Information, Relative Performance and Technology Abandonment: Evidence from a FDA Safety Communication

Guy David
University of Pennsylvania

Bingxiao Wu
Rutgers University

Abstract

This paper studies the role of relative task-specific skill in explaining the heterogeneity in technology abandonment decisions in reaction to negative information shocks. We model how physicians respond to new information regarding the effectiveness of medical technology based on their joint distribution of skills and their patients' characteristics. We test the model's predictions and find physicians to alter their procedural mix towards open procedures in response to negative news related to the use of minimally invasive techniques. However, this effect was nearly muted among physicians with higher relative performance on minimally invasive procedures, highlighting a key impediment to technology abandonment.

Key Words: Physician Relative Performance, Technology Abandonment, Negative Information Shocks

1. INTRODUCTION

The importance of division of labor and the forces limiting its extent introduced by Adam Smith's (1776) classic work received great attention in the economics literature. 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). While a second stream of 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 medical specialty was shown to be important (Baumgardner 1988; David & Helmchen 2011). Regardless of why the division of labor is incomplete, the command of multiple skills by a single individual may result in a set of competencies, where substitutable tasks represent alternative ways to achieve the same goal. For example, physicians often face a choice between different courses of medical treatment in treating a given condition.[1]

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 choice (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 on the counterfactual task by that same individual is not observable and empirical identification relies on assumptions on the joint distribution of skills across tasks and individuals. The cases described above provide a unique opportunity to observe individual performance on 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).

Furthermore, holding the joint distribution of task-specific skills fixed, we are able to study shocks to the appropriateness of tasks by means of negative information. Negative news regarding treatment effectiveness should alter the choice of treatment and may lead to gradual or swift

[1] Other examples include plumbers testing pipes for leakage using either a water gauge or an air pressure gauge; electricians performing installation of electrical wiring using either a joint box system or a looping system; and hygienists performing teeth cleaning using either hand-scaling or an ultrasonic. In all these cases, the choice of task depends on the characteristics of the environment and the professional's task-specific skill.

abandonment of the technology in question. 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). The work presented here argues that variation in relative skill across alternative procedures is linked to the extent of technology abandonment across surgeons.

In this paper, we use a specific clinical setting where surgeons perform either minimally invasive or open hysterectomies for women who were diagnosed with uterine fibroids and related diseases and had to have their uterus removed. We measure both patient appropriateness and surgeons' relative skill in performing the two procedures. We show that after the Food and Drug Administration (FDA) issued a safety communication, which lowered the attractiveness of minimally invasive hysterectomies, surgeons with relatively higher skill in open versus minimally invasive procedures were far more likely to alter their procedural mix away from minimally invasive hysterectomies. Bekelis *et al.* (2017) makes the important distinction between the term "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 and shows the estimation results. Section 6 concludes.

2. BACKGROUND

We explore physicians' choice between two broad surgical alternatives for hysterectomy surgery.[2] 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.[3] This type of surgery 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.[4] This allows us to study relative performance, its impact on the choice of procedure, and its role in affecting surgeons' response to negative information shocks.

In Dec 2013, a petition to the FDA was filed calling to end the use of power morcellation during MI hysterectomies, as it was a potential cause for cancer. Following a widespread national news coverage, the FDA published a safety communication on April 2014 discouraging the use of laparoscopic power morcellation during MI hysterectomy. The communication highlighted the risk of spraying malignant tissue in the event of an unsuspected sarcoma.[5]

This setting provides a unique opportunity to study physicians' heterogeneous response in reaction to negative news for a number of reasons. First, 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. Moreover, most physicians were caught by surprise by the FDA announcement, which lowered the attractiveness of MI hysterectomies. Second, the alternative procedure, open hysterectomy, is not risk free and is associated with relatively higher postoperative complications and was shown

---

[2] To our knowledge, this is the first economic study that uses choices between two surgical treatments to study physician practice variation. Previous literature often focuses on C-section versus normal delivery for childbirth (Currie & MacLeod 2017), or intensive therapy versus medical management for heart attack patients (Chandra & Staiger 2007; Currie *et al.* 2016).

[3] The MI hysterectomy normally uses a device called power morcellator to slice up the uterus, allowing doctors 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 post-operative complications.

[4] This might occur for many reasons, including the need for better visualization of the pelvis, or controlling bleeding during the procedure.

[5] Uterine sarcoma is a hidden type of cancer which has no obvious symptoms, and there are no tests or exams to detect them. Many cases of uterine sarcoma are diagnosed during or after the hysterectomy for what's thought to be benign fibroid tumors.

to be sensitive to physician skill. This further increases the potential heterogeneity in physician compliance decisions and implies the importance of examining relative skill.

Consistent with studies of many different medical technologies, who 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 hysterectomy and the corresponding increase in the share of open hysterectomy following the FDA safety communication is well documented in the medical literature (Barron *et al.* 2015; Harris *et al.* 2016; Wright *et al.* 2016). However, these studies do not provide an explanation for the high degree of heterogeneity in physicians' response to the news.

3. THEORETICAL FRAMEWORK

In this section, we develop a theoretical framework for understanding the role of technology-specific skills in determining procedural mix as well as altering in 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, in their appropriateness for different patients and levels of physicians' performance. 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 level.

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 to 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:

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

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 which 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 $(s_{ij}, e_i) = (s_{ji} + e_i)^q, p, q \in \mathbb{R}_+$[6]. Higher values of p correspond to higher levels of loss from mismatch, while higher values of q correspond to a lower impact of the mismatch. This formulation assumes perfect substitutability between 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

[6] 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.

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

Note that higher level 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 mismatch. This is represented using the following value function, V_j:

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

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

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

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

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

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

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

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, 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:

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

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

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

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

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

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.[7] 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 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.[8] 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.

The second empirical implication of the model is from equation (2). Since $\frac{d\bar{k}}{de_{MI}} > 0$, 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.

Third, equation (3) shows our main prediction: physicians with higher relative skills are less sensitive to the information shock, and are less likely to abandon MI and switch to O.

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

4. DATA AND MEASURES

Our empirical setting focuses on the surgical choice 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.[9] 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 FDA warning exogenously lowers the perceived attractiveness of MI hysterectomy for all physicians, allowing us to evaluate physician response to information shock.

Our data includes hospital inpatient discharges and outpatient visits for all patients who received either minimally invasive or open hysterectomies between January 2012 and September 2015 in the state of Florida.[10] 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 *Patient Appropriateness Measures*

To construct a measure of patient appropriateness for a minimally invasive hysterectomy, 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,

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

In Equation (5), X_{it} is a vector of patient characteristics. Following (Harris *et al.* 2016), we control for a set medical risks recorded for conditions present on admission, including the Charlson

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

[10] 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.

Comorbidity index (CCI) (Charlson *et al.* 1987),[11] severe pelvic adhesion, indicator of malignancy, morbid obesity, cervical dysplasia, and whether the principal diagnosis is uterine fibroid. Other patient characteristics are 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. We also include the year-quarter fixed effects $Time_t$ to capture the overall trend of the comparative effectiveness of the two procedures over time. The estimates are presented in Appendix Table A1. The model fits the data well, with a pseudo R-squared of 0.29. Patient *MI appropriateness* is then measured using fitted values from the logit regression: $MI\ appropriateness_{it} = \hat{G}(X_{it}\Phi + Time_t\Upsilon)$.

