Working Paper Series
Congressional Budget Office
Washington, D.C.

Does Doctors’ Experience Matter in LASIK Surgeries?

Juan M. Contreras
U.S. Congressional Budget Office
Juan.contreras@cbo.gov
Tel:1-202-226-9099

Beomsoo Kim
Department of Economics
Korea University
kimecon@korea.ac.kr
Tel: 82-2-3290-2204

Ignez M. Tristao
Inter-American Development Bank
ignezt@iadb.org
Tel: 1-202-623-1647

March 2010

2010-01

(An earlier version of this document was published as
CBO Working Paper 2007-12.)

CBO's Working and Technical Papers are preliminary and are circulated to stimulate discussion and critical comment. The papers are not subject to CBO's formal review and editing process. The analysis and conclusions expressed are those of the authors and should not be interpreted as those of the Congressional Budget Office. References in publications should be cleared with the authors. Papers in this series can be obtained from www.cbo.gov/publications.

Does Doctors' Experience Matter in LASIK Surgeries?

ABSTRACT

In this paper, we use a longitudinal census of laser in situ keratomileusis (LASIK) eye surgeries collected directly from patient charts to examine the learning-by-doing hypothesis in medicine. LASIK surgery has precise measures of presurgical condition and postsurgical outcomes. Unlike other types of surgery, the impact of unobservable underlying patient conditions on outcomes is minimal. Individual learning by doing is identified through observations of surgical outcomes over time, based on the cumulative number of surgeries performed. Collective learning is identified separately, through changes in a group adjustment rule determined jointly by all the surgeons in a structured internal review process. Our unique data set overcomes some of the measurement problems in patient outcomes encountered in other studies and improves the possibility of identifying and separating the impact of learning by doing from other effects. We cannot conclude that the outcome of LASIK surgery improves as an individual surgeon's experience increases, but we find strong evidence that experience accumulated by surgeons as a group in a clinic significantly improves outcomes.

Key words: learning, experience, LASIK, volume

JEL codes: I10; I12; I18

We are very grateful to the Maryland Population Research Center at the University of Maryland for providing financial support to fund this research. Beomsoo Kim acknowledges the grant support of Korea University. We are thankful to the administration and surgeons of the Clínica Oftalmológica de Antioquia (CLOFAN) for their help and collaboration, especially to Dr. Luis Fernando Pelaez, Dr. Mauricio Jaramillo, and Dr. Leonardo Orjuela for their support during our data collection in Colombia and for answering our multiple questions about the procedure and the specifics of the clinic operation. We also would like to thank William Evans, Seth Sanders, John Rust, and Mark Duggan for helpful discussions and advice during the elaboration of this work and Chris Ruhm, David Ribar, Stephen Holland, Chris Swann, Doug Hamilton, Bob Dennis, Jeff Kling, Se-oh Oh and Jeong Im Lee for providing many helpful comments. The analysis and conclusions expressed here are those of the authors and should not be interpreted as those of the Congressional Budget Office.

1. INTRODUCTION

How much does experience enhance a physician's skills and improve patient outcomes? Although economists have tried hard to empirically determine the existence of learning by doing in medicine and discover the mechanism behind this phenomenon, it has proved elusive.

In this paper, we address those questions using a unique data set of LASIK refractive eye surgeries–operations with well-defined eligibility criteria, precise measures of underlying conditions and postsurgery outcomes, and minimal postsurgical care. We took the data directly from individual patients' charts as part of a two-year longitudinal census of LASIK surgeries from one of the largest ophthalmologic clinics in Colombia.

We find strong evidence that experience accumulated by doctors as a group in a clinic substantially improves outcomes in LASIK surgery. By contrast, we find that the role of an individual surgeon's experience in improving patient outcomes is insignificant in a statistical sense.

The hypothesis of learning by doing postulates that workers can learn through experience in the production process, resulting in economies of scale in future production. Learning curves have been studied and documented in many areas, including economics and operations management. Over the last 30 years, many studies in the medical field have analyzed the relationship between the physician's (or hospital's) volume of surgical procedures and patient outcomes. Halm et al. (2002) reviewed 135 papers studying this correlation in the health care industry between 1980 and 2000. Approximately 70 percent of the papers reported a statistically significant association between higher volume and better outcomes. Based on this and other studies, the *Washington Post* wrote that it is now a sacrosanct belief in the medical world that patients fare better if operated on by a doctor who frequently performs the same operation.[1] Researchers often have interpreted this correlation as evidence of a causal relationship–that "practice makes perfect." Moreover, this view, that "practice makes perfect," provides support for the expansion of regional or specialized medical centers to facilitate learning by doing.[2] Debates on whether society can derive more benefit from expensive care in regional or specialized centers, such as critical care or cardiovascular surgical services, than from distributing those services among small clinics, depend on the existence and magnitude of learning by doing (Thompson et al., 1994; Menke and Wray, 2001). If learning by doing exists in sizable magnitude in medicine, society might benefit from regionalized facilities even after taking into account any additional costs resulting from difficulty of access.

There are, however, other possible explanations for the observed correlation between volume and outcome. One of them is selective referral. Primary-care physicians may refer their surgery patients to more skilled surgeons, causing those surgeons to show higher volume and better outcomes because of innate ability, not more experience. [3] If selective referral is indeed the cause of this relationship of high volume and better outcome, then regionalization could increase the price of medical care by reducing competition without improving outcomes. Determining the existence and magnitude of a surgeon's learning curve is at the heart of this issue.

---

[1] Boodman, Sandra G. "High Volume, High Quality?" Washington Post, October 28, 2003.

[2] In this context, regionalization of medical care entails the financing of larger medical centers instead of many small ones. See Luft et al. (1979) and Rathore et al. (2006).

[3] Selective referral can also reflect the fact that patients themselves select specific doctors using their primary care physician as a patient's agent. On the other hand, a high risk patient may look for a surgeon with a better record (high ability) because their marginal benefit of being treated by a better doctor is higher than the marginal benefit for a low risk patient. If the econometrician cannot observe the patient risk, the performance of the surgeons with higher ability would be underestimated.

The other possibility is that doctors treat patients selectively. Dranove et al. (2003) found that physicians select relatively healthy patients to increase their probability of success when these success rates are to be observed by the public. If only high-volume doctors have the luxury of this selection, then high volume doctors may treat a higher proportion of lower-risk patient than their low-volume peers. Not controlling completely for patient risk could lead to an omitted variable bias on the estimate of the relationship between high volume and better outcomes.

The challenge of measuring a surgeon's learning curve is twofold: data limitations and hard-to-measure outcomes. Data limitations have prevented previous studies from observing the evolution of a surgeon's learning curve over time. Most studies have relied on annual (or, at best, quarterly) volume data for each surgeon or hospital because the exact day and time of each surgery is not available.[4] However, in our data we measure a surgeon's experience precisely from the cumulative number of procedures performed and the exact day and time (with a precision of seconds) of each procedure.

The difficulty of defining and measuring patient outcomes constitutes another challenge. Postoperative mortality is commonly used as an indicator of an adverse outcome because it is accurately recorded. But mortality is an extreme outcome, and relying on that alone is an inadequate means of capturing both qualitative information and the full range of outcomes.[5] In addition, the effect of a surgeon's skill is hard to isolate from a patient's underlying conditions, which may affect outcomes.[6] Even with the most detailed data, some relevant underlying patient

---

[4] A few recent studies mentioned in the next section used individual level data. For example, Huckman and Pisano (2006), who use data from the Pennsylvania Health Care Cost Containment Council (PHC4) and include patient-level records for every individual receiving coronary artery bypass graft (CABG) at a hospital in Pennsylvania in 1994 and 1995. That procedure, however, is not performed by a single surgeon but instead by a surgical team, making it difficult to separate individual learning from collective learning.

[5] For example, morbidity or quality of life might also be important outcomes of a medical procedure.

[6] For example, in coronary surgery some conditions like smoking behavior or a history of heart disease can affect

conditions are not recorded. And lastly, patient outcomes also depend on other factors, such as the quality of the surgical team and the postsurgical care provided by the hospital staff. Thus, it is difficult for researchers to measure an individual physician's learning curve.

Another challenge is to separate collective learning from individual learning in the measurement process. Hospitals are complex organizations, and surgery involves not only an individual surgeon but also a surgical team and hospital staff. For example, coronary artery bypass graft (CABG) surgery usually requires a cardiac surgeon, a surgical assistant or a surgical team, a nursing team, an anesthesiology team, and a perfusionist.[7] In those situations, researchers cannot determine whether the final outcome is the result of one particular surgeon or the surgical team. Patient outcomes could depend on team coordination, which is acquired through group learning. It is also possible that final outcomes improve through the interactions of surgeons.[8] For instance, a cardiac surgeon might learn from his or her peers' experience while interacting with them.

LASIK surgery has common characteristics with other surgical procedures. For example, two key inputs to any surgery are the surgeon's skill and the technology. The surgeon's skill is important in LASIK in developing the surgical plan and executing that plan during surgery. The laser machine is also important because the laser ablation depends on the machine's technology.

---

the outcome of the surgery.

[7] http://health.yahoo.com/respiratory-resources/health-professionals-involved-with-coronary-artery-bypass-graft-surgery/healthwise--ue4707abc.html

[8] Huckman and Pisano (2006) suggest two explanations for the existence of firm-specific performance of workers. The first explanation is the complementarity between a worker and the human, physical or organizational assets held by a given firm. They cite as an example the familiarity between members of a team and cite Pisano et al. (2001), who note that "well-developed surgical teams are often capable of performing procedures with minimal verbal communication between members." The second explanation stresses the influence of an individual within a particular organization, for example, a surgeon demanding specific resources from that hospital for his or her own procedures. Both cases relate to the observation of outcomes from specific teams rather than outcomes from individuals independent from the team.

LASIK surgeries also have particular characteristics that allow us to mitigate the measurement limitations mentioned above. First, the outcome of this surgery is measured precisely and is relatively unaffected by unobserved underlying patient conditions.[9] Second, the procedure is performed by only one surgeon, so we can measure the effect of each surgeon's experience on outcomes. Third, patients require almost no postsurgical care. Such care can complicate the measurement of the effect of a surgeon's experience on outcomes. And fourth, we observe a clear mechanism of group interaction which allows us to measure collective learning separately from individual learning.

In addition, our data is free from selective referral. Patients in most of cases are randomly assigned to a doctor in order to equalize the workload across doctors who are equal shareholders of the clinic.[10] This special feature of the data allows us not to worry about selective referral, which has been very hard to tease out in previous literature. Most of the previous studies were not able to deal with selective referral, although some, like Gowrisankaran et al. (2006), addressed this issue assuming that people would choose hospitals based on proximity and not on selective referral, an assumption that sometimes might not hold.

This paper is divided into seven sections. Section 2 provides a brief review of the literature. Section 3 gives some background on LASIK surgery. Section 4 describes the data, the measurement of outcomes, and our empirical methodology. Section 5 presents and discusses our

---

[9] LASIK surgery is not recommended for everyone with poor visual acuity. The presences of corneal malformations, irregular astigmatism, or a thin cornea are generally contraindications for refractive surgery. LASIK also is not recommended for patients with an autoimmune disease (e.g., lupus, rheumatoid arthritis) or immunodeficiency (e.g., HIV). Some doctors also do not operate on patients younger than 18 years old or with diabetes. For details, see Pallikaris and Siganos (1997) and Food and Drug Administration (FDA) guidelines on laser surgeries, http://www.fda.gov/cdrh/LASIK/when.htm.

