

NBER WORKING PAPER SERIES

PHYSICIAN PRACTICE STYLE AND PATIENT HEALTH OUTCOMES:
THE CASE OF HEART ATTACKS

Janet Currie
W. Bentley MacLeod
Jessica Van Parys

Working Paper 21218
<http://www.nber.org/papers/w21218>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2015

The authors gratefully acknowledge the Florida Agency for Health Care Administration (AHCA) for supplying the data and Dr. Andrew Doorey, Dr. Shunichi Homma, participants at the June 2015 Society of Labor Economists conference, and two anonymous referees for helpful comments. We thank the Center for Health and Well being at Princeton University for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Janet Currie, W. Bentley MacLeod, and Jessica Van Parys. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Physician Practice Style and Patient Health Outcomes: The Case of Heart Attacks
Janet Currie, W. Bentley MacLeod, and Jessica Van Parys
NBER Working Paper No. 21218
May 2015, Revised February 2016
JEL No. I11

ABSTRACT

When a patient arrives at the Emergency Room with acute myocardial infarction (AMI), the provider on duty must quickly decide how aggressively the patient should be treated. Using Florida data on all such patients from 1992-2014, we decompose practice style into two components: The provider's probability of conducting invasive procedures on the average patient (which we characterize as aggressiveness), and the responsiveness of the choice of procedure to the patient's characteristics. We show that within hospitals and years, patients with more aggressive providers have consistently higher costs and better outcomes. Since all patients benefit from higher utilization of invasive procedures, targeting procedure use to the most appropriate patients benefits these patients at the expense of the less appropriate patients. We also find that the most aggressive and responsive physicians are young, male, and trained in top 20 schools.

Janet Currie
Princeton University
316 Wallace Hall
Princeton, NJ 08544
and NBER
jcurrie@princeton.edu

Jessica Van Parys
Department of Economics
Hunter College
Jessica.Van.Parys@dartmouth.edu

W. Bentley MacLeod
Department of Economics
Columbia University
420 West 118th Street, MC 3308
New York, NY 10027
and NBER
bentley.macleod@columbia.edu

One of the most controversial issues in medicine is whether providers should be evaluated in terms of their adherence to simple metrics. Such metrics have become increasingly popular as a way to improve the quality of health care. For example, under the Affordable Care Act, Accountable Care Organizations are judged on the basis of criteria including: The fraction of patients who receive influenza immunization, tobacco screening, and other forms of screening, as well as whether patients with coronary artery disease are taking appropriate medications, and other such metrics.

In addition, it is thought that better metrics could help to control costs. Findings from the Dartmouth medical atlas project show that in high spending regions of the country, patients receive consistently more treatment. For example, Sutherland, Fischer, and Skinner (2009) found that “discretionary decisions by providers seem to account for most of the regional variation in spending,” and that outcomes were not better in high spending areas of the country. The strong implication is that costs could be reduced without sacrificing outcomes by reigning in the discretion of medical providers (Garber and Skinner, 2008).

However, providers often argue that decisions about treatment should be tailored to the needs of individual patients, and that this type of sensitivity to patient characteristics cannot be captured through adherence to simple rules. In a recent New York Times editorial cardiologist Sandeep Jauhar argued that “... guidelines and checklists are unpopular among most American providers. Instead of being allowed to deliver “patient-centered” care, many providers feel they are being co-opted by regulations... Guidelines are supposed to assist and advise. But all too often, recommended care in certain situations becomes mandated care in all situations.” (New York Times, Dec. 11, 2014). Moreover, the use of guidelines raises significant issues about how meaningful protocols are to be established.

In this paper, we decompose provider practice style into two dimensions, whether providers are more or less aggressive in their use of invasive treatments (a mean effect) and whether there are differences in the extent to which providers tailor their treatment decisions to the characteristics of individual patients (a slope effect). We then ask whether these differences matter for costs and health outcomes. Specifically, we use a rich data set of Florida patients arriving at the Emergency Room (ER) with acute myocardial infarction (AMI), i.e. heart attack, between 1992 and the second quarter of 2014. When such a patient arrives at the ER, the clinicians on duty must quickly decide on a course of action, so that there is less than usual scope for selecting a provider.¹ We focus on within hospital-year variation in provider behavior and the sample is limited to hospitals that have experience performing invasive cardiac procedures to ensure that there is in fact a viable choice to be made.