Figure 1 plots the distribution of *MI appropriateness* for those who received MI hysterectomies and for those who received open hysterectomies. The figure shows that among patients who received MI hysterectomies, density is concentrated at appropriateness levels between 0.7 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 0.8 (among whom 93% received MI hysterectomies), and *appropriate for open* if the MI appropriateness is below 0.3 (among whom 79% received open hysterectomies).[12]

4.2 *Physician Skill Measures*

To construct empirical measures of physician skill, we follow Harris *et al.* (2016) in defining a negative outcome as at least one major postoperative complication.[13] Appendix Table A2 presents the proportion of patients with postoperative complication rates separately for patients who received open and MI hysterectomies, both unadjusted and adjusted by patient characteristics. As expected, open hysterectomy involves higher probability of postoperative complications compared

---

[11] 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.

[12] Results are robust when we use alternative thresholds, including MI appropriateness above 0.7 for "*appropriate for MI*", and below 0.4 for "*appropriate for open.*"

[13] 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.

to MI procedures with risk adjustment shrinking the difference in complication rates between the two procedures, as open hysterectomies are typically performed on higher risk patients.

Next, we construct measures of physician absolute and relative procedural skill using two different methods. The first set of skill measures relies on risk-adjusted performance. We estimate a logistic regression model (Equation 6) to risk-adjust the incidence of any postoperative complication using observations before the FDA warning. The estimates are shown in Appendix Table A3.[14]

$$\Pr(\text{Complication}_{it}) = G(\alpha MI_{it} + X_{it}\Phi + MI_{it} \times X_{it}\Upsilon + Time_t\Psi + \theta_0) \quad (6)$$

For each patient i who receives hysterectomy in year-quarter t, Complication_{it} is an indicator variable for whether the patient has any 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 absolute skill of performing MI or open hysterectomy is measured using the difference between the predicted complication rate and the actual complication rate among her patients who are treated using the specific procedure. Similar to Currie and MacLeod (2017), relative skill is calculated as the difference between the two skill measures.

One concern with the above measure is that the choice of procedure is potentially endogenous to the physician's skill. We, therefore, employ a second measure of absolute skill by calculating the proportion of patients who do not have any postoperative complications among his or her patients who are appropriate for each procedure. Relative skill is again measured by the difference between MI skill and open skill.[15] We use this measure as our preferred measure of skill, as the appropriateness measure only depends on the underlying risk factors, and is unlikely to suffer from procedure choice endogeneity. The two relative skill measures are highly correlated.[16]

---

[14] Table A3's sample size is smaller than in Table A1 because, while we used all patients in the pre-FDA warning period to predict the probability of receiving MI and construct the MI appropriateness measure, in the rest of the paper we use the restricted sample as discussed in Section 4.2.
[15] This method is similar to Currie & MacLeod (2017).
[16] Correlation between the two relative skill measures is 0.48.

In order to compute these two relative skill measures, we restrict our sample to include 657 physicians who have treated at least one patient deemed *appropriate for MI* and one patient deemed *appropriate for open* (as defined in Section 4.1), and have performed at least one MI and one open hysterectomy, in the pre-information shock period.[17] Our sample includes 44,558 patients treated by these 657 physicians, spanning the eight calendar quarters before (2012 Q1-2013 Q4) and the six calendar quarters after (2014 Q2-2015 Q3) the information shock. We exclude the interim period—2014 Q1 because it is after the national news release (Dec 2013) and before the formal release of FDA warning (April 2014).

Table 1 presents the summary statistics of the final sample. Both skill measures exhibit considerable variation across physicians, with a mean of 0.026 and a standard deviation of 0.220 for the first measure, and a mean of 0.244 (given that complications are more frequent for patients appropriate for open than those appropriate for MI) and a standard deviation of 0.281 for the second measure. For the first measure, the 1st percentile is -0.422, while the 99th percentile is 0.733; for the second measure, the 1st percentile is -0.279 and the 99th percentile is 1. We normalize both skill measures using z-scores in the empirical analysis, for ease of interpretation.

4.3 *Heterogeneity in Relative Skill and Relative Performance*

To verify that the observed variation in relative performance represents heterogeneity in human capital and is not determined haphazardly, we use simulation focusing on the measure based on patient appropriateness - our preferred measure of relative skill.

The underpinning of the simulation is as follows. We calculate the pre-period complication rate among patients appropriate for open surgery (28%) and among patients appropriate for a minimally invasive surgery (3%). This provides us with a measure of overall relative skill in our surgeon sample. For each physician j, we observe her actual performance in open procedures - $Open_j$ (i.e., the non-complication or success rate among her patients who are *appropriate for open*) and her performance in MI procedures - MI_j (i.e., the non-complication rate among her patients who are *appropriate for MI*). We also observe the number of patients who are appropriate for open and MI procedures for each physician in our sample, N_j^{open} and N_j^{MI}.

[17] This excludes physicians who either treated only MI-appropriate patients or only open-appropriate patients in the pre-period. This restriction resulted in losing about 30% of patients, mostly treated by low volume physicians.

The goal of the simulation is to compare actual performance to the case were relative skill is fixed. Note that, we allow surgeons to exhibit heterogeneity in absolute skill. To achieve this we set the relative skill of each physician to the overall relative skill in our sample. Since the relative skill is being held constant across surgeons, any variation in the ranking of relative skill from the simulation should reflect variations in either absolute skill or luck.

Specifically, we impute physician j's MI skill by adding the overall relative skill to her observed open skill and as an alternative, we impute physician j's open skill by subtracting the overall relative skill from her observed MI skill.[18] Note that for both the observed and imputed measures, a significant proportion of physicians have zero MI complication rates (i.e., when treating patients who are *appropriate for MI*). These physicians' relative skill reflects solely their success in open procedures. We therefore exclude these physicians from the simulation analysis. This exclusion restricts our sample to 241 physicians who have non-zero imputed MI complication rates (designated as the *imputed-MI sample*) and 125 physicians who have non-zero observed MI complication rates (designated as the *imputed-Open sample*).

Next, for each physician j, we randomly draw N_j^{open} and N_j^{MI} uniform numbers and calculate the share where the random number is greater than $Open_j^{observed}$ and $MI_j^{imputed}$ (or $Open_j^{imputed}$ and $MI_j^{observed}$). This gives us the simulated open and MI skill: $Open_j^{simulated}$ and $MI_j^{simulated}$.

We performed 10 and 50 iterations with very similar results. Finally, we calculated the average simulated open and MI skill and ranked all physicians based on their average simulated skills for open and MI and generate. Based on these simulated open ranking and MI ranking, we calculated the simulated relative ranking by taking the difference between the simulated MI ranking and open ranking.[19]

In Figure 2, we show the scatter plot of the simulated MI vs. open ranking against the observed MI vs. open ranking as well as the overlay of the kernel densities of the simulated and the observed relative skill ranking using 50 iterations for simulation. Panel 1 shows results from the *imputed-*

[18] We set the imputed MI skill to one if it is greater than one; similarly, we set the imputed open skill to zero if it is less than zero.

