

Avdic, Daniel; Lundborg, Petter; Vikström, Johan

Conference Paper

Learning-by-Doing in a High-Skill Profession when Stakes are High: Evidence from Advanced Cancer Surgery

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik
- Session: Health II, No. B10-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Avdic, Daniel; Lundborg, Petter; Vikström, Johan (2014) : Learning-by-Doing in a High-Skill Profession when Stakes are High: Evidence from Advanced Cancer Surgery, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Health II, No. B10-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100278>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Learning-by-Doing in a High-Skill Profession when Stakes are High: Evidence from Advanced Cancer Surgery^a

by

Daniel Avdic^b, Petter Lundborg^c and Johan Vikström^d

12th February, 2014

Abstract

Although learning-by-doing is believed to be an important source of productivity growth there is limited evidence that production volume affects productivity in a causal sense. We document evidence of learning-by-doing in a high-skill profession where stakes are high; advanced cancer surgery. For this purpose, we introduce a novel instrument that exploits closures and openings of entire cancer clinics which has given rise to sharp and exogenous changes in cancer surgical volume performed at Swedish public sector hospitals. Using detailed register data on more than 100,000 episodes of advanced cancer surgery our results suggest positive effects of surgery volume on survival. In addition, we provide evidence on the mechanisms through which these improvements occur. We also show that results are not driven by changes in patient composition or by other changes at the hospital level.

Keywords: hospital volume, learning-by-doing, cancer surgery, survival, causal effect.

JEL-codes: I11, I12, I18, L11.

^aWe thank Jerome Adda, Tabea Bucher-Koenen, Tomas Buchmueller, Christina Felfe, Martin Huber, Per Johansson, Martin Karlsson, Nikolaj Malchow-Moller, Paolo Pertile and seminar participants at MEA-Munich, IFAU-Uppsala, University of St. Gallen, University of Southern Denmark, Lund University, 12th Journées LAGV in Aix-de-Provence, The Economics of Disease workshop in Darmstadt, Maastricht, University of Duisburg-Essen, CBS-Copenhagen, University of Gothenburg, ESPE 2013 and EALE 2013. Financial support from the Swedish Council for Working Life and Social Research (DNR 2004-2005 and 2009-0826) is gratefully acknowledged. Vikström acknowledges support from the Tom Hedelius and Jan Wallander foundation.

^bDepartment of Economics, Uppsala University, UCLS and IFAU, SE-751 20 Uppsala, Sweden. E-mail: daniel.avdic@ifau.uu.se

^cDepartment of Economics, Lund University, and IZA. E-mail: petter.lundborg@nek.lu.se

^dIFAU and UCLS. E-mail: johan.vikstrom@ifau.uu.se

Table of contents

1	Introduction	3
2	Cancer surgery and Learning-by-doing	11
3	Data.....	14
4	Empirical strategy.....	16
4.1	Descriptive background.....	16
4.2	The instrument	17
4.3	Econometric model	21
5	Main results	23
5.1	The volume-outcome relationship	23
5.1.1	Learning to avoid mistakes	25
5.2	Alternative interpretations	26
6	Extensions and mechanisms.....	31
6.1	Effect dynamics and learning	31
6.2	Learning-by-doing versus scale effects	32
7	Conclusions	37
	References	40
	Tables and figures	45

1 Introduction

Learning-by-doing is believed to be an important source of productivity growth (Arrow, 1962; Lucas, 1988). The intuition for this is simple — the more individuals or organizations perform a specific task, the better they become at doing it. Evidence for learning-by-doing has been obtained by linking greater accumulated production volumes to reductions in unit labor costs or, more directly, to quality improvements in a variety of activities, including Kibbutz farming, ship building, car manufacturing, airplane flight control, nuclear plant operation reliability, iron works and even pizza making (Wright, 1936; Lundberg, 1961; Darr *et al.*, 1995; Benkard, 2000; Jovanovic and Nyarko, 1995; Thornton and Thompson, 2001; Thompson, 2001; Levitt *et al.*, 2012).

While it is widely believed that greater production volume facilitates learning-by-doing, and thereby productivity, in almost all sectors of the economy, there is surprisingly little causal evidence showing so (Thompson, 2012).¹ Estimating the causal effect of volume has proven difficult for a number of reasons. The perhaps most obvious problem is that conventional measures of experience and tenure are likely to be correlated with unobserved factors which, in turn, are associated with productivity. Moreover, and as noted already by Wright (1936), a negative relationship may arise if more tooling and standardization of procedures means that a firm's ability to use less skilled labor increases.²

In this paper, we focus on learning-by-doing in the health care sector. Apart from the general difficulties in estimating learning-by-doing effects, the specific case of the health care sector brings additional identification problems; first, differences in outcomes between high- and low-volume hospitals may be explained by *selective referral* where high-

¹This is in contrast to the large literature showing how incentives affect productivity in a variety of settings, see, for example, Currie *et al.* (1995); Lazear (2000); Croxson *et al.* (2001); Hamilton *et al.* (2003); Gaynor *et al.* (2004); Bandiera *et al.* (2007, 2009); Acemoglu and Finkelstein (2008); Hossain and List (2009); Propper and Van Reenen (2010); Propper *et al.* (2010); Bloom *et al.* (2011a); Clemens and Gottlieb (2014).

²In a recent overview of the literature, Thompson (2012) states that: “Unsurprisingly, since the much greater part of the empirical learning curve literature predates the wide use of instrumental variable techniques (Angrist and Krueger, 2001), the body of literature offering reliable support for the standard formulation is smaller than is generally supposed.”

quality hospitals has higher volumes simply because they attract more patients — i.e. a reverse causality mechanism. Secondly, individuals with different health characteristics may self-select into areas with different hospital sizes, such as in rural and urban areas. Third, local health shocks, such as localized outbreaks of contagious diseases, affecting both hospitalization incidence and health outcomes, could induce a spurious relationship between cumulated number of treatments and their outcomes. For policy purposes, while being observationally equivalent to learning effects, these competing explanations obviously have very different policy implications than those drawn from a learning-by-doing hypothesis.

The medical literature has tried to account for selection problems by controlling for observable patient characteristics such as age, gender, and reported health conditions.³ A notable contribution is Birkmeyer *et al.* (2003) who, using a nationwide Medicare sample, found a positive relation between volume and survival in cancer and cardiovascular surgery. In order to also account for unobserved heterogeneity across hospitals, some researchers have utilized longitudinal data to estimate fixed-effects models (Hamilton and Hamilton, 1997; Hamilton and Ho, 1998).⁴ The limitation of these types of models is that they only account for time-invariant unobserved heterogeneity across groups, whereas e.g. selective referral may be driven by trends in hospital quality. To overcome such endogeneity concerns, two U.S.-based studies have used the number of hospitals within a certain radius as an instrument for hospital volume and applied instrumental variables techniques (Gaynor *et al.*, 2005; Gowrisankaran *et al.*, 2006).⁵ Since market structure

³Halm *et al.* (2002) review 135 medical studies on the volume-outcome relationship for surgical procedures of which about 70 percent of the studies find a significant and positive correlation. In a summary article in New England Journal of Medicine, Kizer (2003) concludes that the strongest relations between volume and outcome have been found for AIDS treatment, surgery for pancreatic cancer, esophageal cancer, abdominal aortic aneurysm, and congenital heart disease.

⁴In neither of these studies were there any evidence of a causal volume-outcome relationship after controlling for hospital fixed effects, hence supporting the selective referral hypothesis.

⁵Gaynor (2006) finds positive effects of volume on survival following heart surgery using the number of CABG-offering hospitals operating in a specific radius around the hospital of treatment as an instrument for surgical volume. Gowrisankaran *et al.* (2006) use hospital distance interacted with hospital and patient characteristics as an instrument and finds positive effects of volume on survival after heart surgery and repair of abdominal aortic aneurysm.

in a health care system that allows for private providers is an endogenous outcome of a competitive process, however, the number of hospitals in an area is not necessarily exogenous. In such a setting, hospitals may base their entry and exit decisions both on underlying patient characteristics and on the current and expected volume and quality of existing hospitals in the same area.⁶

With the limitations of the existing literature in mind, we provide new evidence of learning-by-doing by studying the causal effect of production volume on quality outcomes in advanced cancer surgery. In order to deal with the endogeneity of production volume, we introduce a novel instrument that generates quasi-experimental variation in the number of cancer surgeries performed in Swedish public hospitals over the last two decades. Specifically, the proposed instrument exploits regional variation in closures and openings of entire hospital cancer clinics over time, which generates sharp changes in the surgical volume performed at nearby cancer clinics. Importantly, since in-patient care in Sweden is organized by the public sector, we can be sure that closures and openings of cancer clinics are not driven by market forces but rather by political decisions unrelated to hospital quality and underlying population health indicators — a conjecture supported by supplementary analyses of the data.

For the purpose of our study, we have collected detailed data on *all* cancer surgeries performed at Swedish hospitals between 1998-2007 — in total more than 100,000 episodes. The data includes rich and detailed information about type of surgery, date of hospital admission and discharge, prevalence of post-surgery complications and comorbidities for each individual patient. Furthermore, we have linked the cancer surgery data to individual-level register data on mortality and socioeconomic characteristics. We select and include the three most common types of cancers in Sweden — but also in most

⁶Other studies have used the total number of hospital beds as an instrument for hospital volume (Luft *et al.*, 1987; Hughes *et al.*, 1988; Farley and Ozminkowski, 1992; Norton *et al.*, 1998). However, in presence of selective referral it is likely that high-quality hospitals should both be larger and have more beds. Finally, Tsai *et al.* (2006) and Kahn *et al.* (2009) uses distance to a high-volume hospital for each patient as an instrument for the probability of getting treated at a high-volume hospital.

of the Western world — in our analysis; breast cancer, prostate cancer, and colorectal cancers. The three cancer types we study cover more than half of the annual number of cancer surgeries performed at Swedish hospitals and constitute some of the major causes of death in the adult population.⁷

We believe that cancer surgery is particularly well suited for studying learning-by-doing; first, the intuition for a causal volume-outcome is simple: practice makes perfect and the outcome of the surgery is to a large extent dependent on the performance of the individual surgeon. Cancer surgery is also a complex procedure — removal of tumors requires high individual skill — where even small mistakes might result in serious consequences for the patient's health, such as permanent impotence and urinary incontinence following prostate cancer surgery. A surgeon who have performed a large number of surgeries may have more finely developed skills and is better able to deal with potential complications during surgery as well as coping with heterogeneity in the patient population. Moreover, clinics performing large numbers of surgeries may also have better outcomes due to greater accumulated staff experience.⁸

Second, stakes are high in cancer surgery and critical decisions made by surgeons may have lethal consequences for the patient — incentives to learn are therefore strong. Third, focusing on cancer surgery means that we can focus directly on quality improvements in terms of survival following surgery.⁹ Moreover, we are able to study follow-up surgeries and re-admission prevalence as indicators of complications due to mistakes made at the initial cancer surgery. These indicators are also useful in the sense that they provide less dramatic health consequences in contrast to mortality outcomes. Such data on direct mechanisms has, with a few exceptions, been lacking in the previous literature.

⁷For an overview of trends in cancer incidence and mortality, see e.g. Cutler (2008) and Ferlay *et al.* (2013).

⁸This type of organizational learning was proposed already by Arrow (1962), who stated that “...it is the very activity of production which gives rise to problems for which favorable responses are selected over time.”

⁹Focusing directly on quality improvements helps to establish learning-by-doing compared to focusing on costs, since any negative relationship between costs and accumulated production can also be explained by the ability to use less and less skilled labor as more tooling and standardization of procedure is introduced (Thompson, 2012).

Fourth, we can look for heterogeneity in learning-by-doing by comparing results between cancer surgeries of varying degrees of complexity. This allows us to test the hypothesis of whether more learning takes place for more advanced, as compared to more trivial, tasks. As most of the general learning-by-doing literature have focused on learning in the manufacturing industry — which typically consists of manual labor — we can provide more insight of the learning process by estimating separate effects by complexity of the surgery.

Sixth, as cancer surgeries are normally planned months ahead we do not have to worry about potential confounders — such as access to health care — in our analysis. This would be more problematic for more acute types of health events, e.g. acute myocardial infarctions, where any gains from learning due to centralization of in-patient care may be offset by reduced health care access for patients living further away from a hospital.

Our focus on the Swedish public health care system brings some advantages for our empirical analysis. In this setting, we can rule out certain effects of mergers on quality of outcomes that are not mediated through learning-by-doing.¹⁰ Importantly, in a competitive system, consolidation and mergers lead to reduced competition which may negatively affect incentives to improve quality in order to attract patients.^{11,12} This effect thus works opposite to any learning-by-doing effects. In a public health care system, such as the Swedish one, we can rule out such competition effects.

Focusing on the volume-outcome nexus in a public health care system is also of great policy relevance. Most existing evidence for learning-by-doing comes from competitive markets where organizational changes facilitating learning-by-doing mechanisms are likely to be intentionally implemented and exploited by profit-maximizing agents. In

¹⁰A related literature looks at the effects of hospital mergers on various outcomes, such as patient health, staff behavior, and costs, see for instance Ho and Hamilton (2000); Dranove and Lindrooth (2003); Currie *et al.* (2005); Harrison (2011); Gaynor *et al.* (2012a).

¹¹Since US health care consumers face little differences in out-of-pocket expenditures across hospitals it has been suggested that incentives are strong for hospitals to compete for patients on quality dimensions rather than on financial dimensions (Gaynor, 2006).

¹²For studies on the effect of hospital competition, see for instance Kessler and McClellan (2000); Propper *et al.* (2004); Volpp *et al.* (2005); Propper *et al.* (2008); Gaynor *et al.* (2011, 2012b).

the public sector, competition is often heavily restricted and regulations control market entries and exits — implying that efficiency-enhancing organizational changes are more likely to remain unexploited. In many countries, the health care sector is an example of such a heavily regulated market where, at the same time, the stakes are high and in which efficiency gains can be counted in lives saved. Moreover, given the rapidly increasing health care costs experienced by many countries, there is a great need to identify the determinants of productivity in the health care sector. Clearly, evidence of learning-by-doing would have important policy implications, for example with respect to health care concentration policies.

