Spring 2012

Behavioral Economics and Physician Behavior

Allyssa S. Ward

BEHAVIORAL ECONOMICS AND PHYSICIAN BEHAVIOR

by

Allyssa S. Ward

A Proposal Submitted to the Honors Council
For Honors in Interdisciplinary Studies in Mathematics and Economics

September 30, 2011

Approved by:

Adviser: Amy Wolaver

Department Chairperson: Tom Kinnaman

ACKNOWLEDGMENTS:

I would like to thank Professor Amy Wolaver for her devotion to this project. This study would not have been possible without her direction.

TABLE OF CONTENTS:

LIST OF FIGURES

ABSTRACT

This study seeks to answer whether the availability heuristic leads physicians to utilize more medical care than is economically efficient. Do rare, salient events alter physicians' perceptions about the probability of patient harm? Do these events lead physicians to overutilize certain medical procedures? This study uses Pennsylvania inpatient hospital admissions data from 2009 aggregated at the physician level to investigate these questions. The data come from the 2009 Pennsylvania Health Care Cost Containment Council (PHC4).

The study is divided into two parts. In Part I, we examine whether bad outcomes during childbirth (defined as maternal mortality, an obstetric fistula or a uterine rupture) lead physicians to utilize more cesarean sections on future patients. In Part II, we examine whether bad outcomes associated with appendicitis (defined as patient death, a perforated or ruptured appendix or sepsis) lead physicians to perform more negative appendectomies (appendectomies performed when the patient did not have appendicitis) on future patients.

Overall the study does not find evidence to support the claim that the availability heuristic leads physicians to overutilize medical care on future patients. However, the study does find evidence that variations in health care utilization are strongly correlated with individual physician practice patterns. The results of the study also imply that physicians' financial incentives may be a source of variation in health care utilization.

I. INTRODUCTION

In 2009, health care spending in the United States accounted for $8,086 per person or 17.6% of gross domestic product (GDP). Over half of these expenditures were attributed to physician-directed care (physician/clinical services: 21% and hospital care: 31%). From 2010 to 2020, health care spending is predicted to grow faster than GDP on average, and by 2020, health care spending is expected to account for 19.8% of GDP (Centers for Medicare & Medicaid Services). The health care industry makes up an increasingly crucial part of the United States' economy, and evidence shows that this important industry is not always efficient. As documented later in this paper, there are substantial variations in health care utilization trends across the country, and higher utilization is not correlated with better quality of care. Understanding physician behavior can therefore help economists understand why these variations in utilization exist and can help economists determine how to increase efficiency in the health care market.

This study uses concepts from behavioral economics to test whether physicians' personal experiences lead to excessive spending. Specifically, we examine the availability heuristic – the idea the salient or common memories can skew physicians' perceptions of probabilistic events. While experimental evidence has suggested that the availability heuristic impacts physician behavior, this observational study assesses whether the availability heuristic has a substantial impact on variations in care across the nation.

We hypothesize that negative outcomes during childbirth and appendicitis encourage physicians to perform more unnecessary cesarean sections or appendectomies, respectively, than are necessary. We will begin by reviewing some background information on variations in health care spending across the United States. We will then review some literature on the availability heuristic and its applications in health care markets. Next, we will explain the economic theory surrounding our hypotheses. We will then describe our data and methods. Finally, we will discuss the results of the study as well as its limitations and draw some conclusions about the availability heuristic and its relevance to health care spending.

II. LITERATURE REVIEW

II. A. Background on Health Care Variations across the United States

When considering health care spending, we must focus our attention on physician behavior. Physicians play a crucial role in determining health care costs because of asymmetric information. Asymmetric information refers to transactions involving two parties in which one party knows more than the other. When it comes to health care, patients know very little about different diagnoses and treatment options. Physicians, on the other hand, are experts when it comes to medical practice. This imbalance of knowledge is why patients allow physicians to act as their “agents.” In other words, when patients are in need of medical care, they ask physicians to make decisions on their behalves and trust that the physicians will act in their best interest (Folland et al. 2010, 197). Therefore, physicians have significant control over medical decisions and, consequently, health care costs.

Many economists have attributed excessive health care spending to unnecessary variations in physician practices. For instance, Fisher et al. (2003) performed a two-part study on regional variations in Medicare spending in the last six months of a patient's life. They grouped the data into hospital referral regions (HRR). For each HRR, they calculated an End-of-Life Expenditure Index (EOL-EI). The EOL-EI measures differences in Medicare spending that result from physician practices rather than underlying patient illness or relative prices. They studied four different cohorts: patients with acute myocardial infarction (MI), hip fracture, and colorectal cancer, which are all diseases with similar hospitalization rates across the country. They also studied the Medicare Current Beneficiary Survey (MCBS) sample. Within each cohort, they broke the data into five quintiles based on values of the EOL-EI, and used multivariate regression to analyze differences among quintiles. They found that regions in the highest quintile received 61% more Medicare resources than regions in the lowest quintile, and that most of the variation was due to evaluation and management services, tests, radiology, and minor procedures. Additionally, regions in the highest quintile were more likely to utilize specialists as opposed to general practitioners.

In the second part of the study, they examined differences in quality of care among the EOL-EI quintiles. They looked at three different measures of quality of care:

the one-year predicted chance of mortality, changes in functional status according to the Health Activities and Limitations Index (HALex), and satisfaction of care as indicated by a twenty-question survey on the MCBS. They showed that higher utilization did not correlate with higher quality of care. They found that the highest EOL-EI quintile actually had a statistically significantly higher predicted chance of mortality than the lowest quintile. Additionally, they found no statistically significant differences in changes in functional status and no consistent, significant differences in satisfaction of care. To summarize, certain regions of the country spend significantly more on health care than others without attaining better results. This implies that certain regions can lower health care spending without lowering quality of care.

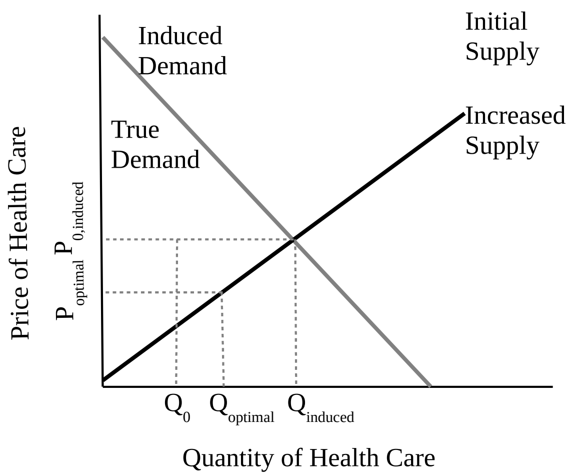
In a more recent study, Fisher *et al.* (2004) studied variations in intensity of care (per patient spending) in academic medical centers (AMC) across the United States for chronically ill Medicare patients. They also divided the AMCs into five groups based on their intensity of care. After verifying that the initial patient health was relatively uniform across all five groups, Fisher *et al.* looked at differences in health care spending across the five groups. They found that per patient spending varied by between 47% to 58% six months after the initial hospitalizations. Most of this variation was due to "supply-sensitive" procedures such as diagnostic tests and imaging and management services rather than surgical procedures. They also found that after the first six months, inpatient use was between 61% and 84% higher for patients in the highest intensity group as compared to the lowest intensity group. Finally, Fisher *et al.* showed that intensity was not related to survival or quality of care (as measured by the Cooperative Cardiovascular Project). In fact, in some cases, higher intensity was actually associated with lower survival! Fisher's work provides evidence that variation in health care spending and utilization across the United States is indeed extensive and unwarranted.

Similarly, the Dartmouth Atlas of Healthcare in Pennsylvania (Bronner et al. 1998) documents significant variations in the use of acute care hospital resources across small areas in Pennsylvania. For instance, utilization trends for elective surgeries are idiosyncratic and vary by individual physicians. Bronner et al. call these idiosyncrasies the “surgical signature.” The atlas also documents large variations in the treatment of diseases for which there are no well-defined standards of care. For instance, rates of coronary artery bypass graft surgery (CABG) varied by a factor of five across regions and rates of cesarean sections (c-sections) varied by a factor of almost four. Again, this study provides evidence that utilization is not linked to patient outcomes.

Bronner et al. also argue that utilization trends are more heavily influenced by the supply of resources available than by the overall illness of the population. Simple linear regression showed a positive correlation between the per capita number of hospital beds and discharge rates for medical conditions with an R-squared value of 0.53. Bronner et al. claim that “geography is destiny” – that patients' treatments are determined by the supply of resources and type of physicians in the area rather than their illnesses.

Bronner et al. are not the first to provide evidence of a correlation between the supply of services and the amount of health care provided. Some health economists believe that physicians alter the demand for health care. In other words, physicians can increase the quantity of care provided without lowering prices, increasing total overall costs in the economy. Figure 1 illustrates the concept of demand-shifting or supplier-induced demand.

Figure 1. Demand-Shifting



Victor R. Fuchs (1978) provided evidence of "demand-shifting" in his study of in-hospital operations across the United States. He hypothesized that an increase in the number of surgeons in an area would increase the number of surgeries performed in the area. His reasoning was that in areas where supply is higher, physicians have to induce demand in order to achieve or maintain a certain level of income. Fuchs looked at utilization data from the Health Interview Survey from 1963 and 1970 in conjunction with physician supply data from the AMA. Using multivariate regression analysis, Fuchs concluded that physicians do in fact shift demand. In fact, each additional surgeon in a particular area was associated with an increase of between 40 to 60 operations each year.

Fuchs also acknowledged that areas with high demand for surgeons might attract more supply. To address this issue, Fuchs predicted the number of surgeons attracted to a particular region based on socioeconomic traits and the potential demand within the region. He found that potential demand had little influence on surgeons' choice of location. Then, using the predicted supply variable in place of actual supply, he once again found evidence of demand-shifting. In the end, he concluded that a 10% increase in the surgeon/population ratio led to a 3% increase in the number of operations in that region, along with an increase in price.

Another study on “demand shifting” or “supplier induced demand” looked at patients admitted to the hospital for childbirth from 1970-1982 according to the National Hospital Discharge Survey (NHDS) (Gruber et al. 1996). During that period, fertility fell by 13.5% in the United States, but the rate of cesarean sections (c-sections) rose by 240%. According to Gruber *et al.*, c-sections require less effort on the part of the physician and reimbursements for c-sections are generally higher than for vaginal deliveries. Gruber et al. therefore hypothesized that a decrease in the fertility rate would lead physicians to utilize more c-sections in order to maintain their income. Using multivariate regression analysis and controlling for patient demographics, hospital size and region, and complications during birth that would make a c-section more likely, Gruber et al. found that there was indeed a statistically significant negative correlation between the rate of fertility and the rate of c-sections used from 1970-1982. A 10% decline in fertility was associated with a 0.97 percentage point increase in the rate of c-sections.

The findings of Fisher et al. (2003, 2004), Bronner et al. (1998), Fuchs (1978), and Gruber et al. (1996) indicate that there is indeed unwarranted variation in physician practices. Fuchs (1978) and Gruber et al. (1996) suggest that this variation may be driven by physicians' financial incentives. However, cognitive biases, specifically the availability heuristic, may be another source of variation. In order to fully understand the healthcare market, we must reject some of the premises of neoclassical economics and borrow some knowledge from behavioral economics. If physicians are indeed subject to the availability heuristic, then perhaps policymakers can enact policies or measures to decrease the use of inappropriate care.

II.B. The Availability Heuristic and Its Relevance to Health Care Spending

The basic model of economic behavior is grounded in the assumption that human beings are rational economic agents that make informed decisions based on marginal costs and marginal benefits. However, empirical evidence of human behavior does not always support this assumption. In their most recent book, *Nudge,* behavioral economist Richard H. Thaler and his colleague Cass R. Sunstein explain that, in fact, human beings are not always rational and do not always make the best decisions. Instead, real humans often struggle with complicated decisions and make systematically biased projections. By accepting that *homo sapiens* are not perfect economic beings (*homo economicus)*, behavioral economists can better understand the decisions that humans make.

Although patients would like to believe that their doctors are infallible, physicians are in fact human. Like all humans, medical doctors are subject to systematic biases that cause a disconnect between their own perceptions and reality. In particular, physicians might make biased decisions that result from the use of heuristics, or “rules of thumb” (Thaler and Sunstein 2009, 17-39).

Tversky and Kahneman (1974) completed the most important work regarding heuristics. They identified three heuristics (representativeness, anchoring and adjustment, and availability), but for the purpose of this study, we will focus our attention on one – the availability heuristic. Availability occurs when we mistakenly judge probabilities based on examples that are salient and easily retrievable. For example, in one of Tversky and Khaneman's experiments, participants were asked to consider all the words in the English language. They were asked to guess whether more words began with the letter “r” or whether “r” more frequently appeared as the third letter of a word. Most people incorrectly chose the former because they could easily think of words that started with “r.” In this experiment, availability led subjects to form incorrect judgments.