[19] Note that when there is a tie in relative ranking, we placed surgeons with higher procedural volume above those with lower procedural volume.

MI sample and Panel 2 shows results for the *imputed-Open sample*. We also show results from 10 iterations in Appendix Figure A1.

If relative skill was not important, observations would line up along the 45^{o} line, very much in the way produced by the simulation, which allows for heterogeneity in absolute skill and randomness, but shuts down heterogeneity in relative performance across procedures. Nevertheless, the results show that the observed variation of relative skill ranking is much greater than that of the simulated one. Observations below the 45^{o} line represent surgeons with relative skill favoring open procedures while observations above the 45^{o} line represent surgeons with relative skill favoring MI procedures. The observed variation is, for the most part, not driven by random chance; instead, it reflects the heterogeneity in relative abilities and human capital, which is the focus of our empirical estimation in the following section.

5. EMPIRICAL ESTIMATION AND RESULTS

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 FDA warning by abandoning MI hysterectomy. Having constructed the physician skill measures in Section 4.2, we define physicians whose relative skill is above the median to be the *Top MI Performers*; those whose relative skill is below the median to be the *Top Open Performers*. We follow this definition in the remainder of the paper.

Figure 2 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*, using the first and the second skill measure in these definitions, respectively. Panel 1 in Figure 2 uses the full sample of 657 physicians. For robustness, we also use the subsample of 125 physicians who have non-zero complication rates when treating patients who are *appropriate for MI* and present the results in Panel 2. Figure 2 shows robust patterns of the reduction of MI utilization and confirm our basic prediction: before FDA warning, *Top Open Performers* are less likely to use MI hysterectomies, but the difference is small; after FDA warning, *Top Open Performers* are more likely to abandon MI, substantially increasing the utilization gap between the two groups of physicians.

One may wonder if the trend in Figure 2 is driven by patient selection over time, i.e., patients who are more appropriate for MI are more likely to choose *Top MI Performers* after FDA warning. To address this concern, we show in Table 2 Panel 1 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 FDA warning, respectively. First, patients treated by *Top Open Performers* do become slightly less appropriate for MI after FDA warning. However, the magnitude of this change is small and does not seem to affect our basic finding. A comparison between Columns (3) and (6) in Table 2 indicates that *Top Open Performers* are more likely to abandon MI after FDA warning, even under the adjustment of patient MI appropriateness.

Panel 2 in Table 2 presents the demeaned relative skills for physicians who treat each cohort of patients. The variation in relative skill as shown here allows us to identify the effect of relative skill on technology abandonment.

Next, we test whether physicians with higher relative skill are more likely to use MI, as predicted in Equation 1, using the two continuous measures of relative skill (as defined in Section 4.2). We employ the following linear probability model for patients who receive hysterectomies before the information shock to test this prediction.

$$MI_{ijkt} = a + b_1 Skill_j + b_2 X_{it} + b_3 County_k + b_4 Time_t + e_{it} \qquad (7)$$

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 physician relative skill. X_{it} are patient characteristics. We also include county fixed effects, $County_k$, to control for characteristics of the location that may affect MI utilization, and 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.

OLS estimation of Equation 7 may lead to biased estimates due to patient selection and measurement error with regards to physician skills. For example, if patients more appropriate for MI choose physicians with higher relative skill favoring MI over open, the resulting estimated effect of relative skill on MI utilization will be biased upward due to unobserved patient characteristics.[20] On the other hand, high MI utilizers who conducted a small number of open cases

[20] Table 2 suggests that the extent to which patient selection biases the result is minimal.

may obtain a high value for open skill simply because postoperative complications are rare. This possibility would bias the coefficient of relative skill downward. To address these identification issues, we instrument for each physician's relative skill using the weighted average skill of all other physicians who perform hysterectomies at the same hospital(s).[21] The argument here is that patient selection, if prevalent, will be less pronounced at the hospital level, compared with the physician level.[22] The statistical issue of measuring physician skill with errors due to small sample size is also alleviated at the hospital level. Previous literature has shown the spillover of practice style and outcomes within the same geographic area (Chandra & Staiger 2007) and the same hospital (Staats *et al.* 2018). The first stage results are presented in Column (2) of Table 3 for relative skill, Column (2) of Appendix Table A4 for MI skill and Column (2) of Appendix Table A5 for open skill. The estimated coefficient and the f-statistics of the excluded instrument in the first stage indicates that colleagues' skills are positively associated with the physician-level skill. The instruments are weak (f-statistics of are below 10) when we use the first skill measure. This issue may be because the first skill measure is conditional on procedure choice, and therefore may mitigate the effect of a bad outcome. The instruments predict well for the second skill measure, with f-statistics above 10. We, therefore, focus on the IV estimations using the second skill measure.

The OLS and IV estimates of Equation 7 are presented in Table 3 Columns (1) and (3), suggesting that higher relative skill (MI relative to open) is associated with higher MI utilization. This result confirms the basic prediction of our theoretic model, showing that patients are sorted into each procedural type based on the relative return. Using our preferred skill measure (i.e., the second skill measure), the IV estimates in Column 3 of Panel 2 suggest that a one standard deviation increase in relative skill is associated with 11.8 percentage points increase (a 17.3% increase) in the probability of utilizing MI before the information shock.

Next, we substitute absolute skill measures of performing MI or open hysterectomies for the relative skill measures in Equation 7. We re-estimate the OLS and IV models and present the

[21] The weights are given by the total number of hysterectomies by each physician.
[22] On average, each surgeon has 13 peer physicians performing hysterectomies at her hospital.

results in Columns (1)-(3) in Appendix Table A4 and A5.[23] Overall, MI utilization appears to be more sensitive to relative skill than to absolute MI or open skill.

Having established the association between relative skill and MI utilization, we turn to assess the extent to which relative skill affects physician technology abandonment in response to a 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.

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

where MI_{ijkt}, $Skill_j$, X_{it}, $Time_t$ and e_{it} have the same definitions as in Equation 7. $Post_t$ indicates whether the patient receives the hysterectomy after 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. The coefficient of interest is b_1, which is expected to be positive for a physician with higher relative skill. Z_j represents physician characteristics other than skill that may also affect technology abandonment. First, we include the physician's average MI share in the pre-period (i.e., 2012-2013), $Share\_MI1213_j$ 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., number of MI hysterectomies performed by the physician) in the pre-period, $Volume\_MI1213_j$, to be a potential factor influencing physician response to negative news (Staats *et al.* 2018).

