DOES PRACTICE MAKE PERFECT: AN EMPIRICAL ANALYSIS OF LEARNING-BY-DOING IN CARDIAC SURGERY

Subramaniam Ramanarayanan[1,2]
Kellogg School of Management
Northwestern University
(Job Market Paper)

November 11, 2006

Abstract

Although extensive research has documented the benefits of increased organizational experience on performance (also known as learning-by-doing), our understanding of how learning occurs is still limited. One of the factors that have been identified to affect the rate at which a firm learns is the competence of individual workers. In this paper I examine a mechanism through which individual workers acquire (or maintain) competence, namely that of experience. Specifically, I analyze whether cardiac surgeons who perform more procedures experience an improvement in performance. In order to do so, I develop an instrumental variables estimation method that addresses the potential endogeneity of surgeon experience (measured by recent procedure volume). As my identification strategy, I consider exogenous shocks to the procedure volume of CABG surgeons in Florida caused by the exit of other surgeons from the same hospital. Using this instrument, I find evidence indicating a strong learning-by-doing effect for cardiac surgeons: an additional procedure a year leads to a reduction in the probability of patient mortality by .05 percentage points. Further, I find this improvement in surgeon performance to be completely transferable across different hospital settings, and find evidence of some economies of scope among the different surgical procedures performed by a cardiac surgeon.

---

[1] I am grateful to David Dranove and Leemore Dafny for their advice and invaluable support and encouragement. I would also like to thank David Besanko, Shane Greenstein, Tom Hubbard, Mike Mazzeo, Hayagreeva Rao, Joel Shalowitz, Scott Stern, seminar participants at the Kellogg School of Management and conference participants at the 2006 meeting of the American Society of Health Economists for helpful comments and suggestions. I would also like to thank Patricia Ledesma for help with formatting.

[2] Email: s-ramanarayanan@kellogg.northwestern.edu

1 Introduction

The effects of organizational experience on quality and costs have been studied in multiple settings. The realization of productivity gains with increased experience is termed learning-by-doing and the presence of a learning curve has been well documented in manufacturing and service firms. However, the extent of our knowledge about why learning occurs is still limited. For example, learning could result from the acquisition of human capital by individual employees or from increased organizational experience that helps standardize production processes. In this paper, I analyze one of the determinants of worker competence in an organizational setting: task experience. I use the term experience here to connote frequency of performing a certain task, as opposed to number of years spent on the job.

In industries employing skilled labor, learning is thought to result mainly from workers becoming more efficient at the tasks they perform through multiple repetitions. I study learning-by-doing at the level of the individual worker in one such setting, specifically that of cardiac surgeons performing Coronary Artery Bypass Graft (CABG) surgeries in hospitals. In particular, I estimate the extent to which a surgeon's recent procedure volume (measured as the number of CABG surgeries performed the previous year) affects performance (measured by patient outcomes). In addition, the setting of this study enables me to test for the degree of specificity (to a firm and to a task) of human capital that is acquired by individuals through learning-by-doing.

Focusing on individual experience provides interesting insights into the learning process of firms. Recent research in the field of organizational learning has focused on the differences in learning rates across organizations (e.g. Pisano et al (2001); Reagans et al (2005)), and proficiency of individual workers has been identified as one of the factors driving this variation. The results from this study can also help shape a firm's strategy on hiring and retention of skilled workers and on allocation of workload across workers. In the health care setting, provider experience has long been used as a proxy for quality by patients, health care providers, surgical accreditation boards and insurers. For instance, the American College of Surgeons states in its guidelines for Coronary Artery Bypass Grafts (CABG) that in order to maintain competence, a surgeon should perform a minimum of 200 procedures per year as the primary operator. Procedure volume also plays a prominent role in hospital marketing brochures and websites, suggesting that it is a metric widely used within the industry to signal quality.

Given the significance attached to provider experience in health care, there is surprisingly little convincing empirical evidence that establishes the presence of learning-by-doing in this setting. Estimates from prior studies are confounded by a fundamental problem of identification: Does experience result in learning or does quality reflect unobserved skills resulting in greater demand, and thus greater experience? In other words, do surgeons improve with experience or are high quality surgeons more experienced because they attract more patients? Existing studies in the healthcare literature assume the presence of a correlation between surgeon experience and quality as being indicative of a learning effect. Failure to account for reverse causality in these papers leads to biased estimates of learning-by-doing.

I use an instrumental variables approach in this study to establish a causal relationship between surgeon procedure volume and patient outcomes. I propose the use of surgeon exit as an exogenous identifier, where the term “exit” connotes any instance where a surgeon stops performing CABG surgeries in Florida. I use data on surgeon characteristics to restrict the set of exitors to surgeons aged 55 and over. Doing so mitigates concerns about possible endogeneity of exit. The rationale behind the instrument is as follows: once exit occurs, patients who would have been treated by the exiting surgeon are now allocated among the remaining (non-exiting) surgeons at the hospital, thereby providing a positive shock to their procedure volumes. By looking at the change in surgeon quality resulting from this shock, I am able to disentangle the learning-by-doing effect from other potential explanations underlying the volume-outcome relationship. I also provide empirical evidence in support of the relevance and exogeneity of the instrument, i.e. I show that the volume shock resulting from exit acts as a good predictor of expected change in (staying) surgeon procedure volumes and is exogenous to changes in surgeon quality.

I make use of a patient-level dataset from the Agency of Health Care Administration (AHCA) in Florida that identifies the unique license number of the operating surgeon. By using the volume of exiting surgeons as an instrument for surgeon procedure volume, I find evidence in support of a strong learning-by-doing effect for CABG surgeons: performing a single additional procedure in the prior year yields a reduction in the probability of an adverse patient outcome of .051 percentage points, which translates to a 1.2% drop (relative to the average mortality rate for CABG of 3.82%). This result is fairly robust to alternate definitions of exit, alternate empirical specifications and to different methods of allocating the procedure volume of exitors among non-exiting surgeons. Also, surgeons performing a large volume of procedures benefit less from additional experience when compared to surgeons with low procedure volume.

I extend this methodology to examine the degree of specificity (to a firm and to a task) of human capital acquired by surgeons as a result of learning-by-doing. In order to test for firm-specificity, I exploit the fact that cardiac surgeons are affiliated with (and perform procedures at) multiple hospitals at the same point in time. This allows me to use hospital-specific measures of surgeon experience[3] and enables me to test whether surgeon performance is affected more by experience at the current hospital than by experience at other hospitals. In order to resolve problems of endogeneity, I instrument for hospital-specific experience using a simple modification of the instrument described above. The estimates from this specification indicate that the improvement in surgeon performance is completely transferable across different hospital settings. Since cardiac surgeons perform procedures other than CABG surgeries, I am also able to test whether experience with non-CABG procedures benefits surgeon performance in CABG by using an estimation strategy similar to the one used to test for firm-specificity. The results indicate that greater experience performing non-CABG procedures does benefit outcomes of CABG patients, but this effect is smaller than the effect of experience performing CABG procedures, i.e. there is some degree of task-specificity to the human capital of surgeons gained from procedure experience.

Taken together, these findings have important implications for managers within the hospital industry, specifically with respect to how to organize the firm so as to better leverage the skills of its professional workers. While allocating procedure volume within a team, it is important for managers to bear in mind that the marginal benefit of additional experience is greater for low volume surgeons when compared to high volume surgeons. Hospitals should

---

[3] Huckman and Pisano (2006) were the first to test for firm-specificity of experience in this setting. In Section 2.3, I outline the main differences between their approach and mine.

ensure that low-volume surgeons get to develop mentoring relationships with more experienced surgeons. The findings on firm- and task-specificity have implications for hiring and retention policies and optimal job design, respectively.

This study contributes to existing empirical studies on learning-by-doing, especially those with a focus on health care (e.g. Hughes et al (1987); Hannan et al (1991); Gowrisankaran et al (2006); Ho (2002)). Most existing studies in this literature fail to adequately account for alternate explanations underlying the observed correlation between experience and outcomes. The primary contribution of this paper is that it is able to establish a causal link between provider experience and patient outcomes by using an instrumental variables technique. Also, the lengthy study period (1998-2003) helps me overcome some of the data limitations experienced by prior researchers. The study also complements recent theoretical and empirical research on firm-specific and task-specific aspects of human capital (e.g. Huckman and Pisano (2006); Gibbons and Waldman (2004)).

The rest of this paper is organized as follows. The next section provides a review of three streams of literature related to this study, and outlines some of the shortcomings of existing empirical research on learning-by-doing in health care. Section 3 describes the data and provides some institutional background. Section 4 outlines the empirical approach followed and describes the instrument in detail. Section 5 presents the main empirical specifications. The final two sections contain a discussion of results and some concluding remarks.

2 Background and Related Research

The present study draws from multiple streams of related research – studies of learning-by-doing in the industrial organization and strategy literatures, human capital theory and volume-outcome studies in the fields of medicine and health economics. I use insights from all these literatures in my model, and aim to make a contribution to each of them.

2.1 The Benefits of Experience

Starting with Wright's (1936) seminal analysis of airframe production, there has been a significant body of empirical and theoretical research aimed at documenting the association between cumulative experience (typically measured by cumulative production volume) and performance improvement across different industries. Some recent settings in which learning curves have been studied include aircraft production (Benkard (2000)), shipbuilding (Argote, Beckman and Epple (1990), Thompson (2001, 2006)), semiconductors (Hatch and Mowery (1998)) and pizza franchises (Darr, Argote and Epple (1995)).[4] These firm-level studies examine the impact of learning-by-doing using unit costs as a measure of performance. I exploit the availability of detailed micro-level data on patients undergoing CABG surgery to study the process of learning at the level of the individual worker. Also, unlike the studies mentioned above, I characterize performance in terms of patient outcomes (or surgeon quality), which, unlike costs, are measured with greater precision.