[10] Patients that arrive with a particular doctor in their mind can be assigned to that doctor, but this case is not common in the clinic we collected the data.

findings, and Section 6 presents various robustness checks. Section 7 concludes with our results and an assessment of the limitations of this study.

2. LITERATURE REVIEW

Of the numerous studies that have analyzed the relationship between volume and outcomes in medical care, most have found a positive correlation between the two. The first analysis of the correlation between volume and outcomes in health care was done by Luft et al. (1979).[11] They compared outcomes from low- and high-volume hospitals for 12 surgical procedures and found that for certain operations mortality decreased with increased hospital volume.

More recently, Halm et al. (2002) reviewed 135 papers studying this correlation in the health care industry between 1980 and 2000. Approximately 70 percent of the papers reported a statistically significant and positive association between higher volume and better outcomes. Ninety of the papers reviewed by Halm et al. examined patient outcomes using the variation in cross-sectional hospital volume used by Luft et al.[12] One shortcoming of this approach is that unobserved hospital characteristics can affect the results. For example, if high-volume hospitals have better technology than low-volume hospitals, then the difference in outcomes may be due to the technology, which is often not factored into the data. In addition, the volume to outcome relationship analyzed in these papers measured only correlations, not causality, because volume may be endogenous to outcomes: better outcomes may lead to higher volume.

---

[11] Learning curves have been studied in other areas. Some examples are Wright (1936) and Asher (1956), who studied the cost-quantity relationships in the aircraft industry. In economics, the first theoretical model of this kind was constructed by Arrow (1962).

[12] This class of studies includes Hughes et al. (1988), Phillips et al. (1995), Jollis et al. (1994), and Kimmel et al. (1995).

To overcome the problem of unobserved heterogeneity in cross-hospital comparisons, some researchers have employed different identification strategies. Ho (2002) examined variation within one hospital over time, using the cumulative annual volume of percutaneous transluminal coronary angioplasty surgeries as a measure of experience. By observing outcome improvements over time, using 13 years of data that tracked annual volumes of coronary surgery for hospitals in California, she found evidence of learning by doing. Unfortunately, she could not control for changes in technology, which would have changed significantly over the 13-year span of the data. As a result, it is unclear from her results whether more experience or better technology improves outcomes over time. In addition, a hospital is a complex organization and postsurgical outcomes depend on multiple factors, such as the surgical team, the operating surgeon's skill, technology, and postsurgical care. Because Ho could not measure the impact of each of these factors separately, the exact reason for the improvement in outcomes is uncertain.

Several researchers have used comparisons among surgeons, based on the assumption that learning by doing occurs at the individual level. Birkmeyer et al. (2003) compared outcomes across high- and low-volume surgeons and found that outcomes had a stronger correlation with the volume of individual surgeons than with hospital-wide volume. However, the direction of causation is unclear. Did surgeon experience improve outcomes (as implied by the learning by doing hypothesis)? Or did better outcomes lead to higher volume (as implied by the competing hypothesis of selective referral)? Unfortunately, the evidence presented in the study cannot distinguish between these two hypotheses.

Recently, several papers have tried to determine the causal relationship between outcomes and volume using instrumental variables (IV) estimation. Gowrisankaran et al. (2006) focused on the outcome of three surgical procedures, using quarterly hospital volume as an explanatory variable. They used the predicted quarterly hospital volume for each procedure as a variable and assumed that, without a selective referral, people would go to the closest hospital. Their results indicate that, for at least two of the three procedures, learning by doing played an important role in explaining differences across hospitals in their risk-adjusted outcomes.

Huckman and Pisano (2006) used patient mortality from CABG surgery to show that the quality of a surgeon's performance at a particular hospital improves significantly with increases in procedure volume at that hospital, but does not significantly improve at other hospitals. In another study, Pisano et al. (2001) provided important insights into the learning process of hospitals by analyzing 660 patients in 16 institutions that implemented a new operation for minimally invasive cardiac surgeries. The study showed that cumulative experience is a significant predictor of learning and that the slope of the learning curve varies significantly across organizations.[13] They suggested that the underlying organizational process is the key to explaining the differences. In a follow-up study, Edmonson et al. (2001) qualitatively examined how different hospital teams implemented minimally invasive cardiac surgery technology.

Because the surgical procedures analyzed in previous studies were performed by a surgery team, it is difficult to separate individual learning from collective learning. For example, the studies by Huckman and Pisano (2006) and Pisano et al. (2001) examined a particular medical procedure where the surgery is performed by a team of doctors and nurses, and surgical outcomes may reflect the effect of the team instead of the effect of the surgeon.[14]

---

[13] Pisano et al. (2001) uses procedure time (time required to perform a surgery) as the measure of learning. Unfortunately, this is not a health outcome. They tried to find a correlation between procedure times and higher postoperative patient mortality rates, but they were not successful because of the low mortality rate due to the small sample size.

[14] They test for the influence versus familiarity of the surgeon with the surgical team. They conclude that familiarity is more important statistically than influence. However, the nature of the surgery and the available data do not give enough statistical power to differentiate between those two explanations

A few recent studies (Huesch, 2009; Vickers et al., 2007; Bridgewater et al., 2004) used individual level data to investigate learning by doing in medicine. Huesch (2009) and Bridgewater et al. (2004) used coronary artery bypass surgery from state of Florida in 1998-2006 and England in 1997-2003 respectively. They identified 57 and 15 new cardiac surgeons who operated during the data period and reached exactly the opposite conclusions. Huesch did not find any evidence of learning by doing but Bridgewater et al. found significant improvement over the first four years of practice. Vickers et al. (2007) examined the recurrence of prostate cancer after radical prostatectomy at four major medical centers involving 72 surgeons and found strong evidence of learning by doing. However, none of the studies above addressed the problem of selection by doctors or by patients.

In comparison with previous studies, this paper offers several advantages for the study of surgical learning by doing. While other studies have used data that identify time of surgery only by year–forcing researchers to use annual volume as a proxy for experience–our data identify the exact time of surgery as well as the day, month, and year.[15] Therefore, we can measure individual experience directly from the first to the *n*th surgery for each surgeon. Moreover, we do not need to be concerned about changing technology, because the most critical technology–the laser machine–did not change during the study period. A group adjustment rule, which was shared by all surgeons and updated over time following an explicit discussion among surgeons, allows us to measure collective learning separately from individual learning. Using LASIK to analyze surgeons' learning by doing has other important advantages, which we will discuss after

---

[15] Except for Pisano et al. (2001), in most studies the most disaggregated unit of data that researchers have used is quarterly volume of procedures in each year.

giving background on this. In addition, the random assignment of patients to doctors in our data strengthens the validity of our findings because they are not confounded by selection issues.

3. BACKGROUND ON LASIK SURGERY

Laser in situ keratomileusis (LASIK) is an elective refractive surgery that improves visual acuity by reshaping the cornea using a special laser.[16] Figure 1 shows how the surgery is performed. First, a surgeon creates a thin flap on the cornea with a special blade called a microkeratome.[17] The flap is folded back, and a laser is used to remove a certain amount of corneal tissue. In this section we explain why LASIK is a good procedure to analyze for individual and collective learning.

3.1 Why LASIK?

LASIK surgery shares many characteristics with other surgical procedures. For example, a surgeon's skill and the technology used are two of the most important factors determining patient outcomes. A surgeon's skill includes the diagnosis of the patient condition as well as the appropriate surgical plan. In addition to these common characteristics, LASIK surgery offers

---

[16] LASIK became very popular due to the fast vision recovery and minimal pain that accompany the procedure. Although there are other techniques for refractive surgery, LASIK is the most popular now in most countries, including the United States and Colombia, where we obtained our data set. Moreover, Colombia has been at the leading edge of developments in refractive surgeries (LASIK being one of multiple options) since Dr. José Barraquer laid down its theoretical and empirical bases in Bogotá in 1948. He provided doctors with knowledge about how much of the cornea had to be left unaltered to provide a stable long-term result, and created the procedure (called keratomileusis) and instrumentation (including the first microkeratome) to cut and reshape the cornea. Later technical and procedural developments included the RK (radial keratomileusis) in the 1970s in Russia by Svyatoslav Fyodorov, the development of PRK (photorefractive keratomileusis) in the 1980s in Germany by Theo Seiler, and finally the introduction in the early 1990s by Italian doctor Luccio Burroto and Greek ophthalmologist Ioannis Pallikaris of laser techniques to reshape the cornea. It was only in 1999 that LASIK was approved by the FDA.

[17] A microkeratome is a tool used to make a flap in the cornea that provides the surgeon access to the corneal stroma in order to reshape it. A smooth flap is crucial for the outcome of the surgery because an irregular surface causes blurred vision.

several advantages for measuring learning by doing. First, as mentioned above, the patient charts contain almost all of the information relevant to the outcome of the surgery, because effects from unobserved underlying patient conditions are minimal in LASIK surgery compared with other types of surgeries.[18]

Second, LASIK offers a unique opportunity to measure the outcome of a surgery precisely because we can accurately measure pre- and postsurgical eyesight. We explain this in more detail in section 4.2.

Third, LASIK, unlike many other surgeries, is performed in most cases by one surgeon. In our data, the operating surgeon is assisted by a nurse and an optometrist, who is responsible for the maintenance of the LASIK machine. Because there is only one optometrist in the clinic and the role of the assisting nurse is minimal, LASIK is an ideal surgery for capturing the surgeon's individual learning curve. In addition, LASIK does not require hospitalization, meaning that postsurgical care has a limited impact on the final outcome, compared with other surgeries that might require a hospital stay of several days.[19]

Collective learning may occur through interaction, coordination, specialization, or knowledge spillover (for example, information exchange or shared experience). In our study, we are able to measure collective learning separately from individual learning by taking advantage of the adjustment rule determined jointly by the surgeons. The collective learning mechanism we

---

[18] For example, outcomes of coronary surgeries, the most extensively studied surgical procedure, can be affected by weight, previous myocardial infarction, previous cardiac surgery, peripheral vascular disease, diabetes, renal function, hypertension, angina, dyspnoea (breathlessness), and smoking (National Adult Cardiac Surgical Database Report 1999-2000 (UK); Roques et al. 1999). In most studies, researchers have not been able to observe underlying conditions such as smoking or history of heart disease.

[19] The mean length of hospital stay after CABG was 8.2 days (Lazar et al., 1995).

observe occurs through sharing information, discussing the outcomes in regular meetings, and updating the group adjustment rule based on the review of the previous outcomes.

3.2 Factors Affecting Outcome in LASIK Surgery

Outcomes of LASIK surgery depend mostly on two inputs: technology and human capital.

With respect to the technology, two kinds affect LASIK surgery: hardware and software. In our paper, the hardware used is a Schwind ESIRIS laser machine.[20] The software, provided by the manufacturer, controls the delivery and fluence of the laser beams that reshape the cornea.[21] The software was upgraded once to enhance the performance of the machine. We control for the software version update in our analysis, and test whether this technological improvement affected patient outcomes.

With respect to human capital, it can accumulate at the individual and collective levels. At the individual level, LASIK surgery involves several ophthalmologic skills.[22] Collective human capital accumulation happens in LASIK because it requires an adjustment rule to

---

[20] LASIK surgical techniques have been evolving rapidly. However, applications of new techniques depend on the use of new machines. For example, eye tracking can improve the precision of the ablation but requires a new machine with eye tracking technology.