Our results suggest rapid gains of increased surgical volume. In our preferred IV-specification our results imply a death rate elasticity with respect to surgical volume of about 0.22. Relating the size of this effect to the variation in hospital volume across hospitals in Sweden we find that increasing the volume of a given hospital from the 25th to the 50th percentile of the hospital volume distribution would imply an estimated decrease in cancer mortality rate with about 4.1 percentage points, or 18.7 percent. We also find, in line with our expectations, that learning effects increases with the complexity of the procedure — learning effects are greater for prostate and intestine cancer than for breast cancer. Moreover, we find that higher surgical volume reduce the probability of another cancer surgery. Given that multiple cancer surgeries for the same patient may indicate e.g. subsequent metastases that the surgeon may have missed when performing the initial surgery, we interpret this result as a surgeon learning effect.

We consider a number of threats to our IV-design. First, even if changes in volume are exogenous one could argue that the estimated effect of volume on mortality runs not only through learning-by-doing but also through changes in patient characteristics. In particular, the exclusion restriction would be violated if the additional patients whom after a closure are referred to a non-closed cancer clinic have a different underlying mortality risk than other patients treated at the same clinic. The estimated effect of an increased

volume may then also reflect a change in the composition of patients. We deal with this potential problem by *i)* including a battery of individual-level health indicators such as medical history and socioeconomic factors in order to capture heterogeneity in health and *ii)* by exploiting an institutional feature of the Swedish health care system in which individuals are assigned hospital based on their geographical location and not by choice. We are thus able to keep the underlying patient population constant by only studying the outcomes of those who belonged to the hospitals catchment area *before* the inflow of additional cancer patients from other catchment areas, ensuring that the studied population remains the same in terms of both observed and unobserved factors. The results are not affected when we impose this restriction.

Second, the exclusion restriction could fail if the closures or openings of cancer clinics cause more general organizational changes at the remaining hospitals. Any improved patient health outcomes at remaining cancer clinics could then be a consequence of these changes rather than from increased learning in cancer surgery. If this would be the case, we would expect patient outcomes to also be improved in other types of surgeries at the same hospital — even in the absence of increased surgical volume for these other types of surgery types.¹³ We look for signs of such organizational changes by exploiting that surgical volume for other surgery types than cancer does not change when cancer clinics close at nearby hospitals. However, if increased volume also correlates with, for instance, greater resources for all types of surgery at the hospital level, we still expect a reduced form effect on the non-cancer surgery types. Utilizing information on another type of advanced surgery — cardiovascular surgery — we find that closures of nearby cancer clinics do not affect the volume of ischemic heart surgeries performed in remaining hospitals. Moreover, the reduced form effect of closures of cancer clinics on the outcomes of heart surgeries is indistinguishable from zero, hence lending support to our exclusion restriction assumption.

¹³Note, however, that such a pattern could also arise if there are substantial productivity spillovers between clinics at the same hospital and must therefore not necessarily reflect a violation of the exclusion restriction.

Third, closures and openings of cancer clinics may affect the technology used in the remaining cancer clinics as well as the composition of surgeons, which may have an effect over and above any learning effects from increased volume. By exploiting detailed data on surgery procedures and the technology used, such as robot-assisted surgery, we find no support for the former hypothesis. Moreover, by exploiting linked employer-employee data from administrative registers we show that changes in surgeon characteristics is an unlikely explanation for our estimated effects. In fact, we show that an increased operation volume, following from closures of nearby cancer clinics, does not lead to any significant increase in the number of surgeons at remaining hospitals. This finding hence suggests an increased case load per surgeon subsequent to the cancer clinic closures.

Finally, closures and openings also affect the distance to the nearest cancer clinic for cancer patients. Even though we expect the distance to the nearest hospital to have less of an effect on patient outcomes for planned cancer surgeries, we collect data on exact distances from the individual patient's place of residence to the nearest hospitals and use this measure as an additional control in our regression model. As expected, controlling for changes in the distance to nearest hospital do not affect our estimates.

To shed light on some of the mechanisms behind our estimated effects, we consider non-linearities in the estimated volume-quality relationship with respect to the existence of a learning threshold, as is often found in the general learning-by-doing literature. Specifically, surgeons in low-volume hospitals may gain important experience from each additional surgery while, at larger hospitals, surgeons may already have had the opportunity to gain experience through a large number of surgeries.¹⁴ In order to study non-linear effects in more detail, we complement the IV-analysis with a hospital fixed-effects analysis, relying on within-hospital changes in volume over time. Our estimation results suggest that the volume effect is more prominent at lower surgical volumes whereas at

¹⁴One may also argue that increased hospital volume below a certain point could have negative impacts on hospital quality. This would occur if a too high volume leads to less efficient information channels, less efficient decision making, to exhausted medical staff and greater coordination problems.

volumes above 140 p.a. there are practically no volume effects at all, hence supporting the hypothesis that learning-by-doing is relatively more important at lower levels of surgical volume.

To sum up, our paper contributes to two strands of literatures. First, we contribute to the general literature on learning-by-doing by providing new evidence of a causal relation between production volume and product quality as well as on the possible mechanisms through which this effect is mediated. Moreover, we contribute with implications for health care policy as our results largely implies that health care consolidation in the context of a public health care sector may lead to important health gains — at least in the context of planned surgery. This is an important finding, not the least with respect to the particular institutional setting we study, as a large number of hospitals in our data operates below the empirical learning threshold we identify.

The remainder of the article proceeds as follows; Section two outlines the basic characteristics of the cancer surgeries we study, with a particular focus on the scope for a learning-by-doing effect. Section three describes the data we use for estimation and section four provides relevant information on the institutional features of the Swedish health care setting and the rationale for why this particular context provides us with plausibly exogenous variation in hospital volume. Section five presents results from our main analysis of the effect of operation volume on cancer survival along with a number of robustness checks to test the sensitivity of our results. Section six discusses potential mechanisms of the estimated results supplemented with empirical analyses in order to assess these mechanisms. Finally, section seven concludes with a brief discussion about the policy implications of our results.

2 Cancer surgery and Learning-by-doing

In our analyses, we focus on the three most common cancer types of the Western world, i.e. breast, prostate and colorectal cancers. Prostate (breast) cancer is the most common

type of cancer among men (women) in Sweden, accounting for 37 (30) percent of cancer incidence in 2003, respectively. Colorectal cancer is the second most common type of cancer for both sexes, constituting about 11 percent of all cancer diagnoses the same year (Socialstyrelsen, 2007). In general, cancer operations are a common surgical procedure in Swedish hospitals and requires high-skilled medical staff in order to be performed.

Breast cancer is commonly diagnosed via so-called triple assessment consisting of a physical examination, mammography/ultrasound and a biopsy. The biopsy usually includes performing a so-called sentinel node biopsy (SNB) where the sentinel node —the first breast lymph node to which cancer cells are most likely to spread — is diagnosed for cancer. After diagnosis, surgical treatment of the cancer may be performed through either a partial, full or a modified radical mastectomy depending on factors such as the size and spread of the tumor and diagnostic results from the SNB. A partial, or breast-conserving, mastectomy is preferred due to its relatively low invasiveness while a full or radical mastectomy is performed for more complicated and aggressive cancer types. We include all three types of surgery in our analysis.

Surgical treatment of prostate cancer normally involves removing the entire prostate via a radical prostatectomy. We include the two most common types of radical prostatectomy; the retropubic and the endoscopic prostatectomy. The procedures differ by the type of strategy the surgeon uses in order to access the tumor — in the retropubic prostatectomy the surgeon makes a large incision in the abdomen (so-called open surgery) while in the endoscopic prostatectomy several small incisions are made in order to insert surgery instruments into the body (so-called minimally invasive surgery). The latter is also called laparoscopic prostatectomy due to its common use of a laparoscope for visualization and can be done either manually or with the assistance of a robot. We also include two additional palliative surgery procedures in the analysis directed towards slowing the disease progression; Orchidectomy which involves surgically removing one or both testicles in order to reduce the amount of testosterone in the body which increases cancer progres-

sion, and transurethral resection of the prostate (TURP) in which a part of the prostate gland is planed away using a resectoscope in order to facilitate urination for end-of-life patients.

Approximately two-third of all colorectal cancers in Sweden are colon-situated while one-third are located in the rectum. We include the three most common colorectal surgeries from each type in our analysis. The specific procedures of each colorectal cancer type is typically chosen depending on where in the intestine the cancer is situated. For rectal cancer, a low anterior resection (LAR) is performed if the cancer is situated in the upper third part of the rectum while an abdominoperineal resection is typically performed if the tumor is located closer to the anus. While the former procedure usually allows for a so-called anastomosis in which the bowels are sewn together again after the surgery, the latter normally result in a colostomy — an opening in the skin of the abdomen to allow body wastes out. A surgical resection of the rectosigmoid colon, known as the Hartmann's procedure, is sometimes also performed in acute cases when there is no possibility of reconnecting the bowels after surgery. In latter years the procedure known as total mesorectal excision (TME) has become popular in surgeries relating to the lower two-thirds of the rectum, replacing abdominoperineal resection surgery due to both lower cancer recurrence rates but also the possibility of avoiding a permanent colostomy as the procedure involves a reconstruction of the rectum. For colon cancer the typical surgical procedure is to perform a colectomy in which a part of the colon is removed and then re-attached. Specifically, we include the right and left hemi-colectomy, referring to the resection of the ascending (right) or the descending (left) colon, respectively, and the sigmoidectomy, involving resection of the sigmoid colon.

Individual and organizational learning-by-doing

Even with general treatment guidelines, surgeons have high discretion in deciding which type of surgical procedure to apply and how to apply it for each patient, for example in deciding how much tissue to be removed (or marginal to keep). Clearly, this involves

making important trade-off decisions, not only before but also during surgery, as available information regarding the tumor is incomplete and constantly updating. Hence, there exists obvious scope for individual skill in executing successful treatments, not only in terms of survival, but also in other aspects of post-operative patient health such as complications from surgery and tumor recurrences.

From the above discussion it is not far-fetched to argue that individual skill might be related to surgical experience — as cancers and human beings are heterogeneous by nature. A greater number of performed surgeries should increase the probability that a previously encountered cancer type will show up again, which, in turn, could reduce initial information deficiencies regarding the particular characteristics of the disease.

Similarly, organizational learning-by-doing may arise from learning and experience at the organizational level. Surgeons performing surgery are assisted by a team of nurses, other specialists (e.g. gynecologists, radiologists, plastic surgeons and orthopedists) and need proper equipment and space. If the organizational environment at the clinic or hospital is poor, for example if surgery teams are badly composed, this will also inhibit the surgeon's individual performance. Hence, greater experience may potentially lead to improved cooperation in such teams as the organization learn over time who works well with who. All in all, we believe that there is clear potential in cancer surgery for both individual and organizational learning-by-doing in isolation of each other as well as possible interaction effects between the two processes on patient outcomes.

3 Data

Our primary data sources consist of the Swedish National Patient Register (NPR), containing population-wide information on all in-patient care in Sweden, and the National Causes of Death register (NCDR), containing information about all deaths for individuals who had a permanent residence in Sweden. The NPR contains individual-level data on date and hospital of admission and discharge, the nature of the admission such as the

length of stay and whether it was acute or planned as well as rich medical information including main and co-diagnoses (through the International Classification of Diseases, ICD) and information about any medical procedures made in relation to the hospitalization (through the National Classification of Surgical Procedures NCSP¹⁵). The NCDR includes information on the date, place and the underlying cause of death. For both registers we have access to data until 2011.

As a major revision of the NCSP in 1997 made comparisons of procedures over time difficult, we chose to sample cancer surgeries from 1998 and onwards. Furthermore, we set 2007 as our last sampling year in order to have at least a four year follow-up period for each patient to evaluate surgical quality. As mentioned previously, we focus on breast, prostate and colorectal cancers and select the three most common surgical procedures within each category. Specifically, as some of the surgery procedures may also be performed for other diseases than cancer (e.g. TURP for prostatic hyperplasia) we restrict our sample to individuals that have a cancer diagnosis at the time of hospitalization.

As our main outcome variable, we use survival four years after surgery, constructed using exact dates of hospitalization and death. We also vary this measure, allowing for long- and short-run survival outcomes. Using the information in the NPR, we additionally construct a number of other quality-related health outcomes such as re-admission prevalence, multiple cancer surgeries and number of days spent in hospital.

Table 1 presents sample statistics of the three included cancer types and the corresponding surgery procedures. Our sample consists of in total 109,761 observations of which 48 percent are breast cancer surgeries, 28 percent are colorectal cancer surgeries, and the remaining 24 percent are prostate cancer surgeries. The table also indicate a slight increase in the number of surgeries and a substantial decrease in four year mortality over time. On average, 39 percent of the patients in our sample die within four years after

¹⁵NCSP consists of fifteen main chapters containing operations organized by organ system, four sub-chapters, and a chapter with additional codes. The Nordic Medico-Statistical Committee (NOMESCO) published the first printed edition of the NOMESCO Classification of Surgical Procedures (NCSP) in 1996.

having undergone surgery. The four year mortality rate is highest for colorectal cancers (56 percent) and lowest for breast cancer (28 percent). Finally, breast cancer patients are on average substantially younger than colorectal and prostate cancer patients.

4 Empirical strategy

4.1 Descriptive background

Before presenting our empirical strategy in more detail, we illustrate the variation we use in our data graphically. The left panel of *Figure 1* plots the raw correlation between the annual number of cancer surgeries and four year survival rates for each included hospital and year. Each dot in the figure corresponds to the average survival rate for a given hospital-year combination. The smoothed average, estimated from a local polynomial regression and indicated by the solid line in the figure, shows that survival rates are higher for hospitals which performs more annual surgeries. This observed pattern is in accordance with a learning-by-doing hypothesis. However, the positive volume-outcome correlation visible in the descriptive figure could clearly also be explained by other mechanisms, such as patient or staff selection from e.g. selective referral or selection in the underlying case-mix of patients in the catchment areas of small and large hospitals.