The effect of experience on individual productivity has been investigated to a lesser extent in these literatures.[5] In the setting I examine, there are a number of ways in which individual experience could matter in determining performance. The skill of the operating CABG surgeon is critical in ensuring a successful procedure. According to the American College of Cardiology[6], experience has a strong influence on a surgeon's cognitive knowledge base and technical skills, both of which determine competence. Experience helps surgeons in selecting the appropriate treatment strategy for patients, and helps them identify and treat complications at an early stage. Experience also has a positive impact on manual dexterity and helps surgeons maintain proper surgical technique. Finally, because of the rare occurrence of adverse outcomes, surgeon competence requires specific training and ongoing experience in managing them so as to be prepared to react optimally when they occur.

2.2 The Theory of Human Capital

The second stream of research relevant to this study examines the nature of accumulation of human capital by individuals. This literature (Becker (1962), Killingsworth (1982)) presents two ways in which individuals accumulate human capital: training (education or on-the-job) and learning-by-doing. In the context studied here, a surgeon may increase her human capital

---

[4] Some of these studies (e.g. Benkard (2000)) also explicitly model the effects of organizational "forgetting".

[5] The literature in psychology has numerous studies on this effect. These studies typically find that individuals complete a task in lesser time and with greater accuracy, the more experience they have with the task. They also find support for diminishing returns of experience. Examples include Newell and Rosenbloom (1981), Delaney, Reder, Staszewski and Ritter (1998).

[6] Refer to Hirshfeld et al (1998) for the complete statement of clinical competence released by the American College of Cardiology.

by investing in (residency) training. She may also accumulate human capital as a result of the skills she gains from performing more procedures.

One aspect of human capital emphasized in this literature from very early on (Becker (1962), Jovanovic (1979)) is the importance of assessing the degree of specificity. Becker (1962) focuses on the dichotomy between general human capital (augments productivity across all firms) and firm-specific human capital (augments productivity in current firm, but not elsewhere). Recent research (Gibbons and Waldman (2004)) has focused on a third category: task-specific human capital, which is defined as being specific to the tasks being performed, as opposed to being specific to a firm. The data in this study allow me to estimate whether human capital acquired by individuals through learning-by-doing is specific to (a) the firms they work in and (b) the tasks they perform. In doing so, this study adds to the empirical evidence establishing learning-by-doing as a source of general or specific human capital. Huckman and Pisano (2006), who study the firm-specificity of performance of cardiac surgeons, is a recent example of one such study. I discuss their approach in detail in the next subsection.

For a cardiac surgeon, general human capital would result from investment in medical school, residency and post-residency training. Another determinant of the level of human capital could be procedure experience, which may help the surgeon develop new skills or maintain old ones. Studies in the human capital literature typically use earnings or wages as their measure of productivity. Since I do not have access to earnings data for surgeons, I use surgeon performance to test for specificity of human capital.

Why may one expect accumulated human capital to be (completely or partially) specific to a firm (hospital) in this setting? Huckman and Pisano (2006) offer some potential explanations, the primary one being that a surgeon tends to develop a degree of familiarity with the rest of the surgical team and other assets at a hospital and this aspect prevents the benefits of experience from being portable across hospitals. On the other hand, if familiarity with assets were not as important as, say, learning how to treat complications, one can imagine a situation where experience adds to both general and firm-specific aspects of human capital, but the general component is much larger.

It is also plausible to think of a scenario where part of the human capital accumulated by surgeons on the job is specific to the tasks being performed, as opposed to being specific to the firm. In this context, I use the term "task" to connote a specific procedure. A cardiac surgeon may perform multiple tasks within a firm: she may perform CABG surgeries and she may also perform other cardiac procedures (e.g. repair of heart valves, heart transplants etc.). In such a situation, some part of the human capital may go unused when the surgeon switches tasks (either within the firm or across firms). Whether surgeon learning transfers across tasks (procedures) depends on the extent to which the skills (of the surgeon or the surgical team) needed to perform these procedures overlap.

Sections 5 and 6 contain the model and results for the tests of these hypotheses.

2.3 Volume Outcome Studies in Healthcare

The final stream of research relevant to this study encompasses papers in the medical and health economics literatures that document a correlation between procedure volume and outcomes at the firm (hospital) and individual (surgeon) levels. While it is difficult to compare findings across studies because of differences in data, disease categories studied and methodology, the general consensus among researchers seems to be that there is a positive relationship between procedure volume (measured at the level of the hospital or the individual surgeon) and patient outcomes.[7]

However, few studies attempt to translate this correlation into a well established causal relationship between procedure volume and outcome. The literature offers two competing hypotheses, with contrasting causal and policy implications that could underlie this observed relationship. The first hypothesis is the "learning-by-doing" or the "practice-makes-perfect" hypothesis. This hypothesis is built on the notion that increased experience results in more finely developed skills which in turn lead to better outcomes.

The second hypothesis is the "selective referral" hypothesis, which postulates that the observed relationship is due to a referral system that directs more patients to high quality providers. Even if patients do not have knowledge about exact mortality rates of a surgeon's

---

[7] See Luft et al (1990) for a comprehensive review of volume-outcome studies, and Halm et al (2002) for a review of volume-outcome studies related to CABG.

former patients, one can think of a scenario where the best surgeons and the best hospitals develop a reputation as being of high quality, and hence attract more patients or referrals from specialists and primary care physicians.

The need to unambiguously distinguish between these two potential explanations arises from the fact that they have contrary implications for policy. “Learning-by-doing” acts as an argument in favor of increased concentration among providers, whereas regionalizing procedures in the face of “selective referral” only leads to reduced competition, without any improvement in outcomes.

A handful of studies have attempted to distinguish between these explanations using data at the level of the firm (hospital) by using instruments that are correlated with procedure volume but unrelated to factors that influence patient outcomes. Luft, Hunt and Maerki (1987) use cross-sectional data to estimate a simultaneous equations model where they use the size of the hospital, teaching affiliation, and the number of appendicitis procedures as instruments in the equation predicting volume. However, to the extent that the quality of different procedures in a hospital may be related or that a teaching hospital may attract sicker patients, it is not immediately clear why these instruments are excluded from the outcome equation.

A few recent studies (Gowrisankaran, Town and Ho (2006), Picone, Trogdon and Trollis (2005)) use predicted volume as an instrument for actual procedure volume (at the hospital level), where predicted volume is estimated from a multinomial hospital choice model based on distance and basic hospital characteristics. The identification in these models comes from the fact that the location of patients relative to hospitals is a key determinant of hospital volume and is assumed to be exogenous to outcomes. These studies conclude in the favor of a strong learning-by-doing effect for hospitals performing cardiac procedures. While working with physician level data (as this paper does), this instrument is not as useful because one needs to sort out the patient’s choice of a hospital from her choice of a physician. Further, these studies cannot test for the presence of any aspects of specificity in learning-by-doing.

A few other studies have used panel data to control for unobserved determinants of hospital quality by including hospital fixed effects in their regressions. Examples of such studies include Ho (2002), Farley and Ozminkowski (1992) and Hamilton and Hamilton (1998). Their estimates, obtained using within-hospital variation in volumes, suggest that hospital volume leads to better outcomes. However, the possibility that changes in within-hospital volume are driven by unobservable elements of hospital quality still cannot be ruled out.

Analyses of volume-outcome effects using data at the surgeon level have tended to be largely correlational – i.e., these studies assume the presence of a positive correlation between provider volume and outcomes as indicative of a learning-by-doing effect. Examples of such studies include Hannan et al (1991) and Hughes et al (1987). Failure to account for selective referral in these studies may bias estimated effects of volume on outcome.

A further problem with correlational studies (that use OLS regression models of patient outcome on provider volume) is that one cannot even sign the direction of the bias on the volume coefficient. The presence of selective referral should serve to make the OLS estimate larger in absolute value when compared to the IV estimate which parcels out this effect. However, if patients who are attracted to high quality surgeons are more ill in ways that are unobservable, it could lead to an increase in bad outcomes for the high quality surgeons, and result in OLS coefficients underestimating the true effect of learning-by-doing (in absolute terms). The only way to determine the direction of bias is by comparing OLS estimates with those obtained by IV regression. Section 6 presents results from such a comparison.

Huckman and Pisano (2006) were the first to examine the important question of firm-specificity of performance by using data on cardiac surgeons in the state of Pennsylvania, from 1994-1995. In particular, they look at whether a surgeon’s experience (measured by procedure volume) at one hospital translates into better outcomes for her patients at other hospitals she operates in. However, they do not control for surgeon fixed effects and fail to account for endogeneity of surgeon volume in their empirical specifications. Their estimates are thus identified off levels implying that unobservable surgeon-specific factors might bias their results. Based on their assumptions, they find a strong learning-by-doing effect and evidence to support firm-specificity of surgeon performance. While the focus of this study is to document the presence of learning effects among surgeons using robust empirical methods, the study setting allows me to make some inferences about the specificity of these effects as well.

Previous studies of individual (surgeon) learning in health care suffer from limits imposed by data as well. Either the panel does not track surgeons over a long enough time period (e.g. Huckman and Pisano (2006)) or the data does not track individual surgeons but only identifies the proportion of procedures performed by high and low volume surgeons (e.g. Hughes et al (1987)). The dataset used in this study overcomes both these limitations and tracks individual surgeons over a six year period.

In summary, the main contributions of this study lie in using detailed micro-level data and robust empirical techniques to shed light on the process by which individuals within organizations attain (or maintain) proficiency at their tasks. In doing so, I aim to contribute to the literatures on organizational learning, human capital management and health economics. In addition, the basic idea underlying the instrument, that of exitor behavior affecting incumbents, is applicable in other settings: for example, one can examine the effects of a firm closure on other firms in a market. The study thus makes a useful contribution to the applied econometrics literature as well.

