Int J Health Econ Manag (2018) 18:25–45
https://doi.org/10.1007/s10754-017-9222-z

RESEARCH ARTICLE

Organizational learning-by-doing in liver transplantation

Sarah S. Stith[1]

Received: 30 May 2016 / Accepted: 17 August 2017 / Published online: 30 August 2017

Abstract Organizational learning-by-doing implies that production outcomes improve with experience. Prior empirical research documents the existence of organizational learning-by-doing, but provides little insight into why some firms learn while others do not. Among the 124 U.S. liver transplant centers that opened between 1987 and 2009, this paper shows evidence of organizational learning-by-doing, but only shortly after entry. Significant heterogeneity exists with learning only evident among those firms entering early in the sample period when liver transplantation was an experimental medical procedure. Firms that learn begin with lower quality outcomes before improving to the level of firms that do not learn, suggesting that early patient outcomes depend on the ability of new entrants to import best practices from existing liver transplant programs. Knowledge of best practices became increasingly available over time through the dissemination of academic research and increasingly specialized training programs, so that between 1987 and 2009, 6 month post-transplant survival rates increased from 64 to 90% and evidence of organization-level learning-by-doing disappeared. The lack of any recent evidence of organizational learning-by-doing implies that common insurer experience requirements may be reducing access to health care in non-experimental complex medical procedures without an improvement in quality.

I would like to thank Tom Buchmueller, Scott Masten, Jeff Smith, Mario Macis, and Steve Leider, for helpful comments. I am also indebted to the many staff members at the University of Michigan Transplant Center for their advice and information, especially Alan Leichtman, John Magee, and Anne Murphy, to Jennifer Wainright at the United Network for Organ Sharing, and to Guerry Johnson, transplant recipient. "This work is supported in part by Health Resources and Services Administration contract 234-2005-370011C. The content is the responsibility of the author alone and does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government". http://optn.transplant.hrsa.gov/data/citing.asp.

Electronic supplementary material The online version of this article (doi:10.1007/s10754-017-9222-z) contains supplementary material, which is available to authorized users.

✉ Sarah S. Stith
ssstith@unm.edu

1 Department of Economics, The University of New Mexico, Albuquerque, NM, USA

Keywords Organizational learning · Learning-by-doing · Liver transplantation · Firm heterogeneity · Firm performance

JEL Classification D24 · D83 · I10 · I11 · L25

Introduction

Kenneth Arrow first formally recognized the importance of organizational learning-by-doing in economics, observing that "it is the very activity of production which gives rise to problems for which favorable responses are selected over time" (Arrow 1962, p. 156).[1] Empirical articles have established its existence across a broad range of manufacturing industries, but in healthcare industries, studies have focused primarily on scale effects rather than organizational learning-by-doing measured from when a provider first begins offering a procedure (e.g., Ho 2002; Gaynor et al. 2005). In addition, despite repeated evidence of the existence of organizational learning-by-doing and documentation of significant heterogeneity across industries (Balasubramanian and Lieberman 2010), only one existing paper considers intra-industry heterogeneity in organizational learning-by-doing (Pisano et al. 2001). Sources of inter-firm/intra-industry heterogeneity remain empirically unidentified in the literature.

In liver transplantation, 6-month post-transplant survival rates increased from 64% in 1987 to 90% by 2009, indicating significant learning occurred during this time period. This paper seeks to identify whether learning-by-doing occurs at the level of the organization, whether it varies across organizations, and if the magnitude of learning-by-doing depends on how much uncertainty exists about how to perform the procedure at the industry level.

As predicted by Arrow (1962) and in accordance with a broad literature in industrial organization, this study finds that organizational learning-by-doing exists in liver transplantation, but only very shortly after entry, approximately within the first 20 patients treated at a center. Extending a sixteen hospital study by Pisano et al. (2001), I find significant heterogeneity in the magnitude of learning-by-doing among the 124 centers in the sample, largely driven by the timing of entry. In accordance with theoretical predications (Jovanovic and Nyarko 1995), the amount of uncertainty at the industry level is a crucial factor in determining the extent of organizational learning-by-doing. In particular, the results show that evidence of learning-by-doing dissipates as liver transplantation advances over time from an experimental procedure to a common procedure with well-established protocols for care, specialized training programs, and better methods of immunosuppression. Organizational learning-by-doing also appears to be most important for mid-range survival, approximately 2–3 months post-surgery. At 2–3 months post-surgery, patient survival has shifted from being driven primarily by the surgical process to being determined by the careful balance of immunosuppression, enough to avoid organ rejection but not so much that the patient succumbs to fatal side effects or opportunistic infections. Although surgical expertise clearly is necessary, the limitations of and uncertainty about the optimal approach to immunosuppression always have been the primary constraints on the feasible duration of survival post-transplant. Correlations from other types of organ transplants support the association between the overall technical uncertainty in how to perform a procedure and the importance of organizational learning-by-doing.

[1] Earlier articles outside of economics already had considered how outcomes improved with practice (e.g., Wright 1936).

The results indicate that with sufficient industry-level learning, i.e., shared knowledge at the industry level, learning-by-doing at the organization level is no longer statistically significant. This directly relates to prior work, which found that learning early in an organization's history occurs primarily through the importation of best practices (Nembhard et al. 2014.) Thus if best practices are easy to acquire, little organization-level learning will exist. Almost all liver transplants are performed at large teaching hospitals with a clear mission to disseminate knowledge throughout the field by conducting research, publishing the results, and training future transplant physicians through transplant fellowships, and other personnel through increasingly specialized programs (e.g., for administrators, coordinators, and nurses.) Technical uncertainty at the industry level affected the availability of best practices and thus created an organization-level learning process in the early years of liver transplantation, but by the midway through the sample period, best practices could easily be acquired, likely through the hiring of skilled personnel.

This study overcomes some of the issues faced in the prior literature on organizational learning-by-doing in health care, in particular with regard to the number of years, providers, and patient-level control variables tracked in the data, the availability of data from entry into a procedure, and institutional and other factors that limit potential endogeneity. Twenty-two years of data are available beginning in September 1987 and include all liver transplants performed in the U.S. One hundred and twenty-four liver transplant centers, or 80% of all liver transplant centers, are followed from entry into liver transplantation. Institutional factors such as the exogenously determined cadaveric donor supply, short organ preservation times (sample mean of 8 h), aggregation of waitlists for donor livers across multiple centers in a geographic area, regulation based on survival outcomes, and limited insurer networks diminish the risk that endogeneity might be driving the results. Center-level fixed effects and extensive patient-, donor-, and match-level variables further control for a wide range of center, patient, donor, and match characteristics, which might affect patient outcomes or lead to an endogenous relationship between cumulative volume and survival.

A potential limitation of this study is that some of what empirically appears to be organizational learning-by-doing may be related to individual-level learning; the data do not include information on the many individual health professionals who interact with a patient during the transplant process. This limits the ability of this study to contribute to understanding the precise mechanism through which learning-by-doing arises in liver transplantation, but it does not detract from the importance of empirically documenting the relationship between center-level cumulative volume and post-transplant survival relationship. It is precisely such a relationship that is assumed to exist in the center-level experience requirements common among insurers and in how the Centers for Medicaid and Medicare Services regulate and provide public information on post-transplant survival outcomes. Studying organizational learning-by-doing with the team rather than the individual as the "unit of administration" is not new to the literature. Manthous et al. (2011) describe critical care as similarly "formulated and delivered by a team," thus making the team the appropriate unit for studying the effects of learning on medical outcomes in that context as well. That study finds that it is the quality of interpersonal interactions rather than individual performance that is crucial.

Although no studies on organizational learning-by-doing in health care have had access to data on all organization members, some studies have compared individual provider- and organization-level learning. Most of these studies find that the majority of learning-by-doing occurs at the organization level (Contreras et al. 2011—LASIK surgery; Pisano et al. 2001—minimally invasive cardiac surgery), and in some cases, no physician-level learning occurs (Huesch 2009—coronary artery bypass surgery) or learning skills are not transferable across hospitals (Huckman and Pisano 2006—cardiac surgery.) These studies focus on types of surgeries involving only one surgeon, so as to isolate the individual learning effect, whereas in liver transplantation multiple surgeons typically participate in a surgery lasting as many as 12 h.[2] In general, overall team size and the complexity of the procedure likely affect how learning-by-doing arises in an organization; David and Brachet (2009) find that in the case of two-person emergency medical response teams individual learning dominates firm-level learning.