Similar to the previous cross-physician analysis, we instrument each physician's skill using the weighted average skill of all other physicians performing at the same hospital to address the issue of dynamic patient selection and measurement errors.[24] The first stage results are shown in Column (7) of Table 3, Appendix Table A4 and A5, for relative skill, MI skill, and open skill respectively. The OLS and IV results for the effect of relative skill on MI abandonment, without and with the control of pre-FDA warning MI share and MI volume, are shown in Table 3, Columns (4)-(6) and Columns (8)-(10), respectively. The results are highly robust with the inclusion of MI share and

[23] In the IV models, physician MI or open skill is instrumented using the corresponding weighted means of skills among all other physicians who perform at the same hospital(s).
[24] Table 2 shows the extent to which dynamic patient selection biases the results is minimal.

MI volume. Using our preferred measure of skill (i.e., the second skill measure), the IV estimates in Columns (8)-(10) of Panel 2 in Table 3 suggest that a one standard deviation increase in relative skill would reduce the probability of MI abandonment by 7.4-9.7 percentage points (an 11%-14% reduction). The effect of absolute skill on MI abandonment from OLS and IV estimations are presented in Appendix Table A4 and A5, Columns (4)-(6) and Columns (8)-(10). These results jointly suggest that MI abandonment responds more to relative skill than to absolute skills, which is consistent with the prediction in our theory (Equation 3).

6. CONCLUSIONS

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 received ample attention in the economics literature; it received 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 interact with the choice of procedural mix and 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, often times she will face a choice between mutually exclusive procedural alternatives, each requires a set of specific skills and confers specific costs and benefits. The physician relative skill level is therefore tied to the notion of implicit opportunity cost in production. If a surgeon is equally skilled at performing open procedures compared with minimally invasive ones, her implicit opportunity cost (or value) of choosing to perform minimally invasive surgery is relatively low. Negative information shocks regarding minimally invasive procedures, in this context, serve to raise the implicit opportunity cost of performing minimally invasive interventions, which in turn lowers their prevalence.

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 very broad definition, which 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 drawn to 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 by (Chandra *et al.* 2011; Chandra *et al.* 2015), we show that a key barrier to translating the diffusion of evidence into the practice of medicine and the procedural choice is skill heterogeneity. Therefore, enforcing the uniform practice of medicine that ignores the heterogeneity in human capital may reduce welfare.

Table 1: Summary statistics of patient and physician characteristics

	Mean	SD
***Panel 1**: Patient Characteristics*		
Received MI	0.659	0.474
Having Any Postoperative Complication	0.105	0.307
Age	49.5	11.7
Emergent	0.017	0.131
Charlson Comorbidity Index	0.176	0.755
Cancer Indicator	0.007	0.090
Pelvic Floor Adhesion	0.067	0.250
Cervical Dysplasia	0.002	0.045
Morbid Obesity	0.030	0.172
Principal diagnosis is Uterine Fibroid	0.393	0.488
Number of Other Diagnoses at admission	1.778	2.571
White	0.684	0.465
Uninsured	0.049	0.215
Medicare	0.168	0.374
Medicaid	0.088	0.283
***Panel 2**: Physician Characteristics*		
MI Skill (Measure 1)	-0.001	0.103
Open Skill (Measure 1)	-0.027	0.211
Relative Skill (Measure 1)	0.026	0.220
MI Skill (Measure 2)	0.964	0.108
Open Skill (Measure 2)	0.720	0.257
Relative Skill (Measure 2)	0.244	0.264
MI Volume (Pre-period)	76	79
Open Volume (Pre-period)	28	30
Total Volume (Pre-period)	103	94
IV – MI Skill of all other physicians in the same hospital(s)	-0.002	0.054
IV – Open Skill of all other physicians in the same hospital(s)	-0.005	0.126
IV – Relative Skill of all other physicians in the same hospital(s)	0.003	0.122
N	44,558	

Notes: Mean and standard deviations are reported. Observations are at the patient level. The sample excludes patients treated by physicians who performed too few MI or open hysterectomies before FDA warning, and who have treated two few patients appropriate for MI or open, for a measure of relative skill to be computed. Measure 1 of physician skill is based on risk-adjusted performance; Measure 2 of physician skill is based on patient appropriateness.

Table 2: Means of MI utilization, MI appropriateness, and physician characteristics by physician relative skill.

	Patients Treated by *Top MI Performers*			Patients Treated by *Top Open Performers*		
	Pre-FDA Warning (2012Q1-2013Q4)	Post-FDA Warning (2014Q2-2015Q3)	Percentage Change in MI Rate	Pre-FDA Warning (2012Q1-2013Q4)	Post-FDA Warning (2014Q2-2015Q3)	Percentage Change in MI Rate
	(1)	(2)	(3)	(4)	(5)	(6)
***Panel 1**: Patient-level Statistics*						
Average MI appropriateness	0.640	0.612		0.630	0.593	
Percentage of patients received MI (Unadjusted)	0.693	0.662	-4.5%	0.671	0.583	-13.1%
Percentage of patients received MI (Adjusted by appropriateness)	0.678	0.670	-1.2%	0.664	0.607	-8.6%
***Panel 2**: Physician-level Statistics*						
Relative skill—Measure 1,	0.086	0.105		-0.097	-0.089	
Demeaned	(0.206)	(0.214)		(0.182)	(0.186)	
Relative skill—Measure 2,	0.248	0.215		-0.214	-0.210	
Demeaned	(0.161)	(0.194)		(0.111)	(.108)	
N	13,456	8,724		13,714	8,664	

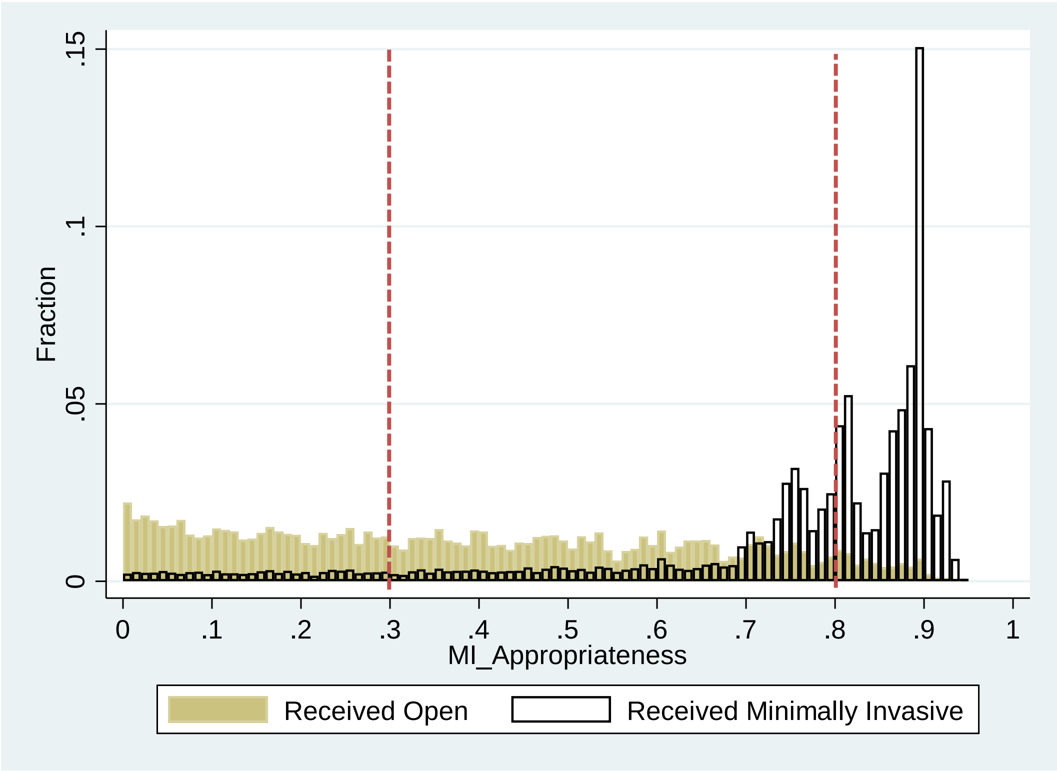
Notes: Standard deviations are in parentheses. We use the preferred measure, the second skill measure which is based on patient appropriateness in the definitions of *Top MI Performers* and *Top Open Performers*.