3 Data and Research Setting

3.1 Selecting a Candidate Procedure

I study the impact of learning-by-doing on quality for surgeons who perform Coronary Artery Bypass Grafts (CABG). Developed in the late 1960s[8], CABG is a risky and invasive surgical procedure that is normally performed on patients with severe or multiple narrowing of the coronary arteries. It is one of the ways[9] in which Coronary Artery Disease, one of the leading causes of death in the US[10], is treated. The procedure involves bypassing a blocked (or narrowed) segment of a heart artery by using a graft from the arm, leg or chest. Hospitals typically have dedicated operating rooms, fitted with specialized equipment and manned by dedicated technicians, in which the procedure is performed.

---

[8] Source: Website of the American Heart Association

[9] Other treatments include medication and angioplasty

[10] Source: National Center for Health Statistics, <http://www.cdc.gov/nchs/fastats/lcod.htm>

The referral process for CABG normally works as follows: a patient experiencing chest pains or shortness of breath starts by visiting a primary care physician who may then refer her to a cardiologist for further treatment and evaluation. The cardiologist evaluates the patient's medical history and symptoms and may perform a cardiac catheterization, a procedure that indicates how well blood is flowing through the vessels that supply the heart muscle. If the catheterization shows abnormal results, the patient is treated according to the extent of blockage in the arteries. Mild to medium blockages are treated using medication or angioplasty[11], while severe cases are referred to a surgeon for bypass surgery.

There are multiple reasons why CABG was selected as a candidate procedure for this study. First, it is a procedure that can be performed only by highly trained and specialized surgeons whose abilities are perceived as being crucial to the quality of care. Given the key role played by the surgeon, it is important to study the factors that affect surgeon performance. Second, CABG is a fairly common procedure with over 470,000 surgeries being performed in the US in 2004.[12] It accounts for ~4-5% of total health expenditure in the US[13], and is hence important in its own right. Third, the procedure has been extensively studied in the health economics literature with the result that there is a commonly accepted and readily available measure of outcomes: in-hospital mortality.[14]

Finally, since CABG is performed by surgical specialists, it is likely that a sizeable proportion of patients are referred to the appropriate hospital or surgeon (by their cardiologist) on the basis of provider quality – implying that "selective referral" could play an important role. Therefore, it becomes all the more important to not rely on correlational studies of volume-outcome while estimating learning-by-doing effects.

3.2 Data

The primary dataset for this study comes from the Hospital Inpatient Data Files provided by the Florida Agency for Health Care Administration (AHCA) for the years 1998-2003. This

---

[11] This procedure is performed by an interventional cardiologist.

[12] Source: 2004 National Hospital Discharge Survey

[13] Ibid.

[14] An alternate measure of mortality used by some volume-outcome studies is 30-day mortality, defined as death occurring within 30 days of admission. The Florida AHCA data only identifies in-hospital mortality, while Medicare data files (which report 30-day mortality rates) typically do not include physician identifiers. So, I proceed with in-hospital mortality as my measure of patient outcome.

data is comparable to patient level discharge data provided by the Health Care Utilization Project (HCUP) and the California Office of Statewide Health Planning and Development (OSHPD). Specifically, AHCA provides information about each discharge including the hospital, physical and demographic characteristics of the patient (e.g. age, sex, payer information), and a comprehensive list of primary and secondary diagnoses. In addition, the data identify the license number of the operating surgeon, enabling me to track procedure volume of individual surgeons over a six year time frame.

I also make use of two secondary datasets that provide me with information on hospital and physician characteristics. I obtain hospital level data from the American Hospital Association (AHA) Annual Survey of Hospitals for the years 1998-2003.[15] I obtain detailed information about each surgeon's training (e.g. medical school/residency program trained at, year of graduation) and draw inferences about the physician's age by linking the license number to an online database provided by the Florida Department of Health.[16] To protect confidentiality, I do not present any physician-specific information.

I restrict attention to surgeons who appear more than once in the panel, and perform at least 5 procedures over the entire panel in order to exclude the effects arising from unrepresentative surgeons, e.g. an Emergency Room surgeon who may perform the occasional surgery. I also exclude hospitals that appear in the bottom one-half percentile of the distribution of hospital procedure volume. The empirical results are robust to these sample restrictions.

4 Using Surgeon Exit to Identify Exogenous Shocks to Procedure Volume

This paper uses an instrumental variables approach to determine the extent to which increased procedure volume for surgeons translates into better outcomes for their patients. An ideal instrumental variable in this case should have the following properties: it should help explain variation in surgeon procedure volume (the endogenous predictor), and have no

---

[15] Since the AHCA dataset does not provide AHA ids for hospitals, I link it to the AHA data using the Medicare (HCFA) id for each hospital and was able to obtain perfect matches for all but one hospital.

[16] This information is available at <http://ww2.doh.state.fl.us/irm00praes/praslist.asp>

causal relationship with surgeon quality except through its effect on procedure volume. One can argue that concern about "selective referral" is mitigated at the surgeon level – this would hold true if patients only choose hospitals (and not individual surgeons) based on quality. However, to the extent that patients are typically referred to providers by other physicians[17] for procedures like the one studied in this paper (CABG), the possibility that higher quality surgeons attract more patients cannot be ruled out.

This paper proposes the use of surgeon exit as an exogenous identifier. I use the term "exit" to refer to any instance where a surgeon stops performing CABG surgeries in Florida.[18] A surgeon may exit because of old age (retirement), death, termination of employment, or relocation to a different state.[19] The rationale is as follows: once a surgeon exits a hospital for exogenous reasons, her procedure volume gets redistributed among non-exiting surgeons at the hospital, giving their procedure volumes a positive shock. If this shock to surgeon volume is not correlated with unobservable determinants of changes in surgeon quality, the use of exit as an identifier is valid. The instrumental variable used in the regressions is defined as a function of the volume of the exiting surgeon(s). I first provide a formal definition of the instrument before discussing some of the underlying identifying assumptions in detail.

4.1 Defining the Instrument

I define surgeon exit as follows: a surgeon exits a hospital if her CABG procedure volume at that hospital drops to zero, and she performs at least 5% of the hospital's CABG procedures in the year preceding exit. The latter part of the definition ensures that surgeons who perform the occasional CABG procedure are not counted amongst exiting surgeons. I consider exit to have taken place only when a surgeon stops performing procedures across all hospitals she was operating in.[20] I further restrict the set of exitors to surgeons aged 55 and over.[21] Doing so increases the probability that a surgeon exited the data because she was retiring and mitigates concerns about endogeneity of exit. In section 6.4, I test the

---

[17] A patient first visits a cardiologist who may then refer her to a cardiac surgeon.

[18] I consider a surgeon to have exited when she stops performing procedures across all hospitals in Florida. I use this definition in order to rule out cases where a surgeon may exit a particular hospital but enter a nearby hospital, thereby taking patients with her.

[19] The data does not identify the cause of exit. However, I use data on the surgeon's age to make some inferences about possible cause of exit, as described in Section 4.1.

robustness of the empirical results by slightly varying the definition of exit and find little change in the conclusions.

The volume of the exiting surgeon is simply calculated as the number of procedures performed by the surgeon in year *t-1*, where *t* denotes year of exit. I assume that the impact of surgeon exit is felt on staying surgeon volumes for one year, i.e. if a surgeon exits in year *t*, then existing surgeons in a hospital experience an exogenous shock to their procedure volumes in year *t* only. The timing of the model works as follows: if a surgeon exits (i.e. stops performing procedures) in year *t*, the non-exiting surgeons at that hospital experience a positive shock to volume in year *t*. This volume shock translates into better patient outcomes for these surgeons in year *t+1*.

While one can use the data to identify instances of surgeon exit, there is no way of using the data to determine how exit volume[22] is allocated among staying surgeons. There are many ways in which patients who would have visited the exiting surgeon may be redistributed across non-exiting surgeons. I discuss three different methods of allocating exit volume across surgeons and test my main specifications under all three scenarios.

Under the first allocation rule, patients of the exiting surgeon are assumed to be allocated equally among all the remaining surgeons at the hospital. The instrument is then computed as the sum of the procedure volumes of all exiting surgeons in that hospital. In other words, all surgeons working at a hospital face the same shock to volume (equal to the above sum of exit volumes), irrespective of their shares prior exit.

One demerit of the instrument as defined above is that it exhibits no variation across surgeons within a hospital in a particular year. In other words, a high volume surgeon stands to gain as much from exit as a low volume surgeon. The second allocation rule addresses this issue by allocating exit volume to staying surgeons in proportion to their share of procedure

---

[20] In a majority of cases, a surgeon exits all hospitals simultaneously and that year is taken to be the year of exit. In a few instances, a surgeon who works across multiple hospitals may exit one hospital but continue working in another. In such cases, I consider the surgeon to have exited only when she stops performing procedures altogether, and the latter year is taken as the year of exit from the data.

[21] The data on surgeon characteristics does not contain the age of the surgeon, but does contain the year in which the surgeon graduated from medical school. I infer the surgeon's age using this information.

[22] I use the term exit volume to refer to the number of procedures performed by an (exiting) surgeon at a hospital the year before exit.

volume within the hospital in year *t-1*, where *t* denotes the year of exit. To the extent that exit volume is allocated to staying surgeons in proportion to their current shares, this volume shock is exogenous to changes in surgeon quality.[23] In specifications that do not have surgeon fixed effects, however, the identification assumption may be violated.