The right panel of *Figure 1* describes our data from a more traditional learning-by-doing outset. Specifically, the figure evaluates the existence of a learning curve by plotting the relationship between *cumulated* surgical volume and survival rates.¹⁶ Each dot in the figure corresponds to the average survival rate as a function of cumulated volume in bins of ten surgeries, pooled over all hospitals and years. Interestingly, we see from the figure that the average survival probability increases substantially with cumulative hospital volume, again suggesting a pattern consistent with a learning-by-doing mechanism. However, while interesting, this analysis does not consider the influence of other confounding

¹⁶Accumulated volume is defined as accumulated volume from 1998 and onwards due to the changes in NPR coding that occurred in 1997 (see the data section).

factors such as time trends from general surgical improvements, serially correlated health or cost shocks affecting both number of surgeries and survival and changes over time in hospital quality and selective referral.

Thompson (2012) concludes in his overview of the learning-by-doing literature and the problems of identifying such effects that; “The solution, of course, is to find settings in which cost shocks do not induce changes in input use, or to find instruments for cumulative output”. In the remainder of the paper we continue our empirical investigation by applying such an instrument, hence shifting focus from discussing general volume-outcome associations towards making causality claims.

4.2 The instrument

In order to motivate our instrument we first provide some relevant background information on the Swedish health care system. In contrast to the U.S., the health care sector in Sweden is highly regulated and the vast majority of hospitals are owned and run by the public sector. In fact, the first and still the only private hospital was started as late as in 1999. Moreover, inpatient health care is organized and mainly financed at the regional level. Sweden is divided into 21 regions, with Stockholm as the largest (with about 2 million inhabitants) and Gotland the smallest (about 60,000 inhabitants). Organizing health care is the single most important responsibility for the regional authorities. For instance, in 2012 about 82 percent of the regional budgets were used for health care.

The regions in Sweden are run by a political council, elected every fourth year. Besides following some general laws and guidelines set by the government, e.g. that health care must be provided to all citizens, the regional authorities are more or less free to organize the health care in their region. In effect, since the hospitals are almost exclusively run by the public sector, the members of the regional board have a great influence in deciding the specific organizational features of the regional health care — in particular the degree of centralization.

We exploit the fact that the substantial regional autonomy of the Swedish health care

system lead to very different responses to a general health care centralization trend taking place in the early 2000s. Specifically, the central government passed a law in 2000 stating that the regional councils were no longer allowed to run budget deficits. This law led many of the regions to discuss and propose centralization measures in order to cut costs and increase efficiency but, due to different outcomes from the political negotiations, these proposals were only implemented — and to different extents — in some of the regions (Lindbom, 2013; Larsson Taghizadeh, 2009). In this paper, we utilize closures of cancer surgery clinics generating discontinuities in the number of cancer surgeries at remaining nearby hospitals within the same region across time. Applying these closures ensure that we utilize volume changes derived from political processes rather than surgical volume swings caused by regional health shocks, patient sorting and case-mix selection in hospital catchment areas. Moreover, as we have access to very detailed data stretching over a long period of time, we are able to perform extensive robustness checks to validate that the closures provide us with exogenous variation for our purpose.^{17,18}

Using the NPR we define a cancer surgery clinic as closed (opened) if the total number of surgeries is two or fewer (greater) during three consecutive years. With this definition we obtain seven closures and two openings during our study period.^{19,20} To supplement the data we also validate the closures/openings using information obtained from local newspapers and political protocols.

In order to illustrate the variation we exploit for estimation the left panel of *Figure 2* displays an indexed measure of the number of cancer surgeries performed at the closed clinics before and after the closure. Clearly, there are striking drops in the number of

¹⁷We are also able to rule out competition effects that are important in some other health care markets (see e.g. Bloom *et al.* (2011b)). The reason is that patients in Sweden are normally referred to a designated hospital (usually the closest one) so that competition between hospitals is essentially non-existent.

¹⁸Note that we only consider closures of cancer surgery clinics — not entire hospitals — which means we can rule out general merger effects as documented by e.g. Gaynor (2006).

¹⁹The two openings of two formerly closed hospitals that was re-opened.

²⁰The closed hospitals are Sophiahemmet (2006), Vrinnevisjukhuset (2003), Motala lasarett (2003), Trelleborgs lasarett (2005), Kristinehamns sjukhus (2000), Säfte sjukhus (2000) and Piteå Älvdals sjukhus (2005) and the opened hospitals are Vrinnevisjukhuset (2007) and Motala lasarett (2007).

surgeries at the closure years. Moreover, the right panel of *Figure 2* illustrates how the closures affected remaining nearby clinics in terms of surgical volume — i.e. the first step of our proposed IV strategy. Specifically, the figure plots the average difference in hospital volume for remaining hospitals in regions with closures and for hospitals in regions without any closures, after adjusting for time and hospital fixed effects. For non-closing regions, the year of closure represents the year a clinic *potentially* could have closed in the region — i.e. a placebo closure.²¹ The closures evidently implied a substantial increase in the number of surgeries at nearby, remaining, hospitals in the region (on average about twenty percent) in contrast to non-closing regions where, as expected, volume was unaffected.

Our identification strategy would be invalidated if the political processes leading up to the closures and openings were influenced by demographic changes such as urbanization, business cycle effects or underlying trends in public health and/or in the population mix. In order to explore this potential threat to identification, *Figure 3* presents level values of certain key variables for regions with at least one opening or closure and unaffected regions, before and after the change took place, respectively. The figures reveal that regions affected by a hospital closure or opening on average have larger populations and higher average morbidity than the corresponding unaffected regions. However, no important group differences in unemployment, share of cancer hospitalizations or cancer mortality is distinguishable.

Due to these group differences in observable characteristics we include hospital and calendar time fixed effects and regional linear time trends²² in the analysis. Including these controls ensure that the variation provided by our instrument will be exogenous unless the timing of the openings/closures are correlated with underlying non-linear time trends from unobserved factors. As is evident from *Figure 4*, showing a plot of the resid-

²¹Since the first and last closure in our data occurred in the beginning and the end of the sampling period, respectively, the difference can only be displayed two years before and after the potential closure year.

²²As a robustness analysis we have also estimated models with linear hospital trends instead of regional trends, resulting in practically unchanged main estimates.

uals obtained from estimating linear regressions including this set of controls, the difference between the groups is now indistinguishable for all variables.²³ Hence, this result lends some credibility to our empirical strategy.

The remaining threat to our instrumental variable approach concerns the potential existence of any direct effects of the closures on our outcomes of interest — i.e. whether the closures affected the surgical quality in remaining cancer clinics through other mechanisms than through changes in operation volume. To investigate this possibility, we collect additional data and perform a series of specification checks. In particular, we test if the closures systematically affected the case-mix in the remaining cancer clinics, since these clinics now had to take on a number of additional new patients. A change in the case mix would be problematic if the new patients treated at the remaining clinics had different underlying health characteristics than other patients. For this reason, we run regressions where we exclude the additional patients from our analysis and only focus on the patients that belonged to the remaining clinics catchment area *before* the increase in volume. Importantly, in this case-mix adjusted analysis we exclude individuals treated at closed clinics both before and after the closure. Hence, systematic hospital closures arising from e.g. different health care quality trends across closed and remaining cancer clinics are *by definition* not a problem in this analysis, since closed units are entirely excluded from the analysis.

Moreover, to assess the exclusion restriction further we; *i*) track all surgeons at hospitals with a cancer clinic closure in order to investigate to what extent the composition of surgeons changed in the remaining hospitals after the closure; *ii*) test for effects from potential organizational changes coinciding with the closures and openings using data on other type of surgeries unaffected by the cancer clinic closures; *iii*) collect data on exact distances from each patient's place of residence to all hospitals in order to examine whether hospital distance affects surgery outcomes; and *iv*) examine whether hospital

²³For ease of comparison we have added the variable means to the residuals.

surgery staff change their procedures (by e.g. changing type of surgery) as a response to the additional inflow of new patients.

4.3 Econometric model

We specify a linear probability model for the probability of surviving at least four years after cancer surgery. Formally, our baseline model for survival of individual i undergoing surgery s at hospital h in calendar year t is:

$$y_{ihst} = \lambda_t + \lambda_h + \lambda_s + \log(volume_{ht})\gamma + X_i\beta_X + \varepsilon_{ihst} \quad (1)$$

Here, $volume_{ht}$ is the number of total cancer surgeries performed at hospital h in calendar year t . The coefficient of interest is γ measuring the effect of an additional surgery on surgical quality, i.e. probability of survival. This baseline model controls for general time trends through calendar time fixed effects λ_t , general differences between different hospitals through hospital fixed effects λ_h , and differences across type of surgery through surgery fixed effects λ_s . A set of individual characteristics X_i , such as gender, age and education level and pre-surgery health factors such as the number of previous hospitalizations and time since last hospitalization, is also included in the model to account for individual heterogeneity in survival probability. Finally, we also include region-specific linear time trends in the model.^{24,25}

We initially follow the traditional medical literature and estimate (1) by OLS. We then proceed by estimating FE-IV models applying the first-stage equation

$$\log(volume_{ht}) = \alpha_t + \alpha_h + \alpha_s + \gamma\delta D_{rt}^{closure} + \beta_X X_i + \varepsilon_{ihst}. \quad (2)$$

where our instrument, $D_{rt}^{closure}$, is implemented as an indicator variable, assuming the

²⁴The health variables are constructed using the data on hospitalizations presented in the data section and the socioeconomic variables are constructed using data from Statistics Sweden on the entire Swedish working age population. For older individuals we use the latest available information.

²⁵As a robustness check we also specify surgical volume in levels.

value one if a cancer surgery clinic in region r closed in year t or earlier.²⁶ This captures the idea that closures generate a shift in surgical volume at remaining clinics within the region. Initially, we restrict δ to be a scalar, implying that a closure is constrained to have the same effect on volume for all regions. We then utilize the variation generated by the instrument in a more flexible way by allowing the volume shifts from the closures to vary across regions by specifying δ as a vector; δ_r . This, unrestricted, version of the instrument captures the idea that effects on remaining hospitals' volume may differ depending on the size of the closing clinic.

The model in (1) is a restricted version of the standard power law formulation of learning-by-doing models.²⁷ In our setting the power law model implies that current surgical quality, Q_{ht} , at hospital h in time period t is related to cumulative prior surgical volume, E_{ht} , through the power law specification, $Q_{ht} = BE_{ht}^\beta$, where β measures the rate of learning and B is a constant. Note that we measure quality (survival) and not production cost so that learning-by-doing implies that β is assumed positive. In this model quality increases with a constant factor each time volume is doubled, so that the quality effect from one additional surgery is lower at higher volumes. Our baseline model is a restricted version of this model where only current volume and not cumulative volume affects current quality. Moreover, our log specification assumes that learning is diminishing with respect to volume, i.e. $\beta < 1$.

In order to investigate the learning dynamics in more detail we replace current volume in (1) with cumulated volume in a subsequent analysis. We explore experience in the last year, last two years and full cumulated volume from 1998 and onwards.²⁸ In these analyses we use lagged versions of our instrument in order to instrument for cumulated volume. The intuition is that the volume shifts induced from the closures cumulates over time as the surgeons gain additional experience for each year since the closure occurred.

²⁶That is, $D_{rt}^{closure} = \mathbf{1}[t \geq t_{closure}|r]$ where $t_{closure}$ is the year of closure in region r . Similarly, if a clinic is re-opened the closure dummy takes the value zero in the re-opening year and all subsequent periods.

²⁷See e.g. Thompson (2012) for a discussion of various organizational learning models.

²⁸As the code classification changed in 1997 we cannot track volume back further than this year.

However, this elaborate analysis stretch the limit of what can be estimated using the variation created by closures and openings, which is also why we mainly focus on the effects of current volume in our result section.

5 Main results

5.1 The volume-outcome relationship

We begin this section by presenting results on the raw volume-quality relationship, estimated from a simple bivariate regression of hospital volume on four year survival. As seen from column 1 of *Table 2*, there is a strong and significant volume-quality association in our data. In particular, the volume estimate implies that a hospital performing twice as many cancer surgeries as another hospital has on average about 5.7 percentage points higher four year survival rates.

We next turn to our instrumental variables strategy. Column 2 of *Table 2* presents the first-stage estimates for our instrument when the closure effect is restricted to be homogeneous across regions. The estimated coefficient on the closure indicator from equation (2) is statistically significant and positive, indicating that the closure of a cancer surgery clinic increases the average number of surgeries at remaining nearby clinics in the region. The standard errors are clustered at hospital level since we estimate individual level effects using variation at the hospital level.²⁹ F-statistics for instrument relevance are reported at the bottom of the table. As can be seen, for the restricted IV specification the F-statistic is below ten and hence does not satisfy the usual criterion for a strong instrument.