The article proceeds as follows. "Theory and hypotheses" section discusses the theory and hypotheses tested in this article, "Institutional framework and data description" section provides a brief institutional overview of how patients obtain transplants and how organs are allocated, and describes the data, "Estimation strategy" section presents the empirical methodology with associated results presented in "Results" section; "Discussion" section discusses the results and "Conclusion" section concludes.

Theory and hypotheses

Although a range of theoretical models of organizational learning-by-doing exist, Jovanovic and Nyarko (1995) develop a Bayesian learning model particularly well-suited to the liver transplantation context, which provides a plausible mechanism for how organizational learning-by-doing would arise and vary across organizations. The learning-by-doing process arises in the Jovanovic and Nyarko (1995) model because the "decisionmaker", i.e., the transplant center, must make decisions based on incomplete information regarding what the optimal decision would be. Decisions could involve staffing, the timing of surgery, or post-transplant dosing of immunosuppressants. After the transplantation process concludes, the transplant center receives a signal regarding how close to optimal its decisions were. The signal in the transplantation context would be the patient's outcomes post-transplant. Although in the context of the model, I discuss only one "signal", the signal would be the ongoing information regarding the patient's wellbeing in the follow-up period beginning with whether the surgery was successful. Information on longer term outcomes will take longer to receive and may be subject to more unexplained and/or random variation than information on shorter term outcomes. After receiving the signal, the transplant center updates its expectation of the optimal set of decisions for the next patient based on the new information gained from the signal. The signal received by the decisionmaker does not clearly indicate what the optimal decisions would have been due to noise in the signal that arises from transitory disturbances such as having a transplant team member unexpectedly absent on the day of transplantation or still poorly understood patient and donor molecular-level genetic incompatibilities. With each subsequent patient treated, the transplant center continues updating its expectation of the optimal approach and its decisions increasingly approximate the ideal set of decisions. Heterogeneity in organizational learning-by-doing in liver transplantation will exist due to variation in both the signal and its variance arising from the high level of diversity in the patient population and the many factors influencing post-transplant survival, from the composition of the surgical team to the quality of patient follow-up care. The model predicts that the amount of learning-by-doing that occurs will depend on the noisiness of the signal received after a treatment period has been completed, the number of decisions involved in each patient's treatment, and uncertainty about the optimal approach for a given transplant patient.

[2] http://www.mayoclinic.org/tests-procedures/liver-transplant/details/what-you-can-expect/rec-20211848. Accessed 01/22/17.

The noisiness of patient outcomes and the large number of decisions made during a patient's treatment did not changed significantly during the sample period. However, the level of uncertainty has changed over time. Later in the panel period decisions will more closely approximate the ideal decision because they will be better educated guesses due to common knowledge regarding liver transplantation that will have been disseminated through research publications, conferences, increasingly specialized training programs, and even the expansion in access to the Internet during this time period. These factors along with transplant-specific technology innovations, such as new and better immunosuppressants and organ allocation algorithms, all act to decrease the uncertainty about the optimal approach by standardizing and simplifying liver transplantation. Further evidencing the development of liver transplantation over time from experimental to mainstream medicine, Medicare extended liver transplant coverage from a case-by-case basis to all Medicare eligible individuals with liver failure other than liver failure caused by Hepatitis B or cancer in 1996, to Hepatitis B patients in 1999, and to patients with hepatocellular carcinoma in 2001 (CMS 2006).

Similarly, I anticipate greater organizational learning-by-doing in longer term survival outcomes than very short-term survival because very short-term survival will depend primarily on surgical expertise, while longer term survival depends on immunosuppression, a task with greater uncertainty and longer delays in feedback than a surgical procedure.

Institutional framework and data description

With end-stage liver disease, only a liver transplant can save the patient's life. After diagnosis, a patient's doctor refers the patient for transplantation. Usually, patients must choose a transplant center within their insurer's preferred provider network.

The decision to waitlist a patient lies with the transplant center. Criteria include the ability to pay for the transplant and long-term immunosuppression, proof that the patient can arrive at the hospital within the limited number of hours in which a donor liver can be used for transplant, and no physical or psychological contraindications. Once on the waitlist, the patient must undergo frequent testing to maintain his or her priority status. The patient then waits to be matched with a donated organ.[3]

While the decision to waitlist a patient lies with the transplant center, cadaveric organs are allocated to patients exclusively through the Organ Procurement and Transplantation Network (OPTN), created by Congress with the National Organ Transplant Act in 1984[4] and operated ever since by the United Network for Organ Sharing (UNOS), a nonprofit contractor. Since 2008, the OPTN has had authority over live organ donation as well (OPTN/UNOS 2008). The OPTN consists of fifty-eight local Organ Procurement Organizations, which each oversee a single exclusive Donation Service Area, containing between one and nine liver transplant centers. These Donation Service Areas are then aggregated into eleven regions.

When a hospitalized patient dies or death is imminent, the hospital must notify the local Organ Procurement Organization, as required by law since 1998 (Department of Health and Human Services 1998). Personnel from the Organ Procurement Organization contact the family of the patient to obtain consent, often even if donation already was authorized. If

[3] Although most patients are waitlisted at some point during the transplant process, living donors can direct their donation to an individual patient, allowing that patient to opt out of the cadaveric donor waitlist. About 4% of transplants in the data used live donors, with the first successful live donor liver transplant performed on November 27, 1989. Directed cadaveric donation is possible, but it is not tracked in the data and rarely occurs according to conversations with staff at a major U.S. transplant hospital.

[4] National Organ Transplant Act of 1984, Pub. L. No. 98-507, Sect. 372, Stat. 2339 (1984).

consent is given, UNOS follows the OPTN liver allocation algorithm to determine the best matches for the organ, based on medical criteria such as blood type, organ size, and disease severity. Time on the waitlist serves as a tie-breaker. In general, livers are first allocated within a Donation Service Area, before being offered regionally, and then nationally. According to the data used in this study, 68% of livers are allocated within the Donation Service Area, 24% are allocated at the regional level, and 7% are allocated nationally.

The OPTN data, obtained through UNOS, include all organ transplants performed in the U.S. since October 1, 1987 (United Network for Organ Sharing 2010).[5] Pursuant to the associated confidentiality agreement, I am not allowed to show results for any individual patients or centers. Based on the actual reported transplant dates, the data panel ranges from September 30, 1987 to December 31, 2009, and includes a total of 101,120 liver transplants. Almost all patients receive only one transplant; the 101,120 transplants documented in the data were performed on 91,626 patients.

Although the period from September 30, 1987 to December 31, 2009 captures the vast majority of liver transplants, 31 of the 155 centers in the data performed transplants prior to September 30, 1987 (Terasaki 1986). Based on news archives, only three centers performed liver transplants prior to 1983 in the U.S., so most missing volume information comes from shortly prior to the inception of my panel. Although the first successful liver transplant was performed in 1967, only fifteen procedures were performed in 1980. By 1986, the annual number of liver transplants had increased to 924. In total, only around 3000 liver transplants were performed before the OPTN began collecting data on all types of organ transplants (Evans 1991).

Although the amount and duration of pre-panel data are limited, the theoretical literature on learning-by-doing, the empirical literature from manufacturing, and the raw data all suggest that most learning-by-doing occurs quite soon after a center enters. Therefore, I include in my analysis only those centers beginning transplantation after September 30, 1987.[6] These 124 centers performed 64,320 transplants on 59,005 patients or 64% of transplants performed and patients treated between September 30, 1987 and December 31, 2009. For the analyses by time period, the sample consists of 72 new centers that entered during the first period with 1145 patients falling within the first 20 transplants performed and 52 centers that entered in the second period with 1154 patients falling within the first 20 transplants performed. Further details on the data construction are available in the "Online Appendix", Section I.