The third allocation rule also leads to variation in the instrument across surgeons within a hospital and also guards against endogeneity. Under this rule, I assign exit volume to each staying surgeon on the basis of the extent of patient zip code overlap with the exiting surgeon. To illustrate the methodology, consider a hospital with three practicing surgeons in year *t-1*, one of whom (surgeon A) exits in year *t*. Table A1 presents the breakdown of the exiting surgeon's volume by the zip code of the patients treated. Rows 2 and 3 of Table A1 present the distribution of patients across those zip codes operated upon by the other (non-exiting) surgeons in the hospital.[24]

Table A1. Zip code breakdown of surgeon procedure volume, year *t-1*

	Number of patients treated by surgeon in zip code, *year t-1*		
	60201	60031	60110
Surgeon A	10	8	6
Surgeon B	3	5	0
Surgeon C	0	1	0

Since surgeon A exits the data in year *t*, her procedure volume in year *t-1* is assigned to the other surgeons in the hospital in year *t*. Table A2 presents the allocation of A's exit volume to the non-exiting surgeons in the year post exit. Note that in Column 1, surgeon C is allocated no patients from zip code 60201 as she does not treat any patients from that location. As a result, the zip code-level exit volume of 10 patients is entirely allocated to surgeon B. In Column 2, the zip code-level exit volume (of 8 patients) is equally divided amongst both staying surgeons as they both treat patients from that zip code in *t-1*. Note that division of exit volume in this zip code is independent of surgeon market share (in that

---

[23] The main specifications include surgeon fixed effects, implying that we look at changes in surgeon quality, and not the absolute level.

[24] Note that the non-exiting surgeons may draw patients from other zip codes as well – this table lists patient zip codes of the exiting surgeon only.

zip code) – this allocation rule tries to ensure that the additional experienced gained by a surgeon from exit is independent of her current procedure volume, which may reflect unobserved components of quality. Finally, patients in non-overlapping zip codes (zip code 60110 in the example above) are allocated equally across all non-exiting surgeons.

Table A2. Calculating exogenous volume shock, year *t*

	Allocation of exit volume, *year t*			
	60201	60031	60110	Total
Surgeon B	10	4	3	17
Surgeon C	0	4	3	7

As mentioned earlier, there is no way of empirically determining the exact mechanism through which patients of the exiting surgeon are allocated. In the remainder of this paper, I present results only using the third allocation rule: the zip code overlap method. The results[25] using the two other allocation rules discussed earlier were qualitatively similar to the main results described in Section 6.

4.2 Why is Surgeon Exit a Plausible Instrument?

I now discuss some of the underlying assumptions in using surgeon exit to identify exogenous changes in surgeon procedure volumes. I test the validity of these assumptions using the data, and include these results in a later section.

The correlation between the procedure volumes of staying surgeons and the volume of exiting surgeons (i.e. the relevance of the instrument) could be mitigated under some conditions. The first possibility is that the hospital hires new surgeons who take over the caseload of exiting surgeons. In that case, the procedure volume of non-exiting surgeons is unaffected by exit. However, the data show that surgeon exit is accompanied by entry (the same year) in only 12% of the cases. Further, an entering surgeon performs only 21 procedures at a hospital in the year of entry, compared to an exiting surgeon who performs 45 procedures on average. Taken together, these facts imply that entry of new surgeons should not affect instrument relevance significantly.

---

[25] Available upon request

The second possibility is that the procedure volume of the exiting surgeon is "lost" to the hospital, i.e. patients who would have visited the exiting surgeon now choose a different hospital instead of visiting another surgeon in the same hospital. However, the data clearly indicate that there is very little variation in hospital volume over time; hospital volume decreases by 2% the year of exit, compared to a 2% annual increase over all years. This implies that surgeon exit does not lead to a significant decrease in a hospital's annual procedure volume.

I now turn to the question of instrument exogeneity. The main specifications include surgeon fixed effects implying that the dependent variable is the change in quality of the surgeon. The identifying assumption I make here is that the cause of surgeon exit is not directly related to changes in (non-exiting) surgeon quality. This assumption ensures that the instrument (which is a function of exiting surgeon volume) is uncorrelated with unobservable determinants of changes in staying surgeon quality. An instance in which this assumption would be violated is if a surgeon exited because she believed her rivals' quality was going to improve in the future. Consider a scenario in which the exiting surgeon's volume is determined by her quality relative to the quality of the staying surgeon. In other words, an exiting surgeon might have higher procedure volume if her rival surgeons were of poor rather than high quality compared to her own quality. This would induce correlation between exit volume and changes in (staying) surgeon quality, thereby confounding identification. I test this assumption using the data and provide the details of this test in Section 6.

5 A Robust Empirical Model of Learning-by-Doing

5.1 Instrumenting for Total Surgeon Experience

In order to estimate the extent of surgeon learning-by-doing, I model surgeon quality as a function of various characteristics of the patient being treated, of the hospital she is treated at, and of the surgeon performing the procedure, including surgeon experience. Specifically, I estimate an instrumental variables regression (using the Stata module *xtivreg*) where observations are at the level of the patient and the endogenous predictor, surgeon experience, is instrumented for. Equations (1.0) and (1.1) represent the first[26] and second stages, respectively, in the instrumental variables estimation procedure. The variables and notation are explained in detail below, starting with the second stage regression equation. In all equations, *i* indexes the patient, *h* indexes the hospital, *p* indexes the surgeon and *t* indexes the time period of observation.

$$\begin{aligned}(Physvol)_{p,t-1} = {} & \alpha_0 + \alpha_1 * (Exitvol)_{p,t-2} + \alpha_2 * X_{i,p,h,t} + \alpha_3 * \theta_p + \alpha_4 * \mu_h \\ & + \alpha_5 * \lambda_{h,t} + \alpha_6 * (Year)_t + \upsilon_{p,t-1}\end{aligned} \tag{1.0}$$

$$\begin{aligned}(Outcome)_{i,p,h,t} = {} & \beta_0 + \beta_1 * (Ph\hat{y}svol)_{p,t-1} + \beta_2 * X_{i,p,h,t} + \beta_3 * \theta_p + \beta_4 * \mu_h \\ & + \beta_5 * \lambda_{h,t} + \beta_6 * (Year)_t + \varepsilon_{i,p,h,t}\end{aligned} \tag{1.1}$$

Second Stage Regression

<u>Dependent variable:</u> The dependent variable in the second stage regression is a measure of surgeon quality. Since analyses are conducted at the level of the patient, I use in-hospital patient mortality as the primary measure for surgeon quality. The main advantage of using in-hospital mortality as a measure of outcome is that there is very little chance of miscoding. In addition, there is sufficient variation in outcomes across surgeons and hospitals which allows me to estimate learning effects with precision. However, it may not completely reflect the health status of the patient post surgery. The dependent variable $\text{Outcome}_{i,p,h,t}$ is a binary variable that is set to one if patient *i* died as a result of a CABG procedure performed by surgeon *p* in hospital *h* at time *t*.

<u>Independent variables – surgeon experience:</u> The primary predictor of interest measures the experience of the surgeon performing the procedure. I proxy for surgeon experience by the total number of CABG procedures performed by the surgeon the previous year, $Physvol_{p,t-1}$. This measure is aggregated across all hospitals the surgeon operates at. In eq.

---

[26] The unit of observation for the first stage regression is the patient (as in the second stage). However, I have suppressed the patient index *i* in the notation for ease of exposition.

(1.1), $Phy\hat{s}vol_{p,t-1}$ denotes the predicted value from the first stage regression, eq. (1.0). The coefficient β_1 tells us the effect of surgeon experience on outcomes, purged of any possible biases arising from endogeneity. A negative sign on β_1 will act as evidence in favor of the learning-by-doing hypothesis. I let surgeon volume enter the specification in a linear fashion above.[27]

While prior theoretical and empirical research on learning-by-doing model experience as having a cumulative effect, I use recent volume to measure surgeon experience mainly because I observe cumulative volume only for a subset of surgeons in the data.[28] The concern over using recent procedure volume instead of cumulative procedure volume is mitigated to an extent by recent research studies (Gowrisankaran, Town and Ho (2006), Gaynor, Seider and Vogt (2005)) that find a significant amount of organizational forgetting among hospitals performing CABG procedures. This implies that experience from the immediate past has a greater impact on outcomes when compared to experience from further before. As a robustness check, I estimate the model using the number of procedures performed by the surgeon in the last *two* years as a proxy for surgeon experience. I also repeat the analysis on a sample that contains only those surgeons who finish their residency training in 1998 or later. In this specification, I use cumulative volume as the proxy for surgeon experience.

I include an interaction term in order to allow for a nonlinear effect of procedure volume on mortality. Specifically, I interact the lagged surgeon volume term, $Physvol_{p,t-1}$, with an indicator for whether the surgeon was a high volume surgeon the previous year.[29] I define a high-volume surgeon as one performing more than 200 procedures that year. The coefficient on the interaction term can be used to infer whether high volume surgeons benefited more (or less) from an exogenous change in experience, when compared to low volume surgeons.

---

[27] In an alternate specification, I used the square root of procedure volume as the measure for volume in order to incorporate nonlinearity. The linear specification was found to have a better fit (in terms of R^2) so I proceed with that for the rest of the specifications. Models using the square root of volume yielded similar conclusions.

[28] Specifically, the data allow me to track all procedures only for surgeons who finish their residency training in 1998 or later.

[29] One concern here, is of course, the potential endogeneity of the interaction term. The use of an indicator for a high volume surgeon (as opposed to using the actual experience of the surgeon) should lessen this concern to an extent.

Independent variables – surgeon characteristics: In order to isolate the effect of surgeon experience on outcomes, I control for other characteristics of the surgeon that determine quality. In the main specifications, I do this by including a vector of surgeon fixed effects (θ_p) in the regression that account for the effect of time-invariant surgeon characteristics (e.g. sex, training etc. and unobservables that are fixed with time) on patient outcomes. I also run a specification that excludes surgeon fixed effects[30]; in this model, I control for the following surgeon characteristics: age, sex and whether the surgeon trained at a foreign medical school.