Next, consider the baseline IV estimate using the restricted IV model presented in column 3 of *Table 2*. The estimated coefficient suggests a sizeable, however not statistically significant, effect of surgical volume on four year survival rate. To improve on inference by increasing precision we switch to the unrestricted version of our instrument

²⁹We have also estimated standard errors clustered at regional level and multi-level standard errors clustered at both hospital and regional level. This exercise shows that clustering at hospital level renders the most conservative inference.

in column 4. The more flexible use of the closure instrument increases the first stage F-statistic by a factor of six resulting in a reduction of standard error of the volume effect point estimate by more than 50 percent. The effect is now highly significant and close to the initial OLS estimates from column 1 of the table. *Table 2* also presents estimates from additional IV specifications using the unrestricted version of our instrument. As patients treated at different hospitals may be heterogeneous with respect to health we add a set of individual-level covariates to the model in column 5. The inclusion of these factors leaves the IV-estimates practically unchanged, suggesting that patients at remaining and closed hospitals have similar observed characteristics. The latter is also confirmed from sample statistics for patients at remaining and closed hospitals one year before the closures presented in *Table 3*.³⁰

In the last column of *Table 2* we also condition on unobservable changes in the case-mix of cancer patients by re-estimating the model on a restricted sample, including only individuals belonging to a remaining hospital's catchment area prior to the closure. The intuition is the following; while we include these additional patients when predicting hospital volume in the first stage equation (our surgical volume measure include all surgeries) we subsequently exclude them in the structural equation. However, even if excluded, the surgeons still perform surgery on these patients so that any learning effects will still be picked up on the outcomes of the included patients.³¹ Excluding patients from closing hospitals have little effect on our volume effect and from this we conclude that our main effects are robust to unobservable changes in the patient case-mix.³²

Is our estimated volume effect also economically significant? The point estimate from our preferred model in the last column of *Table 2* implies that doubling the per annum

³⁰There are some differences with respect to the gender and immigrant mix. Closed hospitals on average have a larger share of female patients and a smaller share of immigrants. These composition differences does not seem to affect the point estimates however.

³¹See section three for a discussion on how patients are assigned hospitals. Descriptive sample statistics show that 86% of the cancer surgery patients are treated at their designated hospital.

³²We have also re-estimated the case-mix adjusted model using volume in levels instead with essentially the same results.

number of surgeries at a hospital will on average increase the survival rate with 4.8 percentage points. The average four year mortality rate in our sample is 0.22 so that the death rate elasticity with respect to surgical volume is about 0.218 — which we regard as a sizable effect. Moreover, relating the estimated effect to the sample variation in hospital surgical volume, we find that increasing the volume of a hospital at the 25th percentile to the 50th percentile of the volume distribution (i.e. from 70 to 130 surgeries p.a.) would decrease the four year mortality rate with about 4.1 percentage points, or 18.7 percent.

Our preferred IV estimate is, while not significantly so, somewhat smaller than the OLS estimate. When interpreting this difference it is important to keep in mind that some factors may bias the OLS estimate upwards and others may bias the estimate downwards. For example, a selective referral story in which high-quality hospitals have a higher volume simply because they attract more patients would imply the former bias, while selective referral in which more severe cases directed to large hospitals will imply the latter bias. Moreover, sorting of patients into areas with varying hospital sizes will also bias the OLS estimate but in an *a priori* unknown direction. Finally, an additional possibility is that the IV estimate captures a LATE effect, i.e. the effect on a subset of the hospital volume distribution.

5.1.1 Learning to avoid mistakes

While an increased survival probability is likely the most important outcome for cancer patients, other important post-surgery outcomes such as subsequent complications are also meaningful quality measures. Moreover, such measures offer an opportunity to study learning effects directly, since an increase in learning should lead to less complications and less re-admissions for additional cancer surgeries.³³ In order to study complications, we use data from the hospitalizations records and construct detailed information on re-admissions and subsequent cancer surgeries. If surgeons can lower re-admission and

³³Note, however, that since the level of invasiveness of the surgical procedure performed may be inversely related to survival probability, survival prospects and post-surgery life quality may sometimes be traded off against each other.

re-operation incidence rates with additional experience, this suggest that greater volume leads to more experience and less mistakes in surgery.

Specifically, we estimate our preferred IV model using the probability of a re-admission within one year and a new cancer surgery within three years from the initial surgery as outcomes.³⁴ The first two columns of *Table 7* present the results from this exercise. The estimate from column 1 suggests a statistically significant reduced probability of a subsequent cancer surgery within three years from the initial surgery. In particular, if hospital surgical volume is doubled, then, on average, the probability of another surgery decreases with about 0.034. With a mean sample probability of follow-up cancer surgery of 0.08, this translates into a sizable effect.³⁵ Finally, the re-admission point estimate also have a negative sign, however indistinguishable from zero at standard levels of significance, which also point in the direction that learning-by-doing not only affects the survival dimension of health care quality but also other aspects of post-surgical life quality.

5.2 Alternative interpretations

Our results so far suggest beneficial effects of an increase in operation volume. While these effects can be interpreted as learning in cancer surgery, there are other interpretations as well. Next, we consider a number of such alternative interpretations.

Organizational changes

If closures of cancer clinics leads to more general organizational changes at hospitals with remaining cancer clinics, this could bias our volume estimate if such changes have a direct effect on survival following cancer surgery. We investigate this by using data on other type of advanced surgeries which were not affected by the cancer clinic closures. Specifically, we utilize in-patient data on the most frequently performed heart surgeries, arguing that

³⁴We have also used other follow-up time periods for the re-admission and re-operation measures which gives similar results.

³⁵These estimates suffer from a competing risk problem since deceased patients cannot undergo additional surgery. However, since we find that higher volumes increases the survival probability then, if anything, our estimates should be biased towards zero.

outcomes following these closely related type of surgeries should also be affected by any general organizational changes, but that the mechanics of this effect would not run through any volume changes (since their volume should be unaffected by a closure of a cancer clinic). Thus, we effectively force the reduced form effect to only include potential direct effects from the closures and not the indirect volume effects from the first stage. Re-estimating the reduced form model using our closure instrument on heart surgery will thus act as an informal test of whether any important organizational changes at the remaining hospitals coincided with the shift in the volume of cancer surgeries.

The last two columns in panel A of *Table 9* present the first-stage and reduced form estimates from estimating our model only on heart surgeries while the first two columns report the baseline cancer surgery results for comparison. The first stage estimate using cancer clinic closures as an instrument for heart surgery volume is small and insignificant indicating, as expected, that the instrument is irrelevant for this particular sample. Moreover, the reduced form estimate is also insignificant and close to zero, implying that there is no direct effect of the closures of cancer clinics on the probability of surviving heart surgery. On the other hand, both the first stage and the reduced form estimates are highly significant for the cancer sample. Taken together, we interpret these results as evidence for the absence of any important general hospital organizational changes that coincided with the cancer clinic closures, which would have violated our exclusion restriction.³⁶

Staff

A closure of a cancer clinic could lead some surgeons to transfer to other hospitals or to become unemployed. At the same time, clinics with a greater inflow of patients may recruit new surgeons in order to meet the increased patient flow. If a large share of the surgeons at closed-down clinics are recruited to other clinics within the same region, our

³⁶Note that the absence of a reduced form effect on heart surgeries also suggests that there is no negative or positive spillover effects from an increase in cancer surgery volume. A negative spillover effect could for instance have resulted from resources being withdrawn for heart surgery in order to meet the increased volume of cancer patients.

exclusion restriction might be violated because of staff compositional changes at the remaining clinics. The estimated positive effect of volume on survival might then simply reflect any greater quality of the new surgeons recruited at the remaining clinics. To address this concern, we gather information on surgeons employed at hospitals who shut down their cancer clinics in order to check to which extent these surgeons migrated to nearby hospitals in the region. To this end we use the Swedish employment register, providing annual information on all hospital employees. Together with information on attained education levels from the Swedish population registry this allows us to identify surgeons at all hospitals and to follow them over time as they switch employment. Moreover, the data also provides background characteristics of each individual.³⁷

We focus on all surgeons at hospitals where a cancer clinic was shut down, distinguishing between surgeons staying at the same hospital and surgeons transferring to other hospitals. In the first row panel A of *Table 8*, we see that 69 percent of all surgeons who were employed at a hospital one year before the hospital closed its cancer clinic remained at the same hospital after the closure occurred. In contrast, 18 and 13 percent of the surgeons transfer to hospitals within and outside of the region in the year of the closure, respectively. In total, however, the transferring surgeons only constitute 2.4 percent of the total number of surgeons in remaining hospital, implying that it is unlikely that the surgeon mix change to any dramatic extent after the closures.³⁸ Even so, we can compare the characteristics of the surgeons who transferred to nearby hospitals to those of already existing surgeons at these hospitals. In the remaining rows of *Table 8* panel A of the table, we see that the labor earnings of surgeons transferring within the same region are higher than the labor earnings of surgeons that stay, perhaps implying that more experienced sur-

³⁷The population registers contains detailed information on type of education. We identify surgeons as individuals with completed medical education specializing in surgery or other surgery related specialties like anesthesia and emergency care. Unfortunately, we are not able to uniquely identify cancer surgeons in the data.

³⁸In Panel B we examine transfers of surgeons three years before the closures. In these years, only six percent of all surgeons at hospitals closing down three years later transferred within the region. Hence, this result strengthens the conclusion that there were no excess transfer of surgeons following the cancer clinic closures.

geons are more likely to leave the closing hospital. However, since the average earnings of transferring surgeons are significantly lower than average earnings of surgeons at the remaining hospitals, where operation volume increases, there is nothing that suggests that the surgeons that transfer have higher than average productivity in comparison with the pool of already existing surgeons. If anything, the opposite is more likely to the case, since transferring surgeons are about the same age on average but have lower earnings. The main message from this exercise is that changes in the composition of surgeons are unlikely to explain our estimates.

Another way of testing for potential changes in the staff composition is to directly estimate the effect from a cancer clinic closure on the number of surgeons employed at remaining hospitals.³⁹ If the increase is negligible we would be less worried that we pick up changes in the staff composition in our estimates. This is indeed what we find, as shown in column 3 of *Table 7*. A closure does not lead to any significant increase in the number of surgeons in the remaining hospitals. This finding goes well in line with the observation that transferring surgeons only constitute 2.4 percent of all surgeons at remaining hospitals. These results also suggest that each individual surgeon has to take on more patients and, hence, that there is scope for learning at the individual level. We will come back to this question when discussing individual versus organizational learning below.

Distance

Shutting down a cancer clinic also means changing the distance to the nearest clinic for some patients. Even though the distance to a clinic should be less important for the outcome of cancer surgeries — which are normally planned months ahead — we investigate if changes in distance, implied by the closures, affect our results. For this, we merge detailed geographical information from Statistics Sweden on the place of residence for

³⁹Note that any additional surgeons recruited do not necessarily have to come from other nearby shut down cancer clinics. By focusing directly on the number of new surgeons recruited, we can include any surgeons, irrespective where they were recruited from.

each patient in our sample down to the level of a single building block. Together with geographical data on all hospitals, this allows us to calculate the exact distance to each hospital. As we only have geographical data for the working-age population, i.e. for individuals younger than 65 years, we proxy hospital distance for older patients using the average distance to the hospital for younger patients living in the same municipality.

In Panel B of *Table 9* we present the IV estimates including hospital distance (in kilometers) as a control variable in the model. We use two different measures of distance; distance to the treating hospital (reported in column 2 of the table) and distance to the nearest hospital (reported in column 3). For comparison, column 1 reports the baseline estimate from *Table 2* without controlling for hospital distance. As the point estimates for the volume effect are similar in all three specifications, we conclude that changes in the distance to hospitals do not affect our results.

Technology and queues

If the inflow of additional patients to remaining cancer clinics also affects the clinic's organizational structure, but not necessarily the hospital's organizational structure, the volume estimates might partly capture the impact of such organizational responses. We investigate two such potentially important mechanisms; changes in the type of surgical procedures performed (e.g. from more advanced and time-consuming procedures to simpler and quicker ones) and changes in waiting times.

To evaluate whether surgeons change their choice of surgical procedure as a consequence of the increase in patient inflow, we use information on the invasiveness of the surgery; partial mastectomies are less invasive but requires more time than full or radical mastectomies and prostatectomies can either be done using the more invasive open, retropubic, surgery or via the more time-consuming, but minimally invasive, laparoscopic procedure. We estimate our IV model separately for each type of cancer using as outcome an indicator variable for whether the surgery was performed using the less invasive procedure. The results presented in panel D of *Table 9* show no significant effects of the

closures on the choice of procedure. In addition, we find no effect of volume on the probability of using a more sophisticated technology; robot-assisted surgery. The latter result is discussed in more detail in Section 6.2, where we investigate potential scale effects.

We look for signs of increased waiting times as volume increase by examining whether the average age of patients is affected by the additional patient inflow. If waiting times increase, we expect to see a higher average age of the patients at the time they undergo surgery.⁴⁰ Panel C of *Table 9* reports the results from re-estimating our preferred IV-model using patient age at surgery as outcome. The point estimate is insignificant and close to zero and hence do not indicate that patients are treated at a later (or earlier) stage due to increased workload. Hence, we conclude that neither selection in surgery procedures nor increased waiting times are likely explanations for the volume-quality pattern we document in this paper.⁴¹

6 Extensions and mechanisms

We have so far documented that greater surgical volume leads to improved surgical quality outcomes in term of both increased survival prospects and fewer post-surgery complications among treated cancer patients. We have also ruled out a number of alternative explanations for our findings. In this section, we further examine the mechanism(s) and potential heterogeneity behind the estimated effects.

6.1 Effect dynamics and learning

In the former section we studied the effect of current volume on current outcomes. This analysis confirmed that recent experience has an important impact on surgical quality

⁴⁰Note, however, that, if anything, we would expect a higher age at surgery to be associated with *lower* survival, meaning that our estimated effect would be downward biased.

⁴¹In addition, we have looked for changes in patient characteristics by replacing our main outcome variable, survival, by various socio-economic and demographic factors. If clinics respond to an increase in volume by prioritizing certain patient groups, we would expect to an effect of volume on the characteristics of the patients treated. We obtain no signs of this, however, when studying the income or education of the treated patients as a function of volume.

in terms of post-surgery survival prospects. We next study the effect dynamics in more detail within our empirical framework by investigating whether learning-by-doing persists over time, i.e. whether cumulative experience matters. To this end, we re-estimate our preferred specification from the last column of *Table 2* and instrument cumulated volume over one, two, and three years using current and lagged closure status during the last one, two and three years as instruments.