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

DV=Received MI	Cross-physician Analysis			Within Physician Analysis						
	OLS	First Stage	IV (2SLS)	OLS			First Stage	IV (2SLS)		
***Panel 1**: Skill Measure 1*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Relative Skill	0.026***		0.083*							
	(0.009)		(0.045)							
Relative Skill*Post				0.025***	0.030***	0.021***		0.055**	0.073***	0.054**
				(0.006)	(0.006)	(0.007)		(0.024)	(0.027)	(0.027)
MI_Share_1213*Post					-0.106***				-0.134***	
					(0.023)				(0.029)	
Log(MI_Volume_1213)*Post						0.009				0.002
						(0.006)				(0.006)
IV: Other's Relative Skill		1.980**								
		(0.924)								
IV: Other's Relative Skill * Post							2.473***			
							(0.846)			
County FE	Y	Y	Y	N	N	N	N	N	N	N
Physician FE	N	N	N	Y	Y	Y	Y	Y	Y	Y
N	27,170	27,170	27,170	44,558	44,558	44,558	44,558	44,558	44,558	44,558
R-squared	0.314	0.177	0.301	0.487	0.488	0.487	0.512	0.486	0.486	0.486
***Panel 2**: Skill Measure 2*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Relative Skill	0.022***		0.118**							
	(0.008)		(0.057)							
Relative Skill*Post				0.025***	0.028***	0.021***		0.074***	0.097***	0.078**
				(0.006)	(0.007)	(0.007)		(0.029)	(0.033)	(0.036)
MI_Share_1213*Post					-0.099***				-0.132***	
					(0.022)				(0.032)	
Log(MI_Volume_1213)*Post						0.009				-0.004
						(0.005)				(0.009)
IV: Other's Relative Skill		1.405***								
		(0.422)								
IV: Other's Relative Skill * Post							1.841***			
							(0.400)			
County FE	Y	Y	Y	N	N	N	N	N	N	N
Physician FE	N	N	N	Y	Y	Y	Y	Y	Y	Y
N	27,170	27,170	27,170	44,558	44,558	44,558	44,558	44,558	44,558	44,558
R-squared	0.313	0.156	0.277	0.487	0.487	0.487	0.476	0.484	0.483	0.484

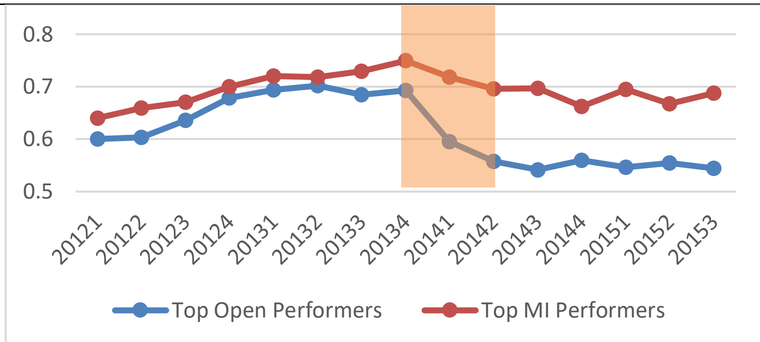
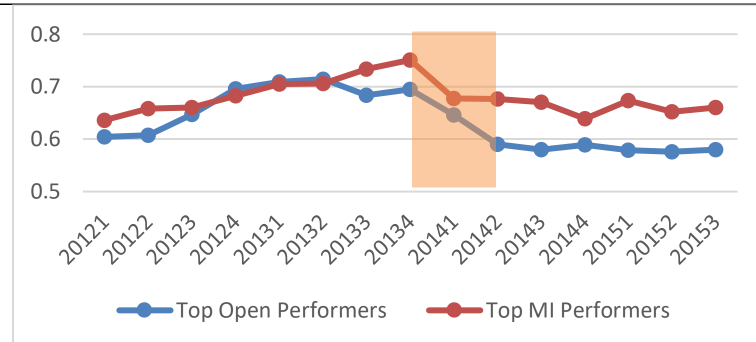
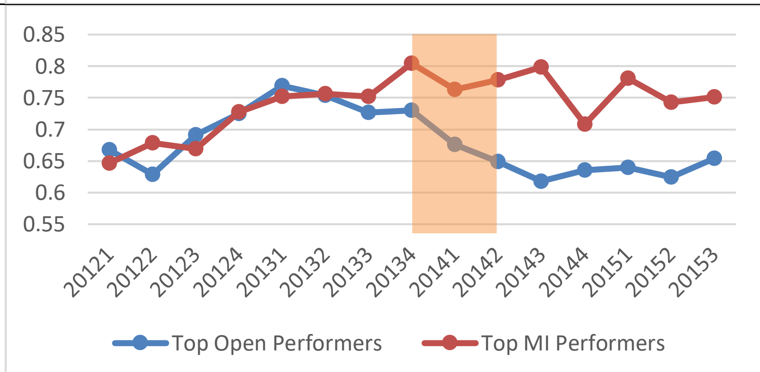
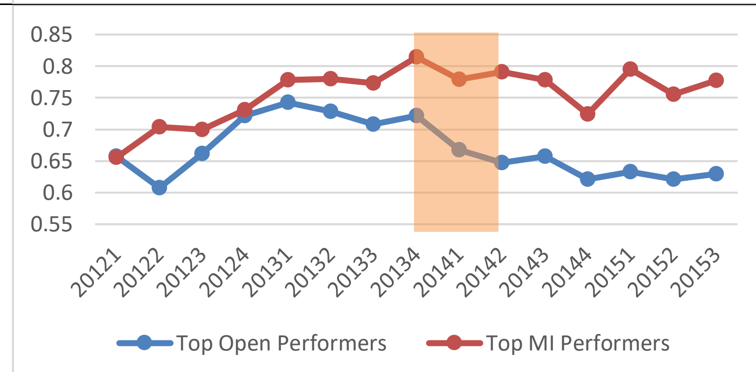
Notes: Results are from linear probability estimations. Relative skills measure the skill level of performing MI hysterectomy relative to Open hysterectomy. All skill measures are z-scored. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

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



Note: MI Appropriateness is measured by the predicted probability of receiving MI from the logistic regression shown in Appendix Table A1.