Independent variables – hospital characteristics: I include a vector of hospital fixed effects (μ_h) to control for systematic intrinsic quality differences across hospitals.[31] Following Ho (2002), I also include a vector of hospital characteristics ($\lambda_{h,t}$) that reflect the hospital's scale, staffing levels and service makeup. These include: the number of general and cardiac care beds, number of FTE registered nurses and licensed practical nurses, number of high-tech services offered (MRI, CT scan, PET scan etc) and indicators for medical school affiliation and having an approved residency program.[32]

Independent variables – patient characteristics: The vector $X_{i,p,h,t}$ includes patient characteristics (e.g. the number of co-morbidities, patient age, sex, an indicator for whether the patient has had a prior CABG), all of which are expected to have an impact on patient mortality. The patient characteristics I include in the regression are: patient age categories, sex, the Charlson co-morbidity index[33], concurrent angioplasty, cardiogenic shock, prior CABG surgery, congestive heart failure, hypertension and an indicator for a heart attack. All specifications also contain year indicators to control for the effect on outcomes of changes in technology that are not captured by other predictors.

---

[30] I do this mainly to facilitate comparison with earlier surgeon level volume outcome studies, that do not include surgeon fixed effects in their specifications.

[31] Note that hospital fixed effects are not perfectly collinear with surgeon fixed effects as surgeons may work across multiple hospitals in the same year.

[32] Since I also include hospital fixed effects in the regression, the explanatory power of these variables is rather limited. An F-test found these variables to have some joint predictive power so I include them in the regression.

[33] The Charlson co-morbidity index reflects the cumulative increased likelihood of one year mortality arising from different categories of co-morbidity such as cancer, diabetes, AIDS etc. The higher the score, the worse the condition of the patient.

First Stage Regression

In the first stage regression (eq. (1.0)), surgeon volume is estimated as a function of the instrument and all exogenous predictors from the second stage. Note that a staying surgeon's procedure volume in year *t-1* is affected by surgeons who exit in year *t-1*. Thus, the instrument allocates exiting surgeons' procedure volume in year *t-2* to staying surgeons in year *t-1*. I use the sum of the exit volumes allocated to the staying surgeon across all hospitals to instrument for past procedure volume of the staying surgeon. The effect of exit on the procedure volume of staying surgeons is assumed to last for a year. Based on the discussion in Section 4, the coefficient α_1 is expected to have a positive sign in eq. (1.0).

5.2 Instrumenting for Specific Aspects of Surgeon Experience

The regression equations for exploring specificity of surgeon experience are set up in a similar manner. I estimate a patient-level instrumental variables regression where the dependent variable is, as before, a measure of surgeon quality (patient outcome). The key difference is that I now use a hospital-specific measure of surgeon experience as the primary predictor. The variable $Physhospvol_{p,h,t-1}$ measures the procedure volume of surgeon *p* at hospital *h* in the year *t-1*, while $Physhospvol_{p,-h,t-1}$ measures the procedure volume of the same surgeon across all other hospitals (apart from hospital *h*) in *t-1*. In Eq. (1.4), the variables $Phys\hat{o}spvol_{p,h,t-1}$ and $Phys\hat{o}spvol_{p,-h,t-1}$ refer to the predicted values of these variables from the first stage regressions.

$$\begin{aligned}(Physhospvol)_{p,h,t-1} &= \alpha_0 + \alpha_1 * (Exitvol)_{p,h,t-2} + \alpha_2 * (Exitvol)_{p,-h,t-2} + \alpha_3 * X_{i,p,h,t} \\ &+ \alpha_4 * \theta_p + \alpha_5 * \mu_h + \alpha_6 * \lambda_{h,t} + \alpha_7 * (Year)_t + \upsilon_{p,h,t-1}\end{aligned} \tag{1.2}$$

$$\begin{aligned}(Physhospvol)_{p,-h,t-1} &= \gamma_0 + \gamma_1 * (Exitvol)_{p,h,t-2} + \gamma_2 * (Exitvol)_{p,-h,t-2} + \gamma_3 * X_{i,p,h,t} \\ &+ \gamma_4 * \theta_p + \gamma_5 * \mu_h + \gamma_6 * \lambda_{h,t} + \gamma_7 * (Year)_t + \eta_{p,-h,t-1}\end{aligned} \tag{1.3}$$

$$(Outcome)_{i,p,h,t} = \beta_0 + \beta_1 * (Phys\hat{h}ospvol)_{p,h,t-1} + \beta_2 * (Phys\hat{h}ospvol)_{p,-h,t-1} + \beta_3 * X_{i,p,h,t} + \beta_4 * \theta_p + \beta_5 * \mu_h + \beta_6 * \lambda_{h,t} + \beta_7 * (Year)_t + \varepsilon_{i,p,h,t} \tag{1.4}$$

Correspondingly, I use hospital-specific instruments for surgeon procedure volume in the first stage regressions (1.2) and (1.3). $Exitvol_{p,h,t-2}$ is calculated as the exit volume allocated (using the extent of zip code overlap) to surgeon *p* at hospital *h* in *t-2,* while $Exitvol_{p,-h,t-2}$ is calculated as the sum of exit volumes allocated to surgeon *p* across all other hospitals (other than *h*) in *t-2*.

If human capital acquired by learning-by-doing were firm-specific, one would expect additional surgeon experience to have a greater benefit on patient outcomes at the hospital where the experience was gained. In other words, if β_1 and β_2 are both negative, and furthermore, β_2 is smaller in absolute value than β_1, the implication is that procedure volume at other hospitals does impact surgeon quality at the hospital under consideration, but not as much as procedure volume at the same hospital. This would be evidence in support of firm-specificity. A stronger form of firm specificity, in which experience is completely non-transferable across firms, would have β_2 equal to zero (or statistically indistinguishable from zero). On the other hand, if performing additional procedures (irrespective of hospital) adds to a surgeon's general human capital, one would expect β_1 and β_2 to be statistically indistinguishable from each other.

I test for task-specificity of surgeon human capital by using a specification similar to the one described above. Instead of using hospital-specific measures of surgeon volume, I now use task-specific measures. Specifically, I divide the total number of cardiac procedures performed by each surgeon into the number of CABG and the number of non-CABG procedures. I construct analogous versions of the instrument and compare the coefficients on the volume terms to make inferences on task-specificity of human capital.

6 Results

Before discussing the regression results, I present some patterns in the raw data. The study uses data on a total of 385 CABG surgeons working at 65 hospitals in Florida over 6 years, with a total volume of 160,210 procedures. The number of surgeons in each year increases from 224 in 1998 to 265 in 2003, while the number of hospitals increases from 57 to 65 over the same time period. Table 1 presents descriptive statistics of some of the key variables used in the specification. These statistics are calculated from patient-level data. The dependent variable, patient mortality, has a mean of 3.82% in the data, which is in line with mortality rates observed in other studies. The patient population is composed of a majority of males (~70%) who were over 65 years old on average. Around 6% of the population had had a previous CABG surgery.

The table also presents surgeon and hospital characteristics (averaged over patient level data). CABG surgeons in the data were overwhelmingly male (almost 98%) and a majority of them graduated from a US medical school (~83%). The mean age of surgeons in the data is almost 40. The average surgeon performs almost 175 procedures annually, across all hospitals. This number drops to 143 when hospital-specific volume is considered. These numbers do not reflect the true distribution of surgeon procedure volume, where one can see a bunch of low-volume surgeons. The average hospital employs around 6.5 surgeons. More than 40% of surgeons work across multiple hospitals during the same time period. The mean for the instrument is rather low (2.95) because of the number of cases in which no surgeon exit is recorded in a hospital (the instrument is assigned a value of zero in this case). In all, I record 33 instances where a surgeon exits the data. 230 non-exiting surgeons in the data are affected at some point in time by surgeon exit and experience a positive shock to their volumes.

6.1 Testing the Identification Assumption

The main identification assumption behind using exitor volume as an instrument is that it is uncorrelated with unobserved determinants of changes in staying surgeon quality. I validate this assumption by testing if the amount of exit volume faced by a surgeon is determined by her quality in the previous period. This reduced form regression is performed at the level of the surgeon-year and uses the instrument as the dependent variable and the lagged risk-adjusted surgeon-level mortality rate[34], $(Mortrate)_{p,t-1}$, as the primary predictor. I also include surgeon characteristics ($\psi_{p,t}$) – age, sex, whether the surgeon is foreign trained – and surgeon-level means of all the patient characteristics described earlier ($X_{p,t}$) as predictors.

$$\begin{aligned}(Exitvol)_{p,t} = \beta_0 + \beta_1 * (Mortrate)_{p,t-1} + \beta_2 * X_{p,t} + \beta_3 * \psi_{p,t} + \beta_4 * \lambda_{h,t} \\ + \beta_5 * (Year)_t + \xi_{p,t}\end{aligned} \tag{1.5}$$

The results, reported in Table 2, indicate that the coefficient on the primary predictor, lagged surgeon quality, is statistically indistinguishable from zero (*p=.984*). This suggests that the amount of exit volume faced by a surgeon is independent of her prior quality, and validates the main identification assumption in the paper.

6.2 Do Surgeons Learn From Experience?

Table 3 provides the main results from first-stage regressions of lagged surgeon procedure volume (across all hospitals) on the instrument, along with all other exogenous predictors from Stage 2. Columns 1 and 2 present the results without and with surgeon fixed effects. The coefficient on the instrument is positive and highly significant indicating that exit volume faced by the surgeon is strongly correlated with actual procedure volume, with t-statistics of 16.62 and 20.43 respectively. The magnitude of the coefficient falls (.61 vs .91) on including surgeon fixed effects, implying that between-surgeon variation was driving some of the effects in column 1. The F-statistic[35] for the instrument significance is substantially larger than the typical recommended thresholds[36], validating use of the instrument.