[21] Laser beam fluency is defined as the amount of energy per pulse that is distributed over a defined area (mJoules/cm2; Machat et al. 1999).

[22] First, a surgeon needs to develop a surgical plan based on the patient's age, degree of eye abnormality (the degree of nearsightedness (myopia), farsightedness (hyperopia), or astigmatism or a combination of the first two with astigmatism), sex, and other factors (Machat et al., 1999). Second, surgeons use a special blade called a microkeratome to make a flap that provides access to the cornea. The surgeon's skill in making a smooth, clean flap is crucial for the surgery because an irregular surface causes blurred vision. The choice of microkeratome type reflects the surgeon's learning because some of the microkeratomes may perform better than others depending on the type of eyesight correction required. After folding the flap and before using the laser to remove the corneal tissue, some surgeons clean the cornea under the flap with sterilized drops, while others prefer to leave it as is to avoid injecting moisture into the corneal tissue. These surgeon-specific preferences and habits affect the final ablation because dryness alters the absorption rate of the laser on the cornea. Therefore, different surgeons might use different surgical plans, based on individual surgical habits, for the same patient to get an identical ablation.

translate the surgical plan into laser machine parameters-- a rule known as nomogram[23]-- that is adjusted by an agreement among all doctors.[24] The surgeons from the clinic where we collected the data use a single adjustment rule. They have a structured process to determine this adjustment rule. They meet regularly to review and discuss patient outcomes and, if necessary, update the adjustment rule. The surgeons changed the adjustment rule twice during the data period we analyzed. This adjustment rule reflects the group cumulative experience due to interaction among surgeons who share the same adjustment rule.

4. EMPIRICAL METHODOLOGY

4.1 Data

The data set used in this study is the population of patients who underwent LASIK surgery in the Clínica Oftalmológica de Antioquia (CLOFAN) in Medellín, Colombia. This clinic opened in July 2003 with a brand new Schwind ESIRIS laser machine and now has the largest market share of LASIK surgeries in Medellín.[25] Before July 2003, the surgeons at CLOFAN performed refractive surgery at two other clinics in Medellín that had older laser technologies. After the new surgical center opened, the surgeons performed all of their LASIK surgeries there. CLOFAN surgeons have an incentive to use their own laser machine because they are all

---

[23] Nomograms are not unique to LASIK. The National Cancer Institute defines a nomogram as a mathematical device or model that shows relationships between things. For example, a nomogram of height and weight measurements can be used to find the surface area of a person, without doing the math, in order to determine the right dose of chemotherapy (http://www.cancer.gov/Templates/db_alpha.aspx?CdrID=439410).

[24] The manufacturer of the laser machine provides an initial adjustment rule based on its test data. However, it is recommended that doctors develop their own adjustment rule based on their surgical habits and the environmental conditions of the clinic. The disadvantage of this approach is that it takes time to accumulate the data needed to formulate the adjustment rule. As described by Machat et al. (1999, p. 67), “The process of developing a LASIK nomogram requires four steps: obtaining patient data, formulating an initial nomogram, entering data into the laser’s computer and evaluating data and outcomes, and making adjustments based on this information.”

[25] The clinic’s market share is estimated at about 57 percent of all refractive surgery procedures done in Medellín.

shareholders of the clinic.[26] In addition, the new Schwind ESIRIS machine is the best available technology in the city.

For our study, we used two years of data from the clinic, from July 2003 to August 2005, encompassing a total of 4,009 surgery cases (surgery on one eye is considered one case).[27] From the patient charts, we collected presurgery eyesight measures, the name of the operating surgeon, and all postsurgery follow-up evaluations. We also recorded the patients' basic demographic characteristics, such as gender, age, marital status, place of birth, occupation, city of residence and neighborhood. The patient charts also include information about the surgery: the date, starting time (with a precision of seconds), operating surgeon, and laser control software version. The group adjustment rule was updated in December 2003 and again in May 2004.[28]

Of the 4,009 surgery cases, we dropped from our analysis the 6.14 percent of patients living abroad. Most of these patients did not stay in Medellín long enough for follow-up evaluations.[29] In addition, we dropped three other categories of patients: those who had had LASIK surgery before–i.e., repeat surgery (341 cases), those who failed to return after the

---

[26] In fact, CLOFAN surgeons need to perform a certain number of surgeries per month to pay the equipment costs and generate a profit.

[27] When a patient needs surgery in both eyes, a surgeon may operate on both eyes on the same day or on each eye separately on different days, after the first eye has completely healed. In the event that surgery was performed on both eyes on the same day, we noted the sequence of the surgery, and this was taken into account in our cumulative measurement. In terms of the outcome measurement, it is irrelevant whether the surgery was done on one eye or on both eyes because each eye was considered a separate operation and the outcomes were measured separately. Conceivably a physician might improve his or her skills between the operation on the first eye of a patient and the operation on the second eye of that patient.

[28] Because we know the month but not the day of the nomogram update, we assume it was on the 15th. We checked the sensitivity of our results by redoing our analysis using the first and the last days of the month. Our results are qualitatively robust.

[29] The majority of patients living abroad were from the United States. Our conversations with the clinic staff confirmed that some patients underwent surgery while visiting family in Colombia. We also learned that most of those patients had already booked their airline ticket and could not stay for the clinic follow-ups.

surgery for follow-up (337 cases), and the optimized refractive keratectomy (ORK) cases (245 cases).[30]

We did not include repeat surgeries because both the presurgery conditions and the goal for the outcome of these cases can be very different from those of an initial surgery. For example, the goal of repeat surgery could be to correct the remaining uncorrected visual impairment. Therefore, doctors might choose a different approach for the follow-up surgery. In addition, the postsurgical outcome from the initial surgery is used as the presurgical visual acuity for the follow-up surgery; therefore, it is no longer exogenous.

Regarding the cases excluded for lack of follow-up, we posit two possible reasons for these patients failure to return: either complete satisfaction or strong dissatisfaction. The second case is unlikely because there are no charges for follow-ups or for additional surgeries if they are needed; thus, we believe the chance of losing observations with adverse outcomes is small. In addition, we checked whether follow-up observations are correlated with time or with the cumulative number of surgeries for each surgeon; we found that they are not.

The final category of excluded cases is the ORK cases. ORK is a LASIK surgery based on custom-treated ablation. Because the surgeon knows the complete map of the cornea, the ORK procedure uses a different set of surgical skills when compared with other LASIK surgeries; also, ORK does not need any adjustment rule. For these reasons, we did not include ORK cases in our analysis.[31]

---

[30] Although the dropped cases are excluded from our analysis, they are relevant to and are included in the calculation of the doctors' experience.

[31] We count ORK when we calculate the cumulative number of surgeries. If we include the ORK cases for the analysis of individual learning, results do not change.

The ideal approach for this study would be to observe each surgeon's entire history of LASIK surgeries. However, our observations include only the LASIK surgeries performed by the surgeons at CLOFAN using the Schwind ESIRIS laser machine.[32] As a result, each surgeon had a different level of experience at the time he or she performed the first surgery with the new laser machine. To control for these time-invariant surgeon-specific differences, we used a surgeon-specific fixed-effect model that could accommodate time-invariant differences such as ability, education, and so forth. Nevertheless, a surgeon needs to learn the characteristics of every new machine and feel comfortable with it in order to obtain good results (Pallikaris and Siganos, 1997). Different laser machines have different characteristics so it is very important for the surgeon to understand the specific laser and software in use to achieve the desired refractive correction[33]. Even though we do not have data for the entire experience history of LASIK surgeries for each surgeon (i.e., their experience with older laser machines), we can still measure the machine-specific learning curve each surgeon had to learn and master the new machine's specific characteristics to obtain good results.

Our data were censored at the time of collection. For example, a patient operated on in July 2003 had two years for possible follow-ups. However, a patient operated on a day before the data collection began had only one day for possible follow-ups. We use 200 days as a window to follow the patient. We dropped the surgeries that occurred less than 200 days from the end of our

---

[32] We tried to obtain the information on the number of surgeries performed by the surgeons before July 2003. However, these surgeries were performed in other clinics and there was no log file to identify the relevant patient charts as was the case for patients using the new machine in the CLOFAN clinic and consequently we were not able to obtain that information from patient charts. The other possibility was to ask doctors about their previous experience performing Lasik surgeries. Even though we interviewed doctors, they simply could not reliably recall the number of LASIK surgeries performed before July 2003.

[33] Those characteristics include laser fluency, pulse frequency, maximum pulse area and ablation pattern (Machat et al, 1999). In particular, each Excimer laser system has a specific energy beam profile (Machat et al., 1999).

data period.[34] We chose 200 days (just over 6 months) because 70 percent of repeat surgeries occurred within 200 days of the first surgery. Figure 2 presents the hazard (or survival) function of having a repeat surgery, this is, the probability that the surgery has to be repeated after the n^{th} day past the first surgery, conditional on not having had a repeat surgery. Figure 2 shows, for example, that a patient has 1.86% probability of having a repeat surgery after 200 days past the first surgery. If we had chosen a shorter window, we would have had more observations with right-censored outcomes. On the other hand, a longer window would probably have yielded more observations with recorded outcomes but would have resulted in fewer observations.

Our data allowed us to test for the existence of learning by doing and to differentiate between individual learning and collective learning. We tested for individual learning by examining surgeons' outcomes as they accumulated LASIK surgeries over time. We also tested for collective learning by examining whether the changes in the laser adjustment rule made by the surgeons improved average patient outcome.

4.2 Measures of Eyesight

A main advantage of using LASIK surgery to examine learning by doing is the availability of precise measures of eyesight to define previous patient condition and surgical outcomes. In this section, we describe how ophthalmologists measure eyesight and explain how we use these measurements in our analysis.

Two different methods are commonly used to examine visual acuity: the Snellen method and refraction measurements. For the first, the patient reads letters of different sizes from a

---

[34] We dropped the LASIK surgeries performed after January 9, 2005, based on this category. We lost 1,143 observations.

distance of 20 feet. The Snellen method measures visual acuity on a scale from 20/10 to 20/800, depending on the letter sizes that the patient can read.[35] Although the Snellen measurement is informative, ophthalmologists need to perform a refraction assessment in order to determine the refractive error and prescribe corrective lens.[36] The results are expressed as measurements of sphere, cylinder, and axis.[37] The sphere and the cylinder are measured in diopters and determine the lens prescription; the axis is measured in degrees and signifies the direction of astigmatism.[38] A negative sphere indicates myopia (near-sightedness) and a positive sphere, hyperopia (far-sightedness). The higher the absolute value of the sphere, the worse the visual acuity. The cylinder reflects the degree of astigmatism.[39] A value of zero indicates a perfect sphere or cylinder, meaning that the patient does not have myopia, hyperopia or astigmatism.[40]

Ophthalmologists use a standard metric called the defocus equivalent to obtain a composite measurement based on the refraction measure of the eye.[41] The defocus equivalent is obtained by the following formula:

$$\text{Defocus equivalent} = |\text{Cylinder}/2 + \text{Sphere}| + |\text{Cylinder}/2|$$

If the refraction evaluation is –2.5 of sphere, –3.5 of cylinder, and 180° of axis (which

---

[35] The Snellen chart cannot measure visual acuity that is worse than 20/800.

[36] Refraction refers to how light waves are bent as they pass through the cornea and lens.

[37] The usual expression is sphere = cylinder * axis. The "=" and "*" do not mean mathematicalequality or multiplication but are conventions used in the field to express the refraction measure.

[38] The value of the axis is not necessary in order to calculate visual acuity.