Salient events can also spark the availability heuristic. For instance, when assessing a patient's risk for suicide, a doctor might consider previous patients' experiences. The doctor may weigh the experiences of patients who committed suicide more heavily than those who were depressed but did not commit suicide because suicide is a salient event (Tversky and Kahneman 1973). Another example is that in the months after 9/11, Americans cut their air travel by 20%. Many Americans began driving more in place of flying because the memory of 9/11 caused them to overestimate the probability of a plane crash. In reality, Americans are approximately 37 times more likely to die in an automobile accident than in a plane accident. In fact, some statisticians estimate that the number of highway deaths due to increased vehicle travel after 9/11 far outweighed the number of people who died on a plane during the 9/11 attacks (Myers 2003). It is this application of the availability heuristic – the relationship between availability and salience – that we examine in this study.

Availability is not always a hindrance to effective decision-making. In fact, more often than not, availability likely serves as a useful tool in health care because it allows physicians to diagnose common or severe diseases more quickly. However, if physicians rely too heavily upon availability, there can be negative consequences for both the patient's health and health care spending.

Doctors themselves admit to falling victim to cognitive biases. In his book *How Doctors Think*, Jerome Groopman, M.D. (2007) describes some biases and pitfalls that he and his colleagues faced when making medical decisions. His own examples and commentary imply that the availability heuristic influences physician behavior.

One example involves Harrison Alter, M.D.. Alter was working in the emergency department in Tuba City, AZ when he met a Navajo woman who complained that she was having trouble breathing. The woman was running a low fever, and she told Alter that she had taken a few aspirin to deal with her symptoms. There had been widespread episodes of pneumonia in the area, so Alter diagnosed the woman with viral pneumonia and referred her to a general internist. Later that day, the internist discovered that the woman did not have viral pneumonia, but rather, she was suffering from aspirin toxicity. Alter had fallen victim to the availability heuristic (Groopman 2007, 63-66).

A second example is of a radiologist who once missed a breast cancer and was sued. After the incident, the particular radiologist became much more aggressive in trying to catch breast cancer. Of all the mammographies that he examined, he called back about 15% to 16% of the patients for a biopsy (as opposed to the average of about 10% to 11%). This caused unnecessary emotional distress for the women whose tumors turned out to be benign, and certainly created unnecessary medical costs (Groopman 2007, 188). In this example, it is clear that the availability heuristic can influence physicians in ways that negatively affect both patient outcomes and health care costs.

Finally, another case of a missed diagnosis suggests that the availability heuristic may affect physicians (Redelmeier 2005). The case involved a patient who was complaining of body aches and was running a slight fever. The physician diagnosed the patient with an upper respiratory tract infection, took some blood cultures and sent him home. Later, the blood cultures revealed bacteria in the patient's blood. The patient did not have an upper respiratory tract infection, but rather, he was suffering from osteomyelitis, an infection in his bones. Redelmeier argues that the availability heuristic caused this misdiagnosis. Since upper respiratory tract infections are very common, examples of patients with upper respiratory tract infections easily came to the physician's mind. Luckily, the follow-up test saved the patient's life.

The aforementioned anecdotes are consistent with Croskerry's (2009) "Universal Model of Diagnostic Reasoning." Croskerry argues that physicians use two types of reasoning when diagnosing patients. System 1 processing refers to an intuitive approach, which relies heavily upon the experience of the physician. System 1 processing is characterized by the use of heuristics, and is generally used when salient features of a diagnosis are present. System 2 processing, on the other hand, refers to a more analytical approach. System 2 processing generally results in less uncertainty, but requires more resources, less boundaries, and more ideal conditions. Ideally, physicians use System 1 processing to quickly diagnose patients, and then use System 2 processing to monitor and sometimes override the System 1 diagnosis. Croskerry warns, however, that physicians often operate under suboptimal conditions, and that "inattentiveness, distraction, fatigue and cognitive indolence may diminish System 2 surveillance and allow System 1 too much latitude." In other words, under suboptimal conditions, the availability heuristic may lead physicians to make diagnoses that are incorrect and potentially costly.

Finally, Milstein and Adler (2003) argue that physicians are influenced by what they call the "unavailability heuristic." They claim that most clinical failures are "subtle and not reported by the media," and that physicians therefore tend to underestimate the likelihood of preventable harm. Milstein and Adler believe that this explains why physicians might continue to use treatments that are ineffective. They do not offer any data to support their position, however.

While the anecdotes above suggest that the availability heuristic affects physicians, there are few econometric studies testing this concept. One existing study examined 227 patients in a university hospital for whom blood cultures were ordered between 1984 and 1985 (Poses et al. 1991). After the blood samples were taken, researchers interviewed the physicians (mostly interns) who had ordered the blood samples and asked them to estimate the probability that the samples would be positive for bacteremia. Researchers also asked the physicians about certain recall variables: an estimate of the proportion of the physician's patients currently receiving antibiotics, the number of patients the physician could recall that had positive blood cultures within the past month, and whether the physician felt that he frequently cared for patients with bacteremia. Poses *et al.* controlled for the severity of the patients' conditions and patient demographic factors. They then used a calibration curve, a receiving operating characteristics (ROC) curve, and multivariate regression analysis to analyze the data. They concluded that the availability heuristic does in fact exist: there was a significant positive correlation between the physician's estimated probability of bacteremia and the proportion of the physician's patients that he estimated were on antibiotics. In addition, physicians who claimed that they frequently treated patients with bacteremia had significantly higher estimates of the probability that a given blood culture would be positive. The study shows that physicians overestimate probabilities of events that are common and easily retrievable. The sample for this study was very limited however. Poses et al. only examined physicians from one hospital and many were interns from the university. The results might differ for more experienced physicians.

A more recent study examined thirty-six residents from Erasmus Medical Centre in the Netherlands, half of whom were in their first year of residency, the other half of whom were in their second year of residency (Mamede et al. 2010). Researchers first presented eight vignettes of clinical cases based on real patients. For each case, the physicians were given a potential diagnosis and asked to rate the probability that the diagnosis was correct. Then, the physicians were shown eight new cases, some of which were similar to those seen earlier, and were asked to form diagnoses as quickly as possible. Diagnoses were then evaluated as fully correct, partially correct, or incorrect, and Mamede et al. used analysis of variance (ANOVA) to analyze the results. They found that second-year residents scored significantly lower on cases that were similar to cases presented in phase 1. Second-year residents incorrectly gave the phase 1 diagnoses because phase 1 diagnoses were easily retrievable in the residents' minds. There were no such results for first-year residents. Additionally, Mamede et al. found that going back and using analytical reasoning allowed second-year residents to score significantly higher in phase 2. However, this study used a limited sample in a hypothetical situation. The results might not hold true in real-life situations with more experienced doctors.

Finally, Camancho, Donkers and Stremersch (2011) examined the availability heuristic's effect on physician prescribing patterns. They studied a panel data set of Dutch general practitioners treating obstructive airways diseases. They examined eight clinically equivalent prescriptions, including a new treatment, Symbicort. Camancho et al. hypothesized that patients who were switched from one treatment to another equivalent treatment due to negative side effects would be salient in the physicians' minds. They believed that salient patients would inhibit physicians from prescribing certain treatments. They then developed a quasi-Bayesian learning model and used a Markov chain Monte Carlo approach to fit the data to the model. Indeed, they found evidence to support their hypothesis. They found that patients who switched prescriptions received seven to ten times more weight in the physician's mind. Camancho et al. concluded that the availability heuristic slows down the adoption of new treatments. This study provides the best evidence of the availability heuristic. However, we would prefer to study whether the availability heuristic affects physicians in the hospital setting.

Currently, most evidence supporting the availability heuristic is anecdotal or uses limited samples in unrealistic contexts. This study uses a more comprehensive sample and studies data about actual physician behavior in a realistic context. The study also provides a more recent analysis of the availability heuristic. Finally, this study adds to existing literature by providing an economic perspective: does the availability heuristic lead physicians to incur more unnecessary costs?

III. THEORY:

This study rests upon both psychological and economic theories. Since physicians often make decisions quickly with limited resources, they rely on heuristics, or mental shortcuts, to form diagnoses and take action. One such heuristic is the availability heuristic. The availability heuristic encourages physicians to make decisions based on easily-retrievable memories. Often, availability allows physicians to form diagnoses quickly and efficiently because they easily recognize diseases that are common or severe.

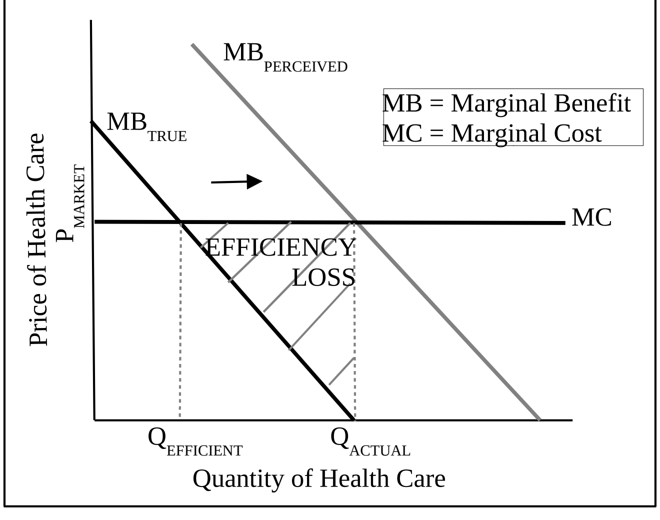
However, in some cases, the availability heuristic may lead to a misallocation of resources. In theory, salient memories lead a physician to overestimate the probability of rare diagnoses or severe patient outcomes. Consequently, the physician overestimates the marginal benefit of certain treatment options. In this case, the physician performs at a point where the marginal cost exceeds the marginal benefit of the care he provides.

For instance, a physician might easily recall a patient who died giving birth because such an event is rare and salient. The physician might then overestimate the probability of maternal death for future patients, overestimating the true marginal benefit of c-sections. He would therefore utilize more c-sections than are economically efficient, leading to deadweight loss.

Figure 2 illustrates this concept. Assume the marginal cost of a c-section – the cost of providing one additional patient with a c-section – is equal to the market price of c-sections. Assume the true marginal benefit – the benefit of providing one additional patient with a c-section – is given by the black line on the left. Optimally, the physician should perform just enough c-sections so that the true marginal benefit equals the true marginal cost. In our example, the physician should optimally perform $Q_{EFFICIENT}$ c-sections.

However, suppose the availability heuristic causes the physician to overestimate the marginal benefit of c-sections. His or her perceived marginal benefit is given by the grey line on the right. The physician will perform enough c-sections so that the marginal cost of the procedure equals his or her perceived marginal benefit. In this example, the physician will perform Q_{ACTUAL} c-sections. He performs more c-sections than are efficient, and the marginal cost of these additional c-sections exceeds their marginal benefit. The total excess cost is the deadweight or efficiency loss.

Figure 2. Efficiency Loss Due to Availability



IV. OUTCOMES OF INTEREST:

It is difficult to test whether the availability heuristic leads physicians to accrue unnecessary health care costs because it is difficult to define unnecessary healthcare costs. Since there is so much variation in physician practices across the country, there is little consensus about which medical procedures lead to higher quality of care and which procedures are actually unnecessary. Nonetheless, in this study we have identified two areas in which unnecessary costs can be identified and in which rare, salient outcomes exist that might skew physicians' perceptions. We examine each of these areas separately.

In Part I of the study, we examine cesarean sections (c-sections). The Agency for Healthcare Research and Quality (AHRQ), a well-respected Federal agency tasked with creating quality metrics to evaluate and analyze hospital care across the United States, has identified c-sections as an "overused procedure" (Department of Health and Human Services 2007). From 1996 to 2009, the national cesarean rate (percentage of all births that were delivered by c-section) rose from 20.7% to 32.3% (Martin et al. 2011). This increase is important from an economic perspective because c-sections are costly. C-sections cost twice as much as vaginal births on average (Baicker et al. 2006). Additionally, in comparison to vaginal births, c-sections are associated with higher risks for maternal rehospitalization sixty days after the initial discharge (Rochelle et al. 2000). It has also been shown that U.S. counties with higher cesarean rates do not have lower rates of infant or maternal mortality (Baicker et al. 2006). There are some patient risk factors that make c-sections necessary in some cases. For instance, older mothers are more likely to have c-sections, and mothers who have had c-sections in the past are also more likely to have c-sections for future children (Triunfo et al. 2009). However, the sharp increase in the national cesarean rate from 1996 to 2009 and the lack of correlation between regional cesarean rates and health outcomes indicate that many c-sections are unnecessary. Studies on supplier-induced demand claim that physicians' financial incentives explain many unnecessary c-sections (Gruber et al. 1996; Triunfo et al. 2009). In this study, we explore whether rare, salient outcomes from physicians' past experiences are another source of c-section overuse.