Estimation strategy

The regression specifications I use in this article parallel those commonly used in the manufacturing literature to study organizational learning-by-doing. In general, those articles use a log-linearized Cobb–Douglas production function as the estimating equation (e.g., Balasubramanian and Lieberman 2010; Levitt et al. 2013). The specification I use approximates a production function with hospital and industry experience, case mix, and time-invariant

[5] The UNOS website states October 1, 1987, but the reported transplant dates in the data start on September 30, 1987.

[6] To briefly describe the differences between "incumbent" centers and centers entering during the panel period: the incumbent centers perform more transplants per year on average (115 vs. 75), are in larger hospitals (average daily census of 536 vs. 494), and are more likely to be government-owned (34 vs. 28%) and less likely to be privately owned (0 vs. 3%). (Ownership and average daily census data were obtained from the American Hospital Association (2009)). The mean year of entry and average 6-month survival rates for incumbents versus entrants are 1983 and 1990 and 85 and 88%, respectively.

hospital-specific factors as inputs in the production of patient survival. I measure hospital experience using categorical measures of cumulative volume since entry, overall changes in industry experience and technology and regulatory shocks using year fixed effects, case mix using donor, patient and match characteristics associated with patient survival, and time-invariant hospital characteristics using center fixed effects. Examples of such characteristics include hiring a particularly accomplished liver transplant surgeon to lead the new center, attempting to start up multiple types of organ transplant centers within a hospital at the same time, or experience transplanting other organs prior to beginning liver transplantation. Given the dichotomous nature of most of the variables included in the regression, simply using a log-linearized Cobb–Douglas production function is not feasible.[7] The categorical cumulative volume variables also fit the data better than using the natural log of cumulative volume.

In order to evaluate the existence of organizational-learning-by-doing, I use the following linear probability model, focusing specifically on early learning, i.e., the within-center change in survival outcomes as cumulative volume increases.[8]

$$\begin{aligned} \mathit{Alive\ at\ 6\ months}_{iht} &= \alpha + \boldsymbol{CumulativeVolume}'_{iht}\boldsymbol{\theta}_{volume} + \boldsymbol{CaseMix}'_{\boldsymbol{iht}}\boldsymbol{\beta}_{casemix} \\ &\quad + \boldsymbol{\gamma_h} + \boldsymbol{\tau_t} + \varepsilon_{iht} \end{aligned} \tag{1}$$

The dependent variable, $\mathit{Alive\ at\ 6\ months}_{iht}$, equals one if patient i treated at transplant center h at time t survives at least 6 months and equals zero otherwise. I use 6-month survival in order to capture the multiple aspects of the transplantation process from surgery through maintenance immunosuppression. For the longer survival duration periods, the effects of learning-by-doing on survival outcomes are less accurately measured due to the influence on survival of other factors exogenous to transplantation and the difficulty in longer term tracking of patients years after surgery. Cumulative volume refers to the cumulative number of patients treated by a given center. Because the relationship between cumulative volume and survival in the raw data does not follow an obvious functional form and the empirical results in manufacturing suggest that organizational learning-by-doing is a short-lived phenomenon, I tested a variety of categories of cumulative volume to assess when learning-by-doing occurs. In the main specification, I group transplants below 50 into the cumulative volume categories of 1–10, 11–20, 21–30, 31–40 and 41–50, which allows me to focus on the period shortly after entry when organizational learning-by-doing is likely to be most evident. The vector $\mathit{CumulativeVolume}_{iht}$ represents the cumulative volume categorical variables.[9]

The control variables include center fixed effects (γ_h), year fixed effects (τ_t) and patient, donor, and match characteristics. Organ transplantation is primarily regulated at the national level, so regulatory changes will affect all centers at the same time. In addition, changes in immunosuppression, insurance coverage, and the training of transplant personnel diffused generally across all centers. A positive secular trend in survival rates also suggests that year fixed effects are required. (Average survival rates increase by 24 percentage points between 1987 and 2009). Including year fixed effects could lead to conservative estimates of organizational learning because cumulative volume increases over time. However, for the reasons just mentioned, the inclusion of year fixed effects is appropriate, even at the expense of possibly underestimating the full learning-by-doing effect. $\mathit{CaseMix}_{iht}$ is a vector

7 In general, the literature on learning-by-doing in industrial organization uses a natural log transformation of a standard Cobb–Douglas production function ($f(x, y) = Ax^{\alpha}y^{\beta}$ (becomes $\ln(f(x, y)) = \ln(A) + \alpha \ln(x) + \beta \ln(y)$, where A is a constant measuring total factor productivity and x and y are factors of production.

8 Results from using a logit model are very similar and are provided in Online Appendix Table OA1.

9 Note that cumulative volume is based on total patients treated, i.e., patients dropped due to missing data are included in the cumulative volume measure.

of controls for patient, donor, and match characteristics that previously have been established by the literature on liver transplantation as important for post-transplant survival (Edwards et al. 1999; Axelrod et al. 2004; Freeman et al. 2008).[10] If new centers face disproportionately lower quality organs, worse matches, and sicker patients than they do later on, what appears to be a learning curve might simply be explained by a less favorable mix of cases treated.[11] It does appear that early patients might be higher risk cases. They are more likely to have a life expectancy of less than 7 days (Status 1), be in an intensive care unit, be on life support at the time of transplant, or have a blood type of O.[12] The organs received by these early patients also tend to have had longer storage times (cold ischemia time) and have traveled a longer distance between the donor and recipient hospital, both factors negatively associated with survival. However, these early transplants have characteristics associated with better survival outcomes as well; younger patients with lower serum creatinine levels (measure of kidney function),[13] shorter wait times, younger donors, and fewer blood type incompatibilities between donor and recipient.

After establishing the existence of organizational learning-by-doing, I explore the existence and magnitude of heterogeneity in learning-by-doing among centers by expanding Eq. (1), interacting cumulative volume with the center fixed effects and omitting the constant.

In order to test whether the amount of learning-by-doing decreases with the availability of industry-level knowledge, I used three approaches. First, I split my sample into two periods: 1987–1993 and 1994–2009 to analyze how organizational learning-by-doing changed as liver transplantation became mainstream medicine. The regressions are run separately by period in order to allow the coefficients on the other independent variables to change over time as well, which they do.[14] Second, alternative survival durations are studied continuing with the same

[10] *Casemix* controls for serum creatinine, Status 1 allocation ranking, wait time in days, age, gender, race (African American or other), diagnosis (AHN, biliary atresia, cholestatic liver disease, cirrhosis, metabolic disorder, neoplasm or other), multi-organ transplants, blood type, ICU and hospitalization status at transplant, life support, live donor, donor age, donor gender, donor race (African American or other), donor-recipient blood type compatibility, cold ischemia time, distance from donor hospital to transplant center, and whether a whole or split liver was transplanted. Further details on these variables are available in Online Appendix Table OA2.

[11] Online Appendix Table OA3 details patient, donor, and match characteristics included in the regressions and shows that about half of the means of these variables differ between the first 20 patients and the 21st to 40th patient treated.

[12] As universal donors, donor organs with blood type O can be matched with any other blood type, but type O transplant patients can only receive a type O organ. The allocation policy for Status 1 patients prioritizes these patients in receiving type O organs regardless of patient blood type. For non-Status 1 patients, organs are first matched to patients with identical blood types. Some studies indicate that patients with blood type O may have slightly worse access to organs due to the demand for donor organs of blood type O by all blood types. These studies document longer waiting times for patients with blood type O (Freeman and Edwards 2000; Barone et al. 2008).

[13] As liver disease progresses, it negatively affects the kidneys leading the elevated serum creatinine levels. The current liver allocation algorithm uses serum creatinine as a measure of disease progression in prioritizing patients for cadaveric donor organs.

[14] For example, improvements in immunosuppression mean an incompatible blood type match between donor and recipient matters less for survival than it did before the introduction of the new drugs. (An incompatible blood type between patient and donor in the early period is associated with a statistically significant three percentage point reduction in 6-month survival, but in the later period the effect is positive and insignificant.) The prevalence of most patient, donor, and match characteristics also changes over time and all the control variables differ statistically from each other between the two periods based on a two-sided t-test except for the following three recipient characteristics; diagnosis of adenomatous hyperplastic nodule, being African-American, and having a blood type O (See Online Appendix Table OA4).