The results from this exercise are given in panel A of *Table 4*. First-stage F-statistics are reported for both the joint inclusion of all instruments and for each instrument separately. While the joint inclusion of the instruments is always highly significant, the three-year lagged instrument is somewhat weak, resulting in more imprecise estimates. Nevertheless, the estimated volume effects show an interesting pattern with greater quality effects from cumulated, relative to current, volume. This result hence suggests that cumulative experience may have an additional impact on current surgical quality above and beyond the immediate experience acquired by the surgeon.

To further investigate the relationship between current and cumulative experience and surgical quality we include both current volume and cumulated volume as two separate explanatory variables in panel B of *Table 4*. While the standard errors are high we see an interesting pattern from the point estimates reported in the table; as experience is accumulated over several years, the estimated proportion of the effect on current experience diminishes in relation to the cumulated experience. This effect pattern is expected in a situation where learning from experience is, at least partially, kept and not forgotten over time, see e.g. Benkard (2000).

6.2 Learning-by-doing versus scale effects

Besides learning, our estimated volume effects could in principle also arise from scale effects, in particular from the utilization of more advanced technology that allow for scale economics. Note that scale effects are likely of less importance in advanced cancer surgery compared to other more standardized procedures. In the end, successful cancer

surgery to a large extent hinges on the skills of the individual surgeon. If, however, larger surgical volume means that clinics can exploit scale economics — so that any fixed costs will be distributed over a greater number of surgeries — the clinic could afford more expensive technologies which could potentially could improve surgery outcomes. We investigate potential scale effects by using complementary information on perhaps the most obvious candidate for such a technology; robot assisted surgery.

A surgery robot is an expensive and advanced type of technology which could potentially lead to improved surgery outcomes. We conjecture that, if scale effects exist, one indication for their existence would be an increased usage of surgery robots in remaining hospital cancer clinics after a nearby cancer clinic closes. In our sample, robots are only used in the treatment of prostate cancers and we estimate the probability that a cancer surgery is performed with the assistance of a surgery robot both for the aggregated sample and separately for the relevant surgery types, using our preferred IV specification.⁴² The results are reported in *Table 5*, showing no statistically significant increase in the probability of using a robot after a nearby clinic closes.

Another way to empirically separate learning-by-doing and scale effects would be to study the effect dynamics in the remaining clinics in a region in which a closure occurred. Hypothetically, volume effects from learning-by-doing should emerge gradually as surgeons accumulate experience while scale effects, on the other hand, should be static and shift quality immediately (at least for some technologies). While the evidence obtained from estimating cumulative volume effects gave us some initial support for the learning-by-doing hypothesis, we now take the analysis one step further by estimating the reduced form of our restricted model, allowing for a time-varying effect of the closures in order to capture their impacts one and two years after the closure occurred. If the effect persists several years after the closure occurred, we conjecture that it cannot be driven solely by scale effects. *Table 6* presents the results from this exercise; the first column report the

⁴²On average about five percent of all prostate cancer surgeries are performed with the assistance of a robot in our sample.

baseline reduced form point estimate for all years while column 2 separates this effect into the first and the second year after closure. Clearly, as the point estimate is much higher for the second year after the closure we conclude that it is unlikely that scale effects could explain all of the volume effect we identify in the data.

Complexity of the surgery

Learning is likely to be more rapid for complex rather than for simple tasks. We investigate this conjecture by estimating the volume effects separately for each cancer type, making use of the fact that breast cancer surgery is generally considered a less complicated procedure than prostate and colorectal cancer surgeries. This is also visible from the significantly higher average four year mortality for the latter two types of cancer, as compared with breast cancer. If the effect of volume is smaller for breast cancer surgery, this is consistent with a learning interpretation. Also note that, if any scale effects exist, they should also exist for less complicated procedures.

The results from separate estimation of the different cancer types are presented in *Table 10* where the results show a positive, however non-significant (due to decreased sample sizes), volume effect for all three types of surgeries. Even if imprecisely estimated, the lower point estimate for breast cancer surgery — compared to colorectal and prostate cancer surgeries — indicates a positive relationship between complexity and learning. This finding is expected if learning is more rapid for more advanced tasks.

Non-linear effects

We have so far focused on linear volume effects, where our IV estimate captures a weighted average effect for hospitals of different sizes. Non-linear effects are hence difficult to investigate using this strategy. In order to analyze non-linearities, we instead exploit a hospital fixed-effects specification and relying on within-hospital variation over time to

estimate the volume effects.⁴³ This way we are able to specify a more flexible model by dividing volume into volume bins of twenty and estimate separate effects for each bin; i.e. we estimate

$$y_{ihst} = \lambda_t + \lambda_h + \lambda_s + volume_{ht}^{bin} \gamma_{bin} + X_i \beta_X + \epsilon_{ihst} \quad (3)$$

where $volume_{ht}^{bin}$ is a set of indicator variables for each volume bin, where $bin = 1 - 20, 21 - 40, \dots, 480 - 500$. The estimated coefficient vector, γ_{bin} , normalized so that the coefficient for the first volume bin is zero, is plotted in *Figure 5* along with a local polynomial smoothed line. The figure reveals interesting non-linearities of the volume effect, resembling a traditional learning curve. Specifically, the volume gradient is high at low volumes, diminishes over the volume distribution and reaches an empirical “learning threshold” for hospital volumes over 140 surgeries p.a. after which there are practically no volume effects left in our data.⁴⁴ As a large proportion of hospitals in our data operates below the threshold of 140 surgeries, this result lends some support for the idea that concentration of cancer clinics may have a substantial effect on the quality of cancer surgery in Sweden.

Individual or organizational learning

To what extent do our estimated learning effects reflect experience embodied in individual surgeons and to what extent do they reflect experience obtained at the level of the organization?⁴⁵ This question relates closely to the individual vs. firm-specific human capital dichotomy in the labor economics literature. In our application, this distinction would mean that important sources of productivity are lost as surgeons leave the hospital

⁴³In this empirical setup, we need to assume that patients do not sort themselves into different hospitals based on hospital quality trends. However, as we are only interested in the volume-quality relationship over the volume distribution, and not the average volume effect, this may be a more innocuous assumption in this setting.

⁴⁴Note that we do not constrain the effect to be diminishing in this analysis by using the log transformation of volume in the regression model.

⁴⁵See also the discussion on learning-by-doing in surgery in section three.

if individual human capital is the driving force behind the learning effects. On the other hand, if learning is primarily at the organization level, the productivity loss would mainly be carried by the individual surgeon.

While an increase in the number of operations per surgeon following from a closure of a nearby cancer clinic means that there is potential for learning at the individual level, organizational learning might operate simultaneously. In order to separate individual from organizational learning in our data, we therefore follow Levitt *et al.* (2012), who argue that if organizational learning-by-doing is important, learning should be the same at organizations with high and low employee turnover. The intuition is simple; if the entire organization learns from experience, then the fact that employees are constantly replaced at high turnover organizations should not affect the relationship between volume and quality whereas, if experience is mainly embodied in individual employees, we should observe less learning in organizations with a high degree of turnover.

We test the turnover hypothesis using supplementary information from employment records, which allows us to compute each surgeon's tenure. Based on average surgeon tenure, we separate hospitals into groups of low and high turnover and subsequently estimate the fixed-effects model from the previous section separately for each hospital group. The result, shown in *Figure 6*, clearly shows that the volume effects are mainly driven by the low turnover clinics, consistent with a mechanism where the learning effects are mainly driven by individual surgeon experience. As suggested by the graph, there is no evidence of any positive relationship between volume and survival for hospitals with high turnover. This is in great contrast to the line showing the corresponding relationship for hospitals with a low turn-over; here we see a rapidly increasing survival as volume increases, which levels off at about 140 surgeries, consistent with the results shown in *Figure 5*. The results are different, however, in that the magnitude of the effect is much greater at hospitals with low turnover. In sum, these results are in line with an interpretation of our estimates as reflecting that learning to an important extent is embodied in

individual surgeons.

Production costs

Finally, besides improving patient health outcomes, it is also important for health care authorities to know whether health care production costs are affected by learning-by-doing. One important source of hospital costs is the length of the hospital admissions associated with the surgery. Increased learning might reduce costs for a hospital due to shorter admission spells as the organization become more efficient at performing surgery. We analyze this by estimating our IV model using information on the number of days in hospital for each patient in association with the surgery as an outcome. From column 4 of *Table 7* we see that the number of hospital days does not change significantly when volume increases. We conclude that this important aspect of health care production cost is not affected to any important extent by increased efficiency from learning.

7 Conclusions

This paper documents a causal relationship between production volume and quality in advanced cancer surgery, using data on more than 100,000 treatment episodes in Swedish public hospitals. We find that an increase in surgical volume increases survival rates and leads to fewer post-surgical complications. The effect is greater at the lower end of the volume distribution and for more complex procedures, consistent with a learning-by-doing hypothesis in which experience from treating highly heterogeneous patients plays a fundamental role in the learning process. We show that the effects are not driven by selective referral or self-selection, changes in the patient population, changes in operation procedures or technology, or by organizational changes at the hospital level.

Our paper contributes to several strands of literatures; first, our findings relate to the medical literature on the operation volume-outcome nexus. By exploiting the specific features of the public health care system in Sweden, we are able to overcome some of the

empirical difficulties that earlier studies have struggled with. In contrast to the U.S. health care sector, where market structure — and thus hospital volume — is an endogenous outcome of a competitive process, the institutional context of the Swedish health care sector provides us with variation in hospital volume that is not driven by such forces. Instead, closures of cancer clinics, generating large shifts in surgical volume, are in our context driven by political and bureaucratic considerations that we show are unrelated to underlying population health and cancer morbidity.

Second, our findings relate to the general literature on learning-by-doing and productivity growth. We contribute to the small literature that provides causal evidence of learning and on some of specific mechanisms through which it operates. In contrast to the previous literature's focus on learning in predominantly manual manufacturing industries, we provide evidence of rapid learning in a high-skill context with high stakes. Moreover, we focus directly on quality improvements, such as survival rates, subsequent surgeries and re-admissions, whereas most previous literature has focused narrowly on production costs.

Finally, our results have implications for the organization of health care in publicly provided health care systems. Most previous studies on the volume-outcome nexus in the health care sector have taken place in the context of a competitive system where competition effects and determinants of hospital size are interconnected in complicated ways. Our results also contrast to some of the recent findings in the literature on hospital mergers. Gaynor *et al.* (2012a), for instance, find no positive effect of mergers on quality outcomes in the British NHS. It is important to note that the literature on hospital mergers usually focus on mergers of *entire* hospitals, whereas we focus on one particular activity where there are strong reasons to believe that mergers can lead to rapid learning. When entire hospitals merge, there is an increase in the production of all kinds of operations, also in those where there is no scope for learning, or for those in which volume may exceed some upper threshold where quality may instead decline. Moreover, and as suggested by

Gaynor *et al.* (2012a), hospital mergers reduce competition, which may have a negative effect of quality outcomes. In the Swedish public health care sector, such competition effects can be ruled out.

Our findings support recent organizational changes in the Swedish health care system (and in other countries) where mergers in specific areas, such as certain complicated types of surgery and transplanting, have recently been implemented with the aim of improving quality (Socialstyrelsen, 2005). We find that many cancer clinics operate at volume levels where there is substantial scope for learning, suggesting that policy makers can improve patient outcomes by encouraging mergers and consolidation in cancer surgery. One way of accomplishing such changes is to enforce minimum volume thresholds that clinics are not allowed to operate below. Such thresholds have already been called for by researchers and policy-makers alike (Epstein, 2002; Shahian and Normand, 2003). The results obtained in this paper lend some support in favor of such policies.

References

- ACEMOGLU, D. and FINKELSTEIN, A. (2008). Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector. *Journal of Political Economy*, **116** (5), 837–880.
- ANGRIST, J. and KRUEGER, A. (2001). Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives*, **15** (4), 69–85.
- ARROW, K. (1962). The Economic Implications of Learning by Doing. *Review of Economic Studies*, **29** (3), 155–173.
- BANDIERA, O., BARANKAY, I. and RASUL, I. (2007). Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment. *Quarterly Journal of Economics*, **122** (2), 729–773.
- , — and — (2009). Surgeonocial Connections and Incentives in the Workplace: Evidence from Personnel Data. *Econometrica*, **74** (4), 1047–94.
- BENKARD, L. (2000). Learning and Forgetting: The Dynamics of Aircraft Production. *American Economic Review*, **90** (4), 1043–1054.
- BIRKMEYER, J., STUKEL, T., SIEWERS, A., GOODNEY, P., WENNBERG, D. and LUCAS, F. (2003). Surgeon Volume and Operative Mortality in the United States. *New England Journal of Medicine*, **349** (22), 2117–2127.
- BLOOM, N., EIFERT, B., MAHAJAN, A., MCKENZIE, D. and ROBERTS, J. (2011a). *Does Management Matter? Evidence from India*. Working Paper 16658, NBER.
- , PROPPER, C., SEILER, S. and VAN REENEN, J. (2011b). *The Impact of Competition on Management Quality: Evidence from Public Hospitals*. Discussion Paper 983, CEP.
- CLEMENS, J. and GOTTLIEB, J. (2014). Do Physicians’ Financial Incentives Affect Treatment Patterns and Patient Health?, Forthcoming in *American Economic Review*.
- CROXSON, B., PROPPER, C. and PERKINS, A. (2001). Do Doctors Respond to Financial Incentives? UK Family Doctors and the GP Fundholder Scheme. *Journal of Public Economics*, **79**, 375–398.
- CURRIE, J., FARSI, M. and MACLEOD, B. (2005). Cut to the Bone?: Hospital Takeovers and Employment Contracts. *Industrial and Labor Relations Review*, **58** (3), 471–494.
- , GRUBER, J. and FISCHER, M. (1995). Physician Payments and Infant Mortality: Evidence from Medicaid Fee Policy. *The American Economic Review*, **85** (2), 106–111.
- CUTLER, D. M. (2008). Are We Finally Winning the War on Cancer? *Journal of Economic Perspectives*, **22** (4), 3–26.

