opportunity to observe the individual performance of alternative tasks and hence their joint distribution of task-specific skills (i.e., in the context of the Roy model, these cases allow us to observe the fisherman hunt and the hunter fish).

In addition, our analysis relates to the growing emphasis on the eradication of low-value care. The definition of what constitutes low-value care varies from a narrow one, where low-value care is simply care that confers no benefit or is associated with risk that is greater than the expected benefits to the patient, to a broader definition which encompasses care that could be avoided by substituting equally cost-effective or superior alternatives. Even with a very broad definition that includes alternative treatment, the emphasis on what hinders abandonment of low-value care is placed on antiquated practices, invalid science, or supplier-induced demand. Our findings suggest that low-value care is plausibly linked to the notion of relative procedural skill, in that the existence of a cost-effective alternative produces value only to the extent that this alternative can be provided successfully. This suggests that in order to steer physicians away from low-value care, particular importance should be placed on raising their competency and comfort level with alternative interventions.

Finally, the dynamics that followed the negative informational shock in our study highlight the role of skill heterogeneity in guiding a differential reaction to this uniform and public information. This suggests that our results speak directly to policies that aim to integrate evidence-based medicine with individual practice. Echoing the work of Chandra *et al.* (2011) and Chandra *et al.* (2015), we show that a key barrier to translating the diffusion of evidence into the practice of medicine and procedural choice is skill heterogeneity. Therefore, enforcing a uniform practice of medicine that ignores heterogeneity in human capital may reduce welfare.

While the evidence provided in this paper reveals an important mechanism that leads to physician practice variation, this study is not without limitations. First, we use a relatively small sample size in order to construct physician skill measures. While our results provide a preponderance of evidence in the case of hysterectomies, future research, aimed at physicians' technology abandonment decision in the context of additional procedural choices will be essential in corroborating our findings. Second, since we are unable to observe the outcomes of patients who ended up not receiving a hysterectomy after the FDA announcement, it is difficult to assess the full welfare implications of the informational shock. Finally, our study focuses solely on the short-term response of physicians (i.e., within 2 years after the FDA announcement). While this period is ideal for studying the role of relative individual skill on technology abandonment decisions, as it preceded the introduction of competing morcellation products and top-down healthcare system level abandonment rules, long-term responses may be important as well. In the long run, physicians may alter their investments in human capital that may affect their relative skill while producers may invest in the creation of new surgical devices or techniques to reduce the chance that morcellation spreads cancer. This highlights once again the role that alternative procedural options in abandoning existing ones.

Acknowledgement

We are grateful to the editor and two anonymous referees for their valuable comments and suggestions. We would like to thank participants at the 7th Conference of American Society of Health Economists, 17th Annual International Industrial Organization Conference, and at the seminar series of Rutgers University, Peking University, University of Pennsylvania, Tulane University, George Washington University, Hunter College, and University of California San Francisco. We declare no financial or personal relationship that could cause a conflict of interest regarding this article.

APPENDIX A: Appendix Tables and Figures

Appendix Table A1. Logistic regression for the prediction of MI appropriateness (i.e., the probability of receiving MI).

Pr (MI=1)	Coefficient (SE)
Age	-0.02*
	(0.01)
Age_squared	0.0002**

	(0.0001)
Emergent	-0.39***
	(0.10)
Charlson Comorbidity Index	-0.21***
	(0.02)
Pelvic Floor Adhesion	-0.54***
	(0.06)
Cervical Dysplasia	-0.07
	(0.30)
Morbid Obesity	0.21**
	(0.10)
Principal Diagnosis-Fibroid	0.86***
	(0.09)
Principal Diagnosis- Excessive or frequent menstruation	-0.23*
	(0.13)
Principal Diagnosis-Endometriosis of uterus	1.50***
	(0.16)
Principal Diagnosis-Unspecified symptom associated with female genital organs	1.22***
	(0.11)
Principal Diagnosis-Benign neoplasm of unspecified ovary	1.27***
	(0.12)
Number of Other Diagnoses at admission	-0.51***
	(0.01)
White	0.35***
	(0.03)
Uninsured	-1.12***
	(0.05)
Medicare	0.22***

	(0.06)
Medicaid	-0.67***
	(0.04)
Year-quarter fixed effects	Y
N	36,137
Log Likelihood	-17,258

Notes: This table reports the estimation results from the logistic regression for the prediction of MI appropriateness using the pre-period whole sample. The logistic regression is specified in Section 4.1. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

Appendix Table A2. Non-Complication rate (i.e., proportion of patients with no postoperative complication) by procedural type.

	Pre-FDA Warning (2012Q1-2013Q4)			Post-FDA Warning (2014Q2-2015Q3)		
	Non-Complication Rate (Unadjusted)	Non-Complication Rate (Adjusted)	N	Non-Complication Rate (Unadjusted)	Non-Complication Rate (Adjusted)	N
Received Open	74%	81%	3,749	78%	85%	2,813
Received MI	95%	92%	10,522	96%	92%	6,377

Notes: This table reports the proportion of patients with no postoperative complications for patients who received open and MI hysterectomies, respectively, both unadjusted and adjusted by patient characteristics, as discussed in Section 4.2.

Appendix Table A3. Distribution of patient volume (MI-appropriate and open-appropriate) in the pre-period in the original full sample

No. of Physicians = 394	Number of MI-appropriate Patients in the pre-period	Number of Open-appropriate Patients in the pre-period
Minimum	2	2
1% Percentile	2	2
5% Percentile	2	2
10% Percentile	3	2
25% Percentile	6	3
50% Percentile	13	6
75% Percentile	31	13
90% Percentile	67	26
95% Percentile	101	42
99% Percentile	202	69
Maximum	354	114

Notes: This figure presents the distribution of the number of MI-appropriate and open-appropriate patients in the original full sample.

Appendix Table A4. Summary statistics of patient and physician characteristics of the subsample.

Variables	Descriptions	Mean	SD
Panel 1: Patient Characteristics			
Received MI	Whether the patient received minimally invasive	0.70	0.

	hysterectomy (as opposed to open hysterectomy)		46
Post-time Dummy	Whether the patient received hysterectomies after the FDA safety communication (i.e., after April 2014)	0.40	0.49
Having Any Postoperative Complication	Whether the patient had any postoperative complications	0.12	0.33
Age	Page age	52	13
Emergent	Whether the patient was admitted under an emergent situation	0.017	0.13
Charlson Comorbidity Index	An index capturing comorbidities of patients.	0.43	1.41
Pelvic Floor Adhesion	Whether the patient has pelvic floor adhesion at the time of admission.	0.06	0.23
Cervical Dysplasia	Whether the patient has cervical dysplasia at the time of admission.	0.003	0.05
Morbid Obesity	Whether the patient has morbid obesity at the time of admission.	0.04	0.19
Principal Diagnosis - Uterine Fibroid	Whether the patient's principal diagnosis was uterine fibroid.	0.31	0.46
Principal Diagnosis- Excessive or frequent menstruation	Whether the patient's principal diagnosis was excessive or frequent menstruation.	0.17	0.38
Principal Diagnosis- Endometriosis of uterus	Whether the patient's principal diagnosis was endometriosis of uterus.	0.04	0.19
Principal Diagnosis- Unspecified symptom associated with female genital organs	Whether the patient's principal diagnosis was unspecified symptom associated with female genital organs.	0.02	0.13
Principal Diagnosis- Benign neoplasm of unspecified ovary	Whether the patient's principal diagnosis was benign neoplasm of unspecified ovary	0.04	0.19
Number of Other Diagnoses at admission	The number of other diagnoses of the patient at the	1.9	2.