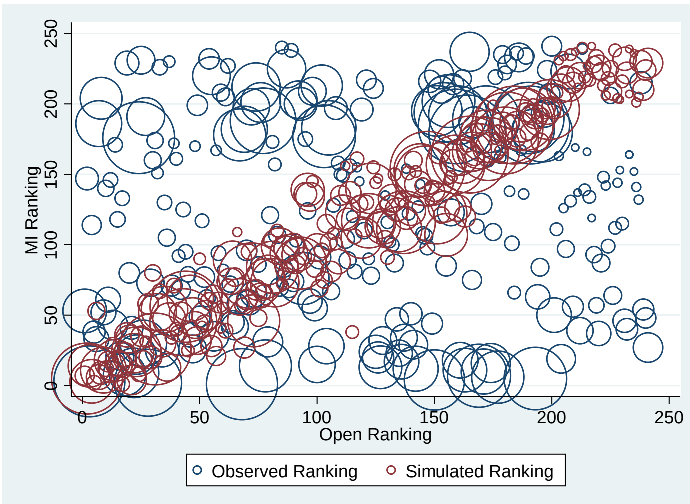
Figure 2: Quarterly MI utilization rate by physician relative skill.

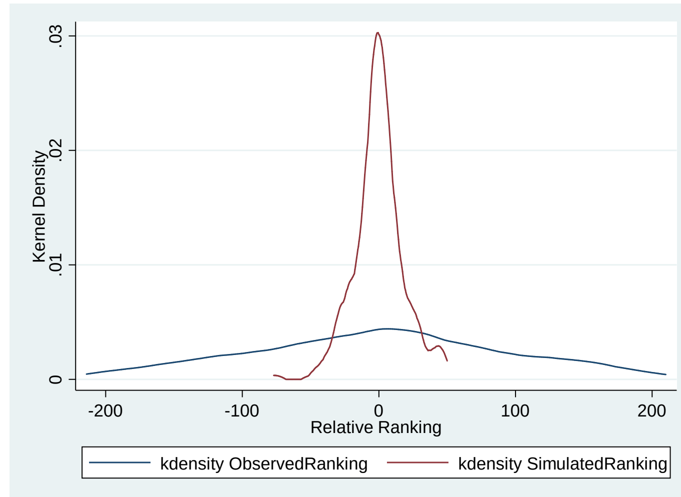
<table>
<tr><td></td><td>Measure 1 of Relative Skill</td><td>Measure 2 of Relative Skill</td></tr>
<tr><td>Panel 1:
Full Sample (No. of Physicians =657)</td><td></td><td></td></tr>
<tr><td>Panel 2:
Subsample excluding physicians who have zero complication rates for patients appropriate for MI.
(No. of Physicians = 125)</td><td></td><td></td></tr>
</table>

Notes: The shaded bar denotes when the national news report started on Dec. 2013, leading to the FDA safety communication released on April 17, 2014.

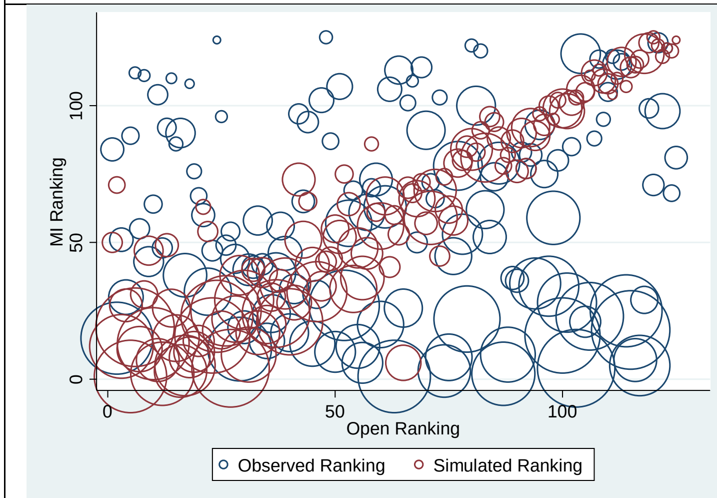
Figure 3: Actual vs. simulated skill ranking – No. of Iterations = 50.

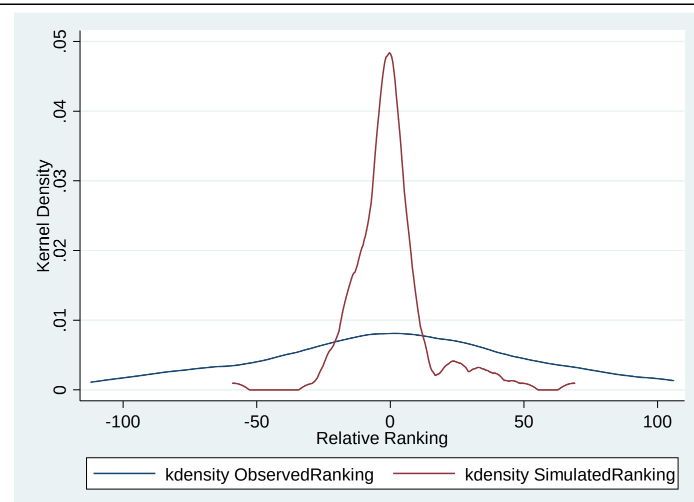
Panel 1: Physicians with non-zero simulated MI complication rates. (No. of Physicians = 241)





Panel 2: Physicians with non-zero observed MI complication rates. (No. of Physicians = 125)





Notes: In the scatter plot graphs, each bubble represents an observed or simulated physician, weighted by the total number of MI and open patients treated by the physician.

Appendix Table A1: Logistic regression model for the probability of receiving MI in the pre-period.

Pr (MI=1)	Coefficient (SE)	Marginal Effect
Age	-0.022**	-0.005**
	(0.010)	
Age_squared	0.000**	0.00006**
	(0.000)	
Emergent	-0.370***	-0.088***
	(0.104)	
Charlson Comorbidity Index	-0.204***	-0.049***
	(0.018)	
Cancer Indicator	-0.371*	-0.089*
	(0.205)	
Pelvic Floor Adhesion	-0.530***	-0.126***
	(0.060)	
Cervical Dysplasia	-0.104	-0.025
	(0.301)	
Morbid Obesity	0.239**	0.060**
	(0.101)	
Fibroid	-0.702***	-0.170***
	(0.028)	
Number of Other Diagnoses at admission	-0.509***	-0.121***
	(0.010)	
White	0.345***	0.082***
	(0.028)	
Uninsured	-1.134***	-0.271***
	(0.051)	
Medicare	0.237***	0.057***
	(0.060)	
Medicaid	-0.673***	-0.161***
	(0.044)	
N	36,137	
Pseudo R2	0.29	

Notes: The regression also includes year-quarter fixed effects. Standard errors in parenthesis are clustered by physician.
*** p<.01, ** p<.05, * p<.1.