In Part II of the study we examine negative appendectomies (NA) – cases in which a patient had his or her appendix removed but was not ultimately diagnosed with appendicitis. In 1997, 15% of patients who underwent an appendectomy were not ultimately diagnosed with appendicitis (Flum and Koepsell 2002). Considering patients whose appendices are removed, those who are negative for appendicitis have longer lengths of stay, higher total charges, higher rates of mortality, and higher rates of infectious complications than those who are positive for appendicitis (Flum and Koepsell 2002). NA is therefore costly. In this study, we seek to answer whether negative past experiences cause physicians to remove more healthy appendices, accumulating more unnecessary costs.

V. DATA:

This study uses Pennsylvania inpatient hospital admissions data from 2009. The data come from the Pennsylvania Health Care Cost Containment Council (PHC4) and consist of one observation for each discharge from all inpatient facilities in Pennsylvania (excluding Veterans Administration Hospitals, Skilled Nursing Facilities and state psychiatric hospitals). The data include demographic information for each patient including age, primary insurance status, and race/ethnicity. The data also classify the severity of each patient's condition into one of five groups based on the predicted chance of mortality at admission according to the MediQual Atlas. The data set also includes each patient's primary diagnosis and up to eight secondary diagnoses, as well as the primary medical procedure performed on each patient and up to five additional procedures.

The original PHC4 data include hospital characteristics such as the number of beds in the hospital and the total number of inpatient hospital discharges during 2009. PHC4 also determined nine geographic regions by which they grouped the data. A map of these regions are included in Appendix A. Finally, the data identify each attending physician by Pennsylvania medical license number.

V.A. Childbirth

For Part I of the study, we created a sample according to the technical specifications of the AHRQ's Inpatient Quality Indicator (IQI) #33, the primary cesarean rate. As previously stated, the AHRQ is a well-respected, Federal agency run by the U.S. Department of Health and Human Services. The AHRQ calculates IQIs to provide objective measures of quality of care. Specifically, IQI #33 measures the use of c-sections in cases in which there were no abnormalities with the fetus and in which the mother had no previous c-sections.

As per AHRQ guidelines, we therefore limited our sample to patients who had given birth, excluding cases of abnormal fetal presentation, preterm births, fetal deaths, multiple gestation and breech births. We also excluded patients who had undergone previous c-sections. Among the patients included in the sample, we identified those who underwent a c-section. We later calculated the cesarean rate as the number of c-sections included in the sample over the total number of patients included in the sample. Our measure of the cesarean rate therefore captures the use of c-sections in patients who were eligible for a non-complicated vaginal birth. In that sense, we might think of our definition of the cesarean rate as the rate of unnecessary c-sections in healthy patients. Nonetheless, there are other maternal risk factors not captured by our definition that could necessitate a c-section. We will herein refer to our measure as the “unnecessary” cesarean rate, but keep in mind that it is an imperfect measure of clinically unnecessary c-sections.

Within our sample, we defined a bad outcome as a maternal death, an obstetric fistula, or a uterine rupture. An obstetric fistula is tissue damage caused by the pressure of childbirth that results in an “abnormal opening between a woman's vagina and bladder or rectum (or both)” (USAID 2009). A uterine rupture occurs when the uterine wall separates from the overlying membrane (Nahum 2010). Both obstetric fistulas and uterine ruptures are rare and serious events that generally require an emergency c-section. Variable definitions for these and other variables are given in Table 1 in Appendix C. A list of ICD-9 codes used to create the c-section and bad outcome variables are shown in Appendix B.

In total, we included 104,311 patients for this study. Of these patients, about 18.2% underwent a c-section. There was only one patient death and there were only 7 obstetric fistulas and 14 uterine ruptures. By our definition, there were therefore 22 bad outcomes in our sample. Most of the patients in the sample were white (69.1%), while 15.1% were black, 6.3% were Hispanic, 2.3% were Asian and 9.0% were neither white, black, Hispanic, nor Asian. Over half of the patients in the sample were between the ages of 20 and 29 (53.0%). About one-third of the patients were between the ages of 30 and 39 (34.1%). 10.6% of patients in the sample were between the ages of 10 and 19, and very few patients (2.2%) were above the age of 40. None of the patients in the sample were below the age of 10 or above the age of 59. Most patients had private insurance (55.9%), while a great deal of patients also had Medicaid as their primary insurer

(40.4%). Few patients had Medicare, another source of government insurance, no insurance or an unknown source of insurance (1.2%, 1.1%, 1.1% and 0.2%, respectively). Finally, the overwhelming majority of patients (99.8%) had no risk at admission. No patients were classified as having maximal risk at admission. Table 1 in Appendix C shows these variable definitions. Table 2 in Appendix C provides mean values of patient characteristics from the sample.

Most of the patients in the sample were admitted to a hospital in the southeastern region (Regions 5,7,8,9 – 62.0%) or the southwestern corner of Pennsylvania (Region 1 – 17.9%). On average, each hospital had 5,686 total discharges for the year and 399 hospital beds. On average, each hospital had less than one (0.54) bad outcome during the year. Table 1 in Appendix C provides these variable definitions. Table 3 in Appendix C provides mean values of hospital characteristics from the sample.

Next, we aggregated the data by physician and quarter, so we had one observation of the “unnecessary” cesarean rate for each physician during each quarter that he or she delivered at least one baby. In total, there were 1,156 physicians in the first quarter, 1,156 physicians in the second quarter, 1,188 physicians in the third quarter, and 1,166 physicians in the fourth quarter with recorded, non-complicated births. We calculated each physician's “unnecessary” cesarean rate for the quarter as the total number of c-sections over the total number of deliveries that were included in the sample. We also identified the physician's “unnecessary” cesarean rate from the previous quarter as well as the total number of bad outcomes from the previous quarter. For bad outcomes, we examined obstetric fistulas and uterine ruptures separately. We also examined all bad outcomes (maternal deaths, obstetric fistulas, uterine ruptures) aggregated together.

The average quarterly "unnecessary" cesarean rate for physicians was about 18.2%, and the average quarterly number of bad outcomes was about 0.005. The standard deviation of quarterly physician "unnecessary" cesarean rates was 10.3 percentage points, indicating substantial variation across physicians and quarters. The average standard deviation of a single physician's quarterly cesarean rate across all four quarters was 5.0 percentage points, indicating that there is not only variation across physicians, but also variation in individual physicians' behaviors over time. These values are shown in Figure 3, as well as in Table 4 in Appendix C.

Figure 3:
Summary of Childbirth Variables at the Physician Level

Variable	Mean	Std. Dev.
Physician's Quarterly "Unnecessary" Cesarean Rate (%)	18.215	10.299
Std. Dev. of Physician's Quarterly Cesarean Rate Across Quarters	5.054	3.039
# Maternal Deaths, Uterine Ruptures or Obstetric Fistulas	0.005	0.069
# Uterine Ruptures	0.003	0.055
# Obstetric Fistulas	0.002	0.039
# Physician Births	22.356	16.893

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

The mean physician "unnecessary" cesarean rate did not fluctuate greatly by quarter (Appendix C: Table 5). Quarter 3 had the lowest mean rate (17.4%), while Quarter 1 had the highest mean rate (18.9%). There was at least one bad outcome in each quarter. Quarter 1 had the lowest average number of bad outcomes per physician (0.001), while Quarters 2 and 3 had the highest average number of bad outcomes per physician

(0.008). No physician had more than one bad outcome per quarter.

We compared quarterly physician “unnecessary” cesarean rates when there was no bad outcome in the previous quarter (17.8%) with quarterly physician “unnecessary” cesarean rates when there was a bad outcome in the previous quarter (19.6%). These values are shown in Table 6 in Appendix C. Although the mean quarterly cesarean rate was higher for physicians with a bad outcome in the previous quarter, the difference is not statistically significant.

Although we would liked to have included physician characteristics such as race and number of years of experience in the study, these variables were not available to us. Nonetheless, we calculated the physician's total number of deliveries as a proxy for physician experience. We also calculated the physician's average patient composition for each quarter: the proportion of the physician's patients who were white, black, Hispanic, Asian or of another race; the proportion of the physician's patients whose primary insurer was a private insurance company, Medicaid, Medicare, another government provider, who were uninsured, or whose insurance status was unknown; the proportion of the physician's patients who were between the ages of 10 and 19, 20 and 29, 30 and 39, 40 and 49 or 50 and 59; and the proportion of the physician's patients who were categorized as no risk, minimal risk, moderate risk or severe risk at admission. Each of these variables was rescaled to range from 0 to 100 as opposed to 0 to 1. Finally, we also calculated mean hospital characteristics for hospitals at which each physician operated during the quarter. Mean, minimum and maximum values of each of these variables appear in Table 4 in Appendix C.

V.B. Appendicitis:

In Part II of the study, we limited our sample to all patients who had either been diagnosed with appendicitis or who had undergone an appendectomy. Among these patients, we defined a case of NA to be any patient who had an appendectomy but who was not ultimately diagnosed with appendicitis. This variable definition is the definition used by Flum and Koepsell (2002). We defined three bad outcomes. First, death was defined to be any patient with a diagnosis of appendicitis who ultimately died. Secondly, sepsis was defined to be any patient with a diagnosis of appendicitis who also had a secondary diagnosis of sepsis or severe sepsis. Finally, we defined a perforated or ruptured appendix to be any patient with a diagnosis of perforated/ruptured appendix. Both sepsis and perforation are rare, severe and dangerous complications of appendicitis (Marks 2011). A list of these variable definitions appears in Table 1 in Appendix C. A list of ICD-9 codes used to create these variables is given in Appendix B.

We included 17,335 patients in our sample. Overall, 94.9% of these patients underwent an appendectomy, but only 78.7% were ultimately diagnosed with appendicitis. The rate NA was 21.3%. This rate NA is higher than the rate NA observed in Flum and Koepsell's (2002) sample, which was 15.3%. This rate NA is also much higher than an estimate of the national rate NA for 2007, which was 8.5% (Seetahal et al.

2011).

Only 0.5% of patients in our sample died from appendicitis. 1.4% of patients in the sample went into sepsis and 12.3% of patients in the sample had a perforated or ruptured appendix. By our definition, 13.8% of patients in our sample experienced at least one bad outcome. Note that some patients had multiple bad outcomes. For instance, some patients had a ruptured appendix and then died. Mean values of patient characteristics are shown in Table 7 in Appendix C.

The overwhelming majority of patients in the sample were white (82.2%). 8.2% of patients were black, 5.1% were Hispanic, 1.2% were Asian, and 4.3% were another race. Most patients in the sample had private insurance (60.3%), Medicare (18.4%) or Medicaid (15.4%). More patients in the sample were male (53.6%) than female (46.3%), which is surprising because previous studies have shown that women of childbearing age are most likely to undergo a negative appendectomy (Flum and Koepsell 2002). Most patients were between the ages of 10 and 69. Finally, most patients were classified as no risk (69.2%) or minimal risk (18.5%) at admission. Only 3.4% of patients were classified as severe or maximal risk at admission. Table 7 in Appendix C lists mean values of these variables at the patient level.

Most patients attended a hospital in the southwestern (Region 1 – 21.9%) or the southeastern corner of the state (Regions 5, 8, 9 – 46.9%). On average, the patients in our sample attended hospitals that treated 4,964 patients over the year and had 377 hospital beds. Mean values of hospital characteristics are shown in Table 8 in Appendix C.

We then aggregated the data by attending physician and quarter. We therefore had one observation for each physician for each quarter that he or she treated at least one case of appendicitis or suspected appendicitis. We included approximately 1,429 physicians in the first quarter, 1,467 physicians in the second quarter, 1,453 physicians in the third quarter and 1,417 physicians in the fourth quarter. On average, each physician only treated 2.2 cases of appendicitis and performed 5.8 appendectomies during each quarter. In a given quarter, the maximum number of appendicitis cases treated and appendectomies performed was 24 and 23 per physician, respectively (Appendix C: Table 9).

We calculated each physician's quarterly rate NA as the number of negative appendectomies he or she performed during the quarter over the total number of appendectomies he or she performed over the quarter. (We multiplied each physician's rate NA by 100 to simplify the interpretation of our results.) In addition, we calculated the total number of bad outcomes – deaths, sepsis cases and perforated/ruptured appendices – that the physician experienced in the quarter. As in the childbirth portion of our study, we examined each of the three bad outcomes separately, as well as all three bad outcomes aggregated together. We also calculated the number of bad outcomes the physician experienced in the previous quarter, as well as the physician's rate NA from the previous quarter. On average, about 21.9% of appendices that physicians in our sample removed were later deemed to be healthy. The standard deviation of physicians' quarterly rates NA was 38.2 percentage points, indicating substantial variation across physicians

Figure 4:
Summary of Appendicitis Variables at the Physician Level

Variable	Mean	Std. Dev.
Quarterly Rate NA (%)	21.918	38.193
Std. Dev. of a Physician's Quarterly Rate NA Across Quarters	4.307	7.675
# Patient Deaths, Sepsis Cases or Perforations	0.393	0.809
# Deaths	0.014	0.116
# Sepsis Cases	0.041	0.208
# Perforations	0.350	0.776
# Appendectomies	5.783	4.536
# Appendicitis Cases Treated	2.247	3.115

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Appendix C. For each of our bad outcomes, the mean quarterly rate NA varied significantly (at the 0.01 significance level) based on whether or not the physician saw a bad outcome in the previous quarter. Using death or sepsis as our bad outcome, we found that a physicians who experienced a bad outcome in the previous quarter had significantly higher quarterly rates NA. On the other hand, using perforation as our bad outcome, we found that physicians who experienced a bad outcome in the previous quarter had significantly lower quarterly rates NA.