	time of admission.		7
White	Whether the patient's race is white.	0.68	0. 47
Medicare	Whether the patient is covered by Medicare.	0.23	0. 42
Medicaid	Whether the patient is covered by Medicaid.	0.08	0. 27
Uninsured	Whether the patient is uninsured.	0.05	0. 21
Panel 2: Physician Characteristics			
MI Skill (Preferred Measure)	The proportion of patients who do not have any postoperative complications among those who are predicted to be appropriate for MI procedure.	0.95	0. 09
Open Skill (Preferred Measure)	The proportion of patients who do not have any postoperative complications among those who are predicted to be appropriate for open procedure.	0.64	0. 22
Relative Skill (Preferred Measure)	Difference between MI skill and open skill.	0.31	0. 22
MI Skill (Alternative Measure)	Risk-adjusted non-complication rate among patients who received MI procedures.	0.96	0. 04
Open Skill (Alternative Measure)	Risk-adjusted non-complication rate among patients who received open procedures.	0.70	0. 20
Relative Skill (Alternative Measure)	Difference between MI skill and open skill.	0.26	0. 20
Number of patients who received MI (Pre-period)	Number of patients who received MI hysterectomies in the pre-period.	133	95
Number of patients who received open (Pre-period)	Number of patients who received open hysterectomies in the pre-period.	45	31

Total number of patients (Pre-period)	Total number of patients who received hysterectomies in the pre-period.	178	103
Number of patients appropriate for MI (Pre-period)	Number of patients who are predicted to be appropriate for MI hysterectomies in the pre-period.	114	91
Number of patients appropriate for Open (Pre-period)	Number of patients who are predicted to be appropriate for open hysterectomies in the pre-period.	36	20
IV: Relative Skill of all physicians in the same type of hospital(s) in the same county excluding the focal hospital	Relative Skill of all physicians in the same type of hospital(s) in the same county excluding the focal hospital	0.21	0.19
N		14,052	

Notes: This table shows the summary statistics for the subsample, where we require physicians to have treated at least 11 MI-appropriate patients and 11 open-appropriate patients in the pre-period. Panel 1 presents the summary statistics for variables about patient characteristics. Panel 2 presents the summary statistics for variables about physician characteristics. Mean and standard deviations are reported. Preferred measures of MI and open skill are the observed non-complciation rate among patients who are predicted to be most appropriate for the procedure. Alternative measures of MI and open skill is the risk-adjusted non-complication rate among patients who received the procedure. Relative skill is the measured by the difference between MI skill and open skill.

Appendix Table A5. Association between physician relative skill and MI utilization

DV=Received MI		
	(1)	(2)
Relative Skill (MI relative to Open)	0.047***	0.026*
	(0.015)	(0.014)
County FE	Y	N
Hospital FE	N	Y
N	14,271	14,271
R-squared	0.33	0.39

Notes: This table shows that conditional on patient characteristics and location or institution, physicians with higher relative skill are more likely to use the MI procedure, as discussed in Section 5.1. Results are estimated from linear probability estimations. Relative skill measures the skill level of performing MI hysterectomy relative to open hysterectomy. All skill measures are z-scored. All regressions control for patient characteristics including age, insurance status, and clinical risk factors. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

Appendix Table A6. The impact of physician (absolute) MI skill and open skill on the utilization of MI hysterectomy.

DV=Received MI	Within-physician Analysis (Pre- vs. Post-period)						
	OLS			First Stage	IV (2SLS)		
Panel 1: MI Skill	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MI Skill *Post	0.016*	0.019**	0.008		0.064**	0.061**	0.061*
	(0.008)	(0.009)	(0.009)		(0.031)	(0.029)	(0.032)
IV: Others' MI Skill * Post				7.557***			
				(1.871)			
Pre-period MI Share * Post	N	Y	N	N	N	Y	N
Log (Pre-period MI Volume) * Post	N	N	Y	N	N	N	Y
Physician FE	Y	Y	Y	Y	Y	Y	Y
F-Statistics on Excluded IV					16	20	16
N	23,461	23,461	23,461	23,461	23,461	23,461	23,461
R-squared	0.424	0.424	0.425	0.482	0.422	0.423	0.423
Panel 2: Open Skill	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Open Skill*Post	-0.025***	-0.025***	-0.015		-0.047**	-0.051**	-0.035
	(0.009)	(0.009)	(0.011)		(0.021)	(0.022)	(0.026)
IV: Others' Open Skill * Post				2.375***			
				(0.802)			
Pre-period MI Share * Post	N	Y	N	N	N	Y	N
Log (Pre-period MI Volume) * Post	N	N	Y	N	N	N	Y
Physician FE	Y	Y	Y	Y	Y	Y	Y
F-Statistics on Excluded IV					6	8	9
N	23,461	23,461	23,461	23,461	23,461	23,461	23,461
R-squared	0.425	0.425	0.425	0.504	0.424	0.424	0.425

Notes: This table shows the impact of physicians' absolute skill levels on their utilization of MI hysterectomy, as discussed in Section 6.1. Results are estimated from linear probability estimations. In columns 5-7, we instrument each physician's MI or open skill using the MI or open skill of other physicians who practice in the "same type" hospitals in the same counties. All skill measures are z-scored. All regressions control for patient characteristics including age, insurance status, and clinical risk factors. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

Appendix Table A7. Robustness Test— Alternative measure for relative skill when controlling for physician fixed effects when estimating patient *MI appropriateness*

DV=Received MI	Within-physician Analysis (Pre- vs. Post-period)						
	OLS			First Stage	IV (2SLS)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Relative Skill*Post	0.024***	0.030***	0.023***		0.056**	0.054**	0.065**
	(0.007)	(0.008)	(0.008)		(0.024)	(0.022)	(0.033)
IV: Others' Relative Skill * Post				1.354***			
				(0.480)			

Pre-period MI Share * Post	N	Y	N	N	N	Y	N
Log (Pre-period MI Volume) * Post	N	N	Y	N	N	N	Y
Physician FE	Y	Y	Y	Y	Y	Y	Y
F-Statistics on Excluded IV					8	17	6
N	28,188	28,188	28,188	28,188	28,188	28,188	28,188
R-squared	0.473	0.473	0.473	0.507	0.472	0.473	0.471

Notes: This table shows the estimation results for the robustness test in Section 6.3, where we controlled for physician fixed effects when generating the predicted patient MI appropriateness measure. Results are estimated from linear probability estimations. Relative skill measures the skill level of performing MI hysterectomy relative to Open hysterectomy. In columns 5-7, we instrument each physician's relative skill using the relative skill of other physicians who practice in the "same type" hospitals in the same counties. All skill measures are z-scored. All regressions control for patient characteristics including age, insurance status, and clinical risk factors. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

Appendix Table A8. Robustness Test—Alternative measure for relative skill based on risk-adjusted performance.

