The Impacts of Task Repetition and Temporal Breaks in Production on Human Capital and Productivity

Jason Hockenberry
University of Iowa

Hsien-Ming Lien
National Cheng Chi University

Shin-Yi Chou
Lehigh University and National Bureau of Economic Research

The productivity of firms has been shown to decline after breaks in production. The literature suggests that one cause of this decline is the depreciation of human capital of individual workers. We examine the productivity of cardiac surgeons and hospitals to determine whether the length of or level of task repetition prior to breaks in production affects the productivity of individuals. We find decreases in surgeon productivity that are directly attributable to the length of the break between surgeries, with little evidence that this effect is mitigated by the level of task repetition prior to the break.

I. Introduction

Most economists would agree that a break in the production of a firm should have a deleterious impact on the productivity of the firm, in part because the individual employees will forget some of the details of the tasks that are involved in their work. One would expect the extent of the impact of a break in production to be proportional to the experience of the firm's workers prior to the break in production and to the length of the break. While this expectation seems logical and testable, the empirical literature examining the direct impact of the experience of the individual workers and the length of the break in production on firm productivity is thin. Previous work by Mincer and Ofek (1982)

Hockenberry and Chou thank the Martindale Center at Lehigh University for financial support. Lien is grateful to the National Health Research Institutes for financial support (NHRI-EX96-9204PP) and for providing the data and to the National Science Council for financial support (NSC 95-2627-H-004-001-MY2). We would like to thank seminar participants at Lehigh University, University of Iowa, the College of New Jersey, St. Cloud State University, and American Society of Health Economists 2008 meetings and Stephen Snyder and Todd Watkins for insightful comments. We also thank two anonymous referees and the editors Michael Grossman and Isaac Ehrlich for their comments and suggestions.

[*Journal of Human Capital,* 2008, vol. 2, no. 3]

investigated the relationship between the unemployment stints of workers and the depreciation of their human capital.[1] However, this line of study was directed more at the impact of breaks in productive activity on the earning power of the individual worker and did not measure productivity directly. The authors found that workers earned lower real wages upon reentry into the workforce, and the assumption made was that their skills had diminished or had become somewhat obsolete, and therefore, by definition, their human capital had depreciated.

Another line of related literature involves the investigation of organizational forgetting and its impact on firm productivity. The idea of organizational forgetting is partially based on breaks in production leading to the human capital depreciation of individual workers, but to this point this line of literature has not explicitly addressed the extent of individuals' human capital depreciation or its impact on productivity.

In this paper we investigate the impact of human capital depreciation on productivity. We hypothesize that there are two components to how quickly human capital depreciates. The primary component in our hypothesis is that temporal distance (defined as the amount of time elapsed) between the execution of production tasks has a direct impact on productivity. This incorporates both the ideas of human capital depreciation and of forgetting that were mentioned above. The previous literature has assumed that there is a temporal component to both of these phenomena; however, the empirical literature is surprisingly sparse on the temporal nature of individual human capital and the impact of the temporal distance in production on the retention of this human capital.

The second component of human capital depreciation that we study is the task repetition effect. This concept is loosely based on the idea of a learning curve or learning by doing because it implies that workers (and organizations) gain some knowledge or skill (human capital) through increased experience in production tasks (repetition). This increased capital in turn leads to lower costs or, alternatively stated, higher productivity.

The impact on human capital of an increase in temporal distance in production (what may alternatively be referred to as a temporal break in production) may be mitigated by the level of repetition of a task before the temporal break in production. In plain terms, the amount of practice one has had leading up to the break in production may affect how far productivity falls upon one's return to production tasks. By performing a task more often, workers are able to maintain their skills and, therefore, a high level of productivity. If task repetition helps maintain individual workers' skills, then longer breaks in production

[1] Their paper was actually one of many in an ongoing discussion in the literature that seems to have begun with Ben-Porath (1967) and included, but was not limited to, Mincer and Polachek (1974, 1978) and Sandell and Shapiro (1978).

should lead to subsequent periods of lower productivity as workers return to production. In this paper, we try to capture whether the temporal distance and the level of task repetition prior to the temporal breaks affect productivity and, if they do, to also capture the extent of this effect.

In this study, we will examine the direct impact of temporal distance between production tasks and task repetition effects on the productivity of cardiac surgeons performing serious and complex procedures. This particular work environment is particularly suited to examining the question of the impact of temporal distance in production and task repetition on productivity, for a few reasons. It is better suited to examining the nature of human capital depreciation than the early work on unemployment spells because the surgeons are generally all engaged in the same surgical tasks, which may not be true of workers, even if they are employed in the same industry or even at the same plant. Examining cardiac surgery also has an advantage over the unemployment literature because we can measure productivity directly; thus there is no need to infer the reduction in productivity through lower wages, as is the case in the unemployment literature.

Another advantage is that production schedules in a surgical environment are not always regular. This means that surgeons do not arrive at work every day and employ the human capital that is being maintained or has been accumulated through the task repetition of the surgical procedure in which they are specialized. Compare this environment with manufacturing production lines that have fairly regular production schedules in which workers are regularly employed in the tasks for which they are specialized. It becomes clear that the surgical environment provides natural breaks in production and levels of task repetition that occur more often and vary more in length than in other industries. Thus, the cardiac surgical environment can be used to investigate and distinguish between the impact of the temporal distance in production and task repetition on human capital and, thus, productivity.

Our study can also add to the discussion on the impacts of human capital retention and depreciation on organizational productivity. By using the cardiac surgical environment to investigate these relationships, we have the ability to examine the impact of the experience of the individual worker and organization simultaneously and in detail.[2] The data we use in this study include detailed information for all surgeons and hospitals in Taiwan involved in the performance of two serious and relatively common cardiac surgical procedures. This data set is a quite

[2] The experience of the hospital in this case is a proxy for the human capital stock of the hospital as an organization. The relevant human capital in this type of study is embedded in both the surgeons who work at a hospital and the experience of the hospital support staff in caring for patients who have undergone specific procedures. By looking at the volume of the surgeon and the volume of the hospital simultaneously, we can distinguish between the effects of the two.

detailed and relatively large panel, so we are able to control for unobserved heterogeneity arising from both hospitals and surgeons. We find evidence that temporal distance in production affects productivity, which implies that human capital depreciates over time. We also find evidence that this depreciation and its impact on productivity are experienced despite the level of task repetition prior to the breaks. This study also suggests that this temporal nature of human capital has a larger impact on the productivity of a hospital organization than does the overall experience of the hospital organization at large. This contributes to the discussion on organizational forgetting by suggesting that organizational forgetting, for which evidence exists in other industries, may be due in part to the human capital of individuals depreciating during breaks in production.[3] Plainly stated, it appears that individual surgeons do forget and do so relatively quickly, and this appears to have an impact on productivity.

II. Background and Literature

In this paper, we seek to add to the literature on human capital and forgetting in a couple of different ways. Our means of investigating the human capital maintenance and depreciation of both individuals and organizations is by focusing on production in the cardiac surgical environment. The human capital in the case of surgery is embedded in the skills and abilities of the surgeons and hospital staff who care for the patients during and after surgery.[4] The fundamental ideas of human capital maintenance and depreciation are grounded in the learning curve literature. Thus, we first draw parallels between the cardiac surgery environment and the study of manufacturing learning curves begun by Wright (1936).

The theory of the learning curve is based on the premise that an organization benefits from increases in productivity as the workers in the firm gain experience in producing the good—that is, as they increase their human capital. Thus, the parallel we draw is that hospitals and/or surgeons should be more productive as they gain experience in performing a specific surgery. There may be some concerns with drawing this parallel. Luft, Hunt, and Maerki (1987) point out that the studies of the volume-outcome relationship do not actually capture the learning curve of hospitals or surgeons, or, at the very least, what they estimate

[3] Organizational forgetting is the phenomenon in which firms experience "jumps" (temporary reductions in productivity) in the learning curve as they increase their experience in production. The cause of this temporary reduction in productivity is thought to arise, at least in part, from breaks in production.

[4] By this reasoning, the human capital of the organization is the aggregation of the individuals' human capital. The exact nature of this aggregation is not important to this study.

is not directly comparable to what economists term a learning curve.[5] They claim the two may not be directly comparable because the learning curve studies in manufacturing examine the relative improvements in productivity (as measured by labor hours needed to produce a product) from the first unit produced onward. They point out that this is not exactly how volume-outcome researchers capture the relationship between hospital and/or surgeon volume and patient outcome because the data that volume-outcome researchers examine are usually collected long after a procedure has been introduced at a hospital and/or when a surgeon has already begun his career.

While this issue does present some interesting points of difference, it is still not difficult to draw a parallel between the two as long as some reasonable assumptions are made. Because we are using similar data as those used in the volume-outcome literature, we must assume that at some point the hospitals and surgeons did experience learning when a procedure was introduced at that particular hospital or when the surgeon began his career. Thus, the point at which we are measuring task repetition is near the area of the learning curve at which the surgeons are approaching the asymptotic minimum of their learning curves. If we assume this, then using recent surgical volume (of a procedure that is well established) as a measure of ''units produced'' will just show the behavior of this curve around its minimum.