Finally, we calculated average hospital and patient characteristics for the physician during the quarter. We calculated the proportion of each physician's patients that were white, black, Hispanic, Asian or another race. We also calculated the proportion of each physician's patients that had private insurance, Medicare, Medicaid, other government insurance, unknown insurance or that were uninsured. We calculated that proportion of the physician's patients that were male or female and between the ages of 0 and 4, 5 and 9, 10 and 19, 20 and 29, 30 and 39, 40 and 49, 50 and 59, 60 and 69, 70 and 79, and 80 and older. These are the same sex/age bands that were used by Flum and Koepsell (2002). We calculated the proportion of the physician's patients that were classified as no risk, minimal risk, moderate risk, severe risk or maximal risk at admission according to the MediQual Atlas. Finally, we calculated the average hospital size (total hospital discharges and number of hospital beds) for hospitals where the physician operated during the quarter, as well as the hospital region. All of these variables were rescaled to take on values ranging from 0 to 100 rather than 0 to 1. Mean, minimum, and maximum values of these variables appear in Table 9 in Appendix C.

VI. METHODS:

Initially, we used multivariate ordinary least squares (OLS) regression analysis to examine the relationship between physicians' practice patterns (quarterly cesarean rates or quarterly rates NA) and bad outcomes from the previous quarter. An OLS regression was appropriate for this study because each of our two dependent variables were continuous random variables ranging from 0 to 100. Equation 1 shows the general form of the model.

[1] $\text{MEASURE OF OVERUSE}_{n,t} = \beta_0 + \beta_1\text{BAD OUTCOME}_{n,t-1} + \beta_2 Z_n + \varepsilon$

where the subscript n represents the individual physician, subscript t represents the quarter, β_0, β_1, and β_2 are the coefficients to be estimated, MEASURE OF OVERUSE is a measure of the physician's utilization (either the quarterly “unnecessary” cesarean rate or the quarterly rate NA), BAD OUTCOME is the number of bad outcomes the physician saw in the previous quarter (obstetric fistula, uterine rupture, death, sepsis, perforation), Z_n is a vector of the average physician, patient and hospital characteristics for the physician, and ε is the error vector.

Since a Brush-Pagan test revealed heteroskedasticity in each of our two samples (p-value < 0.0001 for both samples), we used robust standard errors in our OLS regressions. Since the availability heuristic decreases over time, we did not examine bad outcomes in quarters beyond the previous quarter

We also considered the idea that individual practice patterns might contribute to the physician's utilization of medical care. Each physician has certain immeasurable tendencies that affect his or her likelihood of performing a c-section or a negative appendectomy and that remain consistent over time. We could not control for these factors in the OLS regression model, so we used a fixed effects regression model to examine the data. In the fixed effects model, we essentially controlled for each individual physician to capture the important factors for which we could not control in the OLS regression model. The fixed effects model is given by Equation 2:

$$[2]\quad \text{MEASURE OF OVERUSE}_{n,t} = \beta_0 + \beta_1\text{BAD OUTCOME}_{n,t-1} + \beta_2 Z_n + \beta_3 X_n + \varepsilon$$

where n represents the physician, t represents the quarter, MEASURE OF OVERUSE is either the physician's quarterly "unnecessary" cesarean rate or quarterly rate NA, BAD OUTCOME is the number of bad outcomes from the previous quarter, Z is a vector of average patient characteristics, ε is the error vector, and X is the panel variable which represents the fact that we controlled for each individual physician. Again, we used robust standard errors.

Finally, we decided to examine how much influence each physician has in determining whether or not his or her patient receives a c-section or a negative appendectomy. To answer this question, we used our patient-level data and logistic regression to test the model given by Equation 3.

$$[3]\ \text{MEDICAL PROCEDURE}_k = \beta_0 + \beta_1\ \text{PHYSICIAN'S UTILIZATION OF MEDICAL}$$

$$\text{PROCEDURE}_k + \beta_2 Z_k + \varepsilon$$

where the subscript k is used to indicate each individual patient, MEDICAL PROCEDURE is 1 if the patient received the medical procedure of interest (either a c-section or a negative appendectomy) and 0 otherwise, PHYSICIAN'S UTLIZATION OF MEDICAL CARE is the patient's attending physician's quarterly “unnecessary” cesarean rate or quarterly rate NA *excluding the patient of interest,* Z is a vector of other patient and hospital characteristics, and ε is the error vector.

In all three of our models, we tested the following hypotheses:

H_0: $\beta_1 = 0$

H_a: $\beta_1 > 0$

In other words, in the first two models, we hypothesized that bad outcomes from the previous quarter would lead physicians to perform more “unnecessary” c-sections or more negative appendectomies in the current quarter. In the third model, we hypothesized patients whose physicians have higher utilization rates would be more likely to receive an unnecessary c-section or a negative appendectomy than similar patients whose physicians have lower utilization rates.

VII. RESULTS:

VII.A. Childbirth:

The OLS regression results provide no evidence for the availability heuristic.

Figure 5:
OLS Regression Results for Physician's "Unnecessary" Cesarean Rate as Predicted by the Number of Bad Outcomes from the Previous Quarter and the Physician's Previous "Unnecessary" Cesarean Rate

Variable:	Coefficient	Std. Error
Uterine Rupture, Obstetric Fistula or Patient Death	-1.561	1.873
Uterine Rupture	-0.667	2.008
Obstetric Fistula	-2.883	3.983
Physician's "Unnecessary" Cesarean Rate from Previous Quarter	0.439	0.040

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
Note: None of the relationships are statistically significant at the 0.10 significance level.
Note: Using robust standard errors.

Figure 6:
Fixed Effects Regression Results for Physician's "Unnecessary" Cesarean Rate as Predicted by the Number of Bad Outcomes from the Previous Quarter

Bad Outcome Defined As:	Coefficient	Std. Error
Uterine Rupture, Obstetric Fistula or Patient Death	-2.525	2.527
Uterine Rupture	-2.633	3.214
Obstetric Fistula	-2.367	4.112

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
Note: None of the relationships are statistically significant at the 0.10 significance level.
Note: Using robust standard errors.

the variation is simply due to physician discretion.

We decided to further investigate this phenomenon. How much influence do physicians have in determining whether or not patients receive c-sections? The results from the logistic regression are shown in Figure 7 and in Table 14 in Appendix D. Indeed, we found a statistically significant relationship (at the 0.01 significance level) between the physician's quarterly cesarean rate and the patient's probability of having a c-section. Consider two patients of the same race, age and risk status at the same hospital. The patient whose physician has a one percentage point higher quarterly "unnecessary" cesarean rate has a 2.5% higher predicted chance of having a c-section. Our results therefore suggest that physician practice patterns play a major role in determining health care utilization. However, the availability heuristic does not appear to explain variations in physician practice patterns. It is still unclear what drives such variations.

Figure 7:
Logistic Regression Results for Patient's Probability of Having an "Unnecessary" C-Section as Predicted by Her Physician's "Unnecessary" Cesarean Rate and Her Insurance Status

Variable	Odds Ratio	Std. Error
Physician's Quarterly "Unnecessary" Cesarean Rate	1.025	0.001
Medicaid*	0.791	0.019
Other Government Insurance (excluding Medicare or Medicaid)*	0.710	0.077
Uninsured*	0.739	0.079

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
Note: All relationships are statistically significant at the 0.01 significance level.
*Odds ratios in comparison to private insurance.

The physician's cesarean rate is not the only factor that influences a patient's chance of having a c-section, however. As we would expect, a patient's age, risk at admission and complications during childbirth also impact her chance of having a c-section. Another interesting finding is that patients with Medicaid, other government insurance, or no insurance have statistically significantly lower predicted chances of having a c-section than patients with private insurance (Figure 7). The difference is between 20% and 30%! This is interesting because Medicaid and the uninsured typically compensates physicians less than private insurers (Cowen 2010). The findings therefore suggest that physicians' financial incentives might alter their tendencies to utilize c-sections.

VII.B. Appendicitis:

Our OLS regression results indicate that bad outcomes in the previous quarter might affect a physician's rate NA in the current quarter (Figure 8 and Appendix D: Table 15). There is a statistically significant (at the 0.05 level) positive correlation between the number of perforations that a physician treated in the previous quarter and the physician's rate NA in the current quarter. All else equal, physicians with one additional perforation from the previous quarter have about a 0.8 percentage-point higher rate NA in the current quarter. However, the number of appendicitis-related deaths or appendicitis-related sepsis cases a physician experienced in the previous quarter is not significantly correlated with the physician's quarterly rate NA. Therefore, the OLS regression results imply that the availability heuristic might negatively impact a physician's performance, but only under certain circumstances.

Figure 8:

OLS Regression Results for Physician's Quarterly Rate NA as Predicted by the Number of Bad Outcomes from the Previous Quarter and by Total Hospital Discharges where the Physician Operates

Variable:	Coefficient	Std. Error
# Patient Deaths, Sepsis Cases or Perforations	0.699 *	0.410
# Patient Deaths	1.373	4.135
# Sepsis Cases	-2.290	2.061
# Perforations	0.832 **	0.419
Total Hospital Discharges (100 Patients)	0.103 ***	0.037

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Note: *, **, *** indicate statistical significance at the 0.10, 0.05 and 0.01 significance levels, respectively.

Note: Using robust standard errors.

treats and the number of negative appendectomies he performs.

On the other hand, our fixed effects regression results indicate that the number of patient deaths a physician witnesses in the previous quarter is positively correlated with his or her quarterly rate NA (at the 0.05 significance level). When a given physician has seen one death in the previous quarter, his or her quarterly rate NA is about 3.6 percentage points higher than when he or she has seen zero deaths in the previous quarter. This correlation suggests that a patient death (which is the most severe of our three bad outcomes) remains salient in the physician's mind and might cause him or her to remove more healthy appendices than he or she normally would. An appendicitis-related death appears to be an event that triggers the availability heuristic. However, the magnitude of the impact is relatively small. Recall that the physicians in our sample performed a maximum of 23 appendectomies per quarter. This means that a death from the previous quarter would only lead a physician to perform up to 23*0.03642 = 0.8 additional negative appendectomies on average.

Figure 9:
Fixed Effects Regression Results for Physician's Quarterly Rate NA as Predicted by the Number of Bad Outcomes from the Previous Quarter

Bad Outcome Defined as:	Coefficient	Std. Error
# Patient Deaths, Sepsis Cases or Perforations	0.475	0.422
# Patient Deaths**	3.632	1.622
# Sepsis Cases	-1.840	1.402
# Perforations	0.586	0.444

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
Note: ** indicates statistical significance at the 0.05 significance level, respectively.
Note: Using robust standard errors.

Figure 10:
Logistic Regression Results for Patient's Probability of Having a Negative Appendectomy as Predicted by the Physician's Rate NA and Patient Insurance Status

Variable	Odds Ratio		Std. Error
Physician's Quarterly Rate NA	1.056	***	0.001
Uninsured*	0.465	**	0.148

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Note: **,*** indicate statistical significance at the 0.05 and 0.01 significance levels, respectively.
*Odds ratio is in comparison to private insurance.

would like to have controlled for the exact length of time between one bad outcome and the next procedure a physician performed. Instead, we estimated the timing of events by the quarter in which they occurred. This estimation may have skewed our results. For instance, immediately after witnessing an obstetric fistula, physicians may actually perform more c-sections, but the effect might wear off by the next quarter. Our results show no correlation between past obstetric fistulas and future c-sections, but there may in fact be an immediate effect that we could not capture with our data.

Additionally, we had limited knowledge about the physicians in our sample. It seems plausible that less experienced physicians might be more susceptible to the availability heuristic than their more experienced counterparts. For instance, the first uterine rupture a physician witnesses is likely more salient than the the tenth uterine rupture he treats. Additionally, physicians with different specialties and physicians of different ages, ethnicities and backgrounds might respond differently to negative events. Since we could not control for these factors in our OLS regressions, our OLS regression results are subject to omitted variable bias. Although these factors are implicitly controlled for in our fixed effects regressions, we would like to have been able to control for more physician characteristics in our OLS regressions. Additionally we would like to have examined interaction effects among physician characteristics, bad outcomes and future physician behavior. Future studies should include such characteristics.

A third limitation was our lack of clinical knowledge. In our quest to unveil the availability heuristic, we chose events that we thought to be rare and salient – patient deaths, obstetric fistulas, uterine ruptures, sepsis and perforated appendices. However, these events might not have been salient enough to trigger the availability heuristic. There may be other clinical situations in which availability is more prevalent.