DV=Received MI	Within-physician Analysis (Pre- vs. Post-period)						
	OLS			First Stage	IV (2SLS)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Relative Skill*Post	0.024***	0.025***	0.017*		0.091**	0.086**	0.088*
	(0.008)	(0.008)	(0.009)		(0.040)	(0.038)	(0.049)
IV: Others' Relative Skill * Post				0.941***			
				(0.246)			
Pre-period MI Share * Post	N	Y	N	N	N	Y	N
Log (Pre-period MI Volume) * Post	N	N	Y	N	N	N	Y
Physician FE	Y	Y	Y	Y	Y	Y	Y
F-Statistics on Excluded IV					15	17	6
N	23,461	23,461	23,461	23,461	23,461	23,461	23,461
R-squared	0.449	0.449	0.449	0.455	0.444	0.445	0.444

Notes: This table shows the results of the robustness test in Section 6.3, where we use the alternative measures for physician skill based on risk-adjusted performance. Results are estimated from linear probability estimations. Relative skill measures the skill level of performing MI hysterectomy relative to Open hysterectomy. In columns 5-7, we instrument each physician's relative skill using the relative skill of other physicians who practice in the "same type" hospitals in the same counties. All skill measures are z-scored. All regressions control for patient characteristics including age, insurance status, and clinical risk factors. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

Appendix Table A9. Robustness Test—Alternative Sample (Whole Sample without Dropping Physicians in the Bottom Quartile in MI-appropriate and Open-appropriate Volume in the Pre-period)

DV=Received MI	Within-physician Analysis (Pre- vs. Post-period)						
	OLS			First Stage	IV (2SLS)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Relative Skill*Post	0.027***	0.032***	0.026***		0.039**	0.037**	0.041**
	(0.007)	(0.007)	(0.008)		(0.017)	(0.016)	(0.021)
IV: Others' Relative Skill * Post				1.800***			
				(0.388)			
Pre-period MI Share * Post	N	Y	N	N	N	Y	N
Log (Pre-period MI Volume) * Post	N	N	Y	N	N	N	Y
Physician FE	Y	Y	Y	Y	Y	Y	Y
F-Statistics on Excluded IV					22	24	25
N	30,254	30,254	30,254	30,254	30,254	30,254	30,254
R-squared	0.471	0.472	0.471	0.515	0.471	0.472	0.471

Notes: This table shows the result of the robustness test in Section 6.3, where we use the whole sample to estimate the results. Results are estimated from linear probability estimations. Relative skill measures the skill level of performing MI hysterectomy relative to Open hysterectomy. In columns 5-7, we instrument each physician's relative skill using the relative skill of other physicians who practice in the "same type" hospitals in the same counties. All skill measures are z-scored. All regressions control for patient characteristics including age, insurance status, and clinical risk factors. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

Appendix Table A10. Heterogeneity by patient MI-Appropriateness.

DV=Received MI	OLS		
	(1)	(2)	(3)
Patient MI-Appropriateness*Post	0.043***	0.048***	0.037***
	(0.009)	(0.010)	(0.010)
Pre-period MI Share * Post	N	Y	N
Log (Pre-period MI Volume) * Post	N	N	Y
Physician FE	Y	Y	Y
N	23,461	23,461	23,461
R-squared	0.441	0.441	0.441

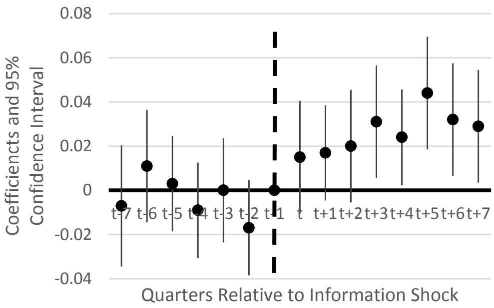
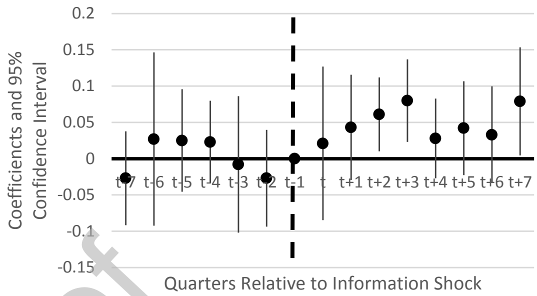
Notes: This table presents the heterogeneity analysis result in Section 6.4, showing that physicians are more likely to decrease the use of MI hysterectomies for patients who have relatively lower appropriateness for MI procedures. Results are estimated from linear probability estimations. All regressions control for patient characteristics including age, insurance status, and clinical risk factors. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

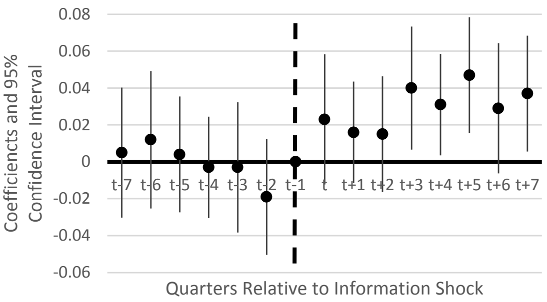
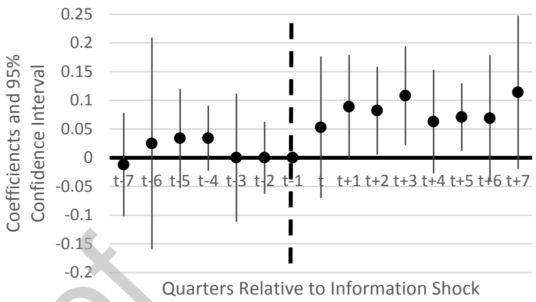
Appendix Table A11. Evidence on extensive margin – Change in hysterectomy volume after the information shock

	Pre-FDA Warning (2012Q1-2013Q4)	Post-FDA Warning (2014Q2-2015Q3)	Change
Panel 1: Quarterly Hysterectomy Volume (Both MI and Open)			
All Physicians	1784	1532	-14.10%
Top MI Performers	882	738	-16.40%
Top Open Performers	902	794	-11.90%
Panel 2: Quarterly MI Hysterectomy Volume			
Top MI Performers	664	549	-17.40%
Top Open Performers	652	514	-21.10%
Panel 3: Quarterly Open Hysterectomy Volume			
Top MI Performers	219	189	-13.40%
Top Open Performers	250	280	11.80%
Panel 4: Share of MI Hysterectomies			
Top MI Performers	0.75	0.74	-1.10%
Top Open Performers	0.72	0.65	-10.30%

Notes: This table shows the change in the number of patients who receive hysterectomies, as well as the share of MI hysterectomies, after the information shock (FDA announcement), as discussed in Section 6.4. Top open performers are defined as physicians whose relative skill is above the median (weighted by patient volume); top MI performers are defined as physicians whose relative skill is below or equal to the median (weighted by patient volume). The sample contains 23,416 patients treated by 215 physicians from Jan. 2012 to Dec. 2013 (pre-period) and Apr. 2014 to Sep. 2015 (post-period).