[39] Ophthalmologists and optometrists use negative and positive signs as a custom. All the measures in our data set use the negative sign norm.

[40] In order to measure the cornea and obtain values for the sphere, the cylinder, and the axis, the surgeon has several options; one is to measure the cornea directly with an automated refractometer; another is to try several combinations of lenses to correct the vision. These exams can be conducted with the eye muscles relaxed using eye drops (dilated measures) or without the use of drops.

[41] Spherical equivalence (SE) is used more widely and is calculated as |cylinder/2| + |sphere|. However, SE can be misleading because it does not fully consider the amount of astigmatism. In addition, it can have a negative sign not associated with any specific meaning. Holladay et al. (1991) proposed the defocus equivalent to overcome these shortcomings of SE.

means a myopia of 2.5 diopters and astigmatism of 3.5 diopters, measured on the 180° axis), the defocus equivalent will be $|(-3.5/2) + (-2.5)| + |(-3.5)/2| = |-4.25| + |-1.75| = 6$. The defocus equivalent for a perfect eye is zero. In our data set we use the defocus equivalent to measure visual acuity because the Snellen test cannot measure severe visual acuity problems.

Visual acuity is recorded several times for each patient. After surgery, surgeons typically evaluate a patient's eyesight with the Snellen measure; for patients that demonstrate good eyesight, surgeons often do not perform a refraction evaluation. Based on this criterion, 21 percent of cases were evaluated using only the Snellen measurement. To predict a value for the defocus equivalent for the observations that used only the Snellen measurement, we used observations for which both Snellen and refraction measurements were reported.[42]

4.3 Measures of Outcomes

We use three outcome measures. The first outcome is the postsurgical defocus equivalent, measured in the last follow-up observed within 200 days after the surgery. The second measure is a dummy for the success or failure of the procedure. We consider a surgery to be successful if the defocus equivalent is within ±1.0 diopters of the desired result.[43] The third outcome measure is an indicator of whether the patient needed at least one repeat surgery.[44]

Table 1 presents descriptive statistics of the data. There were 28 doctors in the clinic, and the average number of surgeries per doctor through January 19, 2005, with the Schwind ESIRIS

---

[42] A linear regression was used in order to fill out the missing data. The Snellen and refraction measures use different standards for evaluation. Snellen measures how well the patient can see, and refraction measures near- or far-sightedness as well as astigmatism. In addition, Snellen can only measure up to 20/800. In our data set, 60 percent of the observations with a postsurgery Snellen measure, but no refraction measure, had 20/20 or better eyesight. We ran a regression with age, sex, and Snellen measure as independent variables.

[43] Waring (2000) discusses standard measures for reporting refractive surgery outcomes.

[44] It is worth noting that not all surgeries outside ±1.0 diopters of the desired result required a repeat surgery.

laser machine, was 98. The mean age of patients in the sample was 39 years, and 70 percent were female. At the bottom of Table 1, we present the mean and standard deviation of our three measures of outcomes. The mean postsurgery defocus equivalent was 1.1 diopters; 42.9 percent of surgeries had a defocus equivalent higher than 1.0 diopters; 8.4 percent of patients required repeat surgery. There were five procedures based on the presurgery eyesight: myopia, hyperopia, myopic astigmatism, hyperopic astigmatism and mixed astigmatism.[45] The average presurgery eyesight measure was 3.1 diopters for myopia and 4.3 diopters for myopic astigmatism. The presence of astigmatism increases the defocus equivalent. The most common correction is the myopic astigmatism (40 percent of LASIK surgeries). Hyperopia cases show significantly worse postsurgical eyesight measured in defocus equivalent than myopia cases. This occurs in part because patients with hyperopia tend to be older, and in part because hyperopia is harder to correct than myopia. Although hyperopia and mixed astigmatism cases had worse outcomes measured in defocus equivalent, they had lower repeat surgery rates than other type of presurgery conditions such as myopia.

4.4 Econometric Model

The main questions addressed by this paper are whether there is learning by doing in LASIK eye surgery and whether learning by doing happens individually, collectively, or both. We define individual learning by doing as an improvement in surgery outcomes attributable to the surgeon's accumulated experience in performing LASIK procedures. We define group learning as an improvement in surgery outcomes attributable to changes in the group adjustment rule. Our

---

[45] Mixed astigmatism means that the two curvatures in the cornea defined by the sphere and the cylinder measures are irregular and have meridians placed differently from the normal (90º and 180º).

empirical strategy aims to identify and measure these effects. To test for learning by doing effects, we estimate the following equation:

$$(1) \qquad Y_{ijk} = X_j * ß_0 + ß_1 * LCS_k + ß_2 * GAR_2 * Treatment + ß_3 * GAR_3 * Treatment + \upsilon_k + \omega_i + \mu_{ijk}$$

Where Y_{ijk} is the outcome for the surgery on eye i (*left* or *right*) of patient j operated on by surgeon k. For the defocus equivalent outcome, the best possible outcome has a value of zero; a higher number indicates postoperative myopia, hyperopia, and/or astigmatism. When calculating surgeon experience, we consider only the first surgeries in our sample, although we accumulate all LASIK surgeries performed after July 2003 (not only the first surgeries but also repeat surgeries, surgeries for persons living abroad, surgeries with no follow-up, and ORK surgeries). X_j is a vector of patient characteristics such as age, sex, and presurgical defocus equivalent. To control for the nonlinearity of presurgical eyesight, we include the square of the presurgical defocus equivalent.[46]

In equation (1), *LCS* stands for logarithm of cumulative surgeries and GAR_2 or GAR_3 represents the group adjustment rule update to the second or third version. GAR_2 and GAR_3, which capture collective learning, are dummy variables that take the value of 0 before the rule update and 1 after the update. The logarithmic transformation will capture the steeper slope in the earlier experience in case it exists.[47]

If there is individual learning by doing, surgeons should get better outcomes (i.e., defocus equivalent measures closer to zero) in the n^{th} surgery than in the (n-1th) surgery. If the experience

---

[46] The results are also robust to higher-order terms.

[47] We also tried the square root and a quadratic functional form in the vector of learning. Results are robust to all these functional forms.

accumulated by surgeons as a group in the clinic matters, surgeons should get better outcomes after updates in the group adjustment rule.

The group adjustment rule is implemented using an Excel spreadsheet. The inputs are the surgical plan and the outputs are the machine parameters. The initial version of the group adjustment rule came with the machine, and it corrected the cylinder and the sphere in all the astigmatism, myopia and hyperopia cases. The updates of the group adjustment rule only affected the astigmatism cases, meaning that, according to the notation in the model, three types of surgery (myopic astigmatism, hyperopic astigmatism, and mixed astigmatism) were the treated group (*Treatment*=1) and two types of surgery (myopia, hyperopia) were the untreated group (*Treatment*=0). Group learning was identified as a difference in difference estimate of GAR_2**Treatment* or GAR_3**Treatment.*

We expect a negative sign in $ß_1$, $ß_2$ and $ß_3$ when there is individual or collective learning because our measures of outcome are defined as adverse.

We include a surgeon-specific fixed effect (υ_k) to eliminate permanent differences across surgeons. Standard errors are clustered by surgeon to take into account any nonlinear surgeon-specific residuals not captured by this fixed effect. We also include type-of-procedure fixed effects (ω). The surgical methods differ depending on the refractive error. For example, myopia requires laser ablation to flatten the central cornea, whereas hyperopia requires laser ablation to make the cornea steeper (Machat et al., 1999).

In addition, we control for the software version in our regressions. The software for the Schwind ESIRIS laser machine and its upgrade were provided by the manufacturer. We include a dummy variable to capture the upgrade installed in the clinic's machine on June 24, 2004.[48]

In Table 2 we examine the descriptive statistics of the postsurgery eyesight by group adjustment rule (GAR) and by dummy of Treatment. In the first column, outcomes for the astigmatism cases impacted by the group adjustment rule update moved from 1.22 to 0.96 after the second group adjustment rule update, and down to 0.85 after the third adjustment rule update. The third GAR improved the outcome 30% for astigmatism cases. On the other hand, the outcomes for the myopia and hyperopia cases, which were not intended to be impacted by any of the GAR updates, improved only 10% after the final update compared to the initial GAR.

4.5 Random Assignment

Our conversations with surgeons and managers at CLOFAN revealed that a patient is assigned almost randomly to the surgeon. When a patient contacts the clinic for consultation without requesting a specific surgeon (which happens in most cases), the receptionist assigns the patient based on availability. Because each surgeon is an equal shareholder of the clinic, the receptionist assigns patients so that each surgeon has the same workload. Nevertheless, LASIK surgeries were performed disproportionately across surgeons. The reason is that, although LASIK surgery is regarded as a general procedure that can be handled by any surgeon in the clinic, some surgeons specialize in performing certain procedures, such as cataract surgery, and as a result have less time available for LASIK cases. Unfortunately, we cannot observe this effect directly because our data include only LASIK patients.

---

[48] Machat et al. (1999) list operating temperature and humidity as other factors that can affect surgery outcomes. Although we have those variables in our data, they are poorly recorded because they are not regarded as an important part of the patient charts by the surgeon who is the usual record keeper; for example, records show large variations within a day, which is impossible as the operating room is a temperature- and humidity-controlled space. Therefore, we decided not to use those variables in our analysis.

5. RESULTS

If there is a learning curve in LASIK surgeries, outcomes will improve either with the individual surgeon's experience (measured as the logarithm of the cumulative number of surgeries) or with their group experience (measured by using the changes in the group adjustment rule multiplied by dummy of *Treatment*). For all our measures of outcome, this translates into a negative slope for both the logarithm of cumulative number of surgeries and the changes in the outcome as a result of group adjustment rule update, given that better outcomes are associated with smaller numerical values.

In order to illustrate graphically how outcomes evolve over time, Figure 3 presents the average pre- and postsurgical defocus equivalent by calendar month. In reading this figure, it is useful to bear in mind that the group adjustment rule was updated in December 2003 and May 2004 and that the larger the average postsurgical defocus equivalent, the worse the outcome.

This figure shows two important things: first, we observe no significant variation or pattern by the average presurgical defocus equivalent over the data period; and second, we observe a decrease in the average postsurgical defocus equivalent. In fact, with the exception of three isolated months (July 2003, March 2004, and December 2004), the average presurgical defocus equivalent fluctuated mostly between 3 and 3.5 diopters. During the LASIK machine's first month of operation, we observed a high average postsurgical defocus equivalent--that is, more adverse outcomes, although average patient presurgical defocus equivalent was also high in the first month. After that, outcomes improved slightly and remained stable until October 2003; thereafter, there was a slight decrease in average postsurgical defocus equivalent until the last month of our sample, with the exception of March 2004. The average postsurgical defocus equivalent is particularly high in March 2004 because of four surgeries with bad outcomes. Although there is more variation in outcomes after May 2004, the overall level is lower after May 2004 than before.