This study was also limited in that our sample was not representative of the entire United States' physician workforce. This study only used data from Pennsylvania in 2009. The training that physicians from other states receive and the cultures in which physicians from other states operate might impact the way those physicians react to bad outcomes. We cannot necessarily extrapolate the results of this study to other states.

Finally, there may actually be a difference in physicians' quarterly “unnecessary” cesarean rates or physicians' quarterly rates NA based on whether or not they witnessed bad outcomes in the previous quarter, but the difference may have been too small to detect. The smaller the difference, the lower the power of our study and the greater the chance of making a Type II error. In other words, for small differences, the probability that we would fail to detect a difference given one truly exists is high. However, from a practical standpoint, a Type II error is not of much concern. The purpose of this study was to determine whether the availability heuristic is a significant driver of unnecessary medical costs. If the availability heuristic does actually lead to more unnecessary costs, but at a rate that is too small too detect, we would still say that the availability heuristic is not a significant driver of costs.

IX. Conclusions

Experimental evidence has shown that the availability heuristic impacts physician behavior. However, this study shows that effect of the availability heuristic is not very meaningful from a broad, economic perspective. We searched for evidence of the availability heuristic under five different scenarios. We looked at the impact of past uterine ruptures and obstetric fistulas on physicians' cesarean rates, as well as the impact of past appendicitis-related deaths, appendicitis-related sepsis cases and perforated or ruptured appendices on physicians' rates NA. We only found evidence of the availability heuristic in one of these cases. Our results show that when physicians witness an appendicitis-related death, they may perform more negative appendectomies in the following quarter. Still, the magnitude of the increase is quite small – their quarterly rate NA only rises by about 3.6 percentage points on average. Additionally, our OLS regression results contradict these findings, so we are not fully confident that availability heuristic impacts physicians' utilization trends in any way.

We can conclude that the events that trigger the availability heuristic in the health care setting (if such events exist) are quite rare. When they do occur, the magnitude of their impact on physician behavior is relatively small. We conclude that the availability heuristic is not a significant driver of health care costs in the United States.

Nonetheless, this study documents significant variations in physician practices that are not due to patient characteristics. We have shown that physicians play a large role in determining health care costs. A patient's probability of having a c-section is strongly correlated with her physician's “unnecessary” cesarean rate. Likewise, a patient's probability of having a negative appendectomy is strongly correlated with his or her physician's rate NA.

This study shows that the availability heuristic does not seem to have a significant influence on variations in health care practices or health care spending across the United States for c-sections or negative appendectomies. However, we have found some clues as to what might drive such variations. This study provides some evidence that physicians' financial incentives contribute to excessive health care spending. Physicians are more likely to perform unnecessary c-sections on patients with private insurance than they are on patients with Medicaid, other government insurance, or the uninsured (at the 0.01 significance level). Similarly, physicians are more likely to perform negative appendectomies on patients with private insurance than on the uninsured (at the 0.05 significance level). These results are interesting because private insurers typically compensate physicians much better than government insurance programs or the uninsured. These results therefore imply that physicians are more willing to perform unnecessary procedures when they are reimbursed at a higher rate. Policymakers should further investigate physicians' financial incentives and possible reforms in order to contain health care costs in the United States.

APPENDICES:

Appendix A: Hospital Regions Defined by PHC4:

Appendix B: ICD-9 Codes Used for Variable Definitions:	
Variable	**ICD-9 Codes**
C-Section	MS-DRG: 765, 766 ICD-9 Procedure Codes: 740, 741, 742, 740, 7499 Exclude ICD-9 Procedure Code: 7491, 7251, 7252, 7253, 7254 Exclude ICD-9 Diagnosis Codes: 64420, 64421, 65100, 65101, 65103, 65110, 65111, 65113, 65120, 65121, 65123, 65130, 65131, 65133, 65140, 65141, 65143, 65150, 65151, 65153, 65160, 65161, 65163, 65180, 65181, 65183, 65190, 65191, 65193, 65220, 65221, 65223, 65230, 65231, 65233, 65240, 65241, 65243, 65260, 65261, 65263, 65640, 65641, 65643, 66050, 66051, 66053, 66230, 66231, 66233, 66960, 66961, 67810, 67811, 67813, 7615, V271, V272, V273, V274, V275, V276, V277
Birth	MS-DRG: 765, 766, 767, 768, 774, 775 Exclude ICD-9 Procedure Codes: 7251, 7252, 7253, 7254 Exclude ICD-9 Diagnosis Codes: 64420, 64421, 65100, 65101, 65103, 65110, 65111, 65113, 65120, 65121, 65123, 65130, 65131, 65133, 65140, 65141, 65143, 65150, 65151, 65153, 65160, 65161, 65163, 65180, 65181, 65183, 65190, 65191, 65193, 65220, 65221, 65223, 65230, 65231, 65233, 65240, 65241, 65243, 65260, 65261, 65263, 65640, 65641, 65643, 66050, 66051, 66053, 66230, 66231, 66233, 66960, 66961, 67810, 67811, 67813, 7615, V271, V272, V273, V274, V275, V276, V277
Obstetric Fistula	ICD-9 Diagnosis Codes: 6190, 6191, 6192, 6193, 6198, 6199
Uterine Rupture	ICD-9 Diagnosis Codes: 66510, 66511
Appendectomy	ICD-9 Procedure Codes: 47, 470, 4701, 4709
Appendicitis	ICD-9 Diagnosis Codes: 540, 5400, 5401, 5409, 541, 542
Death	Discharge Status = 20
Sepsis	ICD-9 Diagnosis Codes: 99591, 99592
Perforated or Ruptured Appendix	ICD-9 Diagnosis Code: 5400

Appendix C. Tables of Definitions and Summary Statistics

Table 1: Variable Definitions at the Patient Level	
Variable Name	**Variable Definition**
C-Section	1 if patient had a c-section (excluding cases of hysterectomy, abnormal presentation, preterm, fetal death, multiple gestation or breech birth) 0 otherwise
Birth	1 if patient gave birth (excluding cases of abnormal presentation, preterm, fetal death, multiple gestation, or breech birth) 0 otherwise
Obstetric Fistula	1 if patient had an obstetric fistula 0 otherwise
Uterine Rupture	1 if patient had a uterine rupture 0 otherwise
Maternal Death	1 if birth=1 and patient died 0 otherwise
Appendectomy	1 if patient had an appendectomy 0 otherwise
Appendicitis	1 if patient had a primary or secondary diagnosis of appendicitis 0 otherwise
Negative Appendectomy	1 if appendectomy=1 and appendicitis=0 0 otherwise
Appendicitis-Related Sepsis	1 if appendicitis=1 and patient had sepsis 0 otherwise
Perforation	1 if patient had a perforated or ruptured appendix 0 otherwise
White	1 if patient is white 0 otherwise
Black	1 if patient is black 0 otherwise
Hispanic	1 if patient is Hispanic 0 otherwise
Asian	1 if patient is Asian 0 otherwise
Other Race	1 if patient is other race 0 otherwise
Private Insurance	1 if patient has private primary insurance 0 otherwise
Medicare	1 if patient has Medicare as primary insurance 0 otherwise
Medicaid	1 if patient has Medicaid as primary insurance 0 otherwise
Other Government Insurance	1 if patient has other government insurance as primary insurance 0 otherwise
Uninsured	1 if patient is uninsured 0 otherwise
Unknown Insurance	1 if primary insurance for patient is unknown 0 otherwise

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4 2009

Table 1 (continued): Variable Definitions at the Patient Level	
Variable Name	**Variable Definition**
Age X-Y	1 if X ≤ age ≤ Y 0 otherwise
Female Age X-Y	1 if patient is female and X ≤ age ≤ Y 0 otherwise
Male Age X-Y	1 if patient is male and X ≤ age ≤ Y 0 otherwise
No Risk	1 if MediQual predicted chance of death at admission ranged from 0.000 to 0.001 0 otherwise
Minimal Risk	1 MediQual predicted chance of death at admission ranged from 0.002 to 0.011 0 otherwise
Moderate Risk	1 if MediQual predicted chance of death at admission ranged from 0.012 to 0.057 0 otherwise
Severe Risk	1 if MediQual predicted chance of death at admission ranged from 0.058 to 0.499 0 otherwise
Maximal Risk	1 if MediQual predicted chance of death at admission ranged from 0.500 to 1.000 0 otherwise
Region 1	1 if patient attended hospital in one of the following PA counties: Allegheny, Armstrong, Beaver, Butler, Fayette, Greene, Washington 0 otherwise
Region 2	1 if patient attended hospital in one of the following PA counties: Cameron, Clarion, Clearfield, Crawford, Elk, Erie, Forest, Jefferson, McKean, Mercer, Potter, Venango, Warren 0 otherwise
Region 4	1 if patient attended hospital in one of the following PA counties: Centre, Clinton, Columbia, Lycoming, Mifflin, Montour, Northumberland, Snyder, Tioga, Union 0 otherwise
Region 5	1 if patient attended hospital in one of the following PA counties: Adams, Cumberland, Dauphin, Franklin, Fulton, Huntingdon, Juniata, Lancaster, Lebanon, Perry, York 0 otherwise
Region 6	1 if patient attended hospital in one of the following PA counties: Bradford, Lackawanna, Luzerne, Monroe, Pike, Sullivan, Susquehanna, Wayne, Wyoming 0 otherwise
Region 7	1 if patient attended hospital in one of the following PA counties: Berks, Carbon, Lehigh, Northampton, Schuylkill 0 otherwise
Region 8	1 if patient attended hospital in one of the following PA counties: Bucks, Chester, Delaware, Montgomery 0 otherwise
Region 9	1 if patient attended hospital in Philadelphia 0 otherwise
Total Hospital Discharges (100 patients)	Total number of hospital inpatient discharges during 2009 for the hospital the patient attended (in units of 100 patients)
Hospital Bed Count (100 beds)	Total number of hospital beds for the hospital the patient attended (in units of 100 beds)

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4 2009

Table 2:

Summary of Patient Characteristics at the Patient Level for Childbirth Sample

Variable	Mean	Std. Dev.
Patient Outcome:		
Had a C-section	18.215%	0.386
Had a Fistula, Rupture or Death	0.021%	0.015
Obstetric Fistula	0.007%	0.008
Uterine Rupture	0.013%	0.012
Patient Race:		
White	69.083%	0.462
Black	15.170%	0.359
Hispanic	6.276%	0.243
Asian	2.251%	0.148
Other Race	8.993%	0.286
Patient Age:		
Age 10-19	10.628%	0.308
Age 20-29	53.014%	0.499
Age 30-39	34.133%	0.474
Age 40-49	2.222%	0.147
Age 50-59	0.004%	0.006
Patient Primary Insurance Status:		
Private Insurance	55.947%	0.496
Medicare	1.242%	0.111
Medicaid	40.367%	0.491
Other Government Insurance	1.096%	0.104
Uninsured	1.135%	0.106
Unknown Insurance	0.183%	0.043
Patient Risk at Admission (MediQual Atlas):		
No Risk	99.769%	0.048
Minimal Risk	0.216%	0.046
Moderate Risk	0.013%	0.012
Severe Risk	0.001%	0.004
Maximal Risk	0.000%	-
n	104,311	

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Table 3:
Summary of Hospital Characteristics at the Patient Level for Childbirth Sample

Variable	Mean	Std. Dev.
Outcomes at Hospital Level:		
Unnecessary C-Sections (%)*	18.215	4.124
Fistula, Rupture or Death*	0.540	0.832
Hospital Size:		
Total Hospital Discharges (100 patients)	56.856	29.998
Hospital Bed Count (100 beds)	3.994	2.211
Hospital Region:		
Region 1	17.869%	0.383
Region 2	6.367%	0.244
Region 3	3.360%	0.180
Region 4	4.611%	0.210
Region 5	15.301%	0.360
Region 6	5.840%	0.235
Region 7	10.288%	0.304
Region 8	22.215%	0.416
Region 9	14.150%	0.349
n	104311	

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
*Weighted by number of deliveries at the hospital.