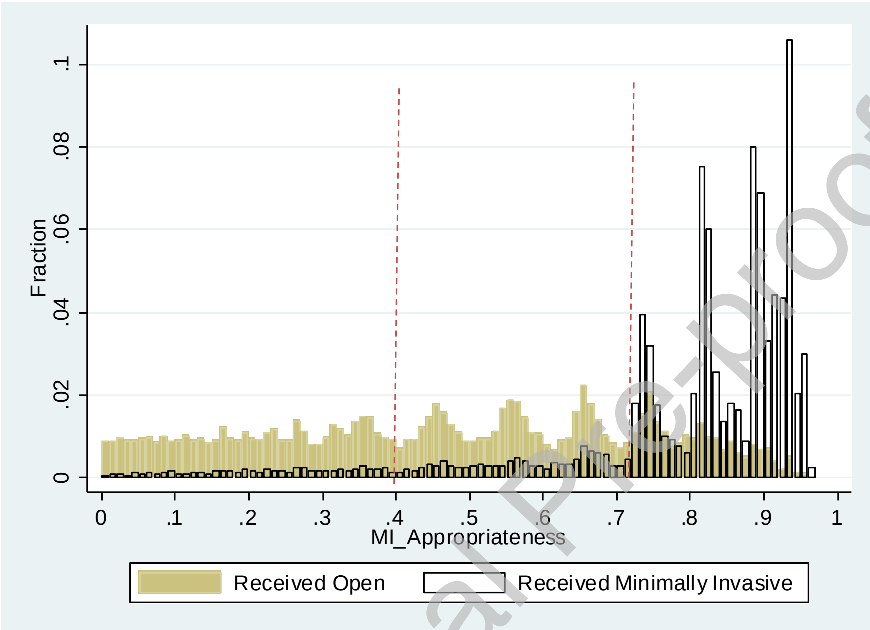
	OLS	IV (2SLS)

Panel 1: Main Sample		

Panel 2: Subsample Excluding Low-Volume Physicians		

Appendix Figure A1. Alternative specification of the event study estimation—using the 4th quarter of 2013 as starting of information shock.
Note: This figure reports the estimated coefficients and 95% confidence intervals of the interaction terms between the standardized physician relative skill and the year-quarter indicators in the event-study specification as discussed in Section 6.2. The time period t refers to the 4th quarter of Year 2013. We set the coefficient of the interaction term to zero for the quarter before the information shock (i.e., the 3rd quarter of 2013). The main sample contains 25,014 patients treated by 215 physicians between Jan. 2012 and Sep. 2015. The subsample contains 14,989 patients treated by 76 physicians between Jan. 2012 and Sep. 2015. Results are estimated from linear probability specifications. The event study regression controls for patient characteristics including age, insurance status, and clinical risk factors, physician fixed effects, and year-quarter fixed effects. Standard errors are clustered by physician.

Appendix Figure A2. Distribution of MI Appropriateness (Robustness test when including physician fixed effects in the estimation of MI Appropriateness.)



APPENDIX B: Construction of the alternative skill measure.

Our alternative skill measure is based on risk-adjusted performance. We adopt this second measure only as a robustness check, in order to mitigate the concern with the interpretation of the first measure. Specifically, we first estimate Equation (A1) to predict the incidence of not having any postoperative complication using observations prior to the FDA warning.

$Pr(\text{Non}Complication_{it}) = G(\alpha MI_{it} + X_{it}\Phi + MI_{it} \times X_{it}\Upsilon + Time_t\Psi + \theta_0)$ (A1)

For each patient i who receives hysterectomy in year-quarter t, $NonComplication_{it}$ is an indicator variable for whether the patient has no postoperative complication. MI_{it} is an indicator variable which equals 1 if the patient receives MI hysterectomy and 0 if the patient receives open hysterectomy; X_{it} are the same set of patient characteristics as in Equation (5). The interaction term between MI_{it} and X_{it} captures differential impacts of risk factors on outcomes by procedural type. $Time_t$ are a set of 14 year-quarter indicator variables.

Each physician's MI skill or open skill is then measured using the risk-adjusted complication rate among his or her patients who are treated using the specific procedure (Equation A2). Relative skill is again measured by the difference between MI skill and open skill.

$RA\_\text{Non}ComplicationRate_{jp} = \frac{Observed\_\text{Non}ComplicationRate_{jp}}{Predicted\_\text{Non}ComplicationRate_{jp}} \times Overall\_\text{Non}ComplciationRate$ (A2)

In Equation (A2), $Observed\_\text{Non}ComplicationRate_{jp}$ represents the observed (unadjusted) non-complication rate for physician j among his or her patients treated with procedure p; $Predicted\_\text{Non}ComplicationRate_{jp}$ represents the predicted non-complication rate for physician j among his or her patients treated with procedure p, predicted from Equation (A1). $RA\_\text{Non}ComplicationRate_{jp}$ is the physician j's risk-adjusted non-complication rate for procedure p.

References

Agha, L., Molitor, D., 2018. The local influence of pioneer investigators on technology adoption: evidence from new cancer drugs. Review of Economics and Statistics 100 (1), 29-44

Baicker, K., Buckles, K.S., Chandra, A., 2006. Geographic variation in the appropriate use of cesarean delivery: do higher usage rates reflect medically inappropriate use of this procedure? Health Affairs 25, W355-W367

Barbash, G.I., Glied, S.A., 2010. New Technology and Health Care Costs—The Case of Robot-Assisted Surgery. New England Journal of Medicine 363 (8), 701-704