---

[34] In order to compute risk adjusted mortality rate for each surgeon, I run a patient-level logit model where the dependent variable equals one if the patient died, and the predictors include all patient characteristics and a vector of year fixed effects. I then sum up the predicted probability of mortality for all patients across a surgeon and divide by the number of patients treated by that surgeon.

[35] In the case of a single instrument, the F-statistic is simply the squared value of the t-statistic.

[36] As a rule of thumb, Staiger and Stock (1997) recommend a first-stage F statistic of at least ten for an instrument not to be considered weak.

Columns 1 through 6 of Table 4 present the results from specifications that estimate the extent of surgeon learning-by-doing. Columns 1, 3 and 5 are estimated without surgeon fixed effects but include a vector of surgeon characteristics.[37] Columns 2, 4 and 6 include surgeon fixed effects to capture the effect of surgeon-specific time invariant factors on mortality.

I present results from OLS regression models (that treat surgeon volume as exogenous) of patient mortality on surgeon volumes in columns 1 and 2, in order to facilitate comparison with existing volume-outcome studies, and to demonstrate the need for instrumenting for surgeon volume. While the first column shows evidence for an inverse relationship between surgeon volume and patient mortality, this cannot be treated as evidence for learning-by-doing because this coefficient contains the influence of "selective referral" as well. Moreover, the coefficient drops drastically in magnitude (and becomes statistically insignificant) once surgeon fixed effects are included, implying that the negative coefficient was not a result of learning-by-doing but was being identified off differences across surgeons.

In order to exclude any influence of the "selective referral" effect, I instrument for surgeon volume and present these results in columns 3 through 6. The result in Column 3 lends strong support to the learning-by-doing hypothesis: the coefficient on volume is negative and significant ($p=.005$). The magnitude of the coefficient implies that performing one additional procedure in the previous year leads to a decline of .05 percentage points in mortality. This represents a decrease of around 1.2% relative to the average value of mortality in the data (.0382), which is indicative of a strong learning-by-doing effect. Column 4 adds surgeon fixed effects to the model, implying that learning effects are now computed only using within-surgeon variation. The coefficient on lagged surgeon volume is still strongly significant ($p=.015$), and the magnitude of the volume coefficient is now slightly larger in magnitude (-.0007 as compared to -.00051 in column 3). The results of a Hausman test (not included here) clearly reject equality between OLS and IV estimates. The coefficient on volume obtained from OLS regression is smaller in magnitude when compared to that obtained from the IV regressions. As discussed earlier, this implies that the type of patients attracted to high quality surgeons tend to be sicker in ways that are unobservable.

---

[37] Models that do not contain surgeon fixed effects have standard errors clustered by surgeon.

In columns 5 and 6, I include an interaction term in order to allow for a nonlinear effect of procedure volume on mortality. The coefficient on procedure volume is still negative and does not change much in magnitude when compared to the earlier specifications (-.00058 vs. -.00051 for the corresponding specification in Column 3). The coefficient on the interaction term, however, is positive and significant in Column 5 and falls just short of statistical significance at the 10% level in the fixed effects specification in Column 6 (*p=.011*). These results indicate that high volume surgeons (defined as those performing 200 procedures a year or above) benefit less from additional experience when compared to their counterparts who do fewer procedures a year.

6.3 Is Experience Specific?

Column 1 of Table 5 presents results from the model in Section 5.3 that tests for the firm-specificity of human capital acquired through learning-by-doing. The estimates reject the hypothesis of firm-specificity. The coefficient on own-hospital volume is negative and significant (*p=.015)* and only slightly larger in magnitude than the coefficient on other-hospital volume, which is also statistically significant (*p=.05*). Further, a t-test is not able to reject equality of the coefficients (*p=.729*) on the two variables measuring procedure volume (at own and other hospitals), indicating that the benefits of procedure experience are portable across hospital settings. This finding is in contrast to the findings of Huckman and Pisano (2006) who find that performance of cardiac surgeons is firm-specific.

To understand the implication of this result, we need to first examine the exact mechanism through which experience affects the human capital of surgeon in this setting. As discussed earlier, increased procedure experience helps surgeons increase their familiarity with any complications that may arise. It is reasonable to believe that these are benefits which will boost surgeon quality no matter where the procedure is performed. Firm-specific settings play an important role as well for reasons stated earlier. However, the empirical results seem to suggest that the general aspect of surgeon human capital that is gained from experience outweighs the specific aspect.

An interesting implication of this result is on the division of rents between surgeons and hospitals in this setting. The finding that surgeon performance is portable across hospitals suggests that surgeons should end up capturing any rents that may accrue from this relationship. Surgeons are not hospital employees, but typically operate as freelance agents who enter into contractual relationships with hospitals. Further, the law forbids hospitals from paying physicians with the objective of directing the physician's patients towards the hospital. One way in which surgeons may capture some of these rents is by forming their own specialty hospitals along with the help of outside investors. Since learning is portable, high quality surgeons are able to leave a hospital but maintain their skills while starting a new practice.

In Column 2, I present results from the specification testing for task-specificity of human capital. The coefficients on both volume terms (CABG procedure volume and non-CABG procedure volume) are negative and statistically significant at the 5% level. However, the CABG volume coefficient is much larger in magnitude (-.00046 vs. -.00009). [38] These estimates indicate that there is some benefit of performing non-CABG procedures on CABG outcomes, but the impact of this experience is smaller than the impact of experience performing CABG procedures. I interpret this as evidence in favor of some degree of task-specificity of human capital. The main implication of this result is that there seems to be some economies of scope across procedures for surgeons. This provides one explanation as to why surgeons do not specialize in performing just one particular procedure.

6.4 Robustness Checks

I test the robustness of the estimates through various specification checks and sample restrictions. Table 6 contains results from these specifications. In all specifications, I focus on the coefficient of lagged surgeon volume which is the main coefficient of interest. In column 1, I test the sensitivity of the results to an alternate definition of the instrument. I relax the restriction on surgeon age used in the definition and designate all surgeons who stop performing procedures (and who perform at least 5% of the hospital's CABG procedures) as exitors. This increases the number of exiting surgeons in the data to 63. The coefficient on surgeon volume is now smaller in magnitude, but still statistically significant (-.00025 vs. -.0007, p=.029).

---

[38] A t-test rejects equality of the two coefficients at the 5% level.

In columns 2 and 3, I address the concern about using recent volume (as opposed to cumulative volume) to proxy for surgeon experience. In column 2, I use the procedure volume of the surgeon over the last *two* years and in column 3, I use the cumulative volume as measures of experience. In both cases I calculate the instrument in an analogous manner. Since I observe cumulative experience only for surgeons who start practicing after 1998, I restrict the sample accordingly in column 3. The coefficient on lagged surgeon procedure volume remains statistically significant (*p=.008*) and largely unchanged when I use the two-year measure instead of the prior year measure (-.00054 vs. -.0007). When I restrict the sample to surgeons who start practicing in 1998 or later (in column 3), the coefficient on lagged procedure volume has the right sign (-.00028 vs. -.0007) but is statistically insignificant (*p=.420*). This is probably an artifact of the smaller sample size in this specification.

Column 4 tests the sensitivity of the results to excluding all the control variables from the model. The coefficient magnitude is similar to that in the original specification (-.00062 vs. -.00070) but there is some loss in precision (*p=.098*). Finally, in column 5, I include surgeon-hospital interaction dummies in the regression instead of including them separately. The model is robust to this specification as well.

I also tested the robustness of the results to the sample restrictions imposed earlier, by including low volume hospitals and surgeons in the sample and found that the main conclusions were not affected. In summary, the results are quite robust to alternate specifications. The only situation in which we lose statistical significance is when the sample is restricted to surgeons who start practicing in 1998 or later. In all other models, our main conclusions remain unchanged.

A possible alternate explanation behind the results is that exiting surgeons chose patients from zip codes whose inhabitants were healthier in unobserved ways. When these surgeons exit, the staying surgeons seem to experience an increase in quality (cause by the rise in procedure volume) when, in reality, they are now operating on patients who are healthier. I address this concern in two ways. First, the data show that patients of exiting surgeons had an average Charlson co-morbidity index of .894 compared to patients of staying surgeons who had an average Charlson index of .882. Since the patients of exiting surgeons seem to be as healthy as the patients of staying surgeons in observable ways, it is reasonable to assume that they do not differ significantly in unobservable ways. Second, the estimates show that increased procedure experience benefits surgeons across hospitals (and thus, across zip codes). This implies that the result is not driven by unobservable differences in patient health across zip codes.

7 Concluding Remarks

In this paper I examine the benefits of experience, one of the mechanisms through which individuals acquire (or maintain) competence in an organizational setting. Specifically, I study whether cardiac surgeons who perform more procedures experience an improvement in performance. In order to do so, I develop an instrumental variables estimation method that addresses the potential endogeneity of surgeon procedure volume. As my identification strategy, I consider exogenous shocks to the procedure volume of existing CABG surgeons in Florida caused by the exit of other surgeons from the same hospital. Using this instrument, I find evidence for a strong learning-by-doing effect for cardiac surgeons: an additional procedure a year leads to a reduction in patient mortality by 1.2%. I also find that procedure experience adds to a surgeon's general human capital, i.e. it has a positive influence on her performance across firms. In addition, I find evidence in support of some degree of task-specificity of surgeon human capital, i.e. a surgeon's experience performing one procedure has a positive effect on patient outcomes in other procedures, but this effect is smaller than the effect of experience performing the procedure in question.