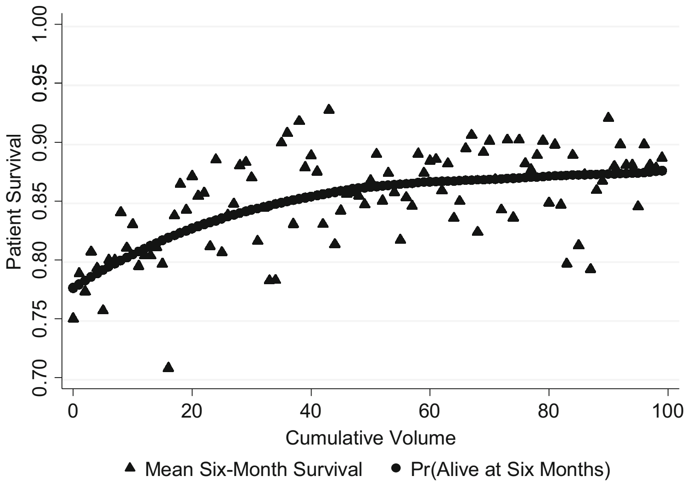

Fig. 1 Six-month survival by cumulative volume. *Notes:* The probability of being alive 6 months post-transplant is predicted by locally weighted regressions as a function of cumulative volume up to 100 patients, using a bandwidth of 0.8, i.e., 80% of the data. The predicted values are graphed by cumulative volume up to 100 patients. The triangles represent the mean 6 month survival across centers for that level of cumulative volume

regression approach as well as depicting the relationship graphically through the use of Kaplan Meier curves. Third, although the data on other types of organ transplants are more limited than those for liver transplantation due to longer pre-data collection histories (kidneys and hearts) or substantially fewer total transplant centers (lungs and intestines), locally weighted regressions are used to compare the correlation between survival and cumulative volume. Kidney and heart transplants are well-established approaches to treating kidney and heart failure while lung, and particularly, intestinal transplants continue to be considered more experimental procedures by providers and insurers.

In all cases, standard errors are clustered at the center level to control for heteroskedasticity and spatial correlation among patients treated at a given center.

Results

Figure 1 presents evidence consistent with organizational learning-by-doing. The scatter plot graphs the relationship between the mean patient survival rate across new centers and cumulative volume. Each triangle represents the mean 6-month survival rate across the 124 centers for that level of cumulative volume. Given the noisiness even in the averaged raw survival rates, I also graph 6-month survival probabilities (predicted using locally weighted regressions) by cumulative volume up to the 100th patient.[15] A clear upward but diminishing trend exists.

Although suggestive, the graph does not control for time-invariant center characteristics, factors affecting all centers in a given time period, and patient, donor and match characteris-

[15] The probability of being alive 6 months post-transplant is predicted by locally weighted regressions as a function of cumulative volume up to 100 patients, using a bandwidth of 0.8, i.e., 80% of the data. The predicted values are graphed by cumulative volume up to 100 patients. The triangles represent the mean 6 month survival across centers for that level of cumulative volume.

Table 1 Main results for learning-by-doing (outcome = alive at 6 months)

	(1)	(2)	(3)	(4)	(5)	(6)
Cumulative volume						
1–10	−0.099***	−0.102***	−0.042**	−0.043**	−0.043**	−0.042***
	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)
11–20	−0.081***	−0.088***	−0.039**	−0.040**	−0.039**	−0.038***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)
21–30	−0.030**	−0.037***	0.004	0.004	0.002	−0.000
	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)	(0.011)
31–40	−0.032**	−0.036***	0.001	0.001	−0.001	−0.002
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
41–50	−0.032***	−0.036***	−0.003	−0.004	−0.003	−0.001
	(0.011)	(0.011)	(0.012)	(0.011)	(0.012)	(0.011)
Observations	59,180	59,180	59,180	59,180	59,180	59,180
R^2	0.003	0.014	0.017	0.020	0.024	0.058
Center fixed effects		Yes	Yes	Yes	Yes	Yes
Year fixed effects			Yes	Yes	Yes	Yes
Donor characteristics				Yes	Yes	Yes
Match characteristics					Yes	Yes
Recipient characteristics						Yes

The omitted cumulative volume category is cumulative volume > 50. Standard errors are clustered at the center level

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

tics. Equation (1) incorporates these factors yielding the results show in Table 1. The reported effects for each of the cumulative volume categories are relative to being the 51st or above transplant recipient at a center. Column [1] of Table 1 presents results from regressing 6-month survival on the cumulative volume categories without any control variables and shows clear evidence of a learning curve. The importance of the effect generally diminishes with increasing volumes and is consistently negative and significant. Given that I am analyzing learning-by-doing or the within-center effect of a change in cumulative volume on survival, I add center fixed effects in Column [2]. The evidence for organizational learning-by-doing becomes even stronger with the inclusion of the center fixed effects, although the center fixed effects contribute little to explaining the variation in the outcome variables.

The inclusion of year fixed effects in Column [3] substantially reduces the size and statistical significance of the coefficients. The large effect on the coefficients from including the year fixed effects suggests that the importance of organizational learning-by-doing may have changed over time and supports that organizational learning-by-doing might be more prevalent earlier in the history of liver transplantation.

In Column [4] I add donor characteristics, in [5] match characteristics and in [6] recipient characteristics. The R^2 more than doubles, but the coefficients on the cumulative volume variables remain essentially unchanged. In other words, it does not appear that patient, donor, and match characteristics are driving the negative coefficients on cumulative volume between 1 and 10 and cumulative volume between 11 and 20.

The coefficients from the complete model in Column [6] of Table 1 imply that almost one additional patient dies within 6 months during a center's first 50 transplants relative to during its later transplants.

Table 2 Heterogeneity in learning-by-doing

Worse or better/ same during first 20 transplants	# of centers	# of patients	Average net effect	Six-month survival: 1–20 (%)	Six-month survival: 21–40 (%)	Six-month survival: overall (%)
Worse	72	41,357	−0.105	75	85	88
Better/same	38	17,703	0.060	91	89	90
Total	110	59,060	−0.048	81	86	89

Net effects are the sum of the main effect of being within the first 20 transplants plus the transplant center-specific effect of being within the first 20 transplants (interaction term). The underlying regression includes center and year fixed effects and patient, donor, and match characteristics

Table 2 summarizes the results from calculating the center-specific effects by interacting *Cumulative Volume* with the center-specific dummies, omitting the constant. For simplicity, I reduce the number of categories to one category for whether or not a patient was within the first 20 transplants performed at a center. (Reducing cumulative volume to one category for patients one to twenty does not affect the explanatory power of the model; I cannot reject the null that the coefficients on the cumulative volume categories over twenty equal zero in any of the columns. See Online Appendix Table OA5).

The two groups of centers are categorized based on whether the net effect on 6-month survival of being within the first 20 patients is negative or if it is positive or neutral at that center. Seventy-two centers improve while thirty-eight do worse or the same after their first 20 patients.[16] The average net effect among those that improve after the first 20 transplants is negative eleven percentage points and for those that get worse or stay the same, it is positive six percentage points. An F-test of the equality of the coefficients of the interaction effects can be rejected with a p-value of zero. These results indicate that heterogeneity in organizational learning-by-doing does exist across centers.

The next two columns of Table 2 give information on how the learning process occurs. Those centers that improve after treating their first 20 patients start off with lower initial survival rates than those that become worse or stay the same after treating their first 20 patients. The gap narrows from a difference of sixteen percentage points for patients 1–20 to a difference of four percentage points for transplants 21–40. The centers that improve gain an average of ten percentage points, while those that get worse or stay the same only drop by two percentage points. The overall survival rates in the last column show only a two percentage point difference overall between those that improve and those that stay the same or get worse during their first 20 transplants.

Figure 2 shows a histogram of the net effect of being within the first 20 patients at a center. Although the effect is fairly evenly distributed around zero, noticeable variation in center-specific learning exists. Both tails have some notable outliers, but the distribution is generally left-skewed.[17]

[16] Because fourteen centers included in these regressions do not reach twenty transplants, I cannot calculate a center-specific learning penalty for those centers.