Barron, K.I., Richard, T., Robinson, P.S., Lamvu, G., 2015. Association of the US Food and Drug Administration morcellation warning with rates of minimally invasive hysterectomy and myomectomy. Obstetrics & Gynecology 126 (6), 1174-1180
Baumgardner, J.R., 1988. Physicians' services and the division of labor across local markets. Journal of Political Economy 96 (5), 948-982
Becker, G.S., Murphy, K.M., 1992. The division of labor, coordination costs, and knowledge. The Quarterly Journal of Economics 107 (4), 1137-1160
Bekelis, K., Skinner, J., Gottlieb, D., Goodney, P., 2017. De-adoption and exnovation in the use of carotid revascularization: retrospective cohort study. BMJ 359, j4695
Berez, J., David, G., Howard, D.H., Neuman, M.D., 2018. Does Bad News Travel Faster? On the Determinants of Medical Technology Abandonment. Journal of Human Capital 12 (4), 569-603
Chan, D.C., Gentzkow, M., Yu, C., 2019. Selection with variation in diagnostic skill: Evidence from radiologists. National Bureau of Economic Research
Chandra, A., Jena, A.B., Skinner, J.S., 2011. The pragmatist's guide to comparative effectiveness research. Journal of Economic Perspectives 25 (2), 27-46
Chandra, A., Khullar, D., Lee, T.H., 2015. Addressing the challenge of gray-zone medicine. New England Journal of Medicine 372 (3), 203-205
Chandra, A., Staiger, D.O., 2007. Productivity spillovers in health care: evidence from the treatment of heart attacks. Journal of Political Economy 115 (1), 103-140
Charlson, M.E., Pompei, P., Ales, K.L., MacKenzie, C.R., 1987. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of Chronic Diseases 40 (5), 373-383
Comin, D., Skinner, J., Staiger, D., 2017. Overconfidence and the Diffusion of Medical Technology. Working Paper, Dartmouth College
Currie, J., MacLeod, W.B., 2017. Diagnosing expertise: Human capital, decision making, and performance among physicians. Journal of Labor Economics 35 (1), 1-43
Currie, J., MacLeod, W.B., Van Parys, J., 2016. Provider practice style and patient health outcomes: the case of heart attacks. Journal of Health Economics 47, 64-80
Currie, J.M., MacLeod, W.B., 2020. Understanding doctor decision making: The case of depression treatment. Econometrica 88 (3), 847-878
Cutler, D., Skinner, J.S., Stern, A.D., Wennberg, D., 2019. Physician beliefs and patient preferences: a new look at regional variation in health care spending. American Economic Journal: Economic Policy 11 (1), 192-221
Cutler, D.M., Sheiner, L., 1999. The geography of Medicare. American Economic Review 89 (2), 228-233
David, G., Helmchen, L.A., 2011. The role of task adhesion in limiting specialization along the medical care continuum. Review of Labour Economics and Industrial Relation 25 (1), 24-44

Dranove, D., Ramanarayanan, S., Sfekas, A., 2011. Does the market punish aggressive experts? Evidence from cesarean sections. The BE Journal of Economic Analysis & Policy 11 (2)
Epstein, A.J., Groeneveld, P.W., Harhay, M.O., Yang, F., Polsky, D., 2013. Impact of minimally invasive surgery on medical spending and employee absenteeism. JAMA Surgery 148 (7), 641-647
Epstein, A.J., Nicholson, S., 2009. The formation and evolution of physician treatment styles: an application to cesarean sections. Journal of Health Economics 28 (6), 1126-1140
Fadlon, I., Van Parys, J., 2020. Primary care physician practice styles and patient care: Evidence from physician exits in Medicare. Journal of Health Economics 71, 102304-102304
Finkelstein, A., Gentzkow, M., Williams, H., 2016. Sources of Geographic Variation in Health Care: Evidence From PatientMigration. The Quarterly Journal of Economics 131 (4), 1681-1726
Harris, J.A., Swenson, C.W., Uppal, S., Kamdar, N., Mahnert, N., As-Sanie, S., Morgan, D.M., 2016. Practice patterns and postoperative complications before and after US Food and Drug Administration safety communication on power morcellation. American Journal of Obstetrics & Gynecology 214 (1), 98. e1-98. e13
Heckman, J.J., Honoré, B.E., 1990. The Empirical Content of the Roy Model. Econometrica 58 (5), 1121-1149
Howard, D.H., David, G., Hockenberry, J., 2017. Selective Hearing: Physician-Ownership and Physicians' Response to New Evidence. Journal of Economics & Management Strategy 26 (1), 152-168
Howard, D.H., Soulos, P.R., Chagpar, A.B., Mougalian, S., Killelea, B., Gross, C.P., 2016. Contrary to conventional wisdom, physicians abandoned a breast cancer treatment after a trial concluded it was ineffective. Health Affairs 35 (7), 1309-1315
Kuo, L.E., Murayama, K., Simmons, K.D., Kelz, R.R., 2017. Variation in the utilization of minimally invasive surgical operations. Annals of Surgery 265 (3), 514-520
Levitz, J., Kamp, J., 2014. Gynecologists Resist FDA Over Popular Surgical Tool. In: Wall Street Journal
MacDonald, G., Marx, L.M., 2001. Adverse specialization. Journal of Political Economy 109 (4), 864-899
Miller, R.A., 1984. Job matching and occupational choice. Journal of Political Economy 92 (6), 1086-1120
Multinu, F., Casarin, J., Hanson, K.T., Angioni, S., Mariani, A., Habermann, E.B., Laughlin-Tommaso, S.K., 2018. Practice Patterns and Complications of Benign Hysterectomy Following the FDA Statement Warning Against the Use of Power Morcellation. JAMA Surgery, e180141-e180141
Noar, M., 1991. Endoscopy simulation: a brave new world? Endoscopy (Stuttgart) 23 (3), 147-149
Rogers, D.A., Elstein, A.S., Bordage, G., 2001. Improving continuing medical education for surgical techniques: applying the lessons learned in the first decade of minimal access surgery. Annals of Surgery 233 (2), 159
Roman, B.R., Asch, D.A., 2014. Faded promises: the challenge of deadopting low-value care. Annals of Internal Medicine 161 (2), 149-150
Rosen, S., 1983. Specialization and human capital. Journal of Labor Economics 1 (1), 43-49
Roy, A.D., 1950. The distribution of earnings and of individual output. The Economic Journal 60 (239), 489-505

Shah, N.D., Montori, V.M., Krumholz, H.M., Tu, K., Alexander, G.C., Jackevicius, C.A., 2010. Responding to an FDA warning—geographic variation in the use of rosiglitazone. New England Journal of Medicine 363 (22), 2081-2084
Siedhoff, M.T., Wheeler, S.B., Rutstein, S.E., Geller, E.J., Doll, K.M., Wu, J.M., Clarke-Pearson, D.L., 2015. Laparoscopic hysterectomy with morcellation vs abdominal hysterectomy for presumed fibroid tumors in premenopausal women: a decision analysis. American Journal of Obstetrics & Gynecology 212 (5), 591. e1-591. e8
Silvennoinen, M., Mecklin, J.-P., Saariluoma, P., Antikainen, T., 2009. Expertise and skill in minimally invasive surgery. Scandinavian Journal of Surgery 98 (4), 209-213
Staats, B.R., KC, D.S., Gino, F., 2018. Maintaining beliefs in the face of negative news: The moderating role of experience. Management Science 64 (2), 804-824
Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65 (3), 557
Sutherland, J.M., Fisher, E.S., Skinner, J.S., 2009. Getting past denial--the high cost of health care in the United States. New England Journal of Medicine 361 (13), 1227
Ubel, P.A., Asch, D.A., 2015. Creating value in health by understanding and overcoming resistance to de-innovation. Health Affairs 34 (2), 239-244
Wennberg, J., Fisher, E., Skinner, J., 2002. Geography and the debate over Medicare reform. Health Affairs (Project Hope), W96-114
Wennberg, J., Gittelsohn, A., 1973. Small area variations in health care delivery: a population-based health information system can guide planning and regulatory decision-making. Science 182 (4117), 1102-1108
Wright, J.D., Chen, L., Burke, W.M., Hou, J.Y., Tergas, A.I., Ananth, C.V., Hershman, D.L., 2016. Trends in use and outcomes of women undergoing hysterectomy with electric power morcellation. JAMA 316 (8), 877-878