A second reason these two relationships are not exactly the same is that most volume-outcome studies examine improvements in surgeon or hospital performance as measured by decreases in mortality associated with increased experience, whereas the traditional learning curve studies measure productivity gains through the reduced man hours it takes to produce each unit of a product. Therefore, the next thing we consider in this context is the definition of productivity. Generally, we do not believe that the productivity of a surgeon is well measured by the number of patients he or she operates on in a day. Instead, we define surgeon productivity to be the additional life preserved, so that a surgeon who extends lives (or at least does not cause them to be shorter) is more productive.[6] Hence, surgeons or hospitals with lower rates of mortality will have higher levels of productivity.

Even with these differences it is still quite easy to draw a nice parallel

[5] The volume-outcome literature is a long and well-developed one that investigates the nature of the procedural/patient volume of health-care providers (hospitals and/or doctors) and the subsequent health outcomes of their patients. Halm, Lee, and Chassin (2002) summarize this literature and show that 70 percent of the studies find a positive relationship between increased volume and better patient outcomes.

[6] The use of the phrase "extends lives" may seem unusual here, but the colloquial use of the phrase "surgeons save lives" is not quite specific enough for what we need to draw a parallel. What surgeons actually do by performing critical surgery is extend life or improve its quality. The type of surgeons we look at actually extend life. It is also important to note that poor performance in the operating room could shorten life, which in our terms is reduced productivity.

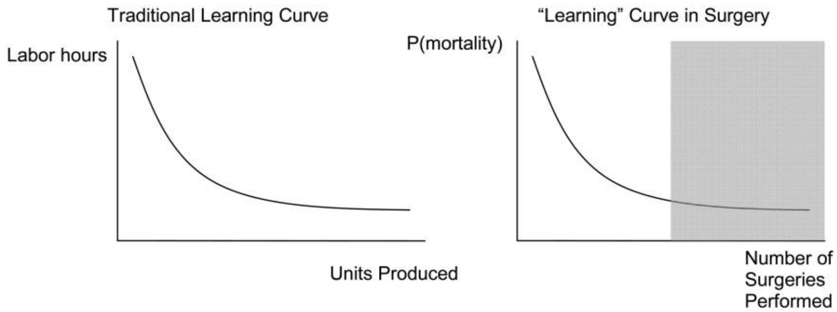

Figure 1.—Traditional learning curve and learning curve in surgery

between our study and the earlier learning curve literature, as long as we are clear on exactly what we are capturing. For the sake of illustration we have included a theoretical graphical comparison of the learning curve we believe occurs for surgery and the general shape of the graphs found in the learning curve literature related to earlier studies in manufacturing (fig. 1). As we stated before, what we are capturing is still related to a learning curve, albeit at a point further along the curve, and is assumed to be somewhere in the shaded region. In addition, the measure of productivity on the *Y*-axis is a little different, since increased productivity in surgery implies reduced probability of mortality within a defined postoperative period for patients, whereas in the previous learning curve studies increased productivity implied reduced man hours per unit of production. Thus, in our study, human capital retention associated with task repetition has a relatively larger influence on patient outcomes than does human capital accumulation through the increased experience provided by an additional surgery. In addressing the task repetition effect, we believe the temporal component of human capital depreciation is an important factor to consider.

Mincer and Ofek (1982) showed that when workers went through labor force withdrawal and subsequently reentered the workforce, their earnings were relatively lower than when they exited. This relationship is founded on the principle that during breaks from work due to spells of unemployment (or being out of the labor force) the workers' skills decay or become obsolete, and so the human capital accumulated during previous employment depreciates; hence we arrive at the idea of human capital depreciation. Thus, the idea of human capital depreciation, while referring to the value of human capital, is based on the failure to retain skills or knowledge previously accumulated through education or experience.

Subsequent to this work, Globerson, Levin, and Shtub (1989) demonstrated in an experimental setting the temporal impact of human capital depreciation on a repetitive data-entry task. The design of the experiment was to start a group of workers on a data-entry task and then, after a given period of time, give a break of differing time lengths to each worker. They then measured the change in the productivity of each worker when he or she returned from the break periods, and they examined the impact of the differing break lengths. Even in the case of this repetitive task and relatively short break periods, the time since the task was last performed was found to affect productivity. This finding is also robust to controlling for the previous experience of each individual, indicating that the differences in postbreak productivity may not be due to how much repetition of the task occurred previous to the break. To us, this finding of temporal distance effects on productivity suggests that the task repetition effect may be outweighed by the temporal component of human capital depreciation. Our question then is, If temporal distance effects can impact the performance of a repetitive task such as data entry, how much more will they impact a more complex set of tasks such as those involved in cardiac surgery?

A second line of literature that is relevant and to which our study makes a contribution is the literature on organizational forgetting. Studies of learning curves and productivity in manufacturing suggest that organizations do indeed experience depreciation of human capital, or what has been aptly termed “forgetting” (Argote, Beckman, and Epple 1990; Argote and Epple 1990; Benkard 2000). This forgetting, while dubbed “organizational,” is actually partially predicated upon the forgetting of individual workers. The theory is that as firms move along the learning curve, breaks in production or changes in the nature of the production tasks cause temporary reductions in productivity. We are interested in the first explanation, breaks in production, as the source of forgetting because this explanation contains a temporal element and is based on individual human capital depreciation.[7] Because surgeons and hospitals perform both planned and unplanned procedures, they are not always on a nice, neat production schedule, as would generally be the case in manufacturing. Thus, the ‘‘production schedule’’ in surgery is irregular and provides significant natural variation in the time since last surgery and an opportunity to investigate the effects of temporal distance and task repetition on human capital retention through their impact on productivity. Again, we find evidence that the human capital depreciation of the individual surgeon is affected by temporal distance and that this temporal component, compared with task repetition, has the relatively larger impact on patient mortality.

To ensure that our findings and implications are robust, it will also be necessary to account for hospital and surgeon heterogeneity, much

[7] Argote et al. (1990), Darr, Argote, and Epple (1995), and Benkard (2000) allude to the fact that the investigation of individual forgetting is important and interesting, but they do not have the data available to them to execute such an investigation. While our productivity measure and industry in this study are not identical to theirs, the individual detail of our results are at least a comment on the relative impact of the individual forgetting on some measure of productivity.

of which may be unobserved to the researcher. Patient preference may lead to selection issues depending on the type of procedures being undertaken (elective/necessary), the nature of the condition being treated (chronic/acute), and even the source of admission (emergency/nonemergency). Even if we were to assume general preferences for patients (or ignore them all together), an additional difficulty arises from the fact that neither the patients nor the procedures being performed are homogenous in any sense (Tsai et al. 2006). Obviously, patient age at the time of surgery varies, as does each patient's relative health, number of confounding conditions, and so on. This gives rise to potential sorting of patients to providers with different volume levels (and hence different levels of task repetition) based on relative health status. Therefore, we account for the unobserved heterogeneity of providers through the use of fixed effects estimation, and we include variables to account for the relative health status of each patient.

III. Data

The data for this study come from Taiwan, and so a description of Taiwan's health-care system is warranted. Taiwan switched to a universal health insurance system in March 1995. The reimbursement system under the National Health Insurance (NHI) program is generous, and the uptake of insurance is quite high, with more than 95 percent of the population being enrolled during any given year across our period of interest. From the patient perspective, the contribution to insurance is relatively minimal and uniform when compared to that in the United States, so the choice of provider is based on the willingness of patients to travel and the information available to them about the quality of each provider. Thus, patient preference is not complicated by differing explicit prices.

After the NHI was instituted, hospitals and clinics wishing to be reimbursed needed to have a contract with NHI, but the hospitals themselves did not become state owned. The labor market for surgeons in Taiwan still operates as a private market, with better surgeons able to negotiate higher salaries and/or bonuses. However, unlike some cases in the United States, surgeons in Taiwan contract to work as employees of the hospital and therefore practice at only one hospital instead of seeking admitting privileges at multiple hospitals. This is convenient because we do not need to concern ourselves with individual differences in surgeon performance stemming from firm-specific performance.[8]

[8] Huckman and Pisano (2006) study firm-specific performance among surgeons in the United States, where doctors often have admitting privileges at several hospitals and will perform operations at multiple locations. With respect to forgetting, this adds another layer of difficulty because one cannot be sure whether the source of differences in surgeon performance is tied to forgetting or to differences in hospitals. Since surgeons in Taiwan

The data for this study were gathered from four different sources in the National Health Insurance Data (NHID), which is maintained by the National Health Research Institutes. We were able to first collect the longitudinal medical claims of patients who underwent the procedures we are interested in between 1998 and 2004. Since nearly the entire population of Taiwan is covered by NHI (varies between 95 percent and 97 percent over the period), we essentially have every one of the procedures of interest performed in Taiwan. The claims records include disease diagnoses, admission dates, discharge dates, and detailed descriptions of medical expenses broken down by category (i.e., room, food, surgery, etc.). This information allows us to create each hospital's and surgeon's volume variables and provides us with the information necessary to create the indicator for the procedure performed and the patient clinical characteristics used in our estimation.[9] An important feature of this data set is that each patient, surgeon, and hospital has a unique identification code that enables one to link these data to other data from NHID.