[17] Because the sample sizes are quite small for some centers, the fixed effects estimates may be imprecise. As a robustness check, due to the small sample sizes at some centers, I use Empirical Bayes' estimation to estimate individual center-specific effects on student outcomes (e.g., Kane and Staiger 2008; Jacob and Lefgren 2005). Although these new estimates suggest less heterogeneity and more learning-by-doing, Guarino et al. (2015) documents via simulation analysis that Empirical Bayes' estimators diminish the variance in the estimated effects but at the price of bias and inconsistency, especially under non-random assignment of patients to transplant centers, which obviously arises in cases at children's hospitals or for patients receiving multiple transplants. A histogram of the shrunken effects is overlaid on Fig. 3 in Online Appendix Figure OA1.

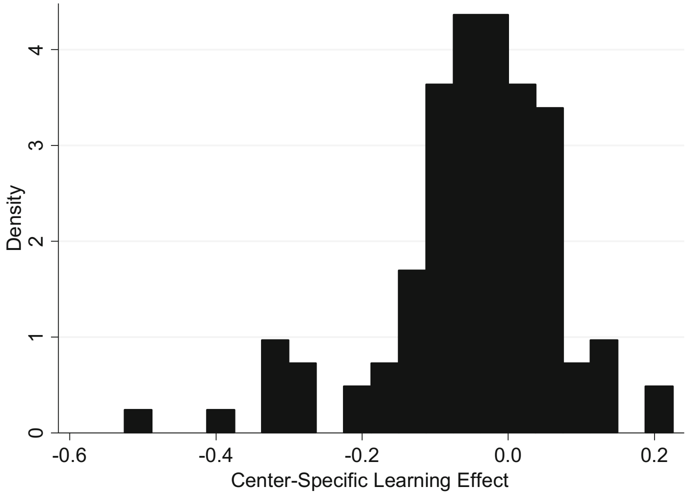

Fig. 2 Center-specific learning. *Notes*: The figure depicts a histogram of the sum of the coefficient on *Cumulative Volume* $\leq$ 20 and the coefficient on the interaction of the center fixed effect and *Cumulative Volume* $\leq$ 20

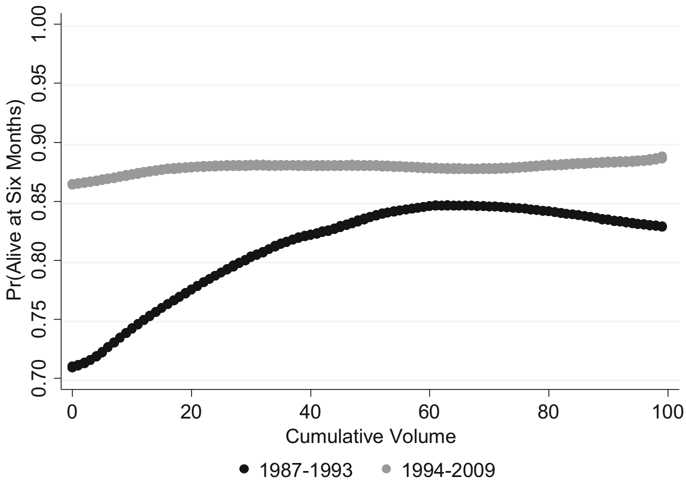

Fig. 3 Six-month survival by cumulative volume by period. *Notes:* The probability of being alive 6 months post-transplant is predicted by locally weighted regressions as a function of cumulative volume up to 100 patients, using a bandwidth of 0.8, i.e., 80% of the data. The predicted values are graphed by cumulative volume up to 100 patients

Figure 3 replicates the locally weighted regression results depicted graphically in Fig. 1 for the two periods (1987–1994 and 1995–2009). Although both periods have close to the same survival rate by patient 60, how they attain that survival rate varies. Pre-1994, a clear learning curve appears that plateaus around 60 patients. In the second period between 1994 and 2009, essentially no evidence of learning-by-doing exists. The slight first period decline in survival rates in centers performing more than 60 transplants can likely be explained by the fact that this figure does not control for patient, donor, and match characteristics, and larger hospitals, especially teaching hospitals, tend to be willing to treat more difficult cases than smaller centers.

Table 3 Learning-by-doing by period (outcome = alive at 6 months)

	1987–1993	1994–2009
Cumulative volume		
1–10	−0.052**	−0.016
	(0.026)	(0.016)
11–20	−0.077***	−0.008
	(0.024)	(0.015)
21–30	−0.008	0.002
	(0.020)	(0.011)
31–40	−0.035*	0.010
	(0.020)	(0.015)
41–50	−0.001	−0.011
	(0.019)	(0.015)
Observations	5782	53,398
R^2	0.139	0.048

The omitted cumulative volume category is cumulative volume >50. All regressions include donor, match, and patient characteristics and center and year fixed effects. Standard errors are clustered at the center level

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

In the first period, the raw survival rate increases from 73 to 82% between the first 20 transplants and transplants 21–40. The second period shows an improvement of only two percentage points, suggesting that any organizational learning-by-doing is likely to occur in the first period.

Table 3 further documents that all learning-by-doing happens in the first period. In the overall sample results in Column [6] of Table 1, 0.8 additional patients die within 6 months as a center treats its first 50 transplants than would have died among fifty patients treated later in that center's history.[18] The cost of inexperience in Period 1 is higher than in Period 2. In Period 1, 1.3 additional patients die from inexperience, but by Period 2, no statistically significant evidence of learning-by-doing exists and the coefficients are much smaller although still negative.[19]

Further exploring the relationship between industry-level uncertainty about the optimal approach and organizational learning-by-doing, Fig. 4 presents Kaplan Meier survival curves and Table 4 summarizes regression results using a range of survival durations across the two periods to show that most of the difference in outcomes associated with learning-by-doing occurs well after surgery and again, only in the first period. In Fig. 4, only patients dying within the first 6 months post-transplant are included. In the first period, the greatest difference appears to exist about 2–3 months post-transplant. If anything, patients in the second period

[18] If each patient within the first 10 patients has a decreased probability of survival of 0.042 percentage points, then the decreased survival probability summed over the group of ten patients would be 0.42 percentage points. Only the first two categories are statistically significant, so on average a center would expect to lose 0.8 more patients during the first 50 transplants versus 50 transplants thereafter.

[19] Using dummy variables for each period and interactions between those variables and cumulative volume less than or equal to 20, gives similar results. The coefficient on the Period 1 dummy variable is large in size and statistically significant, indicating survival improvements by period and the coefficient on the interaction between being within the first 20 transplants and being in Period 1 is essentially the same as in the main period-level results for the single cumulative volume category (cumulative volume = 20) in Table 3. The main effect of being within the first 20 transplants is statistically insignificant with the inclusion of the period-level dummy. Results splitting the second period into 1994–2002 and 2003–2009 are available in Online Appendix Table OA6. Measuring periods based on date of entry for each center rather than data of transplant or limiting the analysis to only new entrants in each period only increases the magnitude of learning-by-doing between 1987 and 1994 (See Online Appendix Table OA7 for more details).

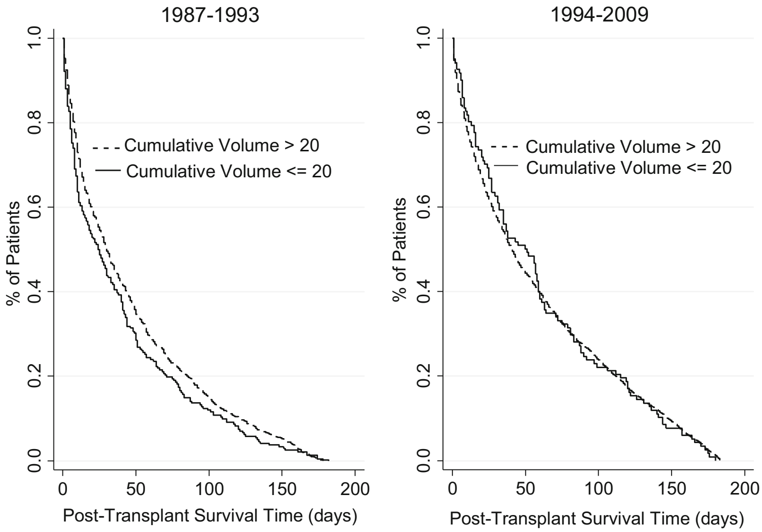

Fig. 4 Kaplan–Meier survival curves by period. *Notes:* The Kaplan–Meier curves graphing survival time for patients dying within the first 6 months post-transplant

appear to have better surgical outcomes at inexperienced centers. However, these graphs do not include any control variables. Table 4 reports the results for the full model including all controls. The coefficients between the two periods are most similar in terms of magnitude and statistical significance for 1 day survival, further suggesting that it is not improvements in surgical care that drives the differences between the two periods. For all other survival periods, the coefficients on cumulative volume less than or equal to 20 are economically and statistically significant in the first period and small and insignificant in the second period.