We also examine patient outcomes as individual experience increases for some doctors. Figure 4 shows postsurgical defocus equivalent changes in the vertical axis and cumulative number of surgeries in the horizontal axis for doctor A. There was a significant improvement in the first 20 surgeries, after which, as individual experience increased, the outcome fluctuated, showing no real trend. For another doctor, reported in Figure 5, there was some deterioration in the first 10 surgeries and then sharp improvement in the next 10 surgeries. Starting from the 30th surgery till the end we observe some fluctuation, but it is hard to tell whether the outcome showed any slope with increased experience. In Figure 6 we saw slight improvement in the 20th surgery, and then the outcome got worse after the first 30 surgeries.[49]

In Table 3, we show the impact of individual learning (measured by the logarithm of cumulative surgeries) on postsurgical eyesight (measured by defocus equivalent). In the first column, we only include the logarithm of cumulative surgeries as an independent variable. As expected, the sign is negative, which means that, as surgeons perform more surgeries, their patients' postsurgical eyesight gets closer to perfect. The magnitude of the coefficient implies that a 1% increase in the number of surgeries would improve the outcome by 0.001 diopters, a

---

[49] It is possible that the logarithmic transformation might not be able to capture individual learning curves: the learning curve may be flat in some portions or the learning speed may be different at certain range of surgeries. We try to examine the importance of these non-linearities by introducing a spline and by estimating a piecewise linear regression in addition to equation (1). The spline we use has knots at surgeries number 10, 20, 30, 60, 90, 120, . . , 210. We used postsurgical eyesight as the dependent variable. The learning curves were not statistically significant in any range. Our results are robust to functional forms that we tried.

minuscule though statistically significant improvement at 1% confidence level. In the second column, we add demographic characteristics as controls. In the third and fourth columns, we include presurgery eyesight to control for the patient's underlying conditions. In column (4), the coefficient for hyperopia cases are positive and statistically significant, meaning that patients with hyperopia have worse outcomes when compared to patients with myopia. As we add more controls, the point estimate for individual learning decreases (in absolute value) from 0.101 to 0.084. This constitutes a very small improvement. In the final column, we control for technological change reflected by the installation of a new version of the software in the same machine. The laser machine software version was upgraded once during our data period.[50] The coefficient for the software update dummy is small and statistically insignificant. However, once we control for this technological change, the individual learning coefficient decreases by 40 percent in terms of magnitude and also becomes statistically insignificant. As a result, we conclude that there is no statistically significant individual learning in LASIK surgery and controlling for technological change can be important.

In all regression specifications, we include a surgeon-specific fixed effect and a type-of-surgery fixed effect to control for time-invariant factors across surgeons and types of procedures. Standard errors are clustered by surgeon to capture nonlinear common factors associated with a particular surgeon.

As we mentioned in an earlier section, we are able to identify clinic-wide collective learning in our data. In Table 4 we use those group adjustment rules as key independent

---

[50] The CLOFAN clinic bought the new software from the manufacturer of the laser machine. Besides some changes in screen appearance, the new software changed the ablation, the transition zone, and the calculation of the depth of ablation. We got this information directly from the manufacturer in the "Schwind Circular letter," Technical Department of Schwind eye-tech-solution GmbH, Rev. 3. August 28, 2003, TM/AL/SB, Germany.

variables, with the initial group adjustment rule as the omitted category. In the first column we control only for the second and the third group adjustment rules in addition to surgeon fixed effects. Both adjustment rules improve the outcome. The second group adjustment rule shows a negative sign, as we had expected, but it is statistically insignificant. The third adjustment rule shows a larger point estimate and is precisely estimated. In comparison to the initial rule, the third adjustment rule that doctors implemented improved the average postsurgical eyesight by 0.24 diopters.

In the second, third, and fourth columns of Table 4, we add demographic controls, presurgical eyesight, and type of eyesight, respectively. The effect of the second group adjustment rule is negative and statistically insignificant throughout the columns and the point estimates were sensitive to controls. On the other hand, the third group adjustment rule shows a robust point estimate and is statistically significant at 5 percent or 1 percent, depending on specifications. In the last column, we include a dummy for changes in the software technology used by the clinic. In this specification, the third group adjustment rule lowers the postsurgical eyesight to 0.16 diopters, which is equivalent to a substantial 15 percent improvement on the average postsurgical eyesight. This point estimate is statistically significant at 10 percent confidence level. Age is consistently significant across specifications, meaning that older patients have significantly worse outcomes. If a patient’s age increases by one year, the postsurgical defocus equivalent is higher by 0.01 diopters. If the patient is a woman, the postsurgical defocus equivalent is 0.1 diopter higher, and this variable shows statistical significance in all columns except the second one.

To control for the nonlinear effect of the presurgical eyesight measure, we include a quadratic term in our model. Both terms in this measure are jointly significant, meaning that when patients have particularly poor presurgical eyesight, the postsurgical outcome is likely to be worse. The software upgrade was statistically insignificant in the last column.

In Table 5, we present the estimates of individual learning and collective learning using the specifications in the last column from Table 3 and Table 4, respectively, where all controls are present. In the third column of Table 5 we consider individual learning and collective learning together in the same regression. As a result, the magnitude of individual learning changed from -0.05 to -0.01, which is a much smaller estimate. The group adjustment rules estimates, however, did not change much after controlling for individual learning. Although the second group adjustment rule improved the outcomes, it is not statistically significant and we cannot conclude it is different from zero. The third group adjustment rule shows a larger point estimate in absolute value, which means a larger improvement compared to the initial adjustment rule, and it is statistically significant at a 10 percent confidence level.

In Table 6, we present the results for two other outcome variables. In the first column, the dependent variable is whether the postsurgical eyesight is greater than 1 diopter (i.e., a failure). When we use postsurgical eyesight as an outcome, a single outlier could have a large effect on the estimates. Using a dummy for success or failure provides a different way of measuring learning that avoids this problem. Individual learning, measured by the logarithm of the cumulative number of surgeries, shows the expected sign but is not statistically significant. The group adjustment rules presented in the second column show a monotonic improvement effect on outcomes, and the third group adjustment rule has an effect on outcomes statistically different from zero at 10 percent confidence level. The third group adjustment rule improved the outcome by 9.1 percentage points. In the third column, when we estimate individual and collective learning simultaneously, the point estimates of the effect of the third group adjustment rule on outcomes is robust. The coefficients measuring the effects on outcomes of individual learning and of the second group adjustment rule changed somewhat but remained statistically insignificant.

In the fourth column we use repeat surgery as the adverse outcome. The overall repeat surgery rate is 8.4 percent. The coefficient of individual experience for repeat surgery is also statistically insignificant and we cannot conclude it is different from zero. Both collective learning variables show a monotonic improvement effect compared to the initial adjustment rule, but they are statistically insignificant. In the last column, when we estimate individual and collective learning together, the magnitude of the third group adjustment rule is robust throughout different specifications, even though it is insignificant.

In Table 7 we examine whether individual learning and collective learning might work differently for easy cases versus severe cases. This may occur because of nonlinear effects of the presurgical eyesight on the outcomes. The upper block of Table 7 repeats the results reported in Table 5 and 6. The middle block shows the results for easy cases, and the bottom block shows the results for severe cases. We consider easy cases to be the ones below the 25th percentile in severity of presurgical eyesight, while severe cases are the ones above the 75th percentile.

In the first column, we use the postsurgical eyesight as the measure of outcome. The severe cases show larger individual learning compared to the easy cases, although they are statistically insignificant. However, the magnitude of the coefficient for the third group adjustment rule is twice the value in the easy cases compared to the severe cases and is statistically significant at 5 percent confidence level. We show in the second column the measure of outcome which looks at a failure of surgery. The results pattern in the second column is similar to the results pattern in the first column. The third group adjustment rule for easy cases is statistically significant at 5 percent and 20 percent improvement in terms of magnitude. In the third column, we show the repeat surgery measure of outcome. The results for this measurement did not have any consistent pattern in previous outcomes. For repeat surgery, the third group adjustment rule improves somewhat for severe cases.

6. ROBUSTNESS CHECKS

6.1. Type-Specific Surgery Volume

We have measured individual learning by calculating the cumulative number of LASIK surgeries for each surgeon. However, different surgery types have different surgical procedures, as is the case for myopia and hyperopia, which have different corneal ablation and different presurgical plan specifications (Machat el al., 1999). This means that the human capital might differ by type of surgery, and it could be acquired separately for each type. For example, it could happen that a doctor with extensive experience and a history of good results with myopic patients may not have good outcomes in his initial operations on hyperopic patients. If this is the case, our previous measure of experience will have measurement error and we may not be able to determine individual learning by doing even though it exists. To test this possibility we measure the type-specific cumulative individual experience instead of the total LASIK experience. We construct this new measure of experience by calculating the number of type-specific cumulative surgeries for each surgeon instead of the total number of cumulative surgeries[51] and run the

---

[51] For example, if a surgeon has performed five myopia surgeries and four hyperopia surgeries, the number of type specific cumulative surgeries would be four for hyperopia and five for myopia, and the total number of Lasik

preferred specification with the logarithm of this new measure as an independent variable. The results are reported in Table 8 and are entirely consistent with the previous results that analyze total LASIK cumulative surgeries. The three outcome variables that we examined for individual learning are still statistically insignificant.

6.2. Interactions Between Individual Learning and Group Learning

Group learning may have different effects on different surgeons depending on their level of experience even though the rule changed for all the surgeons at the same time. More (or less) experienced doctors might benefit more from the second or the third group adjustment rule. To test this possibility, we add an interaction term between the individual learning variable (cumulative volume) and the two group learning rules. We report the results in Table 9. In the first column, the interaction between individual experience and the 2nd group adjustment rule*Treatment is positive and statistically significant different from zero, meaning that, in effect, less experienced doctors benefit more from the 2nd group adjustment rule at 10% confidence level. On the other hand, the coefficient for log(cumulative surgeries)*3rd GAR*Treatment is statistically insignificant. In this regression, the interaction between the individual learning variable and the 3rd group adjustment rule*Treatment is statistically insignificant, and the variable that measures the 3rd group adjustment rule*Treatment alone now becomes statistically insignificant. This might be a result of the small variation left in the data to identify the effects of the 3rd group adjustment rule*Treatment, given that only 29% of the observations have a value of 1 in this case, compared with 52% for the 2nd group adjustment rule*Treatment. The results for “Failure of surgery” in the second column did not change

---

surgeries would be nine.

qualitatively and the interaction term was statistically insignificant. The results in the case of the outcome measure "Need resurgery” also did not change from previous results.

6.3. Possible Issues with Sorting

In section 4.5, we mentioned that the patient sorting is not likely to happen in our data based on anecdotal evidence and institutional detail. In this section, we provide some evidence that confirms that sorting is not likely to exist and, in the event it does exist, we discuss how it might affect our estimates.

Doctors may select patients depending on the case difficulty, and patients may select specific doctors based on their reputation. For example, high-ability doctors may choose to treat just the severe cases if the clinic wants to maximize overall outcomes (doctor sorting), or patients with severe cases of far- or nearsightedness may choose high-ability doctors because their marginal benefit is higher (patient sorting); in both cases, the outcomes for those surgeons would be underestimated. In addition, sorting may happen across doctors and across patients at a specific time (i.e. in a cross section) or sorting might take place over a period of time.

We use variation over-time to identify the effect of learning by doing and control the cross sectional variation in doctors’ outcomes by using fixed effects, meaning that the sorting of specific doctors by patients or the sorting of specific patients by doctors at a specific time (i.e., in a cross section) does not affect our estimates of learning by doing.[52]

---

[52] However, we further checked if selection across surgeons or across patients existed at a specific time by running a regression using the presurgery defocus equivalent as the dependent variable (outcome) and surgeon specific dummies as an independent variable. All the surgeon specific dummies except one were statistically insignificant and even the statistically significant one had a p-value of 0.05, meaning that the presurgery eyesight was not correlated with any specific doctor.

However, sorting over time is a potential source of bias. It may occur if high-ability doctors choose to treat more (or fewer) severe cases as time goes by or if severe patients are able to identify high-ability doctors as times goes by.