Table 4:
Summary of Patient and Hospital Characteristics at the Physician Level for Childbirth Sample

Variable	Mean	Std. Dev.	Minimum	Maximum
Physician Experience:				
Quarterly Physician "Unnecessary" Cesarean Rate (%)	18.215	10.299	0	100
Std. Dev. of Physician's Cesarean Rate across Quarters	5.054	3.039	0	50
Maternal Death, Uterine Rupture or Obstetric Fistula	0.005	0.069	0	1
Uterine Rupture	0.003	0.055	0	1
Obstetric Fistula	0.002	0.039	0	1
Physician Births	22.356	16.893	1	175
Racial Composition of Physician's Patients:				
White	70.667	29.374	0	100
Black	14.303	21.768	0	100
Hispanic	5.971	14.550	0	100
Asian	2.068	5.695	0	100
Other Race	8.604	14.794	0	100
Age Composition of Physician's Patients:				
Age 10-19	10.943	14.292	0	100
Age 20-29	53.532	21.018	0	100
Age 30-39	33.296	21.483	0	100
Age 40-49	2.224	5.394	0	100
Age 50-59	0.006	0.234	0	14
Primary Insurance Composition of Physician's Patients:				
Private Insurance	54.558	30.620	0	100
Medicare	1.360	5.858	0	100
Medicaid	41.404	30.369	0	100
Other Government Insurance	1.183	4.030	0	100
Uninsured	1.295	4.996	0	100
Unknown Insurance	0.159	1.419	0	36
Risk Composition of Physician's Patients:				
No Risk	99.754	1.704	50	100
Minimal Risk	0.229	1.657	0	50
Moderate Risk	0.012	0.258	0	10
Severe Risk	0.005	0.283	0	17
Hospital Size where Physician Operates:				
Total Hospital Discharges (100 patients)	55.287	30.204	2.87	145.46
Hospital Bed Counts (100 beds)	3.866	2.163	0.25	15.72
Hospital Region where Physician Operates:				
Region 1	19.753	39.811	0	100
Region 2	6.325	24.321	0	100
Region 3	3.408	18.146	0	100
Region 4	6.281	24.264	0	100
Region 5	16.962	37.513	0	100
Region 6	5.230	22.266	0	100
Region 7	10.032	30.046	0	100
Region 8	21.058	40.749	0	100
Region 9	10.951	31.212	0	100
n	4,665			

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
Note: Patient variables and hospital regions rescaled to range from 0 to 100 instead of 0 to 1

Table 5:
“Unnecessary” Cesarean Rates and Bad Outcomes at the Physician Level by Quarter

	Quarter 1		Quarter 2		Quarter 3		Quarter 4	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Quarterly Physician “Unnecessary” Cesarean Rate (%)	18.916	17.927	18.328	17.068	17.449	16.124	18.221	16.811
# Maternal Deaths, Uterine Ruptures or Obstetric Fistulas	0.001	0.029	0.008	0.088	0.008	0.087	0.003	0.051
# Uterine Ruptures	0.001	0.029	0.006	0.078	0.003	0.058	0.002	0.041
# Obstetric Fistula	-	-	0.002	0.042	0.003	0.058	0.001	0.029
n	1,156		1,156		1,188		1,166	

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Table 6: Quarterly Cesarean Rates at the Physician Level by Bad Outcomes in the Previous Quarter				
	Quarterly Cesarean Rate if No Bad Outcome in Previous Quarter		Quarterly Cesarean Rate if One Bad Outcome in Previous Quarter	
Bad Outcome:	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Maternal Death, Uterine Rupture or Obstetric Fistula	17.835	0.260	19.559	2.253
Uterine Rupture	17.834	0.260	20.539	2.358
Obstetric Fistula	17.845	0.260	17.599	5.118

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Note: None of the differences are statistically significant at the 0.1 level.

Table 7: Summary of Patient Characteristics at the Patient Level for Appendicitis Sample		
Variable	**Mean**	**Std. Dev.**
Patient Outcomes:		
Had an Appendectomy	94.906%	0.220
Had Appendicitis	78.691%	0.410
Had a Negative Appendectomy	21.309%	0.410
Had a Perforated Appendix, Sepsis or Died	13.770%	0.345
Died	0.479%	0.069
Sepsis	1.431%	0.119
Perforated Appendix	12.258%	0.328
Patient Race:		
White	82.198%	0.383
Black	8.215%	0.275
Hispanic	5.134%	0.221
Asian	1.246%	0.111
Other Race	4.280%	0.202
Patient Primary Insurance Status:		
Private Insurance	60.306%	0.489
Medicare	18.373%	0.387
Medicaid	15.368%	0.361
Other Government Insurance	1.361%	0.116
Uninsured	4.067%	0.198
Unknown Insurance	0.427%	0.065
Patient Sex and Age:		
Female	46.363%	0.499
Age 0-4	0.387%	0.062
Age 5-9	1.719%	0.130
Age 10-19	7.413%	0.262
Age 20-29	7.107%	0.257
Age 30-39	5.913%	0.236
Age 40-49	6.444%	0.246
Age 50-59	6.565%	0.248
Age 60-69	5.042%	0.219
Age 70-79	3.444%	0.182
Age 80+	2.331%	0.151
Male	53.637%	0.499
Age 0-4	0.433%	0.066
Age 5-9	2.348%	0.151
Age 10-19	10.412%	0.305
Age 20-29	8.365%	0.277
Age 30-39	5.896%	0.236
Age 40-49	6.242%	0.242
Age 50-59	7.494%	0.263
Age 60-69	6.392%	0.245
Age 70-79	3.952%	0.195
Age 80+	2.106%	0.144
Patient Risk at Admission (MediQual Atlas):		
No Risk	69.221%	0.462
Minimal Risk	18.545%	0.389
Moderate Risk	8.868%	0.284
Severe Risk	3.182%	0.176
Maximal Risk	0.183%	0.043
n	17,335	

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Table 8: Appendicitis Sample		
Variable	Mean	Std. Dev.
Outcomes at the Hospital Level:		
Rate NA*	22.465%	0.172
Patient Death, Sepsis or Perforation*	27.217	22.037
Hospital Size:		
Total Hospital Discharges (100 patients)	49.644	32.523
Total Hospital Beds (100 beds)	3.771	2.810
Hospital Region:		
Region 1	21.915%	0.414
Region 2	6.236%	0.242
Region 3	4.021%	0.196
Region 4	4.378%	0.205
Region 5	12.847%	0.335
Region 6	6.870%	0.253
Region 7	9.628%	0.295
Region 8	19.879%	0.399
Region 9	14.226%	0.349

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

*Weighted by number of appendicitis/appendectomy patients.

Table 9:

Summary of Patient and Hospital Characteristics at the Physician Level for Appendicitis Sample

	Variable	Mean	Std. Dev.	Minimum	Maximum
Physician Experience:	Quarterly Rate NA	21.918	38.193	0	100
	Std. Dev. of Quarterly Rate NA Over Time	4.307	7.675	**0**	**50**
	# Patient Deaths, Sepsis or Perforations	0.393	0.809	0	8
	# Deaths	0.014	0.116	0	1
	# Sepsis Cases	0.041	0.208	0	3
	# Perforation	0.350	0.776	0	8
	# Appendectomies	5.783	4.536	0	23
	# Appendicitis Cases Treated	2.247	3.115	0	24
Racial Composition of Physician's Patients:	White	82.706	32.633	0	100
	Black	8.628	24.235	0	100
	Hispanic	0.039	0.159	0	1
	Asian	1.142	8.505	0	100
	Other Race	3.550	14.689	0	100
Primary Insurance Composition of Physician's Patients:	Private Insurance	53.504	41.912	0	100
	Medicare	26.879	39.587	0	100
	Medicaid	14.236	28.464	0	100
	Other Government Insurance	1.347	9.322	0	100
	Uninsured	3.539	14.134	0	100
	Unknown Insurance	0.389	5.121	0	100
Sex and Age Composition of Physician's Patients:	Female				
	Age 0-4	0.393	5.497	0	100
	Age 5-9	1.219	8.406	0	100
	Age 10-19	5.325	16.786	0	100
	Age 20-29	6.146	18.658	0	100
	Age 30-39	5.354	17.690	0	100
	Age 40-49	6.192	19.400	0	100
	Age 50-59	7.093	21.286	0	100
	Age 60-69	6.247	20.661	0	100
	Age 70-79	5.123	19.781	0	100
	Age 80+	3.845	17.481	0	100
	Male				
	Age 0-4	0.390	5.303	0	100
	Age 5-9	1.463	8.865	0	100
	Age 10-19	6.920	18.908	0	100
	Age 20-29	6.287	18.007	0	100
	Age 30-39	4.562	15.367	0	100
	Age 40-49	6.188	19.203	0	100
	Age 50-59	9.020	24.231	0	100
	Age 60-69	8.918	25.203	0	100
	Age 70-79	6.066	21.499	0	100
	Age 80+	3.248	15.942	0	100
	n	6,071			

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Note: Patient variables and hospital regions recaled to range from 0 to 100 instead of 0 to 1

Table 9 (continued):
Summary of Patient and Hospital Characteristics at the Physician Level for Appendicitis Sample

	Variable	Mean	Std. Dev.	Minimum	Maximum
Risk Composition of Physician's Patients:	No Risk	56.859	44.383	0	100
	Minimal Risk	23.223	37.054	0	100
	Moderate Risk	14.120	32.076	0	100
	Severe Risk	5.450	21.122	0	100
	Maximal Risk	0.348	5.589	0	100
Hospital Characteristics where Physician Operates:	Total Hospital Discharges (100 Patients)	50.598	33.026	0	145.46
	Hospital Beds (100 beds)	4.019	3.066	0.1	15.72
	Region 1	22.936	42.010	0	100
	Region 2	7.936	26.987	0	100
	Region 3	4.261	20.190	0	100
	Region 4	4.475	20.668	0	100
	Region 5	12.204	32.720	0	100
	Region 6	9.514	29.286	0	100
	Region 7	9.042	28.550	0	100
	Region 8	15.469	36.066	0	100
	Region 9	14.163	34.806	0	100
	n	6,071			

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Note: Patient variables and hospital regions recaled to range from 0 to 100 instead of 0 to 1

Table 10:

Rates NA and Bad Outcomes at the Physician Level by Quarter

Variable	Quarter 1		Quarter 2		Quarter 3		Quarter 4	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Quarterly Rate NA (%)	38.727	47.156	36.886	46.759	36.544	46.856	38.316	46.931
# Death, Sepsis or Perforations	0.381	0.792	0.412	0.824	0.402	0.855	0.377	0.763
# Deaths cases	0.018	0.133	0.013	0.113	0.011	0.105	0.013	0.112
# Sepsis	0.046	0.213	0.031	0.177	0.044	0.210	0.043	0.228
# Perforations	0.330	0.755	0.377	0.800	0.360	0.826	0.332	0.718
n	1,429		1,476		1,453		1,417	

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Table 11:
Mean Physician Rates NA by Whether the Physician Witnessed a Bad Outcome in the Previous Quarter

	Quarterly Rate NA if No Bad Outcome in Previous Quarter		Quarterly Rate NA if at Least One Bad Outcome in Previous Quarter	
Bad Outcome Defined As:	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Perforated Appendix, Sepsis or Death***	42.022	47.807	35.423	46.327
Death***	30.269	43.849	44.856	48.704
Sepsis***	31.218	44.259	43.545	48.525
Perforated or Ruptured Appendix***	41.639	47.735	35.488	46.354

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

*** indicates statistically significant difference at the 0.01 level

Table 12: Robust OLS Regression Results for Effect of Previous Bad Outcomes on Current Quarterly Cesarean Rate at the Physician Level									
Bad Outcome Defined as:	Rupture, Fistula or Death			Uterine Rupture			Obstetric Fistula		
Variable	Coefficient	Std. Err.		Coefficient	Std. Err.		Coefficient	Std. Err.	
Physician Experience:									
Bad Outcome in Previous Quarter	-1.561	1.873		-0.667	2.008		-2.883	3.983	
"Unnecessary" Cesarean Rate in Previous Quarter	0.439	0.040	***	0.439	0.040	***	0.439	0.040	***
Deliveries	0.022	0.018		0.021	0.018		0.022	0.018	
Racial Composition of Physician's Patients:									
Black	-0.029	0.027		-0.029	0.027		-0.029	0.027	
Hispanic	-0.021	0.021		-0.021	0.021		-0.021	0.021	
Asian	-0.125	0.065	*	-0.124	0.065	*	-0.125	0.065	*
Other Race	-0.024	0.038		-0.024	0.038		-0.024	0.038	
Age Composition of Physician's Patients:									
Age 10-19	-0.013	0.030		-0.013	0.030		-0.013	0.030	
Age 30-39	0.029	0.027		0.029	0.027		0.029	0.027	
Age 40-49	0.020	0.056		0.020	0.056		0.019	0.056	
Age 50-59	-0.104	0.380		-0.102	0.380		-0.101	0.377	
Insurance Composition of Physician's Patients:									
Medicare	-0.300	0.083	***	-0.300	0.083	***	-0.299	0.083	***
Medicaid	-0.014	0.017		-0.014	0.017		-0.014	0.017	
Other Government Insurance	0.094	0.098		0.094	0.098		0.094	0.098	
Uninsured	-0.082	0.059		-0.082	0.059		-0.082	0.059	
Unkown Insurance	0.294	0.120	**	0.295	0.119	**	0.294	0.119	**
Risk Composition of Physician's Patients:									
Minimal Risk	0.012	0.189		0.012	0.189		0.012	0.189	
Moderate Risk	-0.193	1.978		-0.193	1.978		-0.193	1.978	
Severe Risk	2.950	78.000		2.950	28.887		2.950	28.538	
Hospital Size where Physician Operates:									
Total Hospital Discharges (100 patients)	0.008	0.020		0.008	0.020		0.008	0.020	
Hospital Bed Counts (100 beds)	0.166	0.259		0.168	0.259		0.165	0.259	
Hospital Region where Physician Operates:									
Region 2	-0.007	0.013		-0.007	0.013		-0.007	0.013	
Region 3	-0.049	0.014	***	-0.049	0.014	***	-0.049	0.014	***
Region 4	-0.022	0.014		-0.022	0.014		-0.022	0.014	
Region 5	-0.031	0.010	***	-0.030	0.010	***	-0.031	0.010	***
Region 6	0.010	0.015		0.010	0.015		0.010	0.015	
Region 7	-0.025	0.011	**	-0.025	0.011	**	-0.025	0.011	**
Region 8	-0.006	0.009		-0.006	0.009		-0.006	0.009	
Region 9	0.007	0.019		0.007	0.019		0.007	0.019	
Constant	10.046	1.794	***	0.000		***	10.054	1.794	***
n	2,477			2,477			2,477		
R-sq	25.60%			25.59%			25.60%		