Figure 5 further supports a relationship between the level of technical uncertainty faced by new entrants and the magnitude of organizational learning-by-doing with evidence from locally weighted regressions using OPTN data on kidneys, pancreases, hearts, lungs, and intestines transplanted between October 1987 and December 2009. Only centers entering during the sample period are included. Kidney transplants are the most common type of transplant, have the longest history and show no evidence of learning-by-doing, while more experimental procedures such as lung and intestinal transplants show more learning-by-doing than the liver transplants studied in the article. Hearts and pancreas transplants fall somewhere in between.

Two common critiques of the literature on learning-by-doing are potential endogeneity in the volume-outcome relationship and whether learning occurs simply by existence or if it requires actual production. If some factor positively influences both cumulative volume and survival, the omitted variable could lead to an endogenous relationship between cumulative volume and survival that looks like learning-by-doing. However, the full model used here, including year and center fixed effects and patient, donor, and match characteristics, greatly reduces and essentially eliminates these possibilities.

The inclusion of center fixed effects means that endogeneity between individual survival outcomes and cumulative volume would have to exist within an individual transplant center, which differs in plausibility from a cross-center comparison context.[20] In addition, decreasing

[20] A common concern in the healthcare literature, termed "selective referral", is that rather than higher volume leading to better outcomes, better outcomes attract more patients. Although this is a clear issue when comparing

Table 4 Alternative survival times (outcomes vary as specified in the table below)

	(1) Overall	(2) 1987–1993	(3) 1994–2009
Alive at 1 day			
Cumulative volume ≤ 20	−0.015***	−0.014*	−0.008
	(0.005)	(0.008)	(0.005)
Observations	59,180	5782	53,398
R^2	0.021	0.091	0.018
Alive at 1 week			
Cumulative volume ≤ 20	−0.025***	−0.042***	−0.004
	(0.007)	(0.012)	(0.007)
Observations	59,180	5782	53,398
R^2	0.037	0.099	0.032
Alive at 3 months			
Cumulative volume ≤ 20	−0.035***	−0.051***	−0.006
	(0.010)	(0.017)	(0.012)
Observations	59,180	5782	53,398
R^2	0.057	0.133	0.048
Alive at 6 months			
Cumulative volume ≤ 20	−0.039***	−0.054***	−0.012
	(0.011)	(0.017)	(0.013)
Observations	59,180	5782	53,398
R^2	0.058	0.138	0.048
Alive at 1 year			
Cumulative volume ≤ 20	−0.036***	−0.060***	−0.001
	(0.011)	(0.018)	(0.014)
Observations	59,180	5782	53,398
R^2	0.110	0.137	0.112
Alive at 3 years			
Cumulative volume ≤ 20	−0.028**	−0.053**	0.003
	(0.012)	(0.024)	(0.017)
Observations	50,521	5782	44,739
R^2	0.112	0.118	0.118

The omitted cumulative volume category is cumulative volume >20. All regressions include donor, match, and patient characteristics and center and year fixed effects. Standard errors are clustered at the center level
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

the potential that apparent learning-by-doing is actually new centers attracting worse cases, I include a wide range of donor, match, and patient characteristics and show in Table 1 that these do little to affect the coefficients on the cumulative volume variables. The fact that all organizational learning-by-doing is evident only shortly after entry means that the omitted

across centers, within center reverse causality seems much less likely, especially with a learning curve that plateaus after twenty transplants and the long wait times for transplants.

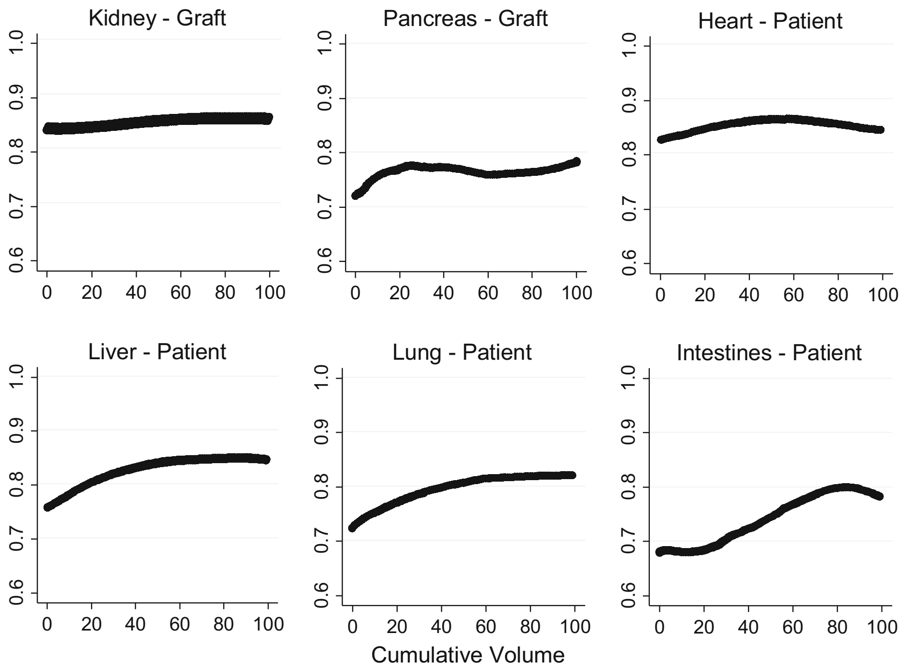

Fig. 5 Six-month survival by cumulative volume by organ. *Notes:* The probability of graft failures for kidneys and pancreas and of being alive 6 months post-transplant for hearts, livers, lungs, and intestines is predicted by locally weighted regressions as a function of cumulative volume up to 100 patients, using a bandwidth of 0.8, i.e., 80% of the data. The predicted values are graphed by cumulative volume up to 100 patients. Although graft survival outcomes will be artificially truncated by death for some kidney and pancreas transplants, I still use the graft survival time for those two organs because graft failure rarely results in immediate death due to the availability of dialysis. Only centers entering after October 1987 are included

variable issue would apply only to the difference between the first 20 patients and those thereafter rather than contributing to a longer term relationship.

Apart from the inclusion of center fixed effects and patient, donor, and match characteristics diminishing the probability of endogeneity leading to inconsistent results, institutional and medical factors make it unlikely that centers select or attract sicker patients early on. Centers have extensive control over the patients they waitlist and choose to transplant (Levenson and Olbrisch 1993; CMS 2007; Stith and Hirth 2016) and are incentivized to maintain high survival rates; centers face the possibility of closure for noncompliance by the OPTN (and CMS after 2007) if the observed survival rate falls too far below the risk-adjusted "expected" survival rate. Furthermore, the possibility that centers systematically select sicker patients shortly after entry is not corroborated by observable case mix or informal discussions with transplant personnel at the University of Michigan Transplant Center and the 2014 World Transplant Congress.

Further reducing concerns about reverse causality in particular, it is difficult for a center or a patient to affect whether or not that patient will be within the first 20 patients treated at that center. Each center determines which patients to waitlist, but the waitlist itself is aggregated at the Donation Service Area level with organs allocated on the basis of the characteristics of the donor organ available, how close a patient is to death without a transplant, and waitlist time. With less than 60% of waitlisted liver transplant candidates receiving a transplant, an average wait time of 224 days, and a liver allocation system that prioritizes the sickest patients, it seems unlikely that these very sick patients or their doctors would turn down a donor liver in order to wait for the transplant center to become more experienced (OPTN 2015).