These findings have implications for managers within the health care industry, specifically with respect to how to organize the firm so as to better leverage the skills of its professional workers. In highly specialized settings such as cardiac surgery, the knowledge retained by individual workers is often key to the successful functioning of the entire organization. While allocating procedure volume within a team, it is important for managers to bear in mind that the marginal benefit of additional experience is greater for low volume surgeons when compared to high volume surgeons. The quality of surgeons with low procedure volumes needs to be actively monitored. Hospitals can benefit from developing mentoring relationships between low-volume (or new) surgeons and more experienced surgeons. The more experienced surgeon can aid the low-volume surgeon in patient selection and also be available to assist in case of complications during the procedure.

The analysis presented in this paper has its share of limitations, one of the main ones being that the effect of hospital procedure volume on outcomes is not modeled. Since physician and hospital services are co-located, one cannot sort out their separate effects so easily. While some of the positive benefits of hospital volume on outcomes are captured by surgeon volume, others may not be. For example, larger hospitals may have a broader range of specialists and thus have experience with more diseases or co-morbidities and may learn to manage these better, leading to an improvement in patient outcomes. To the extent that such effects are not captured by hospital fixed effects and other hospital characteristics in our regression, our estimates of surgeon learning-by-doing will be biased. Since hospital volume affects outcomes in the same way as surgeon volume[39], we can think of the estimates in this study as representing an upper bound on surgeon learning-by-doing.

Another limitation pertains to the analysis of firm-specificity. The analysis assumes that the surgeon decision to operate across multiple hospitals is exogenous to her quality. This may not be true. One can imagine a situation in which surgeons who split their time across hospitals are of, say, higher quality (and are hence able to attract patients across hospitals). In that case, the estimates from the specifications in section 6.2 may be biased. A better strategy would be to instrument for the splitting decision. I leave this avenue open for further investigation.

A final limitation of the study is that CABG may be performed by teams consisting of two or more surgeons. Since the data only identifies the primary operating surgeon, I am unable to factor for the experience of other surgeons in the same surgical team who may have been involved in the procedure.

One of the questions that may arise about this study pertains to what exactly is being measured. Do the estimates reflect the core learning process of surgeons? Or do they represent the process by which surgeons maintain their skills, as opposed to learning new ones? The data does not allow me to disentangle these two effects. However, to the extent that surgeons perfect their procedural skills during (residency) training or the first couple of years of practice itself, most established surgeons can be thought of as already being on the

---

[39] Studies that have tested for the effects of both surgeon and hospital volume (in a correlational manner, without factoring for possible endogeneity biases) have found surgeon volume to be a much stronger predictor of patient outcomes

“flat” portion of their learning curves. Maintaining proficiency of skills by ensuring an adequate volume of procedures remains a recognized concern of surgeons. This concern arises due to the deterioration of skills with lack of practice: a surgeon who has an opportunity to consistently perform procedures over time will be able to maintain her abilities and skill level better than an equally trained surgeon who performs only a handful. Thus, the estimates may well measure the degree of “not-forgetting” as opposed to “learning”.

This study adds to the large and growing theoretical and empirical literature that analyzes learning-by-doing in a variety of settings and industries. While the unit of analysis in most of these studies is the firm, this paper is able to look at learning-by-doing at the level of the individual worker, mainly due to the availability of detailed data. This study is to be viewed as a first step towards understanding the mechanism by which workers acquire and maintain their skills in organizations. While I have documented the existence of learning-by-doing in individuals, the link between individual and organizational learning is yet to be made. The data also lends itself well to exploring other aspects of learning in organizations, e.g. forgetting, learning spillovers etc.

Finally, one can use this setting to study whether the improvement in quality of the surgeon, caused by increased experience, translates into a greater share of patients in the market. More generally, one can quantify the extent to which learning-by-doing contributes to sustainable competitive advantage. I plan to explore these issues in future research.

References

Argote, L., Beckman, S. and Epple, D. 1990. "The Persistence and Transfer of Learning in Industrial Settings", *Management Science,* 36(2), 140-154

Argote, L., Brooks, D., and Regans, R. 2005. "Individual Experience and Experience Working Together: Predicting Learning Rates from Knowing Who Knows What and Knowing How to Work Together", *Management Science*, 51(6), 869-881

Argote, L., Epple, D., Devadas Rao, R., and Murphy, K. 1997. "The Acquisition and Depreciation of Knowledge in a Manufacturing Organization: Turnover and Plant Productivity", *Working Paper*, Carnegie Mellon University

Argote, L., Epple, D., and Devadas Rao, R. 1991. "Organizational Learning Curves: A Method for Investigating Intra-Plant Transfer of Knowledge Acquired Through Learning-by-Doing", *Organization Science*, 2(1), 58-70

Argote, L., Epple, D., and Murphy, K. 1996. "An Empirical Investigation of the Microstructure of Knowledge Acquisition and Transfer Through Learning-by-Doing", *Operations Research,*44(1), 77-86

Becker, G. 1962. "Investment in Human Capital: A Theoretical Analysis", *The Journal of Political Economy*, 70(5), 9-49

Benkard, L. 2000. "Learning and Forgetting: The Dynamics of Aircraft Production", *American Economic Review*, 90(4), 1034-1054

Besanko, D., Doraszelski, U., Kryukov, Y. and Satterthwaite, M. 2005. "Learning-by-Doing, Organizational Forgetting, and Industry Dynamics", *Working Paper*, Northwestern University

Bessen, J. 1998. "Productivity Adjustments and Learning-by-Doing as Human Capital", *Research on Innovation Working Paper*

Darr, E., Argote, L. and Epple, D. 1995. "The acquisition, transfer and depreciation of knowledge in service organizations: Productivity in franchises", *Management Science,* 41(11), 1750-1762

Delaney, P., Reder, L., Staszewski, J.J. and Ritter F. 1998 "The Strategy-Specific Nature of Improvement: The Power Law Applies by Strategy within Task", *Psychological Science*, 9, 1-7

Farley, D. and Ozminkowski, R. 1992. "Volume Outcome Relationships and in hospital mortality: the effect of changes in volume over time" *Medical Care*, 30(1)

Gaynor, M., Seider, H. and Vogt, W. 2005 "The Volume-Outcome Effect, Scale Economies and Learning-by-Doing", *American Economic Association Papers and Proceedings*, May 2005, 243-247

Gibbons, R. and Waldman, M. 2004. “Task-Specific Human Capital”, *American Economic Association Papers and Proceedings*, May 2004, 203-208

Gowrisankaran, G.., Ho, V. and Town, R. 2006 “Causality and the Volume-Outcome Relationship in surgery”, *NBER Working Paper*

Halm E., Lee C. and Chassin, M. 2002. “Is Volume related to outcome in health care? A systematic review and methodologic critique of the literature”, *Annals of Internal Medicine,* 137(6)

Hamilton, B. and Hamilton, V. 1997 “Estimating Surgical Volume-Outcome Relationships Applying Survival Models: Accounting for Frailty and Hospital Fixed Effects”, *Health Economics*, 6: 383-395

Hannan, E., O’Donnell, J., Kilburn, H, Bernard, H., and Yazici, A. 1989. “Investigation of the relationship between volume and mortality for surgical procedures performed in NY state hospitals”, *JAMA*, 262(4), 503-510

Hannan, E., O’Donnell, J., Bernard, H., Lucacik, G. and Shields, E. 1991. “Coronary Artery Bypass Graft surgery: The relationship between in-hospital mortality rate and surgical volume after controlling for clinical risk factors”, *Medical Care*, 29(11), 1094-1107

Hannan, E., O’Donnell, J., Bernard, H., Yazici, A, Lindsey, M.. and Shields, E. 1992. “ A longitudinal analysis of the relationship between in-hospital mortality in New York state and the volume of Abdominal Aortic Aneurysm surgeries performed”, *Health Services Research,* 27(4)

Hatch, N. and Mowery, D. 1998 “Process Innovation and Learning-by-Doing in Semiconductor Manufacturing”, *Management Science*, 44, 1461-1477

Hirshfeld, J.W., Ellis S.G., Faxon, D.P. et al. 1998 “Recommendations for the Assessment and Maintenance of Proficiency in Coronary Interventional Procedures”, *Journal of the American College of Cardiology*, 31(3), 722-743

Ho, V. 2002. “Learning and the evolution of medical technologies: The diffusion of Coronary Angioplasty” *The Journal of Health Economics*, 21(5), 873-885

Huckman, R. and Pisano, G. 2006, “The firm-specificity of individual performance: Evidence from cardiac surgery”, *Management Science*, 52(4), 473-488

Hughes, R., Hunt, S. and Luft, H. 1987 “Effects of Surgeon volume and hospital volume on quality of care in hospitals” *Medical Care* 25(6), 489-503

Jovanovic, B. 1979 “Firm-Specific Capital and Turnover”, *The Journal of Political Economy*, 87, 1246-1260

Killingsworth, M. 1982 “Learning-by-Doing and Investment in Training: A Synthesis of Two Rival Models of the Life Cycle”, *The Review of Economic Studies*, 49(2), 263-271.