To check for sorting over time in our data, we test whether patients' observable characteristics vary over time by surgeon. As mentioned above, one of the advantages of using LASIK surgery is that unobservable underlying conditions are minimal. We ran a regression with presurgical defocus equivalent as the dependent variable (outcome) and with the monthly time trend and the logarithmic cumulative number of surgeries as independent variables for each analysis, including a surgeon-specific fixed effect. The result is statistically insignificant at 5 percent, meaning that the severity of the patients' condition did not change over time or with the individual surgeon's level of experience.[53]

We also observe a different volume of LASIK surgeries by doctor in our data. In section 4.5, we explained that this differential comes from doctors specializing in other types of surgeries. However, this may also happen if high- and low-ability doctors have a very different learning curve--for example, high-ability doctors have a steep learning curve while low-ability doctors have a flat learning curve. If, in addition, high-ability doctors perform more LASIK surgeries than low-ability doctors, then a composition effect will emerge. That is to say, when we observe the initial surgeries, we do so for both high- and low-ability doctors together, but we would only observe high-ability doctors at the highest level of experience. If this is true, we should find individual learning as a result of this compositional effect, which is not the case. We

---

[53] We further tested whether selection over time existed by running a regression of presurgery defocus equivalent as a dependent variable and the total number of LASIK surgeries for the operating surgeon as an independent variable. If in fact no selection across surgeons exists, we would not get a statistically significant coefficient, meaning that the volume of surgeries for a doctor did not have any influence in the severity of the cases treated. We obtained a –0.0009 coefficient with a standard error of 0.0017, which is small and statistically insignificant.

also test this hypothesis by running our preferred specification, keeping the first 100 surgeries (which is the average number of surgeries). None of the main results of the paper changed, and we did not observe any individual learning.

7. CONCLUSIONS AND LIMITATIONS

In this paper, we examine the existence of individual and collective learning by doing for LASIK eye surgeries. LASIK surgery shares many characteristics with other surgeries. In addition, it has some advantages over other medical procedures that allow us to measure individual and collective learning by doing.

In comparison with previous studies, there are two distinguishing features of this paper: first, the use of a longitudinal data set with good measurements of each surgeon's experience and precisely defined medical outcomes, and, second, the analysis of a surgical procedure that allows us to separate individual learning from collective learning. Past studies have used data that often confounded measures of outcomes with unobserved underlying patient conditions, and most of those studies used only annual surgical volume as an indicator of the surgeon's experience. In addition, in past studies it was difficult to isolate the effect of learning by doing from other effects, such as selective referral. It was also difficult to separate the individual from the collective learning effects on outcomes, because those studies analyzed surgical procedures performed by a surgical team rather than by individual surgeons.

The main question addressed in this paper is whether patient outcomes improve with surgeon experience. We use two measures of learning in LASIK procedures. First, we measure individual learning using the cumulative number of surgeries for each surgeon. Second, we measure group learning by using updates of the group adjustment rule, which is the result of a periodic and structured review process by the surgeons. We do not find evidence that as surgeons increase the number of surgeries performed, they obtain better outcomes. However, we do find evidence of collective learning, because outcomes significantly improved after the third group adjustment rule update. An important point is that, if we had not controlled for measures of collective learning or for measures of technological change (reflected by updates in the software of the laser machine), we would have found spurious evidence of individual learning by doing.

We use three different measures of outcomes: postsurgical defocus equivalent, a dummy variable for success or failure, and repeat surgery. The first and second measure of outcome show a consistent and statistically significant 16 percent and 9.1 percent improvement after the third group adjustment rule update, respectively. Patients with easy cases benefited more as a result of the third adjustment rule update. We find that collective learning plays a substantial role in LASIK surgery performance, while individual learning does not.

Our study has some limitations. We analyzed learning by doing for a particular procedure–LASIK eye surgery. The relative importance of individual learning compared with collective learning may change across organizations and across procedures due to the nature of procedure and the setup of the organization. This also implies that our results might not generalize to other surgical procedures because some elements of the human capital may be transferable across surgeries, and other elements may not.[54]

---

[54] Another possible limitation concerning the generalizability of this study is that the data come from one single clinic consisting of 28 doctors with a total number of surgeries of 1,746 after the data was cleaned. In addition, CLOFAN doctors might not be representative even though we do not have any supportive evidence on this.

However, our limitations are shared by other studies of learning by doing in medicine. Previous studies also examined particular medical procedures, in which case their learned skills and human capital may not be transferable to other procedures, and the collective vs. individual learning may follow a different dynamics. For example, it is not clear whether the learning by doing that some studies have found in the case of CABG surgery applies to other procedures or whether the interaction observed in those teams is the same as in other procedures.

Nevertheless, and unlike this study, none of the previous studies were able to measure a doctor's experience from his or her initial surgery, because none of the previous studies had data from the beginning of the procedure, and the doctors observed in those studies had accumulated varying degrees of experience in the procedure. For example, CABG was a relatively rare procedure until the 1980s but since then has become widespread (Gowrisankara et al., 2006). Researchers examined data in the mid 1990s (Huckman and Pisano, 2006) or late 1990s (Birkmeyer et al., 2003) and it is hard to believe that the surgeons observed in these earlier studies were in the initial stages of learning about this procedure.

Notwithstanding these qualifications, we were able to clearly identify individual and collective learning separately, and our results suggest that collective learning may be more important than individual learning for LASIK eye surgeries.

References

Alchian, A. 1963. "Reliability of progress curves in airframe production," *Econometrica*, 31(4):679–694.

Argote, L., Beckman, S.L., and Epple, D. 1990. "Learning curves in manufacturing," *Science*, 247:920–924.

Asher, H. 1956. "Cost-Quantity Relationships in the Airframe Industry," RAND Corporation Study No. R-291.

Arrow, K.J. 1962. "The Economic Implications of Learning by Doing," *Review of Economic Studies*, Vol. 29: 155-73.

Benkard, C.L. 2000. "Learning and forgetting: The dynamics of aircraft production," *American Economic Review*, 90(4):1034–1054.

Birkmeyer, J.D., Siewers, A.E., and Finlayson, E. 2002. "Hospital volume and surgical mortality in the United States," *Annals of Surgery*, 346:1128–1137.

Birkmeyer, J.D., Stukel, T.A., Siewers, A.E., Goodney, P.P., Wennberg, D.E., and Lucas F.L. 2003. "Surgeon volume and operative mortality in the United States," *New England Journal of Medicine*, 349:2117–2127.

Bridgewater, B., Grayson, A.D., Au, J., Hasan, R., Dihmis, W.C., Munsch, C., and Waterworth, P. 2004. "Improving mortality of coronary surgery over first four years of independent practice: retrospective examination of prospectively collected data from 15 surgeons," *British Medical Journal*, 329-421.

Condon, P.I., Mulhern, M., Fulcher, T., Foley-Nolan, A., and O’Keefe, M. 1997. "Laser intrastromal keratomileusis for high myopia and myopic astigmatism," *British Journal of Ophthalmology* 81:199-206.

Dranove, D., Satterthwaite, M, Kessler, D., and McClellan, M. 2003. “Is More Information Better? The Effects of Report Cards on Cardiovascular Providers and Consumers,” Journal of Political Economy, Vol. 111(3): 555-588

Edmondson, A., Bohmer, R. and Pisano, G. 2001. "Disrupted Routines: Team Learning and New Technology Implementation in Hospitals," *Administrative Science Quarterly*, Vol. 46(4): 685-716.

Epstein, A.M. 2003. "Volume and outcome: It is time to move ahead," *New England Journal of Medicine*, 346:1161–1164.

Gillet, P., and Goldblum, K. 2004. "Ophthalmic patient assessment," *The Journal of the American Society of Ophthalmic Registered Nurses*. October-December 23-5

Gruber, H. 1992. "The learning curves in the production of semiconductor memory chips," *Applied Economics*, 24(8):885–894.

Gruber, H. 1994. "Learning by doing spillovers in the semiconductor industry," *Journal of Political Economy*, 102(6):1201–1227.

Gowrisankaran, G., Ho, V., and Town, R.J. 2006. "Causality, learning and forgetting in surgery," Federal Trade Commission, draft, January 2006.

Halm, E.A., Lee, C., and Chassin, M.R. 2002. "Is volume related to outcome in health care? A systematic review and methodological critique of the literature," *Annals of Internal Medicine*, 137(6):511–520.

Hewitt, M. 2000. *Interpreting the volume-outcome relationship in the context of health care quality*. Washington, D.C.: Institute of Medicine.

Ho, V. 2002. "Learning and the evolution of medical technologies: The diffusion of coronary angioplasty," *Journal of Health Economics*, 21:873–885.

Ho, V. 2004. "Certificate of need, volume, and percutaneous transluminal coronary angioplasty outcomes," *American Heart Journal*, 147(3):442–448.

Holladay, J.T., Lynn, M.J., Waring, G.O. III, Gemmill, M., Keehn, G.C., and Fielding, B. 1991. *Archives of Ophthalmology* 109: 70-76.

Huckman, R. and Pisano, G. 2006. "The firm specificity of individual performance: evidence from cardiac surgery," *Management Science*, April.

Huesch, M.D. 2009. "Learning by Doing, Scale Effects, or Neither? Cardiac Surgeons after Residency," *Health Services Research*, 44(6): 1960-1982.

Hughes, R.G., Garnick, D.W., Luft, H.S., McPhee, S.J., and Hunt, S.S. 1998. "Hospital Volume and Patient Outcomes: The Case of Hip Fracture Patients," *Medical Care*, 26(11): 1057-1067.

Jollis, J.G., Peterson, E.D., DeLong, E.F., Mark, D.B., Collins, S.R., Muhlbaier, L.H. and others. 1994. "The relation between the volume of coronary angioplasty procedures at hospitals treating Medicare beneficiaries and short-term mortality," *New England Journal of Medicine*: 331:1625-9.

Keogh, B.E., abd Kinsman, R. "National Adult Cardiac Surgical Database Report 1999–2000," The Society of Cardiothoracic Surgeons of Great Britain and Ireland.

Kimmel, S.E., Berlin, J.A. and Laskey, W.K. 1995 "The relationship between coronary angioplasty procedure volume and major complication," *Journal of American Medical Association.* 274(14).

Lazar, H.L., Fitzgerald, C., Gross, S., Heeren, T., Aldea, G., and Shemin, R. 1995. "Determinants of length of stay after coronary artery bypass graft surgery," *Circulation* 92:20-24.

Luft, H.S., Bunker, J.P., Enthoven, A.C. 1979. "Should operations be regionalized? The empirical relation between surgical volume and mortality," *The New England Journal of Medicine,* 301:1364-1369.

Machat J, Slade S, Probst L. 1999. *The Art of LASIK*, 2nd ed. Thorofare NJ: Slack Incorporated.

Menke TJ, Wray NP. 2001 When does regionalization of expensive medical care save money? *Health Services Management Research,* 14(2):116–124.

Pallikaris I, Siganos D. 1997. *LASIK*. Thorofare NJ: Slack Incorporated.

Philips KA, Luft HS, Ritchie JL. 1995. The association of hospital volumes of percutaneous transluminal coronary angioplasty with adverse outcomes, length of stay, and charges in California. *Medical Care*. 33:502-14.

Pisano, G., Bohmer, R.M.J., and Edmondson, A. 2001. "Organizational differences in the rate of learning: evidence from the adoption of minimally invasive cardiac surgery," *Management Science* 47(6).