The other common critique of learning-by-doing studies is that learning may accrue as a function of the time, i.e., a center's performance improves over time rather than as a function of volume-based experience. Controlling for center age in the regressions only increases the size of the coefficients on cumulative volume without reducing its statistical significance. This is the opposite of what we would expect if learning-by-doing accrues over time rather than with volume. The coefficient on center age becomes insignificant with the inclusion of center and year fixed effects.[21] Because liver transplantation is already a fairly low volume activity, taking a longer time to reach twenty transplants may actually reduce survival due to the absence of sufficient interactions to ensure learning.

Discussion

The results for the overall sample support the existence of learning-by-doing in liver transplantation and that this learning exists only for the set of initial patients treated as found in Pisano et al. (2001) and the manufacturing literature (e.g., Irwin and Klenow 1994; Benkard 2000). When found in other types of organ transplants, organizational learning-by-doing also is isolated shortly after entry.

Significant heterogeneity in the magnitude of learning-by-doing exists across centers with those centers that exhibit learning-by-doing improving their outcomes significantly relative to the change in the performance of those centers that do not improve or that maintain their performance after entry. In general, one would expect learning to be a positive factor; those that learn do better than those that do not. These results, however, indicate that those centers that learn do so because they start off with low survival rates and improve to the level of centers that enter with and maintain high survival rates. Mean reversion could explain some of this effect, but economic theory, the empirical literature documenting organizational learning-by-doing shortly after entry, the magnitude and consistency of the results across specifications, and the isolation of the effect shortly after entry and in the first half of the sample period support that the pattern identified is not random.

Whether a new entrant experiences a learning curve or is able to enter with high survival rates appears to depend fundamentally on three factors, the underlying complexity of the procedure, the existence of knowledge regarding how to successfully perform such a procedure, and the ability of new entrants to obtain this knowledge.

In liver transplantation, the complexity of the underlying procedure has not changed measurably, but the existence of knowledge dramatically increased between 1987 and 2009 as the procedure matured from experimental to mainstream medicine. Even if knowledge exists, new entrants will only be able to benefit from such knowledge if it is easily transferred across organizations, either directly or through hiring of skilled labor. Importation of best practices has been shown to be important for early learning outcomes in caring for heart attack patients, while later learning occurs more through internal creative processes (Nembhard et al. 2014). Given the isolation of the learning penalty to the first transplants after entry, it would follow that external sources of information may be particularly important in liver transplantation.

[21] Online Appendix Tables OA8 and OA9 provide results from regressions measuring center age in quarters and years, respectively.

In transplantation, the flow of knowledge is particularly unimpeded. Almost all transplants are performed at large teaching hospitals where faculty work to produce research and to train new physicians who will go on to work at other institutions, bringing with them their acquired knowledge. Further, because transplantation is a fairly infrequent procedure, clinical trials and training of transplant fellows often occurs across transplant centers in order to include a sufficient number of transplants. Annual conferences and several high impact discipline-specific journals further expedite the flow of information to new entrants. Exogenous factors such as the expansion of internet access during this time period further facilitated knowledge acquisition from external sources. Unlike knowledge obtained through internal processes, knowledge obtained through the adoption of best practices may be less subject to depreciation or forgetting because the information can easily be reacquired. Importation of best practices will be further facilitated by the availability of an experienced and reliable labor force. David and Brachet (2011) find that labor turnover increases organizational forgetting or the inability of the organization to retain learned knowledge over time.

When knowledge is readily shared across organizations, this allows for *industry*-level learning rather than requiring that each individual organization learn from nothing how to perform a procedure. Despite the rather minimal evidence of organizational learning-by-doing in liver transplantation, survival rates increased from 64 to 90% between 1987 and 2009. Even during the second period, when no evidence of organization-level learning-by-doing exists, survival rates continued to increase, indicating that industry-level learning continued to occur as the industry as a whole selected "favorable responses" to problems as they arose (Arrow 1962, p. 156). Underscoring the importance of industry-level learning-by-doing for post-transplant survival, the year fixed effects dramatically reduce the magnitude of the coefficients on cumulative volume in Table 1 with the coefficients on the year fixed effects increasing in size over time, and after 1990, all statistically significantly different from the effect of receiving a transplant during the omitted year (1987).

Comparing the existence of organizational learning-by-doing across types of organ transplants and survival durations shows that the free flow of information alone is insufficient to abolish organization-level learning-by-doing in favor of industry-level learning. For newer, more complex types of transplants (e.g., intestinal and lung transplants), insufficient knowledge exists regarding the optimal approach to these types of transplants, leaving individual new entrants to learn on their own how to improve their outcomes. Similarly, the complexity of immunosuppression relative to surgery means that longer-term outcomes that depend on immunosuppression will be more subject to organization-level learning through trial and error rather than through the importation of best practices (Nembhard et al. 2014).

In summary, industry-level learning-by-doing may completely dominate organization-level learning-by-doing if knowledge is freely shared across organizations. However, the knowledge of how to successfully perform a procedure must exist. Otherwise, the new entrant will still be forced to learn-by-doing, assessing the feedback received from each successive learning opportunity and adapting its approach to subsequent transplants, as described by Jovanovic and Nyarko (1995).

If organizational learning-by-doing from entry does not exist in modern liver transplantation (or other types of organ transplantation), this has policy implications. Most importantly, it brings into question insurer use of experience or volume requirements as a proxy for quality and the negative effects such policies may have for access to transplantation. Transplant centers must perform a certain number of transplants before an insurer will begin providing coverage for transplants at that center. The center must then maintain that annual number of transplants in order to continue to be included in the insurer's network. Within liver transplantation, these vary from eight (Blue Cross Blue Shield)[22] to thirty transplants (Aetna),[23] bringing further into question the correlation of these volume thresholds with quality.

As recently as 2013, 38% of centers performed fewer than 30 transplants in that year, suggesting that at least some patients already face significant restrictions on their ability to access transplantation. Because of the necessity of geographic proximity due to the short organ preservation time (up to 12 h), patients cannot easily switch to a new transplant center if their current center closes down. In addition, transplant centers are already scarce in many parts of the country and limiting patients to only more experienced centers would further restrict access. During the sample period, no liver transplants were ever performed in twelve states.[24] Hearts and lungs have significantly shorter preservation times (4–6 h)[25] and yet are even less available nationwide. Clearly, using minimum experience requirements risks decreasing access to transplantation. It may similarly reduce the ability of individuals to donate their organs if that person dies too far away from a transplant hospital.

On a positive note, the absence of organizational learning-by-doing and existence of high survival rates from entry suggest that in liver transplantation the specialized training programs are effective, knowledge flows readily from research to practice, teams communicate effectively, and newly opened centers are able to attract competent personnel. As a complex medical procedure performed almost exclusively at larger teaching hospitals, the effectiveness and sophistication of its knowledge transfer could serve as a model for other fields. The results also suggest the need to study precisely what it is about teams in liver transplantation that makes them so effective.

Conclusion

Liver transplant centers learn by doing, but only shortly after entry. The importance of learning-by-doing varies among centers with some centers starting at a low survival rate and improving, while others enter with high survival rates. The effect of learning-by-doing increases in contexts with greater technical uncertainty, as measured by differences in the timing of entry, across survival durations, and across types of organ transplants.

These results highlight the importance of focusing on learning-by-doing shortly after entry rather than seeking to identify a constant return to experience, regardless of how experienced an organization is when it first appears in the data. The pattern of early learning-by-doing I identify follows what has been found in many studies in manufacturing, indicating that organizational learning-by-doing is not a phenomenon unique to reducing unit costs or production defects. Organizational learning-by-doing seems to operate quite similarly in complex industries with team-based production processes, such as organ transplants, shipbuilding (Thornton and Thompson 2001), semi-conductors (Irwin and Klenow 1994), air craft (Benkard 2000), auto manufacturing (Levitt et al. 2013), and in healthcare, for minimally invasive cardiac surgery (Pisano et al. 2001) despite vastly different information dissemination mechanisms from proprietary manufacturing practices to medical research published in a highly competitive academic environment. The heterogeneity across centers supports

22 https://www.bcbs.com/sites/default/files/file-attachments/page/Transplants.Selection_0.pdf.