Luft, H., Bunker, J.P., Enthoven, A. 1979. “Should operations be regionalized? The empirical relation between surgical volume and mortality”. *New England Journal of Medicine.* 301(25), 1364-1369

Luft, H, Hunt, S., and Maerki, S. 1987 “The Volume-Outcome Relationship: Practice-makes-Perfect or Selective-Referral Patterns?” *Health Services Research* 22(2), 157-182

Luft, H., Garnick, D., Mark, D., and McPhee, S. 1990. “Hospital Volume, Physician Volume and Patient Outcomes”, Ann Arbor Health Administration Press

Nembhard, D. and Uzumeri, M. 2000. “An Individual-Based Description of Learning within an Organization”, *IEEE Transactions on Engineering Management*, 47(3), 370-378

Picone, G., Trogdon, J. and Jollis, J. 2005 “Hospital Volume and Quality of Care: practice-makes-perfect or selective referral”, *NBER Working Paper*

Pisano, G., Bohmer, M.J., and Edmondson, A. 2001 “Organizational Differences in Rates of Learning: Evidence from the Adoption of Minimally Invasive Cardiac Surgery”, *Management Science*, 47, 752-768

Shafer, S.M., Nembhard, D., and Uzumeri M. 2001. “The Effects of Worker Learning, Forgetting and Heterogeneity on Assembly Line Productivity”, *Management Science*, 47, 1639-1653

Staiger, D. and Stock, J. 1997 “Instrumental Variables Regression with Weak Instruments”, *Econometrica*, 65(3), 557-586

Thompson, P. 2001. “How Much Did the Liberty Shipbuilders Learn? Evidence for an Old Case Study”, *Journal of Political Economy*, 109(1), 103-37

Thompson, P. 2006. “How Much Did the Liberty Shipbuilders Forget?”, *Working Paper*, Florida International University

Thornton, R. and Thompson, P. 2001. “Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding”, *American Economic Review*, 91(5):1350-1368

Wright, T.P. 1936. “Factors Affecting the Cost of Airplanes”, *Journal of the Aeronautical Sciences*, 3(4): 122-128

Table 1: Summary Statistics for Key Variables

	Mean	Std. Deviation	Number of Obs.
Dependent Variable			
Mortality Rate	3.82%	19.17%	160210
Volume Measures			
Total Surgeon Volume, prior year	174.6	82.9	129554
Hospital-Specific Surgeon Volume, prior year	142.9	86.3	126291
Hospital volume	677.4	473.3	160210
Patient Characteristics			
Age	67.5	10.6	160210
% Female	29.4	45.6	160210
Charlson Index	.88	.97	160210
Prior CABG surgery	.06	.24	160210
Concurrent PTCA	.03	.16	160210
Heart Failure	.21	.41	160210
Cardiogenic shock	.02	.14	160210
Hypertension	.01	.10	160210
Hospital Characteristics			
Number of cardiac care beds	23.9	30.6	130371
Number of high-tech services	2.4	.82	130371
Number of FTE Registered Nurses	160.7	.16	160210
Number of FTE Licensed Practical Nurses	86.3	.41	160210
Surgeon Characteristics			
Years since Medical School	21.5	7.8	160210
% Foreign Medical School	17.4	.38	160210
Instrument			
Exit Volume	2.95	10.3	129554

Note: All statistics are calculated using the patient as the unit of observation. The sample contains all patients treated with CABG surgery in Florida for the years 1998-2003. The sample excludes a few low-volume surgeons and hospitals (refer section 3 for details). The discrepancy in the number of observation for surgeon and hospital characteristics is due to missing data. In the case of the volume measures, the discrepancy is because these measures are lagged by one year.

Table 2: Testing the Identification Assumption: Is the Instrument related to Surgeon Quality?

Dependent Variable: Exit volume faced by surgeon

Lagged Surgeon Quality	0.002
	(0.116)
Charlson Index	-1.471**
	(0.588)
Cardiogenic Shock	-2.196
	(4.511)
Concurrent PTCA	1.477
	(3.229)
Hypertension	0.891
	(4.787)
Heart Failure	0.338
	(1.417)
Prior CABG	1.597
	(2.788)
Female	-1.328
	(1.505)
Heart Attack	0.884
	(1.425)
Age	-0.098**
	(0.04)
Hospital Characteristics	Y
Year Fixed Effects	Y
Number of Observations	936

Note: Regression carried out at the surgeon-year level. Standard errors are reported in parentheses.

**** signifies p<.01, ** signifies p<.05 and * signifies p<.1*

Table 3: Relationship between Surgeon Volume and Exit Volume (First Stage)

Dependent variable: Surgeon procedure volume, prior year

	(1)	**(2)**
Exit Volume, prior year	0.914***	0.611***
	(0.055)	(0.03)
Charlson Index	-0.02	0.047
	(0.196)	(0.098)
Cardiogenic Shock	-2.505*	0.516
	(1.36)	(0.683)
Concurrent PTCA	-2.584**	-0.251
	(1.208)	(0.606)
Hypertension	-0.754	1.727*
	(2.018)	(1.012)
Heart Failure	-1.540***	-0.08
	(0.476)	(0.239)
Prior CABG	3.999***	0.672
	(0.823)	(0.413)
Female	-0.645	-0.032
	(0.418)	(0.21)
Heart Attack	-3.776***	0.268
	(0.484)	(0.243)
Age Categories	Y	Y
Hospital Characteristics	Y	Y
Surgeon Characteristics	Y	N
Year Fixed Effects	Y	Y
Hospital Fixed Effects	Y	Y
Surgeon Fixed Effects	N	Y
Number of Observations	107588	107977

Note: Columns 1 and 2 present models without and with surgeon fixed effects, respectively. The difference in number of observations is due to missing data on some surgeon characteristics. Standard errors are reported in parentheses. Models without surgeon fixed effects have standard errors clustered by surgeon.

**** signifies p<.01, ** signifies p<.05 and * signifies p<.1*

Table 4: The Effect of Total Surgeon Experience on Patient Outcome

	Dependent Variable: Did the patient die?					
	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) IV	(6) IV
Surgeon Procedure Volume, prior year	-3.9E-05***	-2.31E-06	-5.08E-04***	-7.0E-04**	-5.9E-04***	-7.8E-04**
	(8.96E-06)	(1.78E-05)	(1.79E-04)	(2.89E-04)	(2.08E-04)	(3.13E-04)
Surgeon Volume (prior year)*High Vol					3.64E-04***	2.3E-04
					(1.38E-03)	(1.42E-04)
Charlson Index	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cardiogenic Shock	0.341***	0.343***	0.342***	0.341***	0.342***	0.341***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Concurrent PTCA	0.010***	0.011***	0.009***	0.010***	0.010***	0.009***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Hypertension	0.028***	0.028***	0.027***	0.029***	0.027***	0.029***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Heart Failure	0.030***	0.030***	0.029***	0.030***	0.029***	0.030***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Prior CABG	0.029***	0.029***	0.031***	0.029***	0.031***	0.029***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Female	0.014***	0.015***	0.014***	0.014***	0.014***	0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Heart Attack	0.005***	0.005***	0.004**	0.006***	0.004***	0.005***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Age Categories	Y	Y	Y	Y	Y	Y
Hospital Characteristics	Y	Y	Y	Y	Y	Y
Surgeon Characteristics	Y	N	Y	N	Y	N
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Hospital Fixed Effects	Y	Y	Y	Y	Y	Y
Surgeon Fixed Effects	N	Y	N	Y	N	Y
Number of Observations	107588	107977	107588	107977	107588	107977

Note: Columns 1 and 2 present OLS models without and with surgeon fixed effects, respectively. Columns 3 through 6 present IV estimates. Standard errors are reported in parentheses. Models without surgeon fixed effects have standard errors clustered by surgeon. The slight discrepancy in the number of observations across models is due to missing data on surgeon characteristics.
**** signifies p<.01, ** signifies p<.05 and * signifies p<.1*

Table 5: Testing for Specificity of Experience

Dependent Variable: Did the patient die?

	(1)	(2)
Surgeon Procedure Volume, prior year, own hospital	-5.73E-04**	
	(2.36E-04)	
Surgeon Procedure Volume, prior year, other hospitals	-5.12E-04**	
	(2.65E-04)	
Surgeon Procedure Volume, prior year, CABG		-4.64E-04**
		(2.28E-04)
Surgeon Procedure Volume, prior year, non-CABG		-9.17E-05***
		(2.53E-05)
Patient Characteristics	Y	Y
Hospital Characteristics	Y	Y
Surgeon Characteristics	N	N
Year Fixed Effects	Y	Y
Hospital Fixed Effects	Y	Y
Surgeon Fixed Effects	Y	Y
Number of Observations	105066	107977

Note: Column 1 presents results from a test of firm-specificity of learning-by-doing. Column 2 presents results from a test of task-specificity. Standard errors are reported in parentheses.

**** signifies p<.01, ** signifies p<.05 and * signifies p<.1*

Table 6: Some Robustness Checks

Dependent Variable: Did the Patient die?	(1) Alt. Def.	(2) 2 Year	(3) Cumul.	(4) No Control	(5) Surg-Hosp
Surgeon Volume, prior year	-2.58E-04**	-5.48E-04***	-2.87E-04	-6.24E-04*	-7.15E-04***
	(1.18E-04)	(2.06E-04)	(3.55E-04)	(3.77E-04)	(2.70E-04)
Patient Characteristics	Y	Y	Y	N	Y
Hospital Characteristics	Y	Y	Y	N	Y
Surgeon Characteristics	N	N	N	N	N
Year Fixed Effects	Y	Y	Y	Y	Y
Hospital Fixed Effects	Y	Y	Y	Y	N
Surgeon Fixed Effects	Y	Y	Y	Y	N
Surgeon x Hospital dummies	N	N	N	N	Y
Number of Observations	107977	81739	13437	129184	107588

Note: All columns present results from IV regressions of patient mortality on total surgeon volume. Column 1 uses an alternate definition of exit. Columns 2 and 3 use volume from the last 2 years and cumulative volume instead of prior year volume as the measure of surgeon experience. Column 4 estimates the model with only fixed effects and no control variables. Column 5 estimates the model with surgeon-hospital fixed effects. Standard errors are reported in parentheses.
**** signifies p<.01, ** signifies p<.05 and * signifies p<.1*