Rathore, S.S., Epstein, A.J., Nallamothu, B.K., and Krumholz, H.M. 2006. "Regionalization of ST-segment elevation acute coronary syndromes care: putting a national policy in proper perspective," *Journal of American College of Cardiology*, 47(7): 1346-9.

Roques, F., Nashef, S.A.M., Michel, P., Gauducheau, E., de Vincentiis, C., Baudet ,E., J., Cortina, David, M., Faichney, A., Gavrielle, F., Gams, E., Harjula, A., Jones, M.T., Pinna, Pintor, P., Salamon, and Thulin, R., L. "Risk factors and outcome in European cardiac surgery: analysis of the EuroSCORE multinational database of 19030 patients," *European*

Journal of Cardiothoracic Surgery.15: 816-823.

Sarrazin, M.V., and Rosenthal, G.E. 2004. "Hospital volume and outcome after coronary angioplasty: Is there a role for certificate of need regulation?," *American Heart Journal*, 147(3):383–385.

Thompson, D.R., Clemmer, T.P., Applefeld, J.J., Crippen, D.W., Jastremski, M.S., Lucas, C.E., Pollack, M.M., and Wedel, S.K. 1994. "Regionalization of critical care medicine: Task force report of the American College of Critical Care Medicine," *Critical Care Medicine*, 22(8):1306–1313.

Vickers, A.J., Bianco, F.J., Serio, A.M., Eastham, J.A., Schrag, D., Klein, E.A., Reuther, A.M., Kattan, M.W., Pontes, E., and Scardino, P. 2007. "The Surgical Learning Curve for Prostate Cancer Control After Radical Prostatectomy," *Journal of the National Cancer Institute*, 99: 1171-7.

Waring, G.O. III. 2000. "Standard graphs for reporting refractive surgery," *Journal of Refractive Surgery*, 16, July/August.

Weatherly, S.L. 2002. "Testing Visual Acuity with the Jaeger Eye Chart. American Society of Nondestructive Testing," available on the Internet at http://www.asnt.org/publications/.

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

Figure 1.
LASIK Surgical Procedure

1-3: Cutting the flap in the cornea with the microkeratome.
4. Folding the flap back.
5. Correcting the corneal tissue through laser ablation.
Source: Allaboutvision.com at: http://www.allaboutvision.com/visionsurgery/lasik.htm

Figure 2.
Cumulative Probability of Repeat Surgeries Conditional on not Having Had Repeat Surgeries

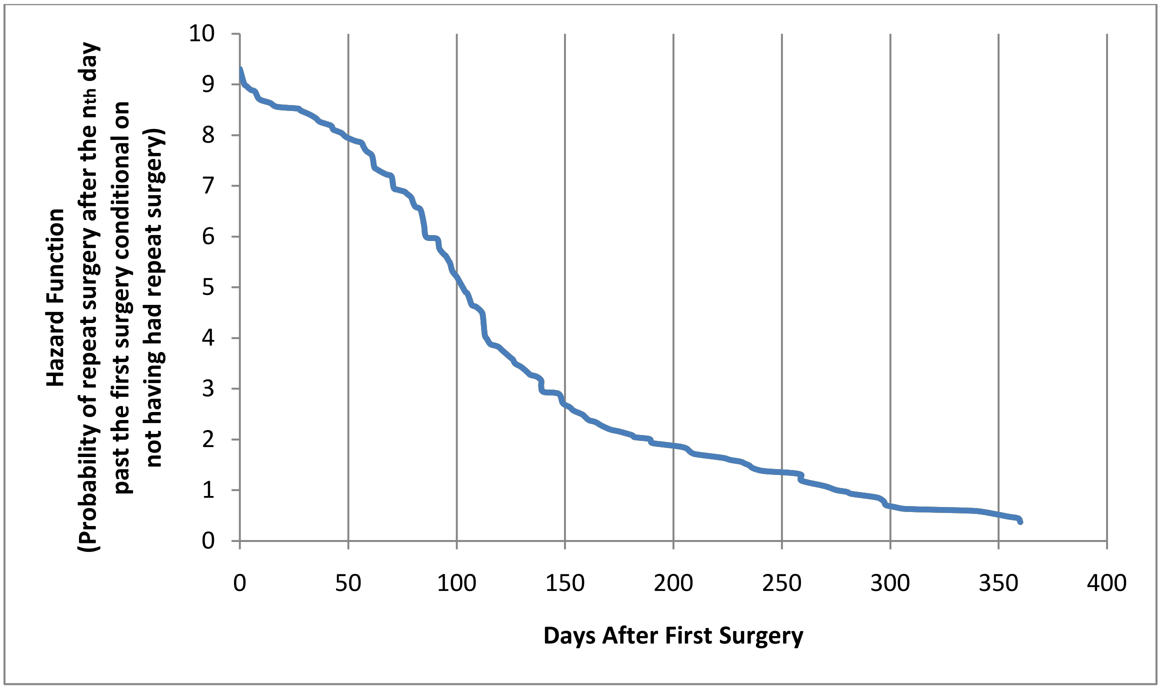
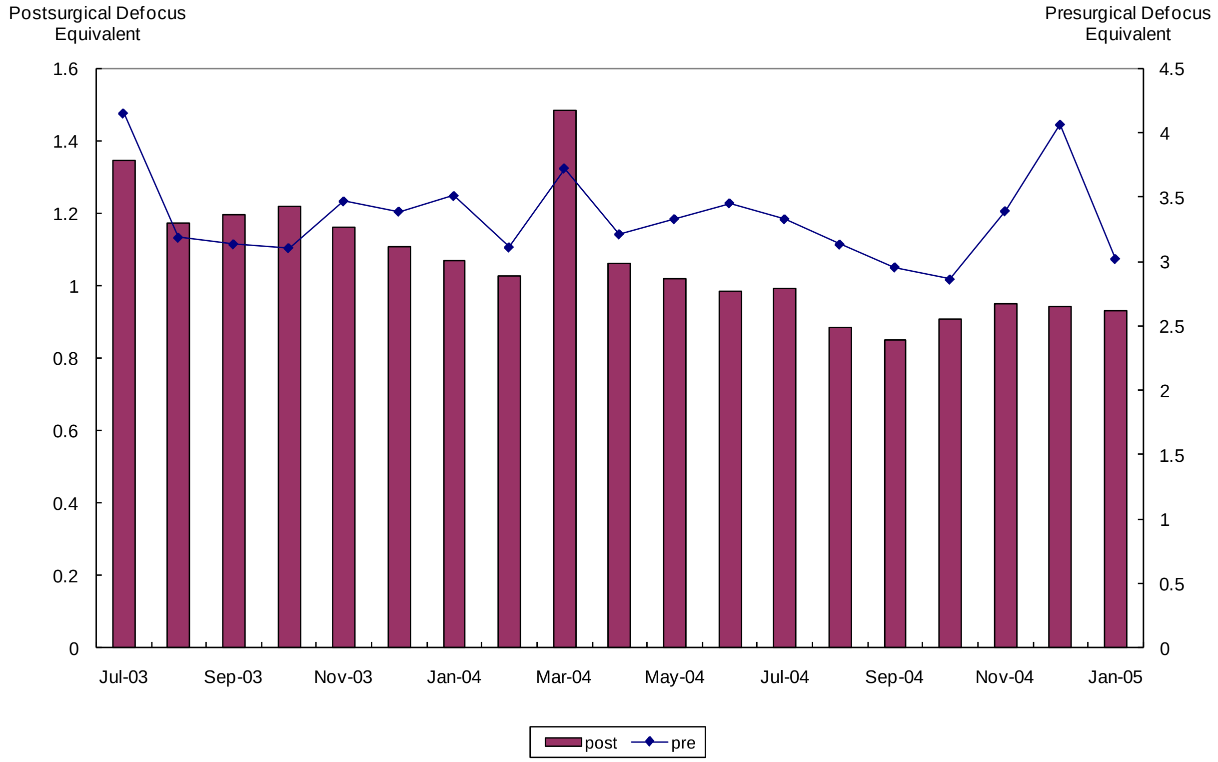


Figure 3.
Presurgical and Postsurgical Eyesight for the Full Sample, by Month*



*Note: The presurgical and postsurgical eyesights are measured as Defocus Equivalent (DE). See section 4.2 in the text for more details

Figure 4.
Postsurgical Change in Eyesight with Individual Cumulative Surgeries, Doctor A

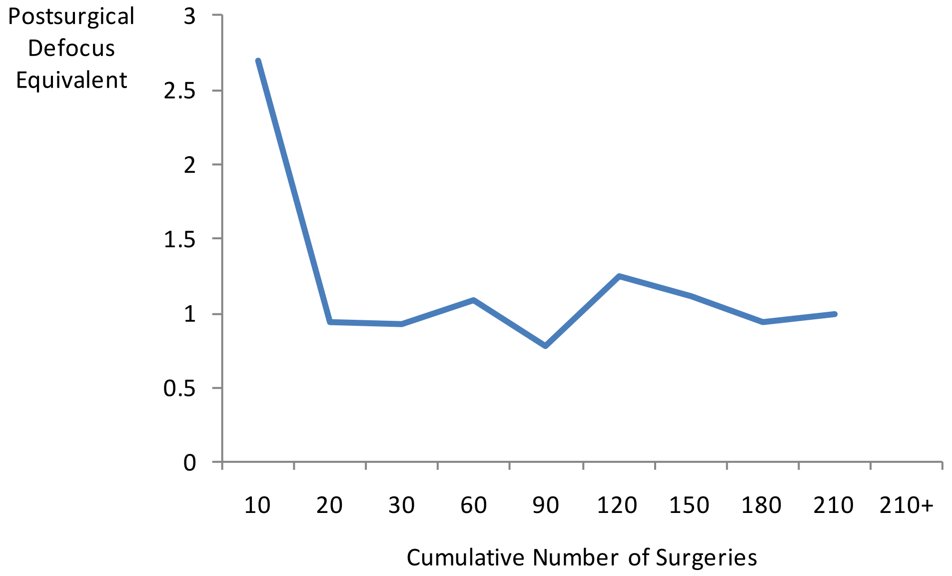


Figure 5.
Postsurgical Change in Eyesight with Individual Cumulative Surgeries, Doctor B

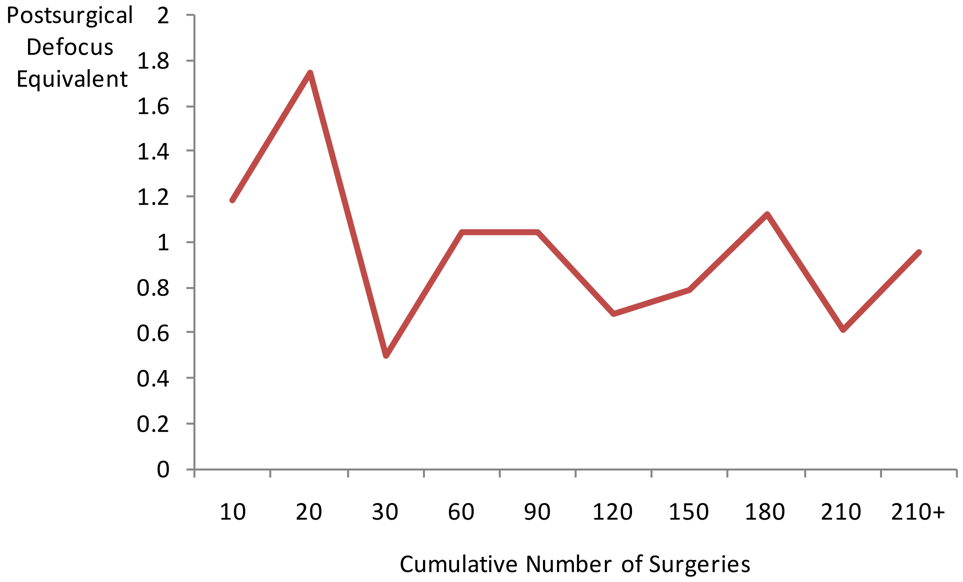


Figure 6.
Postsurgical Change in Eyesight with Individual Cumulative Surgeries, Doctor C