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

*,**,*** indicate statistical significance at the 0.1, 0.05, and 0.01 levels respectively

Table 13: Robust Fixed Effects Regression Results for Effect of Previous Bad Outcomes on Physician Quarterly Cesarean Rate						
Bad Outcome Defined as:	**Death, Rupture or Fistula**		**Uterine Rupture**		**Obstetric Fistula**	
Variable:	**Coefficient**	**Std. Err.**	**Coefficient**	**Std. Err.**	**Coefficient**	**Std. Err.**
Physician Experience:						
Previous Bad Outcome	-2.525	2.527	-2.633	3.214	-2.367	4.112
Patient Race:						
Black	0.008	0.049	0.008	0.049	0.008	0.049
Hispanic	-0.064	0.083	-0.063	0.083	-0.063	0.083
Asian	-0.012	0.078	-0.011	0.078	-0.013	0.078
Other Race	0.034	0.045	0.034	0.045	0.033	0.046
Patient Age:						
Age 10-19	-0.007	0.035	-0.007	0.035	-0.007	0.035
Age 30-39	-0.011	0.030	-0.011	0.030	-0.011	0.030
Age 40-49	-0.026	0.058	-0.026	0.058	-0.027	0.058
Age 50-59	1.869	0.296 ***	1.869	0.296 ***	1.868	0.296 ***
Patient Insurance Status:						
Medicare	-0.324	0.129 **	-0.326	0.129 **	-0.324	0.129 **
Medicaid	-0.028	0.025	-0.028	0.025	-0.028	0.025
Other Government Insurance	0.201	0.117 *	0.201	0.117 *	0.201	0.117 *
Uninsured	0.034	0.075	0.034	0.075	0.035	0.075
Unknown Insurance	0.145	0.133	0.145	0.133	0.145	0.133
Patient Risk at Admission:						
Minimal Risk	0.051	0.098	0.050	0.098	0.051	0.098
Moderate Risk	0.152	0.756	0.152	0.756	0.153	0.756
Severe Risk	0.170	0.066 ***	0.171	0.066 ***	0.170	0.066 ***
Constant	18.800	1.823 ***	18.809	1.823 ***	18.801	1.826 ***
n	2,477		2,477		2,477	
R-sq	1.33%		1.35%		1.36%	

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

Note: We are controlling for each individual physician in this fixed effects model.

*, **, *** indicates statistical significance at the 0.1, 0.05, and 0.01 levels respectively

Table 14:
Logistic Regression Results for Effect of Physician Quarterly Cesarean Rate on a Patient's Probability of Having a C-Section

Variable	Odds Ratio	Std. Err.	
Physician Practice Pattern:			
Physician Quarterly "Unnecessary" Cesarean Rate	1.025	0.001	***
Patient Complication:			
Maternal Death, Uterine Rupture or Obstetric Fistula	17.135	11.100	***
Patient Race:			
Black	1.052	0.034	
Hispanic	0.905	0.042	**
Asian	0.868	0.061	**
Other Race	1.067	0.043	
Patient Insurance Status:			
Medicare	0.953	0.091	
Medicaid	0.791	0.019	***
Other Government Insurance	0.710	0.077	***
Uninsured	0.739	0.079	***
Unknown Insurance	1.197	0.233	
Patient Age:			
Age 10-19	1.069	0.037	*
Age 30-39	0.954	0.021	**
Age 40-49	1.546	0.090	***
Age 50-59	1.445	1.670	
Patient Risk at Admission:			
Minimal Risk	2.090	0.366	***
Moderate Risk	3.560	2.318	*
Hospital Size:			
Total Hospital Discharges (100 patients)	0.999	0.001	
Hospital Bed Counts (100 beds)	1.023	0.010	**
Hospital Region:			
Region 2	1.080	0.045	
Region 3	0.734	0.043	**
Region 4	0.965	0.052	*
Region 5	0.866	0.029	***
Region 6	1.068	0.059	
Region 7	0.884	0.039	***
Region 8	1.015	0.030	
Region 9	0.883	0.045	**

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
*, **, *** indicates statistical significance at the 0.1, 0.05, and 0.01 levels respectively

Table 15:
Robust OLS Regression Results for Effect of Bad Outcomes in the Previous Quarter on Current Quarterly Rate NA at the Physician Level

Bad Outcome Defined As:	Patient Death, Sepsis, or Perforation		Patient Death	
Variable	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Bad Outcomes in the Previous Quarter	0.699	0.410 *	1.373	4.135
Rate NA in the Previous Quarter	0.599	0.023 ***	0.596	0.023 ***
# Appendicitis Cases Treated	-1.892	0.161 ***	-1.836	0.157 ***
Black	-0.016	0.035	-0.016	0.035
Hispanic	-5.257	4.111	-5.081	4.116
Asian	-0.188	0.088 **	-0.187	0.088 **
Other Race	-0.060	0.039	-0.060	0.039
Medicare	0.047	0.028 *	0.047	0.028 *
Medicaid	0.015	0.026	0.016	0.026
Other Government Insurance	-0.010	0.086	-0.011	0.086
Uninsured	-0.032	0.044	-0.033	0.044
Unknown Insurance	-0.222	0.137	-0.223	0.137
Female Age 0-4	0.287	0.250	0.307	0.255
Female Age 5-9	0.004	0.069	0.002	0.069
Female Age 20-29	-0.052	0.045	-0.056	0.045
Female Age 30-39	0.003	0.050	-0.001	0.050
Female Age 40-49	0.107	0.045 **	0.105	0.045 **
Female Age 50-59	0.129	0.041 ***	0.126	0.041 ***
Female Age 60-69	0.116	0.046 **	0.115	0.046 **
Female Age 70-79	0.148	0.055 ***	0.146	0.055 ***
Female Age 80+	0.208	0.058 ***	0.205	0.058 ***
Male Age 0-4	0.661	0.302 **	0.664	0.293 **
Male Age 5-9	0.131	0.057 **	0.134	0.058 **
Male Age 10-19	-0.042	0.036	-0.043	0.036
Male Age 20-29	-0.064	0.038 *	-0.067	0.038 *
Male Age 30-39	0.006	0.039	0.004	0.039
Male Age 40-49	0.122	0.040 ***	0.120	0.040 ***
Male Age 50-59	0.195	0.037 ***	0.193	0.038 ***
Male Age 60-69	0.235	0.041 ***	0.233	0.041 ***
Male Age 70-79	0.195	0.046 ***	0.193	0.046 ***
Male Age 80+	0.175	0.060 ***	0.173	0.060 ***
Minimal Risk	0.042	0.021 **	0.042	0.021 **
Moderate Risk	0.109	0.026 ***	0.109	0.026 ***
Severe Risk	0.103	0.039 ***	0.104	0.039 ***
Maximal Risk	-0.083	0.252	-0.084	0.253

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
*, **, *** indicate statistical significance at the 0.1, 0.05, 0.01 levels, respectively.

Table 15(continued): Robust OLS Regression Results for Effect of Bad Outcomes in the Previous Quarter on Current Quarterly Rate NA at the Physician Level

Bad Outcome Defined As:	Patient Death, Sepsis, or Perforation		Patient Death	
Variable	Coefficient	Std. Error	Coefficient	Std. Error
Total Hospital Discharges (100 patients)	0.103	0.037 ***	0.106	0.036 ***
Hospital Beds (100 beds)	-0.544	0.407	-0.574	0.406
Region 2	0.013	0.022	0.013	0.022
Region 3	0.027	0.022	0.026	0.022
Region 4	0.003	0.025	0.004	0.025
Region 5	0.003	0.017	0.003	0.017
Region 6	0.016	0.022	0.015	0.022
Region 7	0.002	0.018	0.001	0.018
Region 8	0.044	0.017 ***	0.043	0.017 **
Region 9	0.046	0.020 **	0.045	0.020 **
Constant	3.228	2.794	3.745	2.774
n	2,454		2,454	
R-sq	76.72%		76.70%	

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

*, **, *** indicate statistical significance at the 0.1, 0.05, 0.01 levels, respectively.

Table 15 (continued):

Robust OLS Regression Results for Effect of Bad Outcomes in the Previous Quarter on Current Quarterly Rate NA at the Physician Level

Bad Outcome Defined As:	Sepsis		Perforation	
Variable	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Bad Outcomes in the Previous Quarter	-2.290	2.061	0.832	0.419 **
Rate NA in the Previous Quarter	0.595	0.023 ***	0.599	0.023 ***
# Appendicitis Cases Treated	-1.820	0.156 ***	-1.899	0.162 ***
Black	-0.016	0.035	-0.016	0.035
Hispanic	-5.097	4.121	-5.311	4.110
Asian	-0.187	0.088 **	-0.189	0.088 **
Other Race	-0.061	0.039	-0.060	0.039
Medicare	0.047	0.028 *	0.047	0.028 *
Medicaid	0.016	0.026	0.016	0.026
Other Government Insurance	-0.012	0.086	-0.010	0.086
Uninsured	-0.031	0.044	-0.032	0.044
Unknown Insurance	-0.224	0.137	-0.222	0.137
Female Age 0-4	0.315	0.255	0.287	0.249
Female Age 5-9	0.002	0.069	0.003	0.069
Female Age 20-29	-0.055	0.045	-0.051	0.045
Female Age 30-39	-0.001	0.050	0.004	0.050
Female Age 40-49	0.105	0.045 **	0.107	0.045 **
Female Age 50-59	0.127	0.041 ***	0.129	0.041 ***
Female Age 60-69	0.115	0.046 **	0.117	0.046 **
Female Age 70-79	0.147	0.055 ***	0.149	0.055 ***
Female Age 80+	0.205	0.058 ***	0.208	0.058 ***
Male Age 0-4	0.662	0.293 **	0.660	0.304 **
Male Age 5-9	0.133	0.058 **	0.130	0.057 **
Male Age 10-19	-0.043	0.036	-0.042	0.036
Male Age 20-29	-0.067	0.038 *	-0.063	0.038 *
Male Age 30-39	0.004	0.039	0.006	0.039
Male Age 40-49	0.120	0.040 ***	0.123	0.040 ***
Male Age 50-59	0.194	0.037 ***	0.196	0.037 ***
Male Age 60-69	0.233	0.041 ***	0.236	0.041 ***
Male Age 70-79	0.194	0.046 ***	0.196	0.046 ***
Male Age 80+	0.172	0.060 ***	0.176	0.060 ***
Minimal Risk	0.042	0.021 **	0.042	0.021 **
Moderate Risk	0.110	0.026 ***	0.109	0.026 ***
Severe Risk	0.106	0.039 ***	0.103	0.039 ***
Maximal Risk	-0.083	0.252	-0.082	0.252

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

*, **, *** indicate statistical significance at the 0.1, 0.05, 0.01 levels, respectively.

Table 15 (continued):
Robust OLS Regression Results for Effect of Bad Outcomes in the Previous Quarter on Current Quarterly Rate NA at the Physician Level

Bad Outcome Defined As:	Sepsis			Perforation		
Variable	**Coefficient**	**Std. Error**		**Coefficient**	**Std. Error**	
Total Hospital Discharges (100 patients)	0.106	0.036	***	0.103	0.037	***
Hospital Beds (100 beds)	-0.580	0.405		-0.540	0.407	
Region 2	0.013	0.022		0.013	0.022	
Region 3	0.028	0.022		0.027	0.022	
Region 4	0.003	0.025		0.003	0.025	
Region 5	0.002	0.017		0.003	0.017	
Region 6	0.014	0.022		0.016	0.022	
Region 7	0.002	0.018		0.002	0.018	
Region 8	0.042	0.017	**	0.044	0.017	***
Region 9	0.045	0.020	**	0.046	0.020	**
Constant	3.807	2.764		3.163	2.791	
n	2,454			2,454		
R-sq	76.71%			76.72%		

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
*, **, *** indicate statistical significance at the 0.1, 0.05, 0.01 levels, respectively.