We are measuring time since last surgery in days, and since we do not have the actual procedure dates, we simply use admission dates as the surgery date.[10] There are a couple of reasons why this is a valid imputation for the surgical procedures in our study. In the case of planned cardiac procedures, patients are almost always admitted on the morning of the day the surgery is to be performed, or at worse the day before. In the case of emergency procedures due to heart attacks or other serious cardiac symptoms, the procedure is performed as soon as possible to prevent further attacks or symptoms, so the patient will not be in the hospital waiting for a procedure for many days. It is likely that patients who are severely ill—for example, those admitted while having a severe acute myocardial infarction (AMI)—may be in an unstable health state and therefore need to be stabilized before intervention and thus would have the procedure on the day or two following admission. If this were the case, then our assumption that procedure date is the same as admission date would lead to the number of days since a surgeon last performed the surgery being measured with some error throughout the data. In performing our empirical estimation, we did check for whether the assumption that the procedure date was the same as the admission date biased our results, by assigning the patients admitted with AMI a procedure date of admission date plus 1 day (and, in a

are employed at a single hospital, variation in surgeon performance related to changes in recent volume can be attributed to forgetting.

9. These characteristics are listed in tables 2 and 3.

10. To deal with the observations where it is the surgeon's first time performing the surgery, we exclude from our estimation any observations where the surgeon appears in our data for the first time. We also exclude any observations where the surgeon has not performed the surgery in the past 1,000 days, since a break like this is indicative of someone who may have left the profession and returned or changed specialties.

second case, admission date plus 2 days), and it did not affect our results.[11] A discussion of the limitations of our findings related to this measurement error is included in Section VII.

The longitudinal medical claims are then linked to data sets containing NHID health provider information and to patient eligibility files, both of which span 1996–2004. From the provider data sets we extract hospital location, ownership status, and accreditation, the number of beds, and the ages of the surgeons. We use patient eligibility files to extract patient sex, age, and residence at the time of enrollment. The data in this eligibility file are kept in log format so that we are able to track changes in address, enrolled unit, salary, or type. Finally, since NHI is compulsory, we use disenrollment dates to infer the date of death. Since death is the most likely cause of disenrollment for patients who would be undergoing these procedures, these dates are valid for calculating death dates.[12]

Our study focuses on two relatively skill-intensive cardiac procedures, percutaneous transluminal coronary angioplasty (PTCA) and coronary artery bypass graft surgery (CABG). Our sample includes an observation for each time a patient underwent one of these surgical procedures, which are intended to help treat those with symptoms and complications stemming from atherosclerosis of the coronary artery, including angina and AMI.[13] Compared with CABG, PTCA is relatively less invasive, generally requires less time on average to complete, and does not require a patient to be put on a heart-lung machine. Because of these advantages, PTCA is often the first option for treating a blocked coronary artery. Possible complications from either procedure include blood clots, AMI, tearing of the arterial wall, increased stress on the renal system, internal bleeding, and heart arrhythmias, all of which can lead to patient mortality.

Because both procedures are relatively serious, surgery and postsurgical recovery may be more complicated for those patients with specific comorbidities or for older patients. Because the claims data from the NHI include diagnoses codes, we are able to include the patient Charl-

[11] The day on which emergency cardiac patients are admitted has been found to affect their health outcomes, with those admitted on the weekends generally experiencing worse outcomes. These worse outcomes are driven by the fact that these patients are less likely to be treated with a cardiac catheter, percutaneous transluminal coronary angioplasty (PTCA), or coronary artery bypass graft (CABG; Kostis et al. 2007). Thus, we also checked for the presence of effects from the day of the week by inserting a dummy variable for a weekend day of admission, and it was not significant and did not affect the estimation, so it was excluded.

[12] Lien, Chou, and Liu (2004) compare the ending date of coverage with actual death records of stroke patients. For those who died within 1 year after discharge, 90 percent have no differences between these two dates, less than 5 percent have differences larger than a week, and less than 2 percent have differences larger than a month. For further discussion of the data, see Lien, Chou, and Liu (2008).

[13] Atherosclerosis is a hardening of the artery wall that can lead to plaque buildup. Since the coronary artery supplies blood to the heart, this condition is quite serious. This buildup restricts blood flow and can cause chest pain (angina) and may lead to heart attack (AMI).

son indices to account for these confounding illnesses.[14] In addition, access to longitudinal claims allows us to use the medical reimbursements for each patient in the 12 months leading up to the procedure to control for previous intensity of treatment, previous intervention, and relative health (to some degree). The combination of the Charlson indices and the patient age, along with recent medical spending and whether or not the patient has undergone recent surgical intervention for the same condition, provide controls for relative patient risk.[15]

IV. Model and Methodology

The basic theoretical model of the relationship we use in studying surgical productivity is that the probability of a patient experiencing mortality, m, within a defined time period after an operation is a function of the surgical provider's experience in the given procedure, measured by recent volume of the procedure, V, and the characteristics of the patient, X, that may affect his or her relative health. Thus the theoretical equation is stated as

$$P(m) = f(V, X, \varepsilon). \qquad (1)$$

Again, this model parallels the learning curve models in manufacturing, in which the number of labor hours required to produce a unit (the measure of productivity) is a function of the number of units produced by the firm prior to the current unit (the measure of experience). However, this equation is not quite the complete characterization we would like to estimate. To our knowledge, the learning curve studies in manufacturing do not have the ability to examine each individual worker's productivity; they simply measure the productivity of the firm (or production line) based on the experience of that firm (or line). We, however, can separate the individual surgeon's experience from the rest of the hospital staff and can therefore examine whether one person's experience contributes more to the overall productivity in the case of surgery. Thus, V will include both the volume of the surgeon and that of the hospital as separate measures.

The learning curve literature also suggests that temporal distance in production may cause temporary reductions in the level of productivity.[16] Again, this argument is based on the depreciation of human capital occurring over time, often referred to as forgetting. Within this framework, we are interested in the impact of the task repetition before the

[14] Charlson indices are dummy variables equal to one if a patient has specific comorbidity and to zero if they do not.

[15] We define recent intervention as having had a PTCA or CABG within the past year. We limit this definition to the past year because we do not have patients' full medical histories prior to the NHI.

[16] Argote et al. (1990) and Benkard (2000) find this occurring in aircraft manufacturing at Lockheed. Argote et al. also find this to be the case in the Liberty Shipbuilding yards.

break on the mortality outcome of the patients treated immediately after the break. We would also like to capture the relative impact of the depreciation in the surgeon's human capital over time on a patient's mortality outcome. Thus, we also include a measure of the number of days since the surgeon last performed the surgery, D, in our estimation equations.

We start with the linear probability model given by

$$m_{ijkrt} = \alpha + \beta_1 D_{ijt} + \beta_2 \ln V_{ijkt} + \beta_3 S_{jt} + \beta_4 H_{kt} + \beta_5 X_{it} + v_{ijkrt}, \qquad (2)$$

where m is an indicator equal to one if patient i (whose residence is in county r and who is seen by surgeon j at hospital k) dies within 1 month of admission and where D is the time elapsed (in days) since surgeon j last performed the procedure (either PTCA or CABG). This variable for time since last procedure is intended to capture the impact of the temporal depreciation of human capital of the surgeon on the probability of patient mortality. The term $\mathbf{V}$ is a vector containing both the log of surgeon volumes and the log of hospital volumes in the 12 months leading up to the procedure on patient i.[17] The volume variable in this case will capture the relative impact of the task repetition on patient mortality.

The other variables in this equation are meant to control for some of the heterogeneity of both the providers and the patient. The term $\mathbf{S}$ is a vector of surgeon characteristics of surgeon j, $\mathbf{H}$ is a vector of hospital characteristics of hospital k, and $\mathbf{X}$ is a vector of patient characteristics of patient i. We include surgeon and hospital observables to control characteristics of either, which may lead to selection issues or concerns of endogenous variables that would attenuate the validity of our estimation results. For surgeons, we include their age at the time of the operation on patient i and whether or not they had a patient die in the hospital in their last three procedures prior to the operation on patient i.[18] Hospital characteristics include dummies for the number of beds and ownership status. This is similar to the manufacturing studies controlling for observable characteristics of a firm or production line.

In theory, the error term v_{ijkrt} is not truly random, because it is possible that the observable characteristics included in the equation do not capture all the relevant heterogeneity of the surgeons or hospitals, which is likely to be endogenous to the volume of these providers. These

[17] We choose the log-functional form of volume because previous studies suggest that a nonlinear functional form is appropriate here (Cuellar, Lindrooth, and Cross 2006; Gowrisankaran, Ho, and Town 2006; Hockenberry, Lien, and Chou 2008). Logically, this is sensible because the marginal contribution to human capital of the second surgery in the past year is presumably much higher than the marginal contribution of the nth surgery ($n > 2$).

[18] The choice of the most recent three patients was arbitrary. We checked our results for alternate definitions of "recent" in hospital patient death to see whether our results were substantially different with the use of a number of patients other than three most recent, and this did not change our results.

unobservable characteristics of the hospital and surgeon—which may affect the time since the last surgery, recent volume, and mortality—may lead to bias in the estimation and incorrect conclusions and/or policy implications. In addition there may be factors that affect the mortality of patient i dependent upon the county of residence, r, and the time, t, at which they undergo the procedure. Therefore, v_{ijkrt} can be decomposed into

$$v_{ijkrt} = \tau_t + \xi_r + \delta_k + \eta_j + \rho_i + \varepsilon_{ijkrt}. \quad (3)$$

In our first estimation, we account for τ_t and ξ_r through the use of year and county fixed effects. We then proceed to add hospital fixed effects to account for that part of the error term that arises from the unobserved heterogeneity of hospital k, δ_k. Finally, we include surgeon fixed effects to account for that part of the error term arising from the unobserved heterogeneity of surgeon j, η_j. Because patient preferences are likely to be correlated with quality of the hospital k and surgeon j, and since the information they are able to obtain through informal networks will be tied to their area of residence, county r, we assume that the inclusion of the hospital, surgeon, and county fixed effects combined with the observed health status of patient i should account for the unobserved patient preferences.