Table 1. Summary statistics of patient and physician characteristics for the main sample

Variables	Descriptions	Mean	SD
Panel 1: Patient Characteristics			
Received MI	Whether the patient received minimally invasive hysterectomy (as opposed to open hysterectomy)	0.72	0.45
Post-time Dummy	Whether the patient received hysterectomies after the FDA safety communication (i.e., after April 2014)	0.39	0.49
Having Any Postoperative Complication	Whether the patient had any postoperative complications	0.1	0

		1	.31
Age	Page age	50	1 2
Emergent	Whether the patient was admitted under an emergent situation	0.0 15	0 .12
Charlson Comorbidity Index	An index capturing comorbidities of patients.	0.3 0	1 .15
Pelvic Floor Adhesion	Whether the patient has pelvic floor adhesion at the time of admission.	0.0 6	0 .24
Cervical Dysplasia	Whether the patient has cervical dysplasia at the time of admission.	0.0 02	0 .05
Morbid Obesity	Whether the patient has morbid obesity at the time of admission.	0.0 3	0 .17
Principal Diagnosis - Uterine Fibroid	Whether the patient’s principal diagnosis was uterine fibroid.	0.3 9	0 .49
Principal Diagnosis- Excessive or frequent menstruation	Whether the patient’s principal diagnosis was excessive or frequent menstruation.	0.1 1	0 .32
Principal Diagnosis- Endometriosis of uterus	Whether the patient’s principal diagnosis was endometriosis of uterus.	0.0 4	0 .20
Principal Diagnosis- Unspecified symptom associated with female genital organs	Whether the patient’s principal diagnosis was unspecified symptom associated with female genital organs.	0.0 2	0 .14
Principal Diagnosis- Benign neoplasm of unspecified ovary	Whether the patient’s principal diagnosis was benign neoplasm of unspecified ovary	0.0 3	0 .16
Number of Other Diagnoses at admission	The number of other diagnoses of the patient at the time of admission.	1.6	2 .5
White	Whether the patient’s race is white.	0.6 7	0 .47
Medicare	Whether the patient is covered by Medicare.	0.1 8	0 .38
Medicaid	Whether the patient is covered by Medicaid.	0.0 8	0 .27
Uninsured	Whether the patient is uninsured.	0.0 4	0 .20
Panel 2: Physician Characteristics			
MI Skill (Preferred Measure)	The proportion of patients who do not have any postoperative complications among those who are predicted to be appropriate for MI procedure.	0.9 5	0 .10
Open Skill (Preferred Measure)	The proportion of patients who do not have any postoperative complications among those who are predicted to be appropriate for open procedure.	0.7 0	0 .22
Relative Skill (Preferred Measure)	Difference between MI skill and open skill.	0.2	0

		5	.22
MI Skill (Alternative Measure)	Risk-adjusted non-complication rate among patients who received MI procedures.	0.95	0.07
Open Skill (Alternative Measure)	Risk-adjusted non-complication rate among patients who received open procedures.	0.72	0.22
Relative Skill (Alternative Measure)	Difference between MI skill and open skill.	0.23	0.22
Number of patients who received MI (Pre-period)	Number of patients who received MI hysterectomies in the pre-period.	101	87
Number of patients who received open (Pre-period)	Number of patients who received open hysterectomies in the pre-period.	31	30
Total number of patients (Pre-period)	Total number of patients who received hysterectomies in the pre-period.	132	101
Number of patients appropriate for MI (Pre-period)	Number of patients who are predicted to be appropriate for MI hysterectomies in the pre-period.	87	82
Number of patients appropriate for Open (Pre-period)	Number of patients who are predicted to be appropriate for open hysterectomies in the pre-period.	25	21
IV: Relative Skill of all physicians in the same type of hospital(s) in the same county excluding the focal hospital	Relative Skill of all physicians in the same type of hospital(s) in the same county excluding the focal hospital	0.18	0.18
N		23,461	

Notes: This table shows the summary statistics for the main sample of patient-level data. Panel 1 presents the summary statistics for variables about patient characteristics. Panel 2 presents the summary statistics for variables about physician characteristics. Mean and standard deviations are reported. Preferred measures of MI and open skill are the observed non-complciation rate among patients who are predicted to be most appropriate for the procedure. Alternative measures of MI and open skill is the risk-adjusted non-complication rate among patients who received the procedure. Relative skill is the measured by the difference between MI skill and open skill.

Table 2. MI appropriateness and MI utilization by physician relative skill before and after the information shock.

	Patients Treated by *Top MI Performers*			Patients Treated by *Top Open Performers*		
	Pre-FDA Warning (2012Q1-2013Q4)	Post-FDA Warning (2014Q2-2015Q3)	Percentage Change (Post-Pre)	Pre-FDA Warning (2012Q1-2013Q4)	Post-FDA Warning (2014Q2-2015Q3)	Percentage Change (Post-Pre)
	(1)	(2)	(3)	(4)	(5)	(6)

Average *MI appropriateness*	0.669	0.668	No Statistically Significant Difference	0.631	0.644	No Statistically Significant Difference
Percentage of patients receiving MI (Unadjusted)	0.752	0.743	-1.2%	0.723	0.648	-10.4%
Percentage of patients receiving MI (Adjusted by *MI appropriateness*)	0.739	0.732	-0.95%	0.740	0.652	-11.9%
N	7,059	4,427		7,212	4,763	

Notes: This table shows the average patient MI appropriateness, the unadjusted and adjusted (by patient MI appropriateness) MI utilization by Top MI Performers and Top Open Performers, before and after the FDA warning. Top MI Performers are defined as physicians whose relative skill (MI relative to open) is above the median (weighted by patient volume); Top Open Performers are defined as physicians whose relative skill (MI relative to open) is below the median (weighted by patient volume).

Table 3. Main Result: The impact of physician relative skill (MI relative to Open) on the utilization of MI hysterectomy.