Table 16:
Fixed Effect Regression with Robust Standard Errors Results for Effect of Bad Outcomes in the Previous Quarter on Current Quarterly Rate NA at the Physician Level

Bad Outcome Defined As:	Appendix		Patient Death	
Variable	Coefficient	Std. Error	Coefficient	Std. Error
Bad Outcome in Previous Quarter	0.475	0.422	3.632	1.622 **
Black	-0.047	0.050	-0.047	0.050
Hispanic	-0.863	4.268	-0.967	4.279
Asian	-0.058	0.045	-0.057	0.045
Other Race	-0.083	0.042 **	-0.082	0.042 *
Medicare	0.006	0.030	0.006	0.030
Medicaid	0.050	0.026 *	0.051	0.026 **
Other Government Insurance	0.161	0.118	0.159	0.118
Uninsured	0.024	0.043 *	0.023	0.043
Unknown Insurance	-0.160	0.079 **	-0.159	0.079 **
Female Age 0-4	0.055	0.076	0.051	0.076
Female Age 5-9	-0.053	0.060	-0.057	0.060
Female Age 20-29	-0.031	0.043	-0.034	0.043
Female Age 30-39	0.010	0.048	0.007	0.048
Female Age 40-49	0.012	0.041	0.011	0.042
Female Age 50-59	0.002	0.035	0.000	0.034
Female Age 60-69	-0.053	0.039	-0.055	0.038
Female Age 70-79	-0.063	0.055	-0.064	0.055
Female Age 80+	-0.088	0.050 *	-0.089	0.050 *
Male Age 0-4	0.306	0.222	0.307	0.220
Male Age 5-9	-0.086	0.055	-0.088	0.055
Male Age 10-19	-0.032	0.033	-0.034	0.032
Male Age 20-29	-0.052	0.038	-0.054	0.037
Male Age 30-39	-0.053	0.043	-0.055	0.042
Male Age 40-49	-0.010	0.040	-0.011	0.040
Male Age 50-59	0.041	0.038	0.040	0.038
Male Age 60-69	-0.004	0.039	-0.006	0.039
Male Age 70-79	-0.009	0.037	-0.011	0.037
Male Age 80+	-0.019	0.063	-0.020	0.062
Minimal Risk	0.065	0.023 ***	0.065	0.023 ***
Moderate Risk	0.106	0.035 ***	0.106	0.035 ***
Severe Risk	0.130	0.057 **	0.131	0.057 **
Maximal Risk	0.020	0.050	0.022	0.050
Constant	32.879	2.580 ***	33.238	2.508 ***
n	2,491		2,491	
R-sq	15.54%		17.86%	

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
Note: We are controlling for each individual physician in this fixed effects regression.
*, **, *** indicate statistical significance at the 0.1, 0.05, 0.01 levels, respectively.

Table 16 (continued):
Fixed Effect Regression with Robust Standard Errors Results for Effect of Bad Outcomes in the Previous Quarter on Current Quarterly Rate NA at the Physician Level

Bad Outcome Defined As:	Patient Sepsis		Patient Perforated or Ruptured Appendix	
Variable	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Bad Outcome in Previous Quarter	-1.840	1.402	0.586	0.444
Black	-0.046	0.049	-0.047	0.050
Hispanic	-1.254	0.049	-0.875	4.269
Asian	-0.057	0.045	-0.059	0.045
Other Race	-0.081	0.042 *	-0.083	0.042 **
Medicare	0.006	0.030	0.006	0.030
Medicaid	0.051	0.026 **	0.050	0.026 *
Other Government Insurance	0.158	0.118	0.161	0.118
Uninsured	0.024	0.043	0.024	0.043
Unknown Insurance	-0.159	0.079 **	-0.161	0.079 **
Female Age 0-4	0.053	0.075	0.056	0.075
Female Age 5-9	-0.057	0.060	-0.052	0.061
Female Age 20-29	-0.031	0.043	-0.030	0.043
Female Age 30-39	0.008	0.048	0.011	0.048
Female Age 40-49	0.012	0.042	0.013	0.041
Female Age 50-59	0.002	0.034	0.002	0.035
Female Age 60-69	-0.053	0.038	-0.052	0.039
Female Age 70-79	-0.062	0.055	-0.062	0.055
Female Age 80+	-0.090	0.050 *	-0.088	0.050 *
Male Age 0-4	0.299	0.220	0.305	0.222
Male Age 5-9	-0.088	0.055	-0.086	0.055
Male Age 10-19	-0.033	0.032	-0.032	0.033
Male Age 20-29	-0.053	0.037	-0.052	0.038
Male Age 30-39	-0.054	0.043	-0.053	0.043
Male Age 40-49	-0.010	0.040	-0.009	0.040
Male Age 50-59	0.041	0.038	0.042	0.038
Male Age 60-69	-0.004	0.039	-0.003	0.039
Male Age 70-79	-0.010	0.037	-0.009	0.037
Male Age 80+	-0.019	0.062	-0.018	0.062
Minimal Risk	0.065	0.023 ***	0.065	0.023 ***
Moderate Risk	0.105	0.035 ***	0.106	0.035 ***
Severe Risk	0.129	0.057 **	0.129	0.057 **
Maximal Risk	0.014	0.053	0.020	0.051
Constant	33.273	2.532 ***	32.810	2.584 ***
n	2,491		2,491	
R-sq	18.58%		15.17%	

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
Note: We are controlling for each individual physician in this fixed effects regression.
*, **, *** indicate statistical significance at the 0.1, 0.05, 0.01 levels, respectively.

Table 17: Logistic Regression Results for Impact of Physician's Rate NA on Patients' Probability of Having a Negative Appendectomy

Variable	Odds Ratio	Std. Error	
Physician's Quarterly Rate NA	1.056	0.001	***
Black	1.177	0.224	
Hispanic	0.534	0.160	**
Asian	0.939	0.399	
Other Race	1.426	0.445	
Medicare	1.153	0.181	
Medicaid	1.237	0.190	
Other Government Insurance	1.233	0.501	
Uninsured	0.465	0.148	**
Unknown Insurance	0.981	0.778	
Female Age 0-4	5.380	3.884	**
Female Age 5-9	0.644	0.357	
Female Age 20-29	1.223	0.313	
Female Age 30-39	1.229	0.324	
Female Age 40-49	1.457	0.365	
Female Age 50-59	1.503	0.372	*
Female Age 60-69	0.879	0.240	
Female Age 70-79	1.328	0.425	
Female Age 80+	1.043	0.378	
Male Age 0-4	24.612	12.656	***
Male Age 5-9	0.727	0.340	
Male Age 10-19	0.330	0.105	***
Male Age 20-29	0.395	0.130	***
Male Age 30-39	0.582	0.183	*
Male Age 40-49	0.852	0.229	
Male Age 50-59	1.374	0.339	
Male Age 60-69	1.659	0.433	*
Male Age 70-79	1.017	0.317	
Male Age 80+	0.546	0.207	

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009

*, **, *** indicate statistical significance at the 0.1, 0.05, 0.01 levels, respectively.

Table 17 (continued):
Logistic Regression Results for Impact of Physician's Rate NA on Patients' Probability of Having a Negative Appendectomy

Variable	Odds Ratio	Std. Error	
Minimal Risk	3.011	0.383	***
Moderate Risk	6.677	1.187	***
Severe Risk	8.213	2.150	***
Maximal Risk	24.979	25.340	***
Total Hospital Discharges (100 patients)	1.008	0.004	**
Hospital Beds (100 beds)	0.938	0.038	
Region 2	0.753	0.156	
Region 3	0.809	0.200	
Region 4	0.898	0.229	
Region 5	0.945	0.140	
Region 6	1.129	0.241	
Region 7	0.640	0.114	**
Region 8	0.949	0.136	
Region 9	1.009	0.186	
n	10,385		

Source: Pennsylvania Inpatient Hospital Admissions Data, PHC4, 2009
*, **, *** indicate statistical significance at the 0.1, 0.05, 0.01 levels, respectively.

Bibliography

Center for Medicare and Medicaid Spending. NHE projections 2010-2020: Forecast summary and selected tables. Center for Medicare and Medicaid Spending. https://www.cms.gov/NationalHealthExpendData/03_NationalHealthAccountsProjected.asp#TopOfPage (Accessed September 2, 2011).

Fisher, Elliot, David E. Wennberg, Therese A. Stukel and Daniel J. Gottlieb. 2003 The implications of regional variations in medicare spending. Part1: The content, quality and accessibility of care. *Annals of Internal Medicine* 138: 273-87.

Fisher, Elliot, David E. Wennberg, Therese A. Stukel and Daniel J. Gottlieb. 2003 The implications of regional variations in medicare spending. Part 2: Health outcomes and satisfaction with care. *Annals of Internal Medicine* 138: 288-98.

Fisher, Elliot, David E. Wennberg, Therese A. Stukel, and Daniel J. Gottlieb. 2004. Variations in the longitudinal efficiency of academic medical centers. *Health Affairs:* 19-23.

Bronner, Kristen K., Megan McAndrew Cooper, Therese A. Stukel and John E. Wennberg. 1998. The Dartmouth atlas of health care in Pennsylvania. The Center for Evaluative Clinical Sciences Dartmouth Medical School. http://www.dartmouthatlas.org/downloads/atlases/pennsylvania_atlas.pdf (Accessed December 10, 2011).

Fuchs, Victor. 1978. The supply of surgeons and the demand for operations. *The Journal* *of Human Resources* 13: 35-56.

Gruber, Jonathan and Maria Owings. 1996. Physician financial incentives and cesarean section delivery. *The Rand Journal of Economics* 27: 99-123.

Thaler, Richard H. and Cass R. Sunstein. 2009. Biases and Blunders . In *Nudge: Improving decisions about health, wealth, and happiness,* 17-39. New York: Penguin Books.

Tversky, Amos and Daniel Kahneman. 1974. Judgement under uncertainty: Heuristics and biases. *Science* 185: 1124-31.

Tversky, Amos and Daniel Kahneman. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 4: 207-232.

Myers, David G. 2003. Do we fear the right things? *Skeptic* 10: 56-57.

Groopman, Jerome. 2007. *How doctors think.* Boston: Houghton Mifflin Co.

Redelmeier, Donald. 2005. The cognitive psychology of missed diagnoses. *Annals of Internal Medicine* 142: 115-120.

Croskerry, Pat. 2009. A universal model of diagnostic reasoning. *Academic Medicine* 84: 1022-28.

Milstein, Arnold and Nancy E. Adler. 2003. Out of sight, out of mind: Why doesn't widespread clinical failure command our attention?. *Health Affairs* 22: 119-27.

Poses, Roy M. and Michelle Anthony. 1991. Availability, wishful thinking, and physicians' diagnostic judgments for patients with suspected bacteremia. *Medical Decision Making* 11: 159-68.

Mamede, Silvia, Tamara von Gog, Kees van den Berge and Remy M.J.P. Rikers. 2010. Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. *JAMA* 304: 1198-1203.

Camacho, Nuno, Bas Donkers and Stefan Stremersch. 2011. Predictably Non-Bayesian: Quantifying Salience Effects in Physician Learning about Drug Quality. *Marketing Science* 30: 305-20.

Department of Health and Human Services Agency for Healthcare Research and Quality. 2007. Guide to inpatient quality indicators. AHRQ. http://www.qualityindicators.ahrq.gov/Downloads/Modules/IQI/V31/iqi_guide_v31.pdf (Accessed April 1, 2012).

Martin, Joyce A., Brady E. Hamilton, Stephanie J. Ventura and Michelle J.K. Osterman. 2011. Births: Final Data for 2009. National Vital Statistics Report 60: 1-72.

Baicker, Katherine, Kasey S. Buckles and Amitabh Chandra. 2006. Geographic Variation in the Appropriate Use of Cesarean Delivery. *Health Affairs* 25: w355-67.

Lyndon-Rochelle, M., V.L. Holt, D.P. Martin and T.R. Easterling. 2000. Association between method of delivery and maternal rehospitalization. *JAMA* 283: 2411-6.

Triunfo, Patricia and Maximo Rossi. 2009. The effect of physicians' remuneration system on the Cesarean section rate: the Uruguayan case. *International Journal of Health Care Finance Economics* 9: 333-345.

USAID. 2009. Maternal and child health: Obstetric fistula: a devastating injury. USAID. http://www.usaid.gov/our_work/global_health/mch/mh/techareas/fistula.html

(Accessed February 29, 2012).

Nahum, Gerard G. 2010. Uterine rupture in pregnancy. Medscape Reference. http://emedicine.medscape.com/article/275854-overview (Accessed February 29, 2012).

Flum, David R. and Thomas Koepsell. 2002. The clinical and economic correlates of misdiagnosed appendicitis. *Archives of Surgery* 137: 799-804.

Marks, Jay W. 2011. Appendictis. MedicineNet. http://www.medicinenet.com/appendicitis/article.htm (Accessed April 1, 2012).

Seetahal, Shiva A., Oluwaseyi B. Bolorunduro, Trishanna C. Sookdeo, and Tolulope A. Oyetunji. 2011. Negative appendectomy: A 10-year review of a nationally representative sample. *The American Journal of Surgery* 201: 433-7.

Cowen, Tyler. 2010. Following the money, doctors ration care. *The New York Times*, December 11, Economic view, New York edition.