If it is the surgeons' human capital that is the main driver of the level of productivity in surgery, we expect the experience of the hospital, as measured by recent volume, to be a statistically insignificant predictor of patient mortality. Also, if task repetition, as indicated by the volume of surgeries performed over the past year, does contribute to the human capital retention of the surgeon, we expect to find a negative and significant coefficient on the surgeon volume variable. If indeed human capital depreciates as the temporal distance between production tasks (surgeries) grows, we expect to see this temporal effect on productivity manifest as a positive and significant relationship between the number of days since surgeon j's last surgery and the probability of patient i experiencing mortality within 1 month.

Using just the absolute number of days since the last surgery may not provide the comparison we are looking for, so we divide D from equation (2) into a vector of dummy variables for quantiles divided by elapsed day. For PTCA, the five quantiles are 0–2, 3–7, 8–15, 16–21, and 22+ days since last surgery. For CABG, the three quantiles are 0–2, 3–14, and 15+ days.[19] The reasoning behind these breaks is based on the calendar week. The quantiles are narrower in terms of days for our PTCA estimation because the surgery is relatively more common than

[19] There are a couple of reasons for not using identical quantiles for both procedures. The main issue is that CABG is not performed nearly as often as PTCA. On average, surgeons have longer breaks between procedures for CABG, so the finer quantile measures used for PTCA lead to quantile groups for CABG that may have relatively few observations, and interpretation of the results is difficult.

CABG. Some surgeons regularly perform multiple PTCAs on the same day, but this is much more rare in the case of CABG. For CABG, the higher-volume surgeons may perform only two or three of these per week. However, in both cases we use the 0–2-day quantile as a reference group, because these will be the observations of surgeons who specialize in the given procedure and are in their normal routines with higher levels of task repetition. For PTCA, the 3–7-day group captures those surgeons who do not work on weekends and/or who may cluster their scheduled operations on specific days each week. For CABG, the nature of the procedure and the smaller sample size leads to broader quantiles in terms of days, so the 3–14-day group serves to capture the impacts of shorter breaks. The longer measures in both cases are meant to capture the breaks in production not associated with these other common causes of time between surgeries.[20]

The reasons for using this quantile strategy are related to the effects we are trying to capture. An absolute measure of days since last surgery may not be picking up the entire effect we are trying to capture because it only tells us the impact of temporal distance around the conditional mean. Dividing the observations into quantiles based on the days since the surgeon last performed the surgery helps capture any gradient that may exist in the temporal distance effects of human capital depreciation on productivity. With a single measure of days since last surgery, we are only capturing the effect at the mean, which may not yield clear statistical results and may inhibit policy suggestions. However, with day quantiles we can capture the relative impacts of different time breaks since the last procedure performed by the surgeon.

V. Results

We begin by showing the distributions of the number of days since the last surgery performed in figures 2*A* (PTCA) and 2*B* (CABG).[21] It is clear that a large percentage of these operations are performed by surgeons who have performed the surgery very recently. It is also quickly apparent that the distribution of elapsed days is highly skewed for both procedures. Inserting just the number of elapsed days may not capture

[20] We chose cardiac surgeons for this study because these surgeons tend to be very specialized. Admittedly, some of the surgeons in the sample are surgeons with very low volume, which suggests that the procedure is not their specialty, and we do not control for the other types of surgery they may have performed between the surgeries we are examining. We do not address this because we believe there could be conflicting effects related to performing other surgeries. From a theory standpoint, performing other procedures may have effects that reinforce the temporal distance effect (through cognitive interference) or mitigate the temporal distance effect (through maintenance of motor skills and mental acuity associated with other activities). Deciding which activities would have which effect would be rather arbitrary, so we did not explicitly address whether these surgeons are performing other procedures in the interim.

[21] We provide a version of each graph in which we have zoomed in on the graph in the area of the distribution that is of interest.

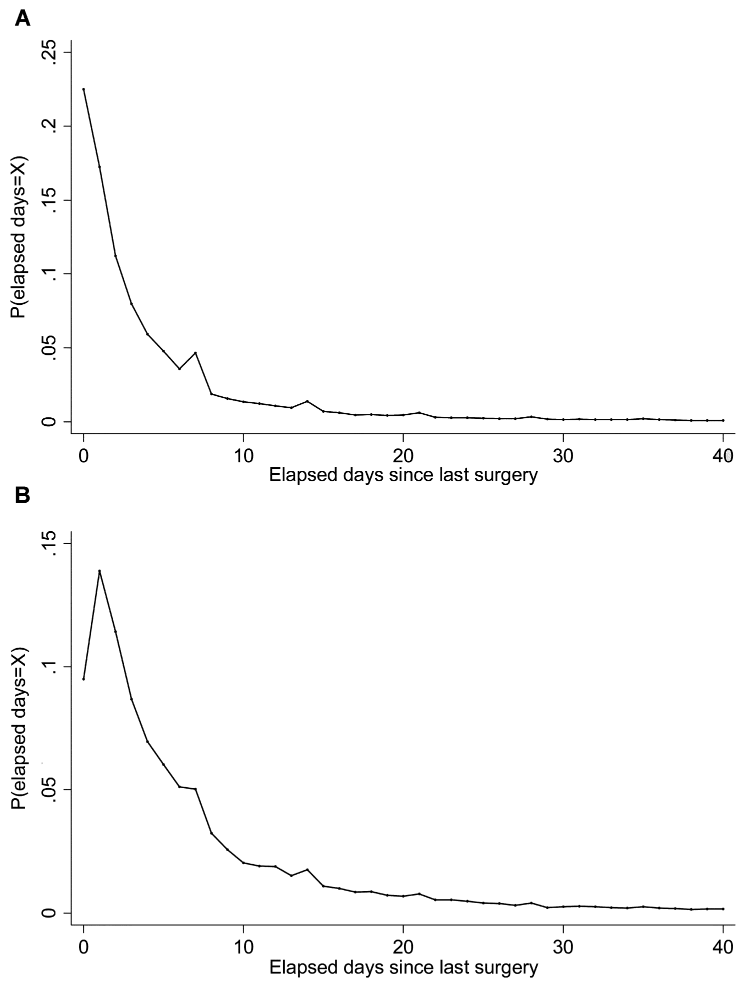

Figure 2.—Distribution of surgeons' temporal distance for PTCA (*A*) and CABG (*B*) procedures.

the temporal impact because the skewness of the data implies that the mean is a poor point to base analysis on, but this is exactly what the ordinary least squares estimation with elapsed days entered as a single variable would do. This is yet another motivation for why we perform estimations using elapsed-day quantile dummies.

Table 1 contains the descriptive characteristics by procedure for the control variables we include as regressors in our estimations, with the exception of the patient Charlson indices. The percentage of patients with each Charlson indicator is broken out by elapsed-day quantile in tables 2 and 3, and the final columns show the total percentage of all patients with the particular indicator. We include these to illustrate any

TABLE 1
DESCRIPTIVE STATISTICS

	PTCA		CABG	
	Mean	Standard Deviation	Mean	Standard Deviation
Provider characteristics:				
% with 250–600 hospital beds	42.2	49.4	31.4	46.4
% with 600+ hospital beds	54.7	49.8	68.0	46.6
% public hospitals	30.9	46.2	38.6	48.7
% nonprofit hospitals	58.3	49.3	56.9	49.5
Surgeon age at time of procedure (years)	42.7	6.0	44.6	8.6
% patients whose surgeon recently had a patient die in the hospital	2.8	16.4	8.7	28.2
Patient characteristics:				
% male	70.0	40.0	75.5	43.0
Age (years)	64.9	11.3	65.6	10.5
Medical expenditures in the past year ($)	4,391.9	3,057.6	13,560.0	9,171.0
% who had an intervention in the previous year	2.44	15.44	.02	1.39
Observations	107,747		20,798	

systematic difference in patient health between those who are treated by doctors who have performed the surgery recently and those who are treated by doctors who have not. The major issue of concern arising from these tables is that those patients suffering from AMI are slightly more likely to be treated by a surgeon who has not performed the procedure within a shorter period, but otherwise the relative health status of the patients is similar. We will address the potential of patient selection in Section VI.