- DARR, E. D., ARGOTE, L. and EPPLE, D. (1995). The Acquisition, Transfer, and Depreciation of Knowledge in Service Organizations: Productivity in Franchises. *Management Science*, **41** (11), 1750–1762.
- DRANOVE, D. and LINDROOTH, R. (2003). Hospital Consolidation and Costs: Another Look at the Evidence. *Journal of Health Economics*, **22** (6), 983–997.
- EPSTEIN, A. (2002). Volume and Outcome — It is Time to Move Ahead. *The New England Journal of Medicine*, **346** (15), 1161–1164.
- FARLEY, D. E. and OZMINKOWSKI, R. (1992). Volume-Outcome Relationships and in Hospital Mortality: The Effect of Changes in Volume over Time. *Medical Care*, **30** (1), 77–94.
- FERLAY, J., STELIAROVA-FOUCHER, E., LORTET-TIEULENT, J., ROSSO, S., COEBERGH, J. W., COMBER, H., FORMAN, D. and BRAY, F. (2013). Cancer Incidence and Mortality Patterns in Europe: Estimates for 40 Countries in 2012. *European Journal of Cancer*, **49** (6), 1374–403.
- GAYNOR, M. (2006). *What Do We Know About Competition and Quality in Health Care Markets?* Working Paper 12301, NBER.
- , LAUDICELLA, M. and PROPPER, C. (2012a). Can Governments do it Better? Merger Mania and Hospital Outcomes in the English NHS. *Journal of Health Economics*, **31**, 528–543.
- , MORENO-SERRA, R. and PROPPER, C. (2011). Competition Could Substantially Benefit Healthcare. *British Medical Journal*, **343**.
- , — and — (2012b). Can competition improve outcomes in uk health care? lessons from the past two decades. *Journal of Health Services Research & Policy*, **17**, 49–54.
- , REBITZER, J. and TAYLOR, L. J. (2004). Physician Incentives in Health Maintenance Organizations. *Journal of Political Economy*, **112** (4), 915–931.
- , SEIDER, H. and VOGT, W. (2005). The Volume-Outcome Effect, Scale Economies, and Learning-by-Doing. *The American Economic Review (Papers and Proceedings)*, **95** (2), 243–247.
- GOWRISANKARAN, G., HO, V. and TOWN, R. (2006). *Causality, Learning and Forgetting in Surgery*. mimeo 17608, University of Minnesota.
- HALM, E. A., LEE, C. and CHASSIN, M. R. (2002). Is Volume Related to Outcome in Healthcare? A Systematic Review and Methodological Critique of the Literature. *Annals of Internal Medicine*, **137** (6), 511–520.
- HAMILTON, B. and HAMILTON, V. (1997). Estimating Surgical Volume-Outcome Relationships Applying Survival Models: Account for Frailty and Hospital Fixed Effects. *Health Economics*, **6** (4), 383–395.

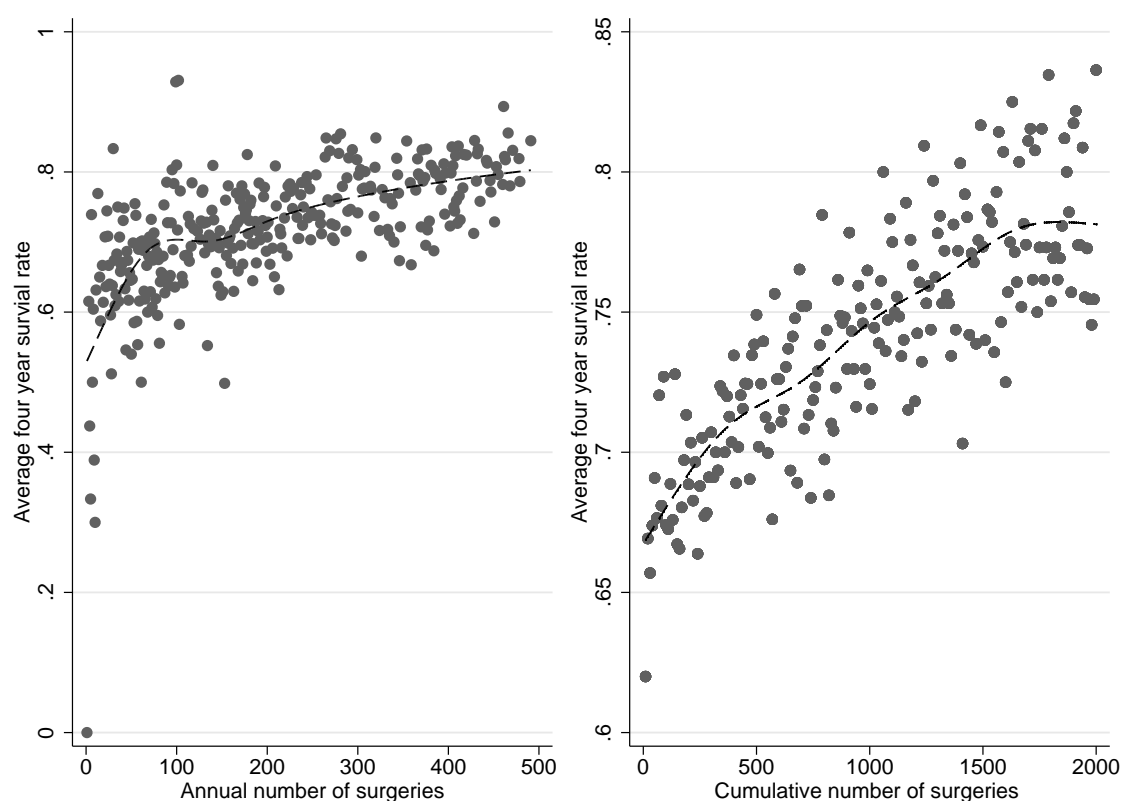
- and HO, V. (1998). Does Practice Make Perfect? Examining the Relationship between Hospital Surgical Volume and Outcomes for Hip Fracture Patients in Quebec. *Medical Care*, **36** (6), 892–903.
- , NICKERSON, J. A. and OWAN, H. (2003). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*, **111** (3), 465–497.
- HARRISON, T. D. (2011). Do Mergers Really Reduce Costs? Evidence from Hospitals. *Economic Inquiry*, **49**, 1054–1069.
- HO, V. and HAMILTON, B. (2000). Hospital Mergers and Acquisitions: Does Hospital Consolidation Harm Patients? *Journal of Health Economics*, **19**, 767–791.
- HOSSAIN, T. and LIST, J. A. (2009). *The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations*. Working paper 15623, NBER.
- HUGHES, R., GARNICK, D., LUFT, H., MCPHEE, S. and HUNT, S. (1988). Hospital Volume and Patient Outcomes. The Case of Hip Fracture Patients. *Medical Care*, **26** (11), 1057–1067.
- JOVANOVIĆ, B. and NYARKO, Y. (1995). A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves. *Brookings Papers on Economic Activity, Microeconomics*, **1**, 247–305.
- KAHN, J., TEN HAVE, T. and IWASHYNA, T. (2009). The Relationship between Hospital Volume and Mortality in Mechanical Ventilation: An Instrumental Variable Analysis. *Health Services Research*, **44** (3), 862–879.
- KESSLER, D. and MCCLELLAN, M. (2000). Is Hospital Competition Socially Wasteful? *Quarterly Journal of Economics*, **115** (2), 577–615.
- KIZER, K. (2003). The Volume-Outcome Conundrum. *New England Journal of Medicine*, **349** (22), 2159–2161.
- KLEIBERGEN, F. and PAAP, R. (2006). Generalized Reduced Rank Tests Using the Singular Value Decomposition. *Journal of Econometrics*, **133** (1), 97–126.
- LARSSON TAGHIZADEH, J. (2009). *Are Client Concerns Invisible and Weak in Sweden? About Protest Movements and their Influence on the County Council's Policy*. Master thesis, Department of Political Science, Uppsala University.
- LAZEAR, E. P. (2000). Performance Pay and Productivity. *American Economic Review*, **90** (5), 1346–1361.
- LEVITT, S. D., LIST, J. A. and SYVERSON, C. (2012). *Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant*. Working paper 18017, NBER.

- LINDBOM, A. (2013). *Waking up the Giant? Hospital Closures and Electoral Punishment in Sweden*. mimeo, Department of Political Science, Uppsala University.
- LUCAS, R. E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, **22** (1), 3–42.
- LUFT, H. S., HUNT, S. and MAERKI, S. (1987). The Volume-Outcome Relationship: Practice-Makes-Perfect or Selective-Referral Patterns? *Health Services Research*, **22** (2), 157–182.
- LUNDBERG, E. (1961). *Produktivitet och Räntabilitet*. Stockholm: Norstedt & Söner.
- NORTON, E. C., GARFINKEL, S. A., MCQUAY, L. J., HECK, D. A., WRIGHT, J. G., DITTUS, R. and LUBITZ, R. M. (1998). The Effect of Hospital Volume on the In-Hospital Complication Rate in Knee Replacement Patients. *Health Services Research*, **33** (5), 1191–1210.
- PROPPER, C., BURGESS, S. and GOSSAGE, D. (2008). Competition and Quality: Evidence from the NHS Internal Market 1991-9. *The Economic Journal*, **118** (525), 138–170.
- , — and GREEN, K. (2004). Does Competition Between Hospitals Improve the Quality of Care? - Hospital Death Rates and the NHS Internal Market. *Journal of Public Economics*, **88**, 1247–1272.
- , SUTTON, M. and WHITNALL, C. (2010). Incentives and Targets in Hospital Care: Evidence from a Natural Experiment. *Journal of Public Economics*, **94**, 318–335.
- and VAN REENEN, J. (2010). Can Pay Regulation Kill? Panel Data Evidence on the Effect of Labor Markets on Hospital Performance. *Journal of Political Economy*, **118**, 222–273.
- SHAHIAN, D. and NORMAND, S. (2003). The Volume-Outcome Relationship: From Luft to Leapfrog. *Annals of Thoracic Surgery*, **75** (3), 1048–1058.
- SOCIALSTYRELSEN (2005). *Nationell Samordning av Rikssjukvården*. Tech. rep., Socialstyrelsen, Proposition to the Parliament 2005/06:73.
- (2007). *Cancervården i Sverige. Kvalitet, Struktur och Aktuella Utmaningar*. Tech. Rep. 2007, Socialstyrelsen.
- THOMPSON, P. (2001). How Much Did the Liberty Shipbuilders Learn? New Evidence for an Old Case Study. *Journal of Political Economy*, **109** (1), 103–137.
- (2012). The Relationship Between Unit Cost and Cumulative Quantity and the Evidence for Organizational Learning-by-Doing. *Journal of Economic Perspectives*, **26** (3), 203–224.

- THORNTON, R. and THOMPSON, P. (2001). Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding. *American Economic Review*, **91** (5), 1350–1368.
- TSAI, A., VOTRUBA, M., BRIDGES, J. and CEBUL, R. (2006). Overcoming Bias in Estimating the Volume-Outcome Relationship. *Health Services Research*, **41** (1), 252–264.
- VOLPP, K., KETCHAM, J., EPSTEIN, A. and WILLIAMS, S. (2005). The Effects of Price Competition and Reduced Subsidies for Uncompensated Care on Hospital Mortality. *Health Services Research*, **40** (4), 1056–1077.
- WRIGHT, T. (1936). Factors Affecting the Cost of Airplanes. *Journal of Aeronautical Sciences*, **3** (4), 122–128.

Tables and figures

Figure 1: The observed relationship between surgical volume and cancer survival



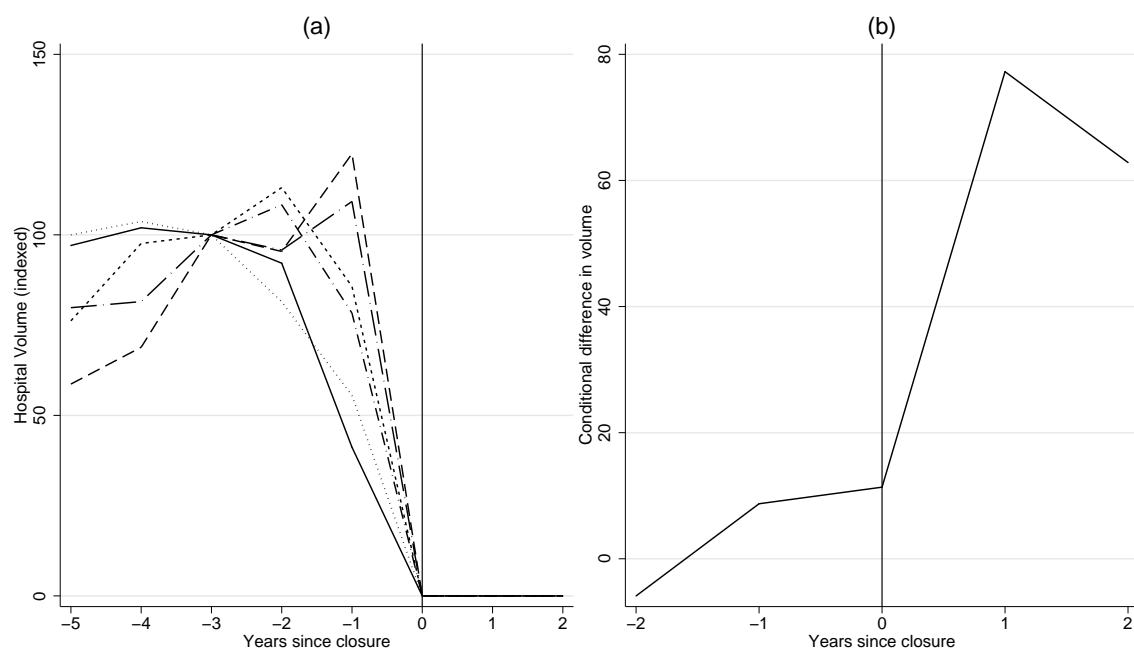
NOTE.—The left panel (a) relates the hospital annual number of cancer surgeries to the probability of being alive four year after a cancer surgery for all individuals in the analysis sample. The right panel similarly plots the average four year survival rate by the cumulative number of surgeries performed at each hospital since 1998 in the sample — i.e. by order of hospital admission date. The dots represent surgical volume averages in bins of ten and the dashed line represent the estimated relationship using local polynomial smoothing techniques.