23 http://www.aetna.com/healthcare-professionals/documents-forms/aetna-institutes-transplant-criteria.pdf. Accessed 04/13/2017.

24 Alaska, Idaho, Maine, Montana, Nevada, New Hampshire, North Dakota, Rhode Island, South Dakota, Vermont, West Virginia and Wyoming.

25 https://optn.transplant.hrsa.gov/learn/about-transplantation/how-organ-allocation-works/. Accessed 07/13/2017.

the limited evidence in the literature that the importance of learning-by-doing can vary substantially across firms. The fact that early learning-by-doing dissipates over time suggests that technology and policy shocks and general industry-level learning may diminish the importance of individual centers' learning-by-doing.

What we now know about organizational learning-by-doing suggests a transient and highly heterogeneous process that relates directly to the amount of non-proprietary knowledge available to new entrants into an industry. In other words, once the body of knowledge commonly available in an industry reaches a certain level, learning-by-doing at the firm-level seems to disappear as best practices no longer require time and effort to acquire. From a health policy perspective, identifying the critical point at which sufficient common knowledge of best practices obviates the need for organization-level learning could save lives if it increases access to complex medical procedures by removing unnecessary insurer experience requirements.

References

American Hospital Association. (2009). *AHA annual survey database (1987–2009)*. American Hospital Association. AHA data on liver transplant centers was obtained through the National Bureau of Economic Research (NBER).

Arrow, K. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, *29*(3), 155–173.

Axelrod, D., Guidinger, M., McCullough, K., Leichtman, A., Punch, J., & Merion, R. (2004). Association of center volume with outcome after liver and kidney transplantation. *American Journal of Transplantation*, *4*, 920–927.

Balasubramanian, N., & Lieberman, M. (2010). Industry learning environments and the heterogeneity of firm performance. *Strategic Management Journal*, *31*, 390–412.

Barone, M., Avolio, A., Di Leo, A., Burra, P., & Francavilla, A. (2008). ABO blood group-related waiting list disparities in liver transplant candidates: Effect of MELD adoption. *Transplantation*, *85*(6), 844–849.

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

Centers for Medicare and Medicaid Services (2006). *National coverage determination (NCD) for adult liver transplantation (260.1)*. Publication number 100-3, version 2, effective 06/19/2006–06/21/2012.

Centers for Medicare and Medicaid Services (2007). Medicare program; hospital conditions for participation: Requirements for approval and re-approval of transplant centers to perform transplants; Final rule. Title 42 code of federal regulations, parts 405, 482, 488 and 498. *Federal Register*, *72*(61), 15201.

Contreras, J. M., Kim, B., & Tristao, I. M. (2011). Does doctor's experience matter in LASIK surgeries? *Health Economics*, *20*, 699–722.

David, G., & Brachet, T. (2009). Retention, learning by doing, and performance in emergency medical services. *Health Services Research*, *44*(3), 902–925.

David, G., & Brachet, T. (2011). On the determinants of organizational forgetting. *American Economic Journal: Microeconomics*, *3*(3), 100–123.

Department of Health and Human Services, Health Care Financing Administration (1998). Medicare and medicaid programs; hospital conditions of participation; identification of potential organ, tissue, and eye donors and transplant hospitals' provision of transplant-related data, final rule. June 22, 1998. 42 CFR Part 482. FR Doc No: 98-16490, 33856-33875. http://www.gpo.gov/fdsys/pkg/FR-1998-06-22/html/98-16490.htm.

Edwards, E., Roberts, J. P., McBride, M., Schulak, J., & Hunsicker, L. (1999). The effect of the volume of procedures at transplantation centers on mortality after liver transplantation. *The New England Journal of Medicine*, *341*(27), 2049–2053.

Evans, R. W. (1991). *Executive summary: The national cooperative transplantation study*. Seattle, WA: Health and Population Research Center at the Battelle-Seattle Research Center.

Freeman, R. B, Jr., & Edwards, E. B. (2000). Liver transplant waiting time does not correlate with waiting list mortality: Implications for liver allocation policy. *Liver Transplantation*, *6*(5), 543–552.

Freeman, R. B., Steffick, D. E., Guidinger, M. K., Farmer, D. G., Berg, C. L., & Merion, R. M. (2008). Liver and intestine transplantation in the United States, 1997–2006. *American Journal of Transplantation*, *8*(4), 958–976.

Gaynor, M., Seider, H., & Vogt, W. B. (2005). The volume-outcome effect, scale economies, and learning-by-doing. *American Economic Review Articles and Proceedings*, *95*(2), 243–247.
Guarino, C., Maxfield, M., Reckase, M. D., Thompson, P., & Wooldridge, J. M. (2015). The evaluation of empirical Bayes' estimation of value-added teacher performance measures. *Journal of Educational and Behavioral Statistics*, *40*(2), 190–222.
Ho, V. (2002). Learning and the evolution of medical technologies: The diffusion of coronary angioplasty. *Journal of Health Economics*, *21*, 873–885.
Huckman, R. S., & Pisano, G. P. (2006). The firm specificity of individual performance: Evidence from cardiac surgery. *Management Science*, *52*(4), 473–488.
Huesch, M. D. (2009). Learning by doing, scale effects, or neither? Cardiac surgeons after residency. *Health Services Research*, *44*(6), 1960–1982.
Irwin, D. A., & Klenow, P. J. (1994). Learning-by-doing spillovers in the semiconductor industry. *The Journal of Political Economy*, *102*(6), 1200–1227.
Jacob, B. A., & Lefgren, L. (2005). Principals as agents: Subjective performance measurement in education. In *NBER working article no. 11463*.
Jovanovic, B., & Nyarko, Y. (1995). A Bayesian learning model fitted to a variety of empirical learning curves. *Brookings Articles on Economic Activity. Microeconomics*, *1995*, 247–305.
Kane, T. J., & Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. In *NBER working article no. 14607*.
Levenson, J. L., & Olbrisch, M. E. (1993). Psychosocial evaluation of organ transplant candidates. A comparative survey of process, criteria, and outcomes in heart, liver, and kidney procedures. *Psychosomatics, 34*(4), 314–323.
Levitt, S., List, J., & Syverson, C. (2013). Toward an understanding of learning by doing: Evidence from an automobile plant. *Journal of Political Economy*, *121*(4), 643–681.
Manthous, C., Nembhard, I. M., & Hollingshead, A. B. (2011). Building effective critical care teams. *Critical Care Medicine*. doi:10.1186/cc10255.
Nembhard, I., Cherian, P., & Bradley, E. H. (2014). Deliberate learning in health care: The effect of importing best practices and creative problem solving on hospital performance improvement. *Medical Care Research and Review*, *71*(5), 450–471.
Pisano, G. P., Bohmer, R., & Edmondson, A. (2001). Organizational differences in rates of learning: Evidence from the adoption of minimally invasive cardiac surgery. *Management Science*, *47*(6), 752–773.
Stith, S., & Hirth, R. (2016). The effect of performance standards on healthcare provider behavior: Evidence from kidney transplantation. *Journal of Economics and Management Strategy*, *25*(4), 789–825.
Terasaki, P. I. (1986). *Clinical transplants 1986*. Los Angeles, CA: UCLA Tissue Typing Laboratory.
The Organ Procurement and Transplantation Network. (2015). Policy 9: Allocation of livers and liver-intestines. In *Organ procurement and transplantation policies* (pp. 93–122). http://optn.transplant.hrsa.gov/resources/by-organ/liver-intestine/. Accessed 25 Oct 2015.
The Organ Procurement and Transplantation Network and United Network for Organ Sharing (2008). OPTN/UNOS living donor committee: Report to the board of directors, November 17–18, 2008 (pp. 3–18).
Thornton, R., & Thompson, P. (2001). Learning from experience and learning from others: An exploration of learning and spillovers in wartime shipbuilding. *American Economic Review*, *91*(5), 1350–1368.
United Network for Organ Sharing. (2010). Standard Transplant Analysis Research (STAR) data files (1987–2009). In *United network for organ sharing*. www.unos.org.
Wright, T. P. (1936). Factors affecting the cost of airplanes. *Journal of Aeronautical Sciences*, *3*(4), 122–128.