DV=Received MI	OLS			First Stage	IV (2SLS)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel 1: Main sample							
Relative Skill * Post	0.030***	0.031***	0.020*		0.049***	0.046***	0.046**
	(0.008)	(0.008)	(0.011)		(0.017)	(0.015)	(0.021)
Pre-period MI Share * Post		-0.037				-0.049	
		(0.038)				(0.038)	
Log (Pre-period MI Volume) * Post			0.020*				0.008
			(0.012)				(0.015)
IV: Others' Relative Skill * Post				1.750***			
				(0.394)			
Physician FE	Y	Y	Y	Y	Y	Y	Y
F-Statistics on Excluded IV					20	26	27
N	23,461	23,461	23,461	23,461	23,461	23,461	23,461
R-squared	0.449	0.449	0.449	0.538	0.449	0.449	0.449
Panel 2: Subsample Excluding Low-Volume Physicians							
Relative Skill * Post	0.030***	0.033***	0.028*		0.064**	0.060**	0.069**
	(0.011)	(0.011)	(0.015)		(0.032)	(0.026)	(0.034)
Pre-period MI Share * Post		-0.054				-0.093	

		(0.055)				(0.068)	
Log (Pre-period MI Volume) * Post			0.004				-0.018
			(0.018)				(0.026)
IV: Others' Relative Skill * Post				1.699***			
				(0.453)			
Physician FE	Y	Y	Y	Y	Y	Y	Y
F-Statistics on Excluded IV					14	26	14
N	14,052	14,052	14,052	14,052	14,052	14,052	14,052
R-squared	0.436	0.436	0.436	0.525	0.435	0.435	0.435

Notes: This table presents the main estimation results in Section 6.1. The main sample includes 23,461 patients treated by 215 physicians; the subsample includes 14,051 treated by 76 high-volume physicians. Results are estimated from linear probability estimations. Relative skill measures the skill level of performing MI hysterectomy relative to open hysterectomy. In columns 5-7, we instrument each physician's relative skill using the relative skill of other physicians who practice in the "same type" hospitals in the same counties. All skill measures are z-scored. All regressions control for patient characteristics including age, insurance status, and clinical risk factors. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

Table 4. Falsification Test: The impact of physician relative skill (MI relative to Open) on patient composition (MI-appropriateness) and patient volume.

Dependent Variable	Patient MI-Appropriateness (Patient Level)			Patient Volume (Physician-Hospital-Year-Quarter Level)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1: Main sample						
Relative Skill*Post	0.001	0.003	0.002	0.236	0.310	0.355
	(0.009)	(0.009)	(0.010)	(0.493)	(0.479)	(0.415)
Pre-period MI Share * Post	N	Y	N	N	Y	N
Log (Pre-period MI Volume) * Post	N	N	Y	N	N	Y
Physician FE	Y	Y	Y	Y	Y	Y
N	23,461	23,461	23,461	3,354	3,354	3,354
R-squared	0.179	0.179	0.179	0.601	0.602	0.602
Panel 2: Subsample Excluding Low-Volume Physicians						
Relative Skill*Post	-0.002	0.000	0.001	0.135	0.354	0.474
	(0.012)	(0.013)	(0.015)	(0.953)	(0.918)	(0.847)
Pre-period MI Share * Post	N	Y	N	N	Y	N
Log (Pre-period MI Volume) * Post	N	N	Y	N	N	Y
Physician FE	Y	Y	Y	Y	Y	Y
N	14,052	14,052	14,052	1,480	1,480	1,480
R-squared	0.165	0.165	0.165	0.557	0.558	0.558

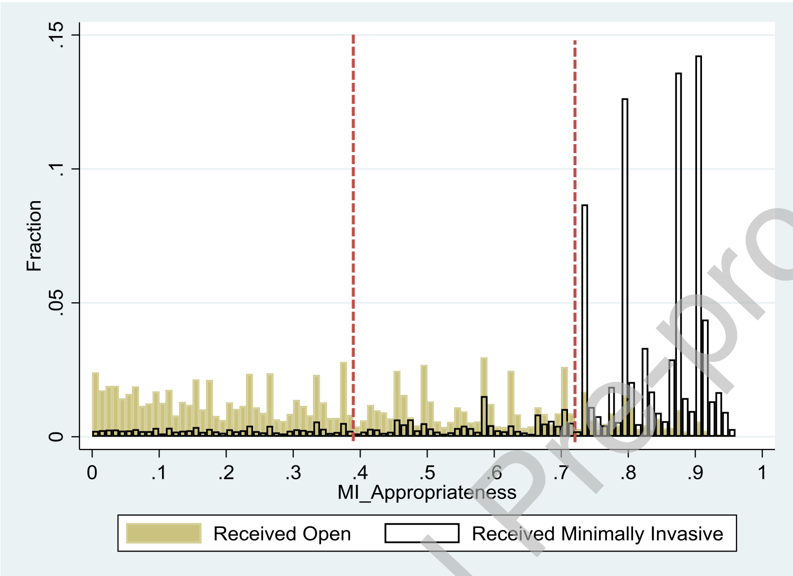
Notes: This table presents the results of the falsification tests in Section 6.2, where we replace the dependent variable in the main analysis with patient MI-appropriateness (Col 1-3) and the physician's patient volume (Col 4-6). The main sample includes 23,461 patients treated by 215 physicians; the subsample includes 14,051 treated by 76 high-volume physicians. Results are estimated from linear probability estimations. Relative skill measures the skill level of performing MI hysterectomy relative to open hysterectomy. All skill measures are z-scored. *** p<.01, ** p<.05, * p<.1.

Table 5. Robustness Test: Controlling for hospital fixed effects.

DV=Received MI	OLS		
	(1)	(2)	(3)
Panel 1: Main Sample			
Relative Skill*Post	0.026***	0.028***	0.017*
	(0.008)	(0.008)	(0.010)
Pre-period MI Share * Post	N	Y	N
Log (Pre-period MI Volume) * Post	N	N	Y
Hospital FE	Y	Y	Y
Physician FE	Y	Y	Y
N	23,461	23,461	23,461
R-squared	0.460	0.460	0.457
Panel 2: Subsample Excluding Low-Volume Physicians			
Relative Skill*Post	0.025**	0.028**	0.024
	(0.011)	(0.011)	(0.015)
Pre-period MI Share * Post	N	Y	N
Log (Pre-period MI Volume) * Post	N	N	Y
Hospital FE	Y	Y	Y
Physician FE	Y	Y	Y
N	14,052	14,052	14,052
R-squared	0.451	0.451	0.445

Notes: This table presents the robust results in Section 6.3 when including the hospital fixed effects in the estimation. The main sample includes 23,461 patients treated by 215 physicians; the subsample includes 14,051 treated by 76 high-volume physicians. Results are estimated from linear probability estimations. Relative skill measures the skill level of performing MI hysterectomy relative to open hysterectomy. All skill measures are z-scored. All regressions control for patient characteristics including age, insurance status, and clinical risk factors. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

Figure 1. Distribution of MI Appropriateness by whether the patient receives MI or open hysterectomy.

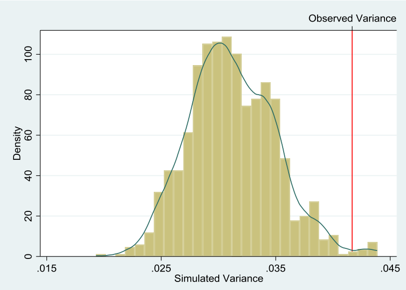


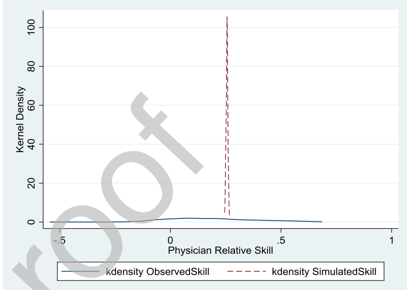
Note: This figure shows the distribution of patient MI appropriateness by whether the patient has received open or minimally invasive hysterectomy in the pre-period. MI Appropriateness is measured by the predicted probability of receiving MI from the logistic regression shown in Section 4.1.