Table 1: Estimation sample statistics

	Frequency	Fraction of sample	Mean age	4 year death rate
<i>Full sample</i>	109,761	1.00	66.26	0.39
1998	9,707	0.09	66.38	0.58
1999	9,914	0.09	66.73	0.55
2000	10,450	0.10	66.69	0.51
2001	10,800	0.10	66.39	0.47
2002	11,225	0.10	66.63	0.43
2003	11,778	0.11	66.22	0.38
2004	12,119	0.11	65.87	0.33
2005	12,568	0.11	66.02	0.28
2006	12,492	0.11	65.83	0.25
2007	8,708	0.08	65.99	0.21
<i>Breast cancer</i>	52,363	0.48	62.03	0.28
Partial mastectomy	25,953	0.24	59.75	0.18
Full mastectomy	10,022	0.09	66.22	0.37
Radical mastectomy	16,399	0.15	63.08	0.39
<i>Intestine cancer</i>	30,626	0.28	70.90	0.56
Right hemi-colectomy	10,903	0.10	73.15	0.58
Left hemi-colectomy	2,086	0.02	70.28	0.53
Sigmoidectomy	5,199	0.05	70.22	0.55
Rectum resection (LAR/TME)	7,342	0.07	68.29	0.47
Hartmann's operation	1,679	0.02	74.41	0.78
Abdominoperineal resection	3,473	0.03	69.03	0.59
<i>Prostate cancer</i>	26,772	0.24	69.22	0.43
Retropubic prostatectomy	11,525	0.11	62.64	0.09
Endoscopic prostatectomy	2,245	0.02	61.53	0.03
Transurethral resection	10,015	0.09	75.79	0.78
Orchidectomy	2,920	0.03	78.59	0.89

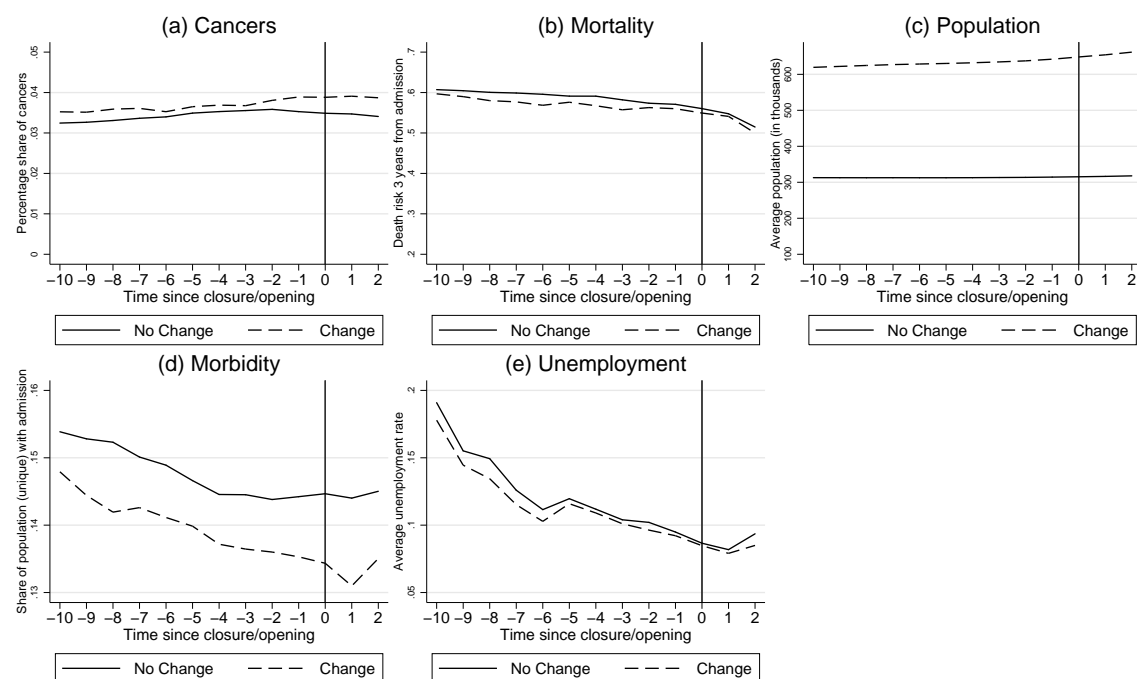
NOTE.— Definition of surgeries are discussed in section 3 of the paper and classified according to the Nordic Classification of Surgical Procedures, NCSP: Partial mastectomy (HAB40) Full mastectomy (HAC20) Radical mastectomy (HAC22) Right hemi-colectomy (JFB30) Left hemi-colectomy (JFB43) Sigmoidectomy (JFB46) Rectum resection (LAR/TME) (JGB00) Hartmann's operation (JGB10) Abdominoperineal resection (JGB30) Retropubic prostatectomy (KEC00) Endoscopic prostatectomy (KEC01) Transurethral resection (KED22) Orchidectomy (KFC10).

Figure 2: Hospital closures and surgical volume in remaining regional hospitals



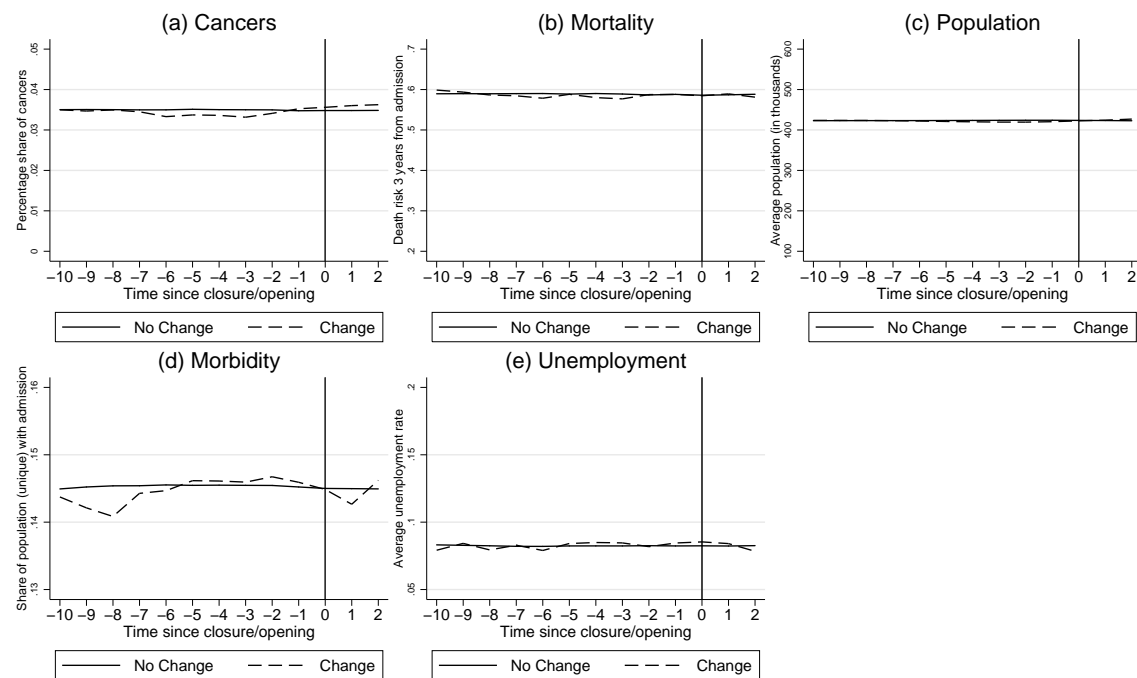
NOTE.—The left panel (a) plots the number of surgeries — indexed by the surgical volume three years before closure (=100) — for each *closing* cancer clinic by years from the closure. The right panel (b) plots the average difference in surgical volume for all *remaining* clinics in each county with a closure (dashed line) and the corresponding average surgical volume for unaffected clinics by years from the closure, adjusted by hospital and calendar time fixed effects. Unaffected regions are used as a *potential* closure for each year a change occurred in an affected region.

Figure 3: Aggregate trends in health and socioeconomic statistics in regions affected and unaffected by a clinic opening or closure



NOTE.—The figure plots the *unconditional* regional aggregate trends in health, population and economic conditions for regions that had an opening or closure (dashed line) and for unaffected regions (solid line) over time since the opening or closure occurred. Unaffected regions are used as a *potential* opening or closure for each year a change occurred in an affected region. The vertical line at zero indicates the year of the closure or opening.

Figure 4: Aggregate trends in health and socioeconomic statistics in regions affected and unaffected by a clinic opening or closure



NOTE.—The figure plots the *conditional* regional aggregate trends in health, population and economic conditions (after adjusting for calendar year and hospital fixed-effects and regional linear trends) for regions that had an opening or closure (dashed line) and for unaffected regions (solid line) over time since the opening or closure occurred. Unaffected regions are used for each year a change occurred in an affected region. The vertical line at zero indicates the year of the closure or opening.

Table 2: OLS and IV estimates of the effect of surgical volume on four-year post-surgery survival probability

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	FS	Restricted IV	Main IV	With controls	Case-mix adjusted
log(volume)	0.057*** (0.005)		0.070 (0.043)	0.055*** (0.018)	0.055*** (0.017)	0.048*** (0.018)
Closure		0.167** (0.077)				
<i>Socioeconomic characteristics</i>						
Age					0.018*** (0.001)	0.018*** (0.001)
Age squared					-0.000*** (0.000)	-0.000*** (0.000)
Female					0.059*** (0.005)	0.060*** (0.005)
Immigrant					0.002 (0.003)	0.003 (0.004)
Unknown origin					-0.012 (0.009)	-0.009 (0.010)
High school					0.013*** (0.003)	0.013*** (0.003)
College					0.023*** (0.003)	0.023*** (0.003)
Unknown educ.					-0.016* (0.009)	-0.016* (0.009)
<i>Previous admissions</i>						
One					0.005 (0.004)	0.006* (0.004)
Two					-0.003 (0.005)	-0.001 (0.005)
Three					-0.014*** (0.005)	-0.013*** (0.005)
Four					-0.006 (0.007)	-0.004 (0.007)
More than four					-0.066*** (0.005)	-0.065*** (0.005)
<i>Time since last admission</i>						
One to three years					0.039*** (0.004)	0.040*** (0.004)
Four to six years					0.060*** (0.004)	0.061*** (0.004)
Seven to nine years					0.068*** (0.005)	0.067*** (0.004)
More than nine years					0.060*** (0.018)	0.062*** (0.018)
Time FE	✓	✓	✓	✓	✓	✓
Hospital FE		✓	✓	✓	✓	✓
Linear trends		✓	✓	✓	✓	✓
First-stage F stat.	-	4.7	4.7	29.3	29.3	199.1
Mean survival	0.78	0.78	0.78	0.78	0.78	0.78
Observations	109,761	109,761	109,761	109,761	109,761	105,028

NOTE.— The table reports point estimates and (standard errors) from estimation of various regression models specified in section five of the paper. Outcome variable in all specifications is four year survival after cancer surgery. Surgical volume is measured in hundreds. Fixed effects in the last two columns also include surgery fixed effects. In columns (2) and (3) the closure instrument is specified as a dummy variable and in columns (3)-(6) as a set of dummy variables. The last column reports estimates from using the case-mix restricted sample excluding everyone with a designated hospital which at some point is either opened or closed. First stage F-statistics, due to Kleibergen and Paap (2006), are reported for columns (2)-(6). Reference categories are male, native, less than high school, no previous admissions and admission the same year. Standard errors are clustered at the hospital level. *, ** and *** denote statistical significance at the 10, 5 and 1 percent levels.

Table 3: Sample statistics for closed/opened and remaining clinics

	Closed/opened clinics	Remaining clinics
<i>Patient characteristics</i>		
Age	66.7 (12.8)	65.3 (12.3)
Post-secondary	0.23	0.27
Immigrants	0.09	0.13
Females	0.74	0.58
<i>Hospital characteristics</i>		
# Cancer surgeries	117 (51)	439 (263)
Days of care	9.2 (32.0)	8.6 (22.5)

NOTE.— The table reports shares/averages for different characteristics of the patient population by type of cancer clinic. Remaining clinics are defined as cancer clinics within the region that remain when a clinic within the same region is closed or opened. Standard deviations are reported in (parentheses).

Table 4: IV estimates of the effect of cumulated surgical volume on four-year post-surgery survival probability

<i>Panel A: Cumulated volume</i>			
	(1)	(2)	(3)
log(current volume)	0.048*** (0.018)		
log(cum. vol. last 2 years)		0.083* (0.046)	
log(cum. vol. last 3 years)			0.082 (0.068)
<i>F statistics First-stage</i>			
F-stat. Level	199.1	23.8	25.0
F-stat. Lag 1		14.5	34.2
F-stat. Lag 2			3.4
F-stat. All	199.1	45.9	117.3
Observations	105,028	95,802	86,272
<i>Panel B: Cumulated and current volume</i>			
	(1)	(2)	(3)
log(current volume)	0.048*** (0.018)	0.050** (0.021)	0.032* (0.017)
log(cum. vol. last 1 year)		0.021 (0.059)	
log(cum. vol. last 2 years)			0.050 (0.048)
<i>F statistics First-stage</i>			
Log volume	199.1	183.7	220.4
Cumulated variable(s)		18.7	127.6
Observations	105,028	95,802	86,272

NOTE.— The table reports point estimates and (standard errors) from estimation of various regression models with cumulative hospital surgical volume estimated using instrumental variable techniques discussed in section five of the paper. Outcome variable in all specifications is four year survival after cancer surgery. In column (1) the volume instrument is specified as a set of dummy variables indicating a closure in the region the same year while in columns (2) and (3) instruments for cumulative volume are both current and two (three) years lagged clinic closures. All models include calendar time, hospital and surgery fixed-effect as well as regional linear trends and are estimated using the case-mix adjusted sample specified in the last column of *Table 2*. First stage F-statistics are reported for all instruments both separately and jointly. Panel A reports estimates when including current volume in the cumulative volume measure while panel B reports results when current and one and two-years cumulative volume are included as separate regressors. Standard errors are clustered at the hospital level. *, ** and *** denote significance at the 10, 5 and 1 percent levels.