In tables 4 and 5, we have listed descriptive statistics for both surgeon and hospital volumes broken down by elapsed-day quantiles. Since we are asserting that the temporal distance between production tasks effects are at least as important as task repetition effects, it is important to illustrate any systematic differences in volumes across the temporal space. Given that the time since the last surgery is going to be somewhat correlated with volumes, increased temporal distance will tend to indicate a lower surgeon volume but should have little or no impact on hospital volume. This is indeed what we observe in tables 4 and 5 when we look at both the mean and median of the volumes. More interesting, though, and important to our assertion of the importance of temporal effects, when we look at the upper end of the surgeon volume distribution and the maximum in each quantile, it appears that there are surgeons who would be considered higher-volume surgeons who occasionally have longer temporal breaks in their practice. Thus, when we perform estimations including both measures of time since last surgery and volumes, the time since last surgery is not simply a redundant

TABLE 2
RELATIVE PATIENT ILLNESS BY ELAPSED-DAY QUANTILE (PTCA)

	Number of Elapsed Days					Full Sample
	0–2	3–7	8–14	15–21	22+	
% of patients within the cohort with:						
Acute myocardial infarction	36.33	38.34	42.32	43.16	43.29	38.24
Peripheral vascular disease	1.72	1.83	1.90	1.88	2.13	1.81
Hypertension	35.61	37.13	34.18	33.28	31.86	35.51
Chronic obstructive pulmonary disease	7.12	7.17	7.64	8.07	7.78	7.27
Liver disease	1.34	1.45	1.84	1.84	1.81	1.47
Type I diabetes	29.85	28.95	28.27	27.53	28.16	29.23
Type II diabetes	10.29	10.04	11.31	12.16	13.90	10.66
Renal disease	4.25	4.31	4.56	4.74	7.48	4.56
Rheumatism	.33	.33	.32	.15	.54	.34
Cancer	1.54	1.48	1.55	1.59	2.04	1.56
Tumor	.20	.21	.17	.31	.48	.22
Dementia	.20	.24	.11	.31	.33	.22
Observations	55,687	29,455	10,333	4,116	8,156	107,747

TABLE 3
RELATIVE PATIENT ILLNESS BY ELAPSED-DAY QUANTILE (CABG)

	Number of Elapsed Days			Full Sample
	0–2	3–14	15+	
% of patients within the cohort with:				
Acute myocardial infarction	29.63	31.21	33.32	31.05
Peripheral vascular disease	4.39	4.53	5.34	4.63
Hypertension	36.68	36.29	32.41	35.71
Chronic obstructive pulmonary disease	8.90	8.57	8.80	8.73
Liver disease	1.56	1.28	1.73	1.46
Type I diabetes	36.00	35.27	33.83	35.26
Type II diabetes	13.79	13.87	16.94	14.41
Renal disease	5.21	5.63	6.31	5.61
Rheumatism	.34	.27	.36	.31
Cancer	1.39	1.77	1.63	1.61
Tumor	.13	.17	.23	.17
Dementia	.16	.24	.18	.20
Observations	7,423	9,975	3,400	20,798

TABLE 4
SUMMARY OF 12-MONTH VOLUME BY ELAPSED-DAY QUANTILE (PTCA)

	Number of Elapsed Days					Full Sample
	0–2	3–7	8–14	15–21	22+	
Surgeon volume:						
Mean	141.26	92.99	58.44	43.29	23.42	107.42
Standard deviation	118.00	72.49	47.19	35.78	25.81	102.53
25th percentile	65	44	26	18	6	41
50th percentile	113	77	48	34	15	81
75th percentile	177	125	78	58	31	141
Minimum	1	2	2	2	1	1
Maximum	712	706	667	477	245	712
Hospital volume:						
Mean	684.79	592.09	490.97	449.25	476.02	615.98
Standard deviation	350.18	361.07	349.39	349.51	374.99	364.26
25th percentile	434	251	179	153	137	288
50th percentile	669	580	427	352	378	617
75th percentile	933	874	748	717	767	889
Minimum	1	2	2	3	1	1
Maximum	1,420	1,421	1,410	1,413	1,410	1,421
Observations	55,687	29,455	103,33	4,116	8,156	107,747

TABLE 5
SUMMARY OF 12-MONTH VOLUME BY ELAPSED DAY QUANTILE (CABG)

	Number of Elapsed Days			Full Sample
	0–2	3–14	15+	
Surgeon volume:				
Mean	81.78	61.00	28.70	63.13
Standard deviation	47.67	36.44	24.20	43.17
25th percentile	48	34	10	31
50th percentile	78	54	23	55
75th percentile	102	84	40	87
Minimum	2	2	1	1
Maximum	265	265	211	265
Hospital volume:				
Mean	190.80	167.82	131.95	170.16
Standard deviation	92.71	93.93	94.29	95.63
25th percentile	122	96	49	97
50th percentile	176	155	122	159
75th percentile	275	235	195	248
Minimum	1	1	1	1
Maximum	371	369	372	372
Observations	7,423	9,975	3,400	20,798

variable; it is indeed capturing the impact of temporal distance conditional on volume.

Our initial estimation strategy is to simply include the number of days elapsed since the last time that surgeon *j*, operating on patient *i*, performed the procedure. The results of this estimation are found in table 6. For each procedure in the table, column 1 shows the linear probability model controlling for county fixed effects only. We use the log of both surgeon and hospital volumes because previous volume-outcome literature suggests that volumes are of a nonlinear functional form in this setting.[22] For PTCA, it seems the temporal impact is the one that is statistically significant, and the impact of the volume of the surgeon and hospital are not different from zero. In the case of CABG, it seems that both the surgeon and hospital volume are the significant factors, and the temporal effects do not matter. However, we note that this model has controlled for little unobserved heterogeneity, and an unsuspecting researcher might make some untenable conclusions from these results.

Column 2 under each procedure in table 6 includes both county and hospital fixed effects. The addition of hospital fixed effects is to control for the unobserved heterogeneity associated with hospital quality. For PTCA, this does little to change the precision or point estimate of the coefficient on the elapsed days since the last procedure. For CABG, the inclusion of hospital fixed effects raises the point estimate of the coefficient on elapsed days, though it is still not significant. Controlling for unobserved heterogeneity at the hospital level appears to reduce the point estimate of the coefficients on both surgeon and hospital volume and leads to the conclusion that they are no longer statistically significant predictors of mortality either.

Column 3 under each procedure in table 6 shows the linear probability specification in which we include county, hospital, and surgeon fixed effects. The inclusion of the surgeon fixed effects should capture the remaining unobserved heterogeneity attributable to the characteristics of the surgeon. As the reader can see, the inclusion of all three fixed effects slightly reduces the coefficient on the elapsed-days variable for PTCA, and thus it is only significant at the 10 percent level, despite the precision of the estimate being only slightly worse. For CABG, the inclusion of these surgeon fixed effects does not seem to render any statistically significant results as in the last estimation. It is important at this juncture to point out that even with what appears to be large sample sizes to some degree, this model is quite restrictive with the simultaneous

[22] We probed for quartile effects in the surgeons' and hospitals' volume as well, but these did not change our main result: that temporal distance has an impact on productivity and task repetition does not.

TABLE 6

Impact of Surgeons' Days since Last Surgery on Mortality

	Mortality within 1 Month[a]					
	PTCA (N = 107,747)			CABG (N = 20,798)		
	(1)	(2)	(3)	(1)	(2)	(3)
Days since last procedure	.000059***	.000060***	.000060**	.000007	.000039	.000104
	(.000023)	(.000023)	(.000028)	(.00006)	(.000058)	(.000075)
Surgeon's 12-month volume	−.001163	−.001249	.00014	−.009428**	−.005365	.003993
(ln)	(.00094)	(.000813)	(.00189)	(.004508)	(.003722)	(.004797)
Hospital's 12-month volume	.001307	−.00019	−.001273	−.009174**	−.003035	−.00751
(ln)	(.000894)	(.001899)	(.002192)	(.004486)	(.006777)	(.006951)
Patient characteristics[b]	X	X	X	X	X	X
Hospital characteristics[c]	X	X	X	X	X	X
Surgeon characteristics[d]	X	X	X	X	X	X
County fixed effects	X	X	X	X	X	X
Hospital fixed effects		X	X		X	X
Surgeon fixed effects			X			X

Note.—Standard errors listed in parentheses after the coefficients are robust clustered by surgeon.

[a] Dependent variable.

[b] Patient characteristics include age, sex, medical expenditures in the past year, dummies for each of the Charlson indices located in tables 2 and 3, and whether or not this is a repeat intervention.

[c] Hospital characteristics include dummies for the number of beds (0–250, 250–600) and public or nonprofit status.

[d] Surgeon characteristics include dummies for surgeons' age group (35–37, 38–40, 41–43, 44–46, 47–49, 50–52, 53–55, or 56+ years) and whether they had a patient die in the hospital in their last three procedures.

** Significance level at 5 percent.

*** Significance level at 1 percent.

inclusion of the fixed effects for hospitals and surgeons.[23] Given that the coefficients are simply measuring deviations from the conditional mean, this estimation strategy may not be capturing all the relevant information we are interested in. As such, conclusions from these particular estimations may be misleading, and so we proceed with our alternate specifications that will yield some clearer conclusions.

In table 7, results are listed for specifications in which we substitute the elapsed number of days variable for a vector of dummy variables for elapsed days since last surgery performed by the surgeon, grouped into day quantiles. The comparison group in each table is the outcomes of the surgeries performed by surgeon j on patient i when surgeon j had last performed the procedure within 2 days. Again, it is important to note that day quantiles are broader for CABG (and thus there is a smaller number of day quantiles) because the sample size is much smaller and because the CABG procedures are not performed as often as PTCA.

Similar to the progression of table 6, column 1 in table 7 includes only county fixed effects, column 2 includes county and hospital fixed effects, and column 3 includes county, hospital, and surgeon fixed effects. We would like to point out here first that the within R^2 is much higher than the between R^2 in the model with all fixed effects, meaning that our model explains more of the within-surgeon variation than the between-surgeon variation. The within standard deviation of temporal distance is 27 days (mean 9.6 days) for PTCA and 31 days (mean 12.7 days) for CABG, suggesting that there is still a fair amount of variation within the surgeon panel.