Appendix Table A2: Complication rate (i.e., the proportion of patients with any postoperative complication) by procedural type.

	Pre-FDA Warning (2012Q1-2013Q4)			Post-FDA Warning (2014Q2-2015Q3)		
	Complication Rate (Unadjusted)	Complication Rate (Adjusted)	N	Complication Rate (Unadjusted)	Complication Rate (Adjusted)	N
Received Open	24.0%	14.5%	8,646	19.6%	12.7%	6,568
Received MI	5.0%	7.1%	18,524	3.7%	5.5%	10,820

Appendix Table A3: Logistic regression of postoperative complication on patient risk factors and procedure type in the pre-period.

DV=Any Postoperative Complication	Coefficient (SE)	Marginal Effect
MI	-1.805***	-0.119***
	(0.077)	
Charlson Comorbidity Index	0.229***	0.015***
	(0.016)	
MI * Charlson Comorbidity Index	-0.074**	-0.005**
	(0.037)	
Cancer Indicator	0.217	0.014
	(0.197)	
MI * Cancer Indicator	0.202	0.014
	(0.433)	
Pelvic Floor Adhesion	-0.052	-0.003
	(0.076)	
MI * Pelvic Floor Adhesion	0.450***	0.029***
	(0.154)	
Cervical Dysplasia	-0.488	-0.032
	(0.479)	
MI * Cervical Dysplasia	0.288	0.019
	(0.946)	
Morbid Obesity	0.130	0.009
	(0.107)	
MI * Morbid Obesity	-0.193	-0.013
	(0.215)	
Fibroid	-0.118**	-0.008**
	(0.058)	
MI*Fibroid	0.484***	0.032***
	(.092)	
Number of Other Diagnoses at admission	0.115***	0.008***
	(0.010)	
MI * Number of Other Diagnoses at admission	0.152 ***	0.010***
	(0.017)	
Age	-0.041***	-0.003***
	(0.012)	
Age_Squared	.0004***	0.00003***
	(.0001)	
White	-0.103**	-0.007**
	(0.045)	
Emergent	0.668***	0.044**
	(0.108)	
Uninsured	0.167***	0.011**
	(0.079)	
Medicare	-0.144*	-0.010*
	(0.076)	
Medicaid	-0.140*	-0.10*
	(0.072)	
N	27,170	
Pseudo R2	0.18	

Notes: The regression also includes year-quarter fixed effects. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

Appendix Table A4: The impact of physician MI skill on the utilization of MI hysterectomy.

DV=Received MI	Cross-sectional Analysis			Within Physician Analysis						
	OLS	First Stage	IV (2SLS)	OLS			First Stage	IV (2SLS)		
***Panel 1**: Skill Measure 1*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
MI Skill	-0.004		-0.038							
	(0.012)		(0.038)							
MI Skill*Post				0.023***	0.024***	0.021***		0.057***	0.057***	0.055**
				(0.008)	(0.007)	(0.008)		(0.022)	(0.022)	(0.022)
MI_Share_1213*Post					-0.088***				-0.091***	
					(0.022)				(0.024)	
Log(MI_Volume_1213)*Post						0.011**				0.009
						(0.005)				(0.006)
IV: Other's MI Skill		5.289***								
		(1.306)								
IV: Other's MI Skill * Post							5.370***			
							(1.078)			
County FE	Y	Y	Y	N	N	N	N	N	N	N
Physician FE	N	N	N	Y	Y	Y	Y	Y	Y	Y
N	27,170	27,170	27,170	44,558	44,558	44,558	44,558	44,558	44,558	44,558
R-squared	0.312	0.171	0.307	0.487	0.487	0.487	0.481	0.486	0.486	0.486
***Panel 2**: Skill Measure 2*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
MI Skill	0.017*		-0.047							
	(0.010)		(0.050)							
MI Skill*Post				0.015*	0.018**	0.013		0.063**	0.062**	0.061**
				(0.008)	(0.007)	(0.008)		(0.026)	(0.025)	(0.026)
MI_Share_1213*Post					-0.094***				-0.113***	
					(0.023)				(0.028)	
Log(MI_Volume_1213)*Post						0.012**				0.009
						(0.005)				(0.006)
IV: Other's MI Skill		4.275***								
		(1.361)								
IV: Other's MI Skill * Post							7.085***			
							(1.607)			
County FE	Y	Y	Y	N	N	N	N	N	N	N
Physician FE	N	N	N	Y	Y	Y	Y	Y	Y	Y
N	27,170	27,170	27,170	44,558	44,558	44,558	44,558	44,558	44,558	44,558
R-squared	0.313	0.194	0.296	0.486	0.487	0.487	0.447	0.484	0.485	0.485

Notes: Results are from linear probability estimations. All skill measures are z-scored. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

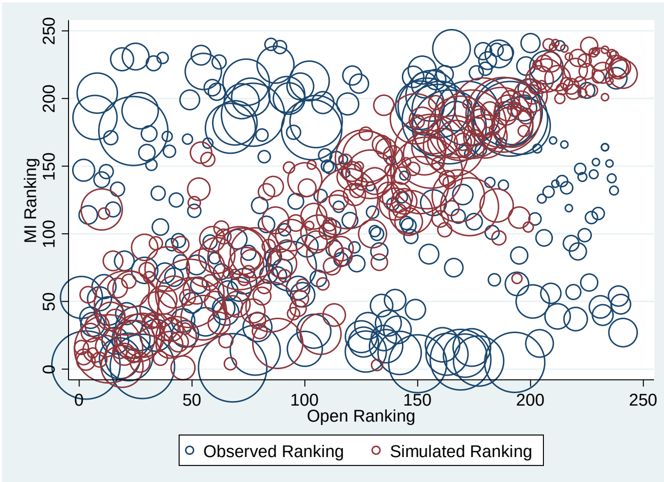
Appendix Table A5: The impact of physician Open skill on the utilization of MI hysterectomy.