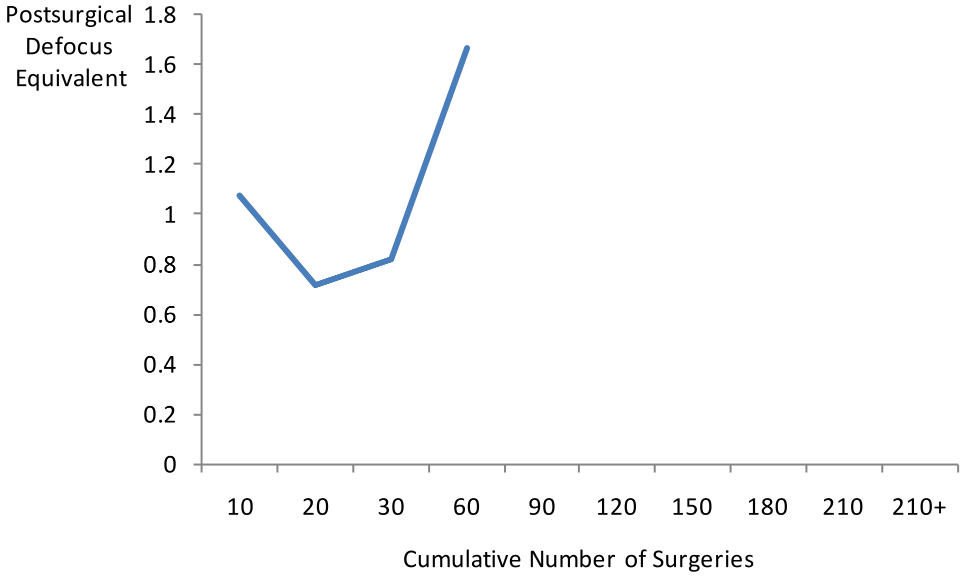


Table 1. Descriptive Statistics

		By Type of Presurgical Eyesight				
	Full Sample	Myopia	Myopic Astigmatism	Hyperopia	Hyperopic Astigmatism	Mixed Astigmatism
Age (years)	38.5	32.0	32.4	49.6	44.9	34.8
	(13.5)	(10.1)	(10.2)	(11.3)	(13.8)	(11.7)
Female (%)	70	70	60	70	70	70
Presurgery Eyesight (DE)	3.32	3.05	4.28	1.96	3.19	2.38
	(2.5)	(2.18)	(2.93)	(1.45)	(2.04)	(1.43)
Postsurgery Eyesight (DE)	1.06	0.90	0.94	1.29	1.17	1.03
	(0.95)	(1.03)	(1)	(0.84)	(0.9)	(0.85)
Defocus Equivalent >1.0 D (%)	42.90	26.46	31.68	64.48	54.52	43.70
Need Resurgery (%)	8.36	9.42	7.53	10.15	8.49	5.88
	(1.9)	(1.5)	(1.6)	(2.2)	(2)	(1.6)
Average Surgeries per Surgeon	98.2					
	(95.4)					
Observations	1746	223	704	335	365	119

( ) standard deviations

The full sample includes LASIK patients operated on from July 1, 2003, when the new laser machine became available, to Jan. 19, 2005.

We keep the 200 days observation period after the operation throughout data. We use the most up-to-date postsurgical eyesight measurement within 200 days after the operation. We drop repeat surgeries from the sample even though we count them when we calculate individual cumulative number of surgeries, because the objective of repeat surgeries might be different.

Table 2. Average Postsurgical Eyesight by Group Adjustment Rule and Type of Surgeries

	Treatment=1	Treatment=0
	Myopic Astigmatism Hyperopic Astigmatism Mixed Astigmatism	Myopia Hyperopia
Initial GAR[a]	1.22	1.20
2nd GAR	0.96	1.11
3rd GAR	0.85	1.08

a.GAR stands for Group Adjustment Rule.

Table 3. Impacts of Individual Learning on Postsurgical Eyesight

	(1)	(2)	(3)	(4)	(5)
Logarithm of Cumulative Surgeries	-0.101	-0.095	-0.091	-0.084	-0.050
	(0.027)	(0.029)	(0.025)	(0.025)	(0.032)
Age		0.011	0.013	0.009	0.009
		(0.002)	(0.001)	(0.001)	(0.001)
Female		0.099	0.105	0.095	0.095
		(0.057)	(0.048)	(0.045)	(0.045)
Presurgical Eyesight			-0.006	0.014	0.013
			(0.032)	(0.033)	(0.033)
Presurgical Eyesight^2			0.008	0.007	0.007
			(0.003)	(0.003)	(0.003)
Myopic Astigmatism				-0.078	-0.070
				(0.086)	(0.084)
Hyperopia				0.263	0.269
				(0.080)	(0.081)
Hyperopic Astigmatism				0.106	0.110
				(0.107)	(0.109)
Mixed Astigmatism				0.069	0.080
				(0.084)	(0.082)
Software update					-0.115
					(0.071)
Observations		1,746	1,746	1,746	1,746

Robust standard errors are clustered by surgeon and included in parentheses.

We use the most up-to-date eyesight measurement within 200 days after the surgery.

A surgeon-fixed effect is included for all the columns.

The presurgery eye condition of myopia is omitted.

Table 4. Impacts of Collective Learning on Postsurgery Eyesight

	(1)	(2)	(3)	(4)	(5)
2nd Group Adjustment Rule	-0.066	-0.014	-0.098	-0.125	-0.126
	(0.059)	(0.055)	(0.056)	(0.074)	(0.074)
3rd Group Adjustment Rule	-0.239	-0.242	-0.196	-0.196	-0.155
	(0.089)	(0.086)	(0.073)	(0.070)	(0.083)
Additional Controls		Age, Gender	Presurgery Eyesight, Presurgery Eyesight^2	Type of Abnormality *	Software update
Observations	1746	1746	1746	1746	1746

See notes for Table 3.

The first group adjustment rule which is the initial adjustment rule provided by the manufacturer was omitted.

* myopic astigmatism, hyperopia, hyperopic astigmatism and mixed astigmatism.

Table 5. Impacts of Learning Vector on Postsurgical Eyesight

	(1)	(2)	(3)
Logarithm of Cumulative Surgeries	-0.050		-0.010
	(0.032)		(0.047)
2nd Group Adjustment Rule		-0.126	-0.111
		(0.075)	(0.109)
3rd Group Adjustment Rule		-0.155	-0.157
		(0.083)	(0.083)
Observations	1746	1746	1746

See notes for Table 3. To save space we do not report estimates of the control variables. The same specification with Table 3 is used.

Table 6. Impacts of Learning Vector on Various Outcomes

	Failure of Surgery			Need Resurgery		
	(1)	(2)	(3)	(4)	(5)	(6)
Log(cumulative surgeries)	-0.013		0.005	-0.011		-0.007
	(0.014)		(0.023)	(0.012)		(0.012)
2nd Group Adjustment Rule		-0.037	-0.045		-0.018	-0.008
		(0.039)	(0.060)		(0.031)	(0.031)
3rd Group Adjustment Rule		-0.091	-0.091		-0.032	-0.033
		(0.048)	(0.048)		(0.025)	(0.025)
Observations	1746	1746	1746	1746	1746	1746

See notes for Table 3. To save space we do not report estimates of the control variables.
The same specification with Table 3 is used.
Failure of Surgery is a dummy variable of 1 if the postsurgery eyesight is larger than or equal to 1 diopters.

Table 7. Impacts of Learning Vector on Various Outcomes, Based on Severity of Cases

	Postsurgical Eyesight (1)	Failure of Surgery (2)	Need Resurgery (3)
Logarithm of Cumulative Surgeries	-0.010	0.005	-0.007
	(0.047)	(0.023)	(0.012)
2nd Group Adjustment Rule	-0.111	-0.045	-0.008
	(0.109)	(0.060)	(0.031)
3rd Group Adjustment Rule	-0.157	-0.091	-0.033
	(0.083)	(0.048)	(0.025)
Observations	1746	1746	1746
		Easy Cases	
Logarithm of Cumulative Surgeries	0.059	-0.008	-0.023
	(0.046)	(0.037)	(0.022)
2nd Group Adjustment Rule	0.009	-0.052	-0.028
	(0.113)	(0.084)	(0.029)
3rd Group Adjustment Rule	-0.316	-0.199	0.022
	(0.121)	(0.070)	(0.029)
Observations	529	529	529
		Severe Cases	
Logarithm of Cumulative Surgeries	-0.143	0.003	-0.005
	(0.148)	(0.050)	(0.035)
2nd Group Adjustment Rule	-0.023	0.062	0.018
	(0.325)	(0.186)	(0.095)
3rd Group Adjustment Rule	-0.177	-0.125	-0.088
	(0.254)	(0.137)	(0.053)
Observations	429	429	429

See notes for Table 3.
Easy cases are defined as those below the 25th percentile of presurgery eyesight. Severe cases are defined as those above the 75th percentile of presurgery eyesight.

Table 8. Impacts of Type-Specific Cumulative Surgeries on Various Outcomes

	Postsurgical Eyesight			Failure of Surgery			Need Resurgery		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(type specific cumulative surgeries)	-0.013		0.037	0.005		0.027	-0.006		0.000
	(0.032)		(0.046)	(0.013)		(0.019)	(0.009)		(0.009)
2nd GAR*Treatment		-0.126	-0.171		-0.037	-0.070		-0.018	-0.018
		(0.075)	(0.108)		(0.039)	(0.051)		(0.031)	(0.032)
3rd GAR*Treatment		-0.155	-0.154		-0.091	-0.091		-0.032	-0.032
		(0.083)	(0.083)		(0.048)	(0.048)		(0.025)	(0.025)
Observations	1,746	1,746	1,746	1,746	1,746	1,746	1,746	1,746	1,746

The type-specific cumulative surgeries of myopia, hyperopia, myopic astigmatism, hyperopic astigmatism, and mixed astigmatism were counted separately for each doctor

Table 9. Impacts of Learning Vector on Various Outcomes, with Interactions Between Individual and Collective

	Postsurgery Eyesight (1)	Failure of Surgery (2)	Need Resurgery (3)
Log(cumulative surgeries)	-0.025	0.009	-0.007
	(0.049)	(0.025)	(0.012)
2nd GAR*Treatment	-0.585	0.092	-0.024
	(0.190)	(0.113)	(0.102)
3rd GAR*Treatment	-0.285	-0.330	0.013
	(0.280)	(0.158)	(0.093)
Log(cumulative surgeries) *2nd GAR*Treatment	0.117	-0.034	0.004
	(0.048)	(0.029)	(0.026)
Log(cumulative surgeries) *3rd GAR*Treatment	0.017	0.056	-0.010
	(0.069)	(0.033)	(0.022)
Observations	1,746	1,746	1,746

See note for Table 5 and 6.