Table 5: IV estimates for the choice to perform robot assisted surgery

	All surgeries	Prostate cancer
	(1)	(2)
log(volume)	0.015 (0.013)	0.041 (0.044)
Observations	105,028	25,693

NOTE.— Outcome is an indicator for the choice to perform robot assisted surgery. All models include calendar time fixed effect, hospital fixed effects, linear regional trends and surgery fixed effect. Case-mix restricted samples excluding everyone with home hospital that is either opened or closed. Standard errors are clustered at the hospital level. *, ** and *** denote significance at the 10, 5 and 1 percent levels.

Table 6: Time-varying reduced form estimates following regional closures of cancer surgery clinics

	(1)	(2)
Closed clinic in region	0.016** (0.007)	
First year after closure		0.013* (0.007)
Second year after closure		0.033** (0.014)
Observations	107,743	107,743

NOTE.— The table reports point estimates and (standard errors) from estimation of the reduced form regression from the empirical model discussed in section five using a restricted sample including only observations for a maximum of two years after a regional cancer clinic closure. Outcome is four year survival after cancer surgery. All models include calendar time, hospital and surgery fixed-effect as well as regional linear trends. Column (1) report the reduced form estimate on post-closure survival rates of a regional cancer clinic closure while column (2) separates this estimate into the first and second year after the closures occurred, respectively. Standard errors are clustered at the hospital level. *, ** and *** denote significance at the 10, 5 and 1 percent levels.

Table 7: IV estimates of the effect of surgical volume on additional outcomes

	(1) New cancer surgery	(2) Re-admission	(3) Number of surgeons	(4) Time in hospital
log(volume)	-0.034** (0.017)	-0.041 (0.027)	2.204 (5.289)	0.344 (0.401)
Mean of outcome	0.08	0.42	82.87	5.91
Observations	105,028	105,028	80,633	105,028

NOTE.— The table reports point estimates and (standard errors) from estimation of various regression models of hospital surgical volume, estimated using instrumental variable techniques discussed in section five of the paper, on different outcomes. New cancer surgery (re-admission) is defined as an indicator for the event of subsequent cancer surgery (hospitalization) within three (one) year(s) after the initial cancer surgery. Number of surgeons is measured as the total number of surgeons in the hospital in which the patient was treated for cancer and time in hospital is measured as the length in days of the admission spell at the time of cancer surgery. All models include calendar time, hospital and surgery fixed-effect as well as regional linear trends and are estimated using the case-mix adjusted sample specified in the last column of *Table 2*. Standard errors are clustered at the hospital level. *, ** and *** denote significance at the 10, 5 and 1 percent levels.

Table 8: Statistics for surgeons at remaining and closing cancer clinics who remain or transfer within or outside closure regions

Panel A: Statistics one year after closure

	Surgeons at closed clinics one year before closure			Remaining hospitals
	Stay	Transfer within region	Transfer outside region	
% of surgeons at "closed"	0.69	0.18	0.13	
Age	45.8	46.0	46.5	45.4
Labor earnings	606,278	611,720	581,665	634,123
Females %	0.49	0.20	0.30	0.36
Married %	0.63	0.64	0.48	0.74

Panel B: Statistics two years before closure

	Surgeons at closed clinics three years before closure			Remaining hospitals
	Stay	Transfer within region	Transfer outside region	
% of surgeons at "closed"	0.88	0.06	0.06	
Age	45.8	45.0	47.2	43.6
Labor earnings	536,698	520,714	534,520	586,588
Females %	0.37	0.57	0.27	0.36
Married %	0.62	0.71	0.60	0.66

NOTE.— The table sample means characteristics for surgeons who were employed at hospitals where the cancer clinic remained and surgeons employed at hospitals where the cancer clinic closed who stayed, transferred within or outside the region, respectively. Panel A and panel B report sample means for surgeons one and three year before the closures, respectively. Labor earnings are measured on an annual basis in SEK (one SEK is approximately equal to 0.15 USD).

Table 9: Robustness analysis: organizational changes, distance to hospital, access to surgery and type of surgery

Panel A: First-step and reduced form estimates for cancer and heart surgeries

	Cancer		Heart	
	FS (1)	RF (2)	FS (3)	RF (4)
log(volume)	0.167** (0.059)	0.012* (0.006)	0.038 (0.057)	-0.002 (0.003)
Observations	109,760	109,760	139,976	139,976

Panel B: IV estimates adjusting for changes in distance to hospital

	Main (1)	Distance to treating hospital (2)	Distance to closest hospital (3)
log(volume)	0.048*** (0.018)	0.049*** (0.018)	0.050*** (0.019)
Distance measure	No	Yes	Yes
Observations	105,028	105,028	105,028

Panel C: IV estimates on timing of surgeries (age of patients)

	Age at surgery (1)
log(volume)	0.061 (0.068)
Observations	105,028

Panel D: IV estimates for choice between less and more invasive procedure

	Partial mastectomy (1)	Laparoscopic prostatectomy (2)
log(volume)	-0.013 (0.039)	0.011 (0.025)
Observations	50,096	25,693

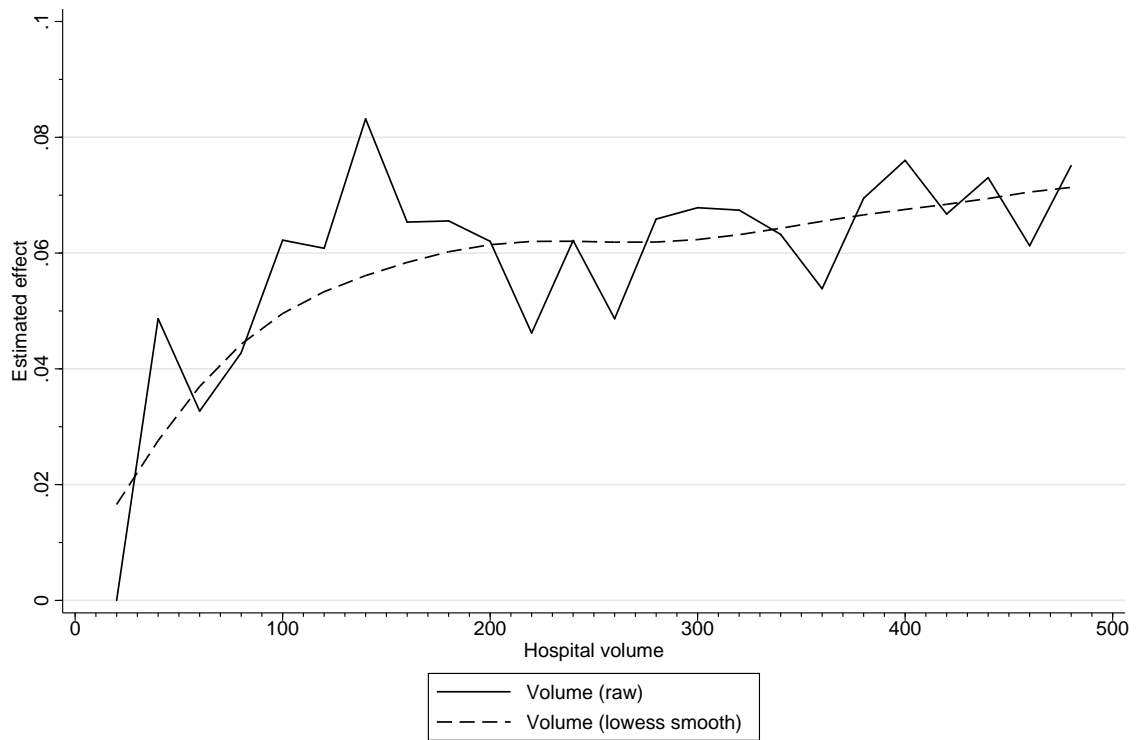
NOTE.— The table reports point estimates and (standard errors) from estimation of various regression models to assess the robustness of the main results from Table 2. Panel A report the first-stage and reduced-form estimates of surgical volume from the cancer and heart surgery samples in which the cancer clinic closure instrument is used for both samples (see the results section). Outcome is four year survival after cancer surgery and all models include calendar time, hospital and surgery fixed-effect as well as regional linear trends. Column (1) of Panel B reproduces the results from our preferred specification in column (6) of Table 2 while columns (2) and (3) additionally include as control variable the distance to i) the *treating* hospital and ii) the *closest* hospital. Panel C reports the result from our preferred specification in column (6) of Table 2 using patient age at surgery as outcome. Finally, panel D also report estimates from our preferred IV specification using the subsample of breast (prostate) cancer surgery in column (1) ((2)) and an indicator variable for whether the performed surgery was coded as a partial mastectomy (laparoscopic prostatectomy) as outcome variable. For all specifications in the table, standard errors are clustered at the hospital level and *, ** and *** denote significance at the 10, 5 and 1 percent levels.

Table 10: IV estimates of the effect of surgical volume on post-surgery survival by complexity of the surgery

	Breast (1)	Colorectal (2)	Prostate (3)
log(volume)	0.017 (0.020)	0.066 (0.087)	0.059 (0.051)
Mean survival rate	0.86	0.62	0.71
First-stage F-stat.	311.9	93.3	55.6
Observations	50,096	29,334	25,693

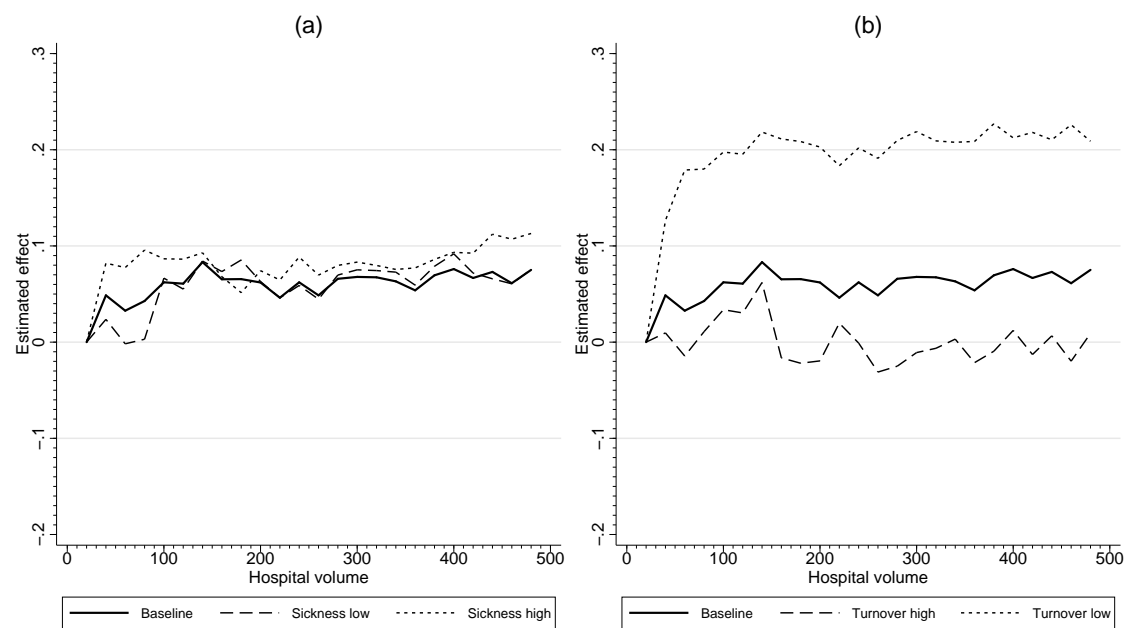
NOTE.— The table reports point estimates and (standard errors) from estimation of our preferred model specification from column (6) of *Table 2* on different subsamples of our main analysis sample. Each column corresponds to the estimation results conditional on cancer type; breast, colorectal or prostate cancer. Outcome is four year survival after cancer surgery. All models include calendar time, hospital and surgery fixed-effect as well as regional linear trends and are estimated using the case-mix adjusted sample estricted sample excluding everyone with a designated hospital which at some point is either opened or closed. Standard errors are clustered at the hospital level. *, ** and *** denote significance at the 10, 5 and 1 percent levels.

Figure 5: Non-linear volume effects on four-year post-surgery survival probability



NOTE.—The figure plots the estimated coefficients of surgical volume on four-year post-surgery survival from a fixed-effects regression model including regional, calendar year and surgery fixed effects, see equation 3 in the results section. The non-linear volume effects are estimated by including dummy variables for each twenty volume bin constructed from the analysis data, using the first volume bin as reference category. The solid line pertains to the non-smoothed relationship while the dashed line is obtained by running locally weighted regressions of the estimated volume coefficients on the corresponding dependent variable.

Figure 6: Non-linear volume effects on four-year post-surgery survival probability by high and low hospital average turnover



NOTE.—The figure plots the estimated coefficients of surgical volume on four-year post-surgery survival from a fixed-effects regression model including regional, calendar year and surgery fixed effects, see equation (3) in the results section. The non-linear volume effects are estimated by including dummy variables for each twenty volume bin constructed from the analysis data, using the first volume bin as reference category. Each line corresponds to a different sample category; the solid line in both plots pertains to the aggregated volume effect from *Figure 5* using the full sample. The dashed (dotted) lines in plot (a) pertain to hospitals with less (more) than median sickness absence rates. The dashed (dotted) lines in plot (b) pertain to hospitals with more (less) than median turnover, estimated by pooling average surgeon tenure across all hospitals and all years in our sample.