There is a stark contrast between the results in column 3 in table 7 and those in table 6. First of all, there is an obvious gradient to the coefficient, which indicates that the probability of a patient experiencing mortality within 1 month increases as temporal distance between surgeries increases. In the case of PTCA, this increased mortality becomes significant for the quantiles in which the time elapsed is greater than 2 weeks (15+ days), and the gradient is immediately apparent and significant for CABG in even less time.[24]

To get a sense of the overall impact of temporal distance on 1-month mortality, we need to compare the coefficients on the temporal distance quantiles to the mean of 1-month mortality for patients having a procedure performed by a surgeon who has performed this procedure in the past 2 days (the 0–2-day reference group). The mean 1-month post-

[23] This is particularly true for CABG, for which the sample size is quite a bit smaller, and the number of fixed effects is not reduced by the same magnitude. We also estimated eq. (2) for both PTCA and CABG without the temporal distance measure included, and volume was not a significant predictor of mortality after inclusion of the full set of fixed effects.

[24] This difference in CABG may be due to the less refined quartile groupings, but finer quartiles yield some groups with very few observations, which complicates estimation and interpretation.

TABLE 7
IMPACT BY ELAPSED-DAY QUANTILE

	Mortality within[a]			
	1 Month (1)	1 Month (2)	1 Month (3)	12 Months (4)
PTCA (N = 107,747):				
Sample mean of the dependent variable	.0205	.0205	.0205	.0727
Sample mean of the 0–2-day reference group	.0166	.0166	.0166	.0634
Days since last procedure:				
3–7	.000271	.000184	.000576	.002534
	(.000995)	(.000992)	(.000985)	(.001989)
8–14	.000889	.001331	.001819	.0018
	(.001624)	(.001622)	(.001646)	(.002754)
15–21	.005436*	.005982**	.006477008893**	.010115**
	(.002959)	(.002958)	(.003015)	(.004817)
22+	.009751***	.010359***	.008565***	.011305***
	(.00267)	(.002737)	(.002833)	(.004296)
Surgeon's 12-month volume (ln)	−.00094	−.00095	.000092	−.000484
	(.000945)	(.000832)	(.001871)	(.002921)
Hospital's 12-month volume (ln)	.001617*	−.000231	−.001193	−.004151
	(.000907)	(.001901)	(.002192)	(.003656)
CABG (N = 20,798):				
Sample mean of the dependent variable	.0643	.0643	.0643	.1446
Sample mean of the 0–2-day reference group	.0503	.0503	.0503	.1269

Days since last procedure:				
3–14	.007642*	.008313**	.008562**	.004413
	(.003903)	(.003899)	(.003905)	(.004713)
15+	.012735**	.014321***	.013499**	.013978**
	(.005583)	(.005485)	(.005392)	(.006904)
Surgeon's 12-month volume (ln)	−.007655*	−.004534	.003094	.004971
	(.004124)	(.003361)	(.004903)	(.007069)
Hospital's 12-month volume (ln)	−.008893008893***	−.002718	−.006557	−.005537
	(.004444)	(.006715)	(.006905)	(.008826)
Patient characteristics[b]	X	X	X	X
Hospital characteristics[c]	X	X	X	X
Surgeon characteristics[d]	X	X	X	X
County fixed effects	X	X	X	X
Hospital fixed effects		X	X	X
Surgeon fixed effects			X	X

Note.—Standard errors listed in parentheses after the coefficients are robust clustered by surgeon.

[a] Dependent variable.

[b] Patient characteristics include age, sex, medical expenditures in the past year, dummies for each of the Charlson indices located in tables 2 and 3, and whether or not this is a repeat intervention.

[c] Hospital characteristics include dummies for the number of beds (0–250, 250–600) and public or nonprofit status.

[d] Surgeon characteristics include dummies for surgeons' age group (35–37, 38–40, 41–43, 44–46, 47–49, 50–52, 53–55, or 56+ years) and whether they had a patient die in the hospital in their last three procedures.

* Significance level at 10 percent.

** Significance level at 5 percent.

*** Significance level at 1 percent.

operative mortality of the 0–2-day reference group is 1.66 percent for PTCA and 5.03 percent for CABG. Thus, the main result, which includes fixed effects for year, county, hospital, and surgeon, suggests that PTCA performed by a surgeon in the 15–21-day quantile is 39 percent more likely to result in mortality within 1 month than a PTCA performed by a surgeon with a 0–2-day break. For CABG, the results using the full set of fixed effects suggest that a surgery performed by a surgeon with a 3–14-day temporal distance since his last procedure is 17 percent more likely to result in mortality within 1 month than a CABG performed by a surgeon who has a 0–2-day temporal distance since his last procedure.[25]

Overall, the magnitude of the impact of such a short temporal distance on the productivity of surgeons appears to be large, indicating a large reduction in the level of human capital of the surgeon, but there is literature suggesting that these large magnitudes are feasible both from the human capital and mortality perspective. Gowrisankaran et al. (2006) find that the additional human capital accumulated by an exogenous increase in surgeon volume fully depreciates within 1 year. In addition, a recent study by Guru et al. (2008) finds that 32 percent of the in-hospital deaths after CABG surgeries in their sample were deemed to be preventable. Thus, while our results initially reflect quite a large effect of temporal distance, these large effects are indeed possible in light of other studies.

We also consider a longer-term measure of patient health outcomes as our measure of productivity, namely, mortality within 12 months rather than 1 month. In column 4 of table 7, we report our results using mortality within 12 months and the full set of fixed effects (year, county, hospital, and surgeon) for comparison to the results using mortality within 1 month with the full set of fixed effects.

For PTCA, the results using 12-month mortality are similar (in significance and sign) to those using mortality within 1 month as the measure of productivity. For CABG, the coefficient on the temporal distance quantile group of 3–14 days is not significant and is much smaller in magnitude when we use mortality within 12 months as our dependent variable. There are a couple of potential explanations for this. The first is that the impact of having a surgeon with a short temporal distance extends the lives of patients for only a few months; therefore, the productivity difference is not well measured by looking at long-term measures of mortality. A second explanation, which is tangential to the first, is that patients treated with CABG may have more systemic health issues, such as diabetes, and their long-term mortality outcomes are less attributable to the surgeon's performance.

The results to this point are indicative of a temporal distance effect

[25] To get these percentages, we take the coefficients on the indicator for the quantile we are interested in and divide by the mean mortality of the 0–2-day reference group for each procedure.

on the human capital depreciation of surgeons, which is at least as important as the maintenance effects of task repetition on human capital retention. This result is robust to controls for unobserved heterogeneity stemming from the unobserved qualities of both the surgeon and the hospital. The volume of the surgeon seems to be less important, but this result may also be driven by the restrictiveness of the model. We then proceed by performing some robustness checks on these results.

VI. Robustness Checks

One issue that could potentially affect our estimation is that of patient selection. The issue is that patients of worse latent health could select or be channeled to the surgeons who have characteristics that are indicative of more ability but are unobservable to us.[26] If volume is indicative at least in part of the presence of more unobserved (to us) ability, this would mean that surgeons with higher volume (and generally lower temporal distance) would be operating on patients with higher odds of nonsurvival ceteris paribus, which would tend to underestimate the impact of both volume and temporal distance. Alternatively, those in worse health may arrive at the hospital in an emergency condition and be assigned the first doctor available, which oftentimes would be a surgeon with relatively low volume and/or a high temporal distance and would tend to overstate the impact of volume and temporal distance. While the surgeon fixed effects will capture the effect of those qualities that are time invariant in the surgeon, some ability of the surgeon or at least the recognition of these abilities (or lack thereof) by patients and/or other care providers who provide advice for patients may vary across time, and this would not be accounted for by the fixed effects. Therefore, we performed two checks to probe for evidence of selection.

First, we ran two separate regressions to see if patient health systematically predicted either temporal distance or surgeon volume. The first regression was temporal distance against all the other covariates in equation (2) from the main set of estimations, and the second regression was surgeon volume against all other covariates in equation (2). If those of greater ability were indeed attracting sicker patients, then one would expect some of the comorbidities indicated by the Charlson indices in the regression to significantly predict lower temporal distance and/or higher volume. Conversely, if sicker patients were somehow being selected to surgeons with lower volume and higher temporal distance, then this would show up as at least some of the Charlson indices significantly predicting lower volume and/or higher temporal distance. The results of these estimations are in table 8. Whereas the other surgeon characteristics listed do seem to be associated with volume and

[26] Ostensibly these would, on average, be surgeons with higher volume and lower temporal distance, although this is not necessarily the case.