Figure 2. Comparing the distribution of simulated relative skill with observed relative skill (No. of Iterations = 1,000)

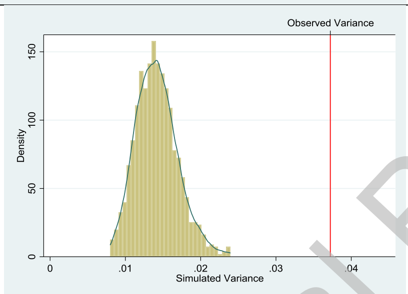
	Distribution of simulated variances of relative skill v.s. observed variance of relative skill	Distribution of average simulated relative skill v.s. distribution of observed relative skill.

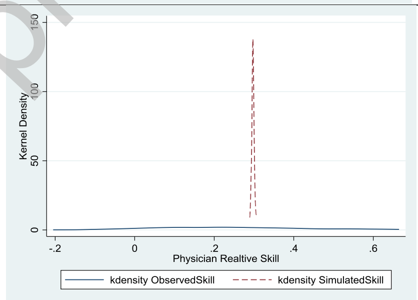
Panel 1: Main Sample





Panel 2: Subsample Excluding Low-Volume Physicians

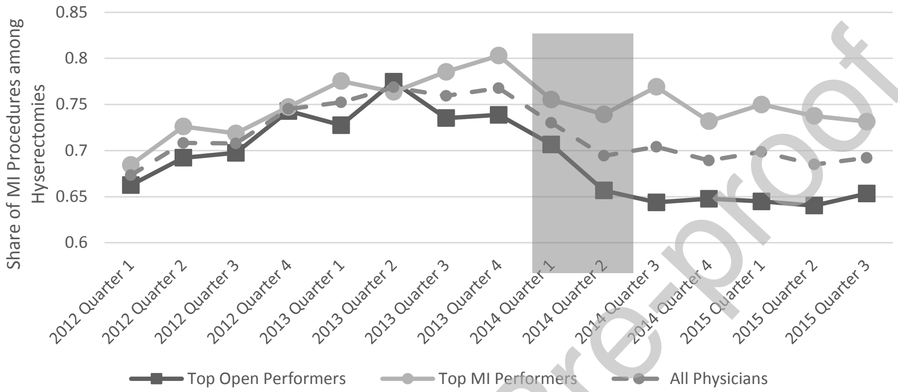




Note: This figure compares the distribution of simulated physician relative skill and that of the observed relative skill as discussed in Section 4.3. The main sample contains 215 physicians; the subsample contains 76 medium to high-volume physicians. In each iteration of the simulation, we randomly assign patients to physicians and calculate the implied relative difference in MI and open skill for each physician. Number of iterations is 1,000.

Figure 3. Share of MI procedures by physician relative skill.

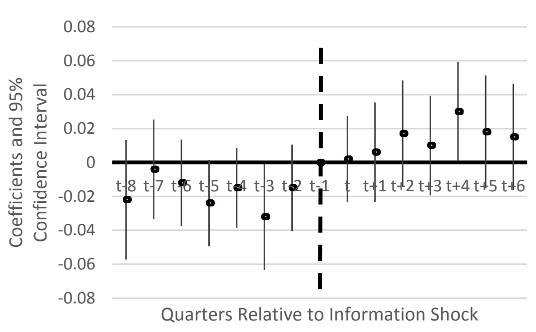
	OLS	IV (2SLS)

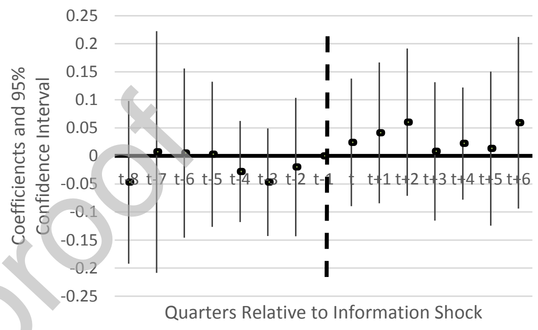


Notes: This graph reports the quarterly MI utilization rate among patients who receive hysterectomies from the raw data. The sample contains 25,014 patients treated by 215 physicians between Jan. 2012 and Sep. 2015. Top open performers are defined as physicians whose relative skill is above the median (weighted by patient volume); top MI performers are defined as physicians whose relative skill is below or equal to the median (weighted by patient volume). The shaded bar denotes when the national news report started on Dec. 2013, leading to the FDA safety communication released on April 17, 2014.

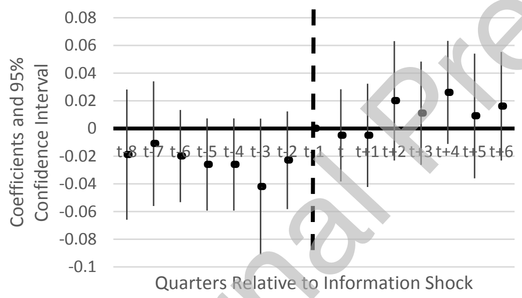
Figure 4. Event study result (quarterly plot): the impact of physician relative skill (MI relative to Open) on the utilization of MI hysterectomy.

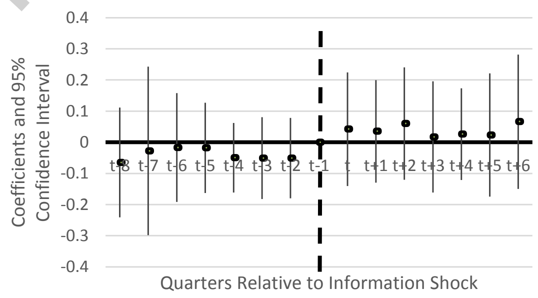
Panel 1: Main Sample





Panel 2: Subsample Excluding Low-Volume Physicians





Note: This figure reports the estimated coefficients and 95% confidence intervals of the interaction terms between the standardized physician relative skill and the year-quarter indicators in the event-study specification as discussed in Section 6.2. The time period t refers to the 1st quarter of Year 2014. We set the coefficient of the interaction term to zero for the quarter before the information shock (i.e., the 4th quarter of 2013). The main sample contains 25,014 patients treated by 215 physicians between Jan. 2012 and Sep. 2015. The subsample contains 14,989 patients treated by 76 physicians between Jan. 2012 and Sep. 2015. Results are estimated from a linear probability specification. The event study regression controls for patient characteristics including age, insurance status, and clinical risk factors, physician fixed effects, and year-quarter fixed effects. Standard errors are clustered by physician.