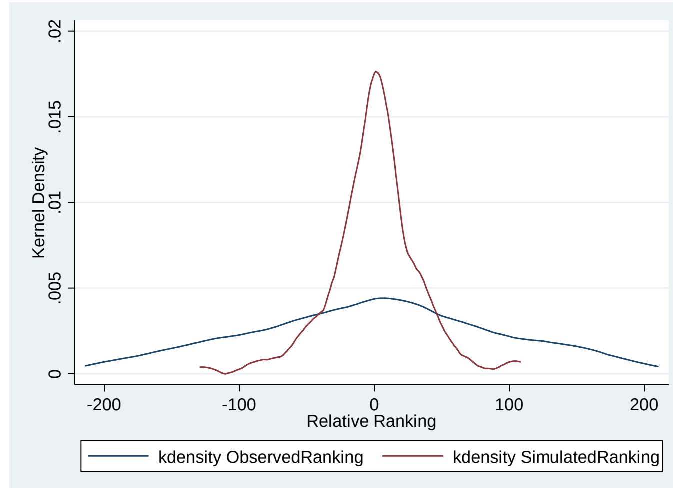
DV=Received MI	Cross-sectional Analysis			Within Physician Analysis						
	OLS	First Stage	IV (2SLS)	OLS			First Stage	IV (2SLS)		
Panel 1*: Skill Measure 1*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Open Skill	-0.029*** (0.008)		-0.074** (0.032)							
Open Skill*Post				-0.015** (0.006)	-0.019*** (0.006)	-0.011 (0.007)		-0.026* (0.016)	-0.039** (0.017)	-0.021 (0.017)
MI_Share_1213*Post					-0.098*** (0.023)				-0.111*** (0.025)	
Log(MI_Volume_1213)*Post						0.011** (0.006)				0.009* (0.006)
IV: Other's Open Skill		2.524*** (0.812)								
IV: Other's Open Skill * Post							2.838*** (0.783)			
County FE	Y	Y	Y	N	N	N	N	N	N	N
Physician FE	N	N	N	Y	Y	Y	Y	Y	Y	Y
F-Statistics on Excluded IV		9.7					13.1			
N	27,170	27,170	27,170	44,558	44,558	44,558	44,558	44,558	44,558	44,558
R-squared	0.315	0.216	0.307	0.486	0.487	0.487	0.533	0.486	0.487	0.487
Panel 2*: Skill Measure 2*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Open Skill	-0.015* (0.008)		-0.104** (0.043)							
Open Skill*Post				-0.019*** (0.007)	-0.021*** (0.007)	-0.015** (0.007)		-0.036* (0.021)	-0.051** (0.023)	-0.030 (0.023)
MI_Share_1213*Post					-0.093*** (0.022)				-0.102*** (0.025)	
Log(MI_Volume_1213)*Post						0.010* (0.005)				0.007 (0.006)
IV: Other's Open Skill		1.800*** (0.393)								
IV: Other's Open Skill * Post							2.083*** (0.369)			
County FE	Y	Y	Y	N	N	N	N	N	N	N
Physician FE	N	N	N	Y	Y	Y	Y	Y	Y	Y
F-Statistics on Excluded IV		21.0					31.9			
N	27,170	27,170	27,170	44,558	44,558	44,558	44,558	44,558	44,558	44,558
R-squared	0.312	0.165	0.280	0.487	0.487	0.487	0.486	0.486	0.486	0.487

Notes: Results are from linear probability estimations. All skill measures are z-scored. Standard errors in parenthesis are clustered by physician. *** p<.01, ** p<.05, * p<.1.

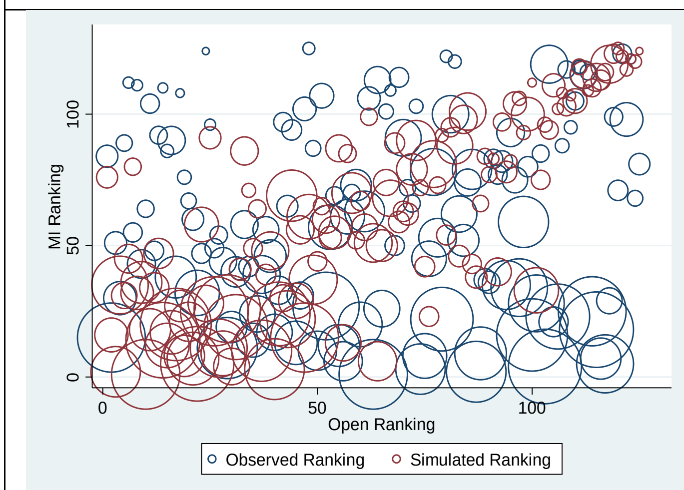
Appendix Figure A1: Actual vs. simulated skill ranking – No. of Iterations = 10.

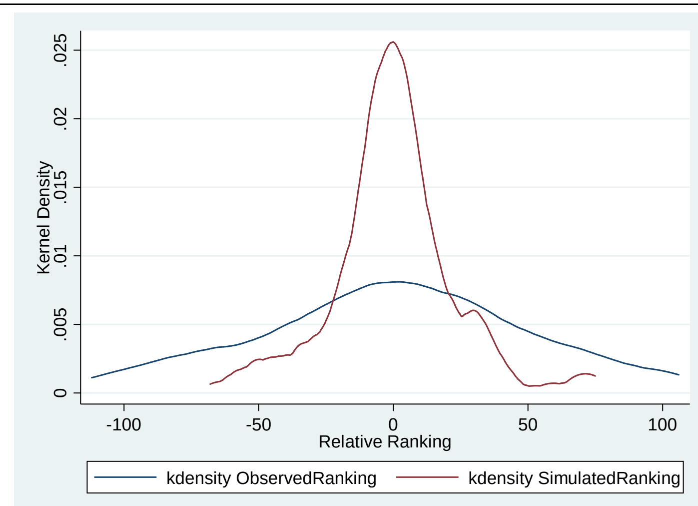
Panel 1: Physicians with non-zero simulated MI complication rates. (No. of Physicians = 241)





Panel 2: Physicians with non-zero observed MI complication rates. (No. of Physicians = 125)





Notes: In the scatter plot graphs, each bubble represents an observed or simulated physician, weighted by the total number of MI and open patients treated by the physician

References

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, 1174-1180

Baumgardner, J.R., 1988. Physicians' services and the division of labor across local markets. *Journal of Political Economy* 96, 948-982

Becker, G.S., Murphy, K.M., 1992. The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics* 107, 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., David, H., Neuman, M., 2018. Does Bad News Travel Faster? On the Determinants of Medical Technology Abandonment. *Journal of Human Capital* (Forthcoming)

Chandra, A., Jena, A.B., Skinner, J.S., 2011. The pragmatist's guide to comparative effectiveness research. *Journal of Economic Perspectives* 25, 27-46

Chandra, A., Khullar, D., Lee, T.H., 2015. Addressing the challenge of gray-zone medicine. *New England Journal of Medicine* 372, 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, 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, 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, 1-43

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

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, 24-44

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, 98. e1-98. e13

Heckman, J.J., Honoré, B.E., 1990. The Empirical Content of the Roy Model. *Econometrica* 58, 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, 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, 1309-1315

MacDonald, G., Marx, L.M., 2001. Adverse specialization. *Journal of Political Economy* 109, 864-899

Miller, R.A., 1984. Job matching and occupational choice. *Journal of Political Economy* 92, 1086-1120

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, 159

Roman, B.R., Asch, D.A., 2014. Faded promises: the challenge of deadopting low-value care. *Annals of Internal Medicine* 161, 149-150

Rosen, S., 1983. Specialization and human capital. Journal of Labor Economics 1, 43-49

Roy, A.D., 1950. The distribution of earnings and of individual output. *The Economic Journal* 60, 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, 2081-2084

Silvennoinen, M., Mecklin, J.-P., Saariluoma, P., Antikainen, T., 2009. Expertise and skill in minimally invasive surgery. *Scandinavian Journal of Surgery* 98, 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, 804-824

Ubel, P.A., Asch, D.A., 2015. Creating value in health by understanding and overcoming resistance to de-innovation. *Health Affairs* 34, 239-244

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, 877-878