TABLE 8
Checks for Patient Selection

	Temporal Distance[a]		Surgeon Volume[a]	
	PTCA (1)	CABG (2)	PTCA (3)	CABG (4)
Provider characteristics:				
Temporal distance			−.05***	−.05***
			(.00)	(.01)
Surgeon volume	−.04***	−.16***		
	(.01)	(.05)		
Hospital volume	.00	.03**	.15***	.34***
	(.00)	(.02)	(.03)	(.04)
Recent patient death	3.43***	6.90***	−1.80***	−1.18**
	(.91)	(1.22)	(.68)	(.57)
Patient characteristics:				
Acute myocardial infarction	−.14	.59	.31	.07
	(.20)	(.54)	(.30)	(.24)
Peripheral vascular disease	−.61	1.92	−1.63	−.78
	(.85)	(1.33)	(1.15)	(.77)
Dementia	−4.26	−4.93	−2.45*	−1.88
	(2.70)	(4.54)	(1.47)	(1.67)
Chronic obstructive pulmonary disease	.59	.83	−.43	−.05
	(.36)	(.83)	(.40)	(.55)
Type I liver disease	−.35	−3.73	.06	.55
	(.73)	(2.52)	(.56)	(1.02)
Type II liver disease	1.33	−3.00	−7.53**	−.95
	(5.20)	(4.71)	(3.33)	(2.92)
Type I diabetes	.11	−.34	.00	.33
	(.19)	(.47)	(.26)	(.26)
Type II diabetes	.54	.82	.18	.64
	(.35)	(.78)	(.46)	(.44)
Renal disease	−.10	.96	.95	.54
	(.66)	(1.41)	(1.01)	(.50)
Rheumatism	.12	1.30	.30	−2.00
	(1.23)	(2.18)	(1.43)	(1.44)
Cancer	−1.01	−1.25	−.95	.56
	(.89)	(.93)	(.75)	(1.19)
Tumor	.39	−2.52	2.58	−2.10
	(1.53)	(2.97)	(1.61)	(3.18)
Hypertension	−.43**	.57	−.32	.22
	(.19)	(.48)	(.34)	(.23)
Observations	107,747	20,798	107,747	20,798
R^2	.014	.037	.309	.427

Note.—All estimations in this table include year, county, hospital, and surgeon fixed effects. Standard errors in parentheses are robust clustered by surgeon.
[a] Dependent variable.
* Significance level at 10 percent.
** Significance level at 5 percent.
*** Significance level at 1 percent.

temporal distance in a significant way that will be discussed in more detail later, the patient level characteristics meant to control for relative illness do not. The *F*-tests on joint significance of all Charlson indices are less than two for all regressions, suggesting that patient health is not systematically correlated with either temporal distance or volume. None of the Charlson indices are significantly correlated with the temporal distance or surgeon volume for CABG, suggesting that the patient-selection problem could be very limited in this case. One of the 13 Charlson indices is significantly correlated with lower temporal distance at the 10 percent level for PTCA (hypertension), and none of these are significantly correlated with higher temporal distance for CABG. Taken together, there is no significant indication whether patient selection will overstate or understate the estimation in table 7.

Second, we estimated the impacts of temporal distance on productivity, using the specification with the full set of fixed effects and dummies for each temporal-distance quantile on the sample of patients who have not been given diabetes diagnoses. Selecting a group of healthier patients allows us to minimize the potential of patient selection. If this subsample yields similar estimations, then it will suggest that the patient selection does not seriously bias our results. We choose patients without diabetes for two reasons. First, the sample size is big enough to allow the inclusion of surgeon fixed effects. Second, whether or not a person has diabetes is an important factor in the choice of treatment of coronary artery disease and the subsequent health outcomes of patients after surgical intervention for coronary artery disease. Analyzing these patients as a separate risk group is common in the medical literature (e.g., see Hannan et al. 2005). The results of this estimation can be found in table 9. For PTCA, the coefficients of 15–21 days and 22+ days are slightly larger than the coefficients on the comparable estimation in column 3 of table 7; for CABG, the results are similarly larger on each of the temporal distance quantiles for nondiabetic patients than those in the comparable estimation of column 3 of table 7. It suggests that patient selection based on latent health, if it exists, is more likely to lead to an underestimation of the effect of temporal distance due to sicker patients being treated by a surgeon with a shorter temporal distance.

The second issue that concerns us is a possible serial correlation of error terms in equation (2). The variable we included to measure whether a surgeon had a recent patient die in the hospital prior to the current patient was a significant predictor of the patient in the current observation dying within 1 month. As we show in table 8, having a recent patient die in the hospital will increase the temporal distance and decrease the volume. This may lead the reader to suspect that a previous poor outcome would increase temporal distance (perhaps the case needs to be reviewed by administration) and potentially lead to longer temporal distance and that there may be something akin to serial correlation

TABLE 9
NONDIABETIC PATIENTS

	Mortality within 1 Month[a]	
	Coefficient	Standard Error
PTCA:		
Days since last procedure:		
3–7	−.00058	(.00111)
8–14	.00281	(.00207)
15–21	.00740**	(.00358)
22+	.00919***	(.00337)
Surgeon's 12-month volume (ln)	−.00114	(.00204)
Hospital's 12-month volume (ln)	−.00262	(.00271)
Observations	71,173	
CABG:		
Days since last procedure:		
3–14	.00900*	(.00521)
15+	.01436**	(.00694)
Surgeon's 12-month volume (ln)	−.00147	(.00643)
Hospital's 12-month volume (ln)	−.00284	(.00962)
Observations	12,337	

Note.—Standard errors are robust clustered by surgeon. Patient characteristics include age, sex, medical expenditures in the past year, dummies for each of the Charlson indices located in tables 2 and 3, and whether or not this is a repeat intervention. Hospital characteristics include dummies for the number of beds (0–250, 250–600) and public or nonprofit status. Surgeon characteristics include dummies for the surgeons' age group (35–37, 38–40, 41–43, 44–46, 47–49, 50–52, 53–55, or 56+ years) and whether they had a patient die in the hospital in their last three procedures.
[a] Dependent variable.
* Significance level at 10 percent.
** Significance level at 5 percent.
*** Significance level at 1 percent.

in poor performance. This would be a concern because this could be the effect that the temporal distance variable is, at least in part, capturing, and thus the actual impact of temporal distance would be overstated.

The estimations in columns 1 and 2 of table 8 show that having a recent patient die in the hospital is indeed associated with an increased temporal distance. While this is a statistically significant relationship, the magnitude of these coefficients is important to consider. The coefficients on this variable are relatively small (in each case it increases temporal distance by less than 1 week). However, given that we are focusing on the temporal distance quantile estimations, we want to know whether this effect is indeed shifting surgeons to other quantiles. In the case of PTCA, the coefficient on the recent patient death indicator is 3.42, and the mean of the temporal distance in the sample is 9.58 days. This means the average patient will have a surgeon with a temporal distance in the third quantile (8–14 days), regardless of whether a recent patient died in the hospital. However, for CABG, the coefficient on the recent patient death indicator is 6.81, and the mean temporal distance in the sample is 12.65 days, so shifting the average patient from a surgeon without a recent patient death to one with a recent patient death would imply that this surgeon would move from the 3–14-day quantile to the 15+-day quantile. While this may raise some concern, the fact that the coefficient on the 3–14-day quantile in the main estimation (table 7) is still significant with the inclusion of this recent patient death indicator suggests that this recent patient death effect may at most be overstating the effect of a break of 15+ days.

A final concern with our estimation is that the higher temporal distance quantiles may have a large number of observations of surgeons with very low volume and that this is driving our results. From a theoretical standpoint, our model is to some extent reliant on the assumption that we are measuring surgeons who have approached the asymptotic minimum of their learning curve and that task repetition serves to keep the surgeon at this minimum, while the depreciation of human capital due to increased temporal distance between tasks leads to deviations from this minimum. Therefore, we estimate a set of regressions as a final robustness check to ensure that our results are not driven by surgeons who have never approached the minimum of their learning curve.

We begin by narrowing our sample to those surgeries performed by a surgeon who has reached some minimum volume measure during the period for which we observe. As our minimum volume measure, we use the median 12-month volume of each procedure during 1998, the first year of our sample (PTCA = 69, CABG = 45). To be clear, our sample in this case includes only those procedures that were performed by a surgeon who at some point prior to the procedure on patient i had a cumulative 12-month volume of procedures greater than 69 for PTCA and 45 for CABG.[27] The results of these estimations using this higher-volume sample and the full set of fixed effects can be found in column 1 of table 10, where the full sample results from table 7 are included in column 3 for reference.

In a second check on this assumption of where we are on the learning curve, we consider a sample that includes only procedures performed by surgeons with a minimum level of task repetition in the 12 months prior to the procedure on patient i. In order to have a large enough sample size to use the full set of fixed effects in the CABG estimation, we exclude those procedures performed by surgeons whose 12-month volume was below the first quartile level of volume for the entire sample for each procedure (as listed in the last column of tables 4 and 5). The

[27] Note that for PTCA this means that the surgeons had to have averaged a little more than one procedure per week at some point in their careers, and for CABG they had to have averaged slightly less than one procedure per week at some point in their careers. Even if their volume drops below this level later in the sample, the procedures they perform remain in our estimation.

TABLE 10
High-Volume versus All Surgeons

	Mortality within 1 Month[a]		
	High-Volume Surgeons		
	Definition 1[b] (1)	Definition 2[c] (2)	Full Sample (3)
PTCA:			
Days since last procedure:			
3–7	−.00032	.00054	.000576
	(.00112)	(.00095)	(.000985)
8–14	.00229	.00201	.001819
	(.00231)	(.0019)	(.001646)
15–21	.01091*	.00938**	.006477**
	(.00617)	(.00397)	(.003015)
22+	.01503**	.01087**	.008565***
	(.00674)	(.00488)	(.002833)
Surgeon's 12-month volume (ln)	.00049	.0017	.000092
	(.00205)	(.00212)	(.001871)
Hospital's 12-month volume (ln)	.00052	.00084	−.001193
	(.00323)	(.00295)	(.002192)
Observations	69,695	85,125	107,747
CABG:			
Days since last procedure:			
3–14	.00548	.00836**	.008562**
	(.00414)	(.00409)	(.003905)
15+	.01948**	.01463**	.013499**
	(.00746)	(.00585)	(.005392)
Surgeon's 12-month volume (ln)	.00028	.00826	.003094
	(.0075)	(.00939)	(.004903)
Hospital's 12-month volume (ln)	−.01007	−.00757	−.006557
	(.00799)	(.00774)	(.006905)
Observations	15,013	17,142	20,798

Note.—Standard errors listed in parentheses after the coefficients are robust clustered by surgeon. Patient characteristics include age, sex, medical expenditures in the past year, dummies for each of the Charlson indices located in tables 2 and 3, and whether or not this is a repeat intervention. Hospital characteristics include dummies for the number of beds (0–250, 250–600) and public or nonprofit status. Surgeon characteristics include dummies for surgeons' age group (35–37, 38–40, 41–43, 44–46, 47–49, 50–52, 53–55, or 56+ years) and whether they had a patient die in the hospital in their last three procedures.

[a] Dependent variable.

[b] Surgeons who at some point prior to the current observation reached a minimum level of volume (69 for PTCA and 45 for CABG).

[c] Surgeons who at the time of the procedure in the current observation had a 12-month volume greater than 40 for PTCA and greater than 30 for CABG.

* Significance level at 10 percent.

** Significance level at 5 percent.

*** Significance level at 1 percent.

results of this estimation can be found in column 2 of table 10. Overall, the results in table 10 suggest our findings are robust to the exclusion of low-volume surgeons in the analysis.

VII. Discussion and Conclusion

The purpose of this study is to address the impacts that temporal distance between production tasks and the level of task repetition have on the human capital and, thus, productivity of cardiac surgeons. Our results suggest that temporal distance does have an impact on productivity, as measured by patient mortality outcomes. From a human capital theory perspective, this implies that temporal distance may be of great importance to the level of human capital the surgeon is able to maintain. Our results are robust to the examination of those surgeons who have a higher level of task repetition prior to the break and who we can reasonably expect are specialists.

Although we have performed some robustness checks to address the validity of our results, caution should be taken until other research confirms the findings in our study. First and foremost, we generally do not know the mechanism behind these temporal breaks for the surgeons. During these breaks, the surgeons could be performing other productive tasks within their field, engaging in professional activities, or taking a much-needed break, which we would generally categorize as activities that would potentially lead to an overall increase in the level of human capital and productivity in the long run. However, these breaks could also be for reasons that, at least temporarily, only serve to distract the surgeon and thus reduce the level of human capital employed (e.g., their own illness or that of a loved one, personal issues such as divorce or a child in trouble, non-performance-related conflicts at work). Even if one assumes that the reasons for the temporal breaks are random, caution in interpreting the results still needs to be taken.

Despite our efforts to address potential concerns with the robustness of the estimation, the measurement error in the productivity measure and the temporal distance between surgeries (stemming from our lack of data on actual death dates and procedure dates), along with the patient-selection issues discussed previously, could be affecting the estimation to some extent. Measurement error in the dependent variable will reduce the precision of our estimates. This is not as much of a concern with relation to our assertion of the importance of temporal distance because these coefficients were still significant; however, it limits the conclusions we can draw about the lack of significance on task repetition. Measurement error in our measure of temporal distance will tend to attenuate the coefficients on temporal distance and could also bias other coefficients, such as surgeon volume (assuming the measurement error in the dependent variable is random). Therefore, conclusions drawn from the lack of significance on volume, which is our measure of the level of task repetition, are tenuous at best. Future research with more detailed data, including a more exact measure of temporal distance and the mechanisms leading to the break in production (i.e., vacation, personal issues, professional development, refresher training, illness), may be able to clarify the exact nature of the impact of temporal distance on productivity and to address some of the other limitations in this study.

Despite the limitations of our data, these findings have implications for a variety of interests. We contribute to the literature on organizational forgetting by demonstrating empirically that it may be that organizational forgetting is largely due to the sum of the forgetting of individuals. Our findings also suggest that the lower rates of pay experienced by workers returning from spells of unemployment found in previous research are in fact justified at least in part by the depreciation/deterioration of their human capital over time. As for policy makers and managers involved in medical licensing and operating-room scheduling, our results suggest that even high-volume surgeons will experience depreciation of their skills over short periods of time and that strategies to mitigate the impact of these effects on the health of patients should be developed. Movements toward the regionalization of major surgical procedures based on the previous volume-outcome literature should also be undertaken with caution because the improvements in outcomes that one expects from consolidating procedures to higher-volume providers may be somewhat lower than expected given that this temporal distance effect is robust to the exclusion of low-volume surgeons and appears to be independent of hospital volume. Perhaps using a moving average number of surgeries per month or quarter is a better measure of how active surgeons are, as it relates to reaccreditation, licensing, and policy making related to market consolidation. Also, hospital managers targeting mortality rates may want to institute oversight policies for all surgeons returning from long professional breaks, regardless of previous experience.

References

Argote, L., S. Beckman, and D. Epple. 1990. "The Persistence and Transfer of Learning in Industrial Settings." *Management Sci.* 36 (2): 140–54.

Argote, L., and D. Epple. 1990. "Learning Curves in Manufacturing." *Science* 247 (4945): 920–24.

Benkard, C. L. 2000. "Learning and Forgetting: The Dynamics of Aircraft Production." *A.E.R.* 90 (4): 1034–54.

Ben-Porath, Y. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *J.P.E.* 75 (4): 352–65.

Cuellar, A. E., R. Lindrooth, and J. Cross. 2006. "Volume Incentives for Inpatient Quality and Patient Safety." Working paper, Dept. Health Policy and Management, Columbia Univ.

Darr, E., L. Argote, and D. Epple. 1995. "The Acquisition, Transfer, and Depre-

ciation of Knowledge in Service Organizations: Productivity in Franchises." *Management Sci.* 41 (11): 1750–62.
Globerson, S., N. Levin, and A. Shtub. 1989. "The Impact of Breaks on Forgetting When Performing a Repetitive Task." *IIE Transactions* 21 (4): 376–81.
Gowrisankaran, G., V. Ho, and R. J. Town. 2006. "Causality, Learning and Forgetting in Surgery." Working paper, Dept. Econ., Univ. Arizona.
Guru, V., et al. 2008. "Relationship between Preventability of Death after Coronary Artery Bypass Graft Surgery and All-Cause Risk Adjusted Mortality Rates." *Circulation* 117 (23): 2969–76.
Halm, E. A., C. Lee, and M. R. Chassin. 2002. "Is Volume Related to Outcome in Health Care? A Systematic Review and Methodologic Critique of the Literature." *Ann. Internal Medicine* 137 (6): 511–20.
Hannan, E. L., et al. 2005. "Long-Term Outcomes of Coronary-Artery Bypass Grafting versus Stent Implantation." *New England J. Medicine* 352 (21): 2174–83.
Hockenberry, J., H.-M. Lien, and S.-Y. Chou. 2008. "Learning, Forgetting and Productivity in Cardiac Surgery." Working paper, Dept. Health Management and Policy, Univ. Iowa.
Huckman, R. S., and G. P. Pisano. 2006. "The Firm Specificity of Individual Performance: Evidence from Cardiac Surgery." *Management Sci.* 52 (4): 473–88.
Kostis, W., et al. 2007. "Weekend versus Weekday Admission and Mortality from Myocardial Infarction." *New England J. Medicine* 356 (11): 1099–1109.
Lien, H.-M., S.-Y. Chou, and J. T. Liu. 2004. "Hospital Competition under Universal Health Insurance: The Stroke Treatment in Taiwan." Working paper, Dept. Public Finance, Nat. Cheng Chi Univ.
———. 2008. "Hospital Ownership and Performance: Evidence from Stroke and Cardiac Treatment in Taiwan." *J. Health Econ.* 27 (5): 1208–23.
Luft, H. S., S. S. Hunt, and S. C. Maerki. 1987. "The Volume-Outcome Relationship: Practice Makes Perfect or Selective Referral Patterns?" *Health Services Res.* 22 (2): 157–82.
Mincer, J., and H. Ofek. 1982. "Interrupted Work Careers: Depreciation and Restoration of Human Capital." *J. Human Resources* 17 (1): 3–24.
Mincer, J., and S. W. Polachek. 1974. "Family Investments in Human Capital: Earnings of Women." *J.P.E.* 82 (2; suppl.): S76–S108.
———. 1978. "Women's Earnings Reexamined." *J. Human Resources* 13 (1): 118–34.
Sandell, S. H., and D. Shapiro. 1978. "The Theory of Human Capital and the Earnings of Women: A Reexamination of the Evidence." *J. Human Resources* 13 (1): 103–17.
Tsai, A. C., M. Votruba, J. Bridges, and R. D. Cebul. 2006. "Overcoming Bias in Estimating the Volume-Outcome Relationship." *Health Services Res.* 41 (1): 252–64.
Wright, T. P. 1936. "Factors Affecting the Cost of Airplanes." *J. Aeronautical Sci.* 3 (4): 122–28.