Using a rich set of observable patient characteristics including the age, sex, detailed diagnosis code, previous history of AMI, and comorbidities of the patient, we show that there is substantial variation across providers in the extent to which observable patient characteristics affect choice of procedure. Some providers are much less likely to use invasive procedures on the oldest and sickest patients, while others appear to pay little attention to these factors. Perhaps surprisingly in light of concerns about over-treatment and excessive spending, we find that providers who treat everyone aggressively have the best outcomes. In contrast, teams that follow the prevailing norms in teaching hospitals where providers are less likely to perform

¹ When such a patient arrives at the ER, clinicians must quickly decide on a course of action. For example, if a patient was to be treated with drugs (e.g. thrombolysis) the goal would be a door-to-drug time within minutes, while if they were treated with angioplasty, the goal would be a “door-to-balloon” time within 90 minutes (Zafari and Yang, 2014).

invasive procedures on the “least appropriate” patients, reduce costs but achieve poorer health outcomes for these patients.

We also ask how the aggressiveness and responsiveness of the providers’ decision-making varies with their characteristics. We find that the most aggressive and responsive providers are young, male, and graduated from top-20 medical schools. In addition, Spanish speaking providers tend to be more aggressive but less responsive to patient condition.

Overall, our results suggest that it is possible to identify important elements of provider practice style using routinely collected administrative data and to relate practice style to patient outcomes. In turn, it may be possible to use this knowledge to refine treatment protocols in order to cost-effectively improve patient outcomes.

The rest of the paper is laid out as follows. Section 2 presents background information. Section 3 covers data and methods. Results are presented in Section 4, and section 5 concludes.

2. Background

Tsai et al. (2010) ask whether the treatment of AMI is consistent with guidelines from the American Council of Cardiology, and concluded that compliance is “low to moderate” suggesting that there is a great deal of room for improvement. The literature offers many possible reasons for “mistakes” in the choice of procedure to be systematically biased in favor of aggressive and expensive treatment (Chandra et al., 2012). One common explanation for faulty decision making is “defensive medicine,” the idea that providers perform unnecessary procedures in order to protect themselves from lawsuits. However, the extent to which defensive medicine drives practice style remains controversial. For example, Baicker et al. (2007) argue

that there is little connection between malpractice liability costs and provider treatment of Medicare patients.

There is also a substantial literature arguing that providers are swayed by financial incentives (e.g. Gruber, Kim, Mayzlin, 1999; Gruber and Owings, 1996), and especially by fee-for-service reimbursement models that can incentivize providers to increase the number of services rendered (McClellan, 2011).

A third possibility is that providers are influenced by the decisions of those around them. Chandra and Staiger (2007) study the choice of surgery vs. medical management of cardiac patients. Knowledge spillovers are the main theoretical driver of regional variation in procedure use in their model. Providers in areas that specialize in surgery are assumed to become better at surgery and worse at medical management, and vice-versa. Their model raises the possibility of mismatch between patients and providers. All patients in high surgery areas will be more likely to have surgery, even if medical management would be more appropriate for some of them. An implication of their model is that patients who are good candidates for medical management will fare more poorly with high surgery providers than they would have with less aggressive providers.

What matters in our application is not only whether providers have high or low average levels of invasive procedures, but also the extent to which they tailor their decisions to the characteristics of the individual patient, a factor that has not been considered earlier. Note, that since we are looking at variation in procedure use within hospital-years, our work is complementary to the Chandra and Staiger research that looks at variation in mean procedure rates across regions.

Doyle et al. (2010) suggest that variations in practice style may arise because some providers may just be more competent than others. They study a setting in which patients arriving at a large medical center are randomly assigned to one of two medical groups. One of the groups is affiliated with a prestigious medical school while the other is not. They find that providers from the better medical school systematically conduct fewer tests and have lower costs even though both groups have similar patient outcomes. We will build on this work by examining the relationship between a provider's practice style and whether they attended a top 20 medical school (using the U.S. News and World Reports rankings).

Patient preferences are often cited as a fifth potential reason for medically unnecessary procedure use. In an innovative study using vignettes from patient and provider surveys, Cutler et al. (2013) assess the hypothesis that regional variations in procedure use are driven by differences in patient demand across areas. They conclude that patient demand is a relatively unimportant determinant of regional variations and that instead the main driver is provider beliefs about appropriate treatment that are often unsupported by clinical evidence.

Finkelstein et al. (2014) address the same question using longitudinal Medicare claims data that allow them to track the same patients as they move through different health care markets. They suggest that about half of the observed variation in procedure use is due to supply-side factors, while half is due to patient-level, or demand-side. However, they conclude that much of the variation in patient demand is driven by exogenous patient health, and so does not simply reflect patient tastes for procedures. These findings agree with those of Cutler et al. (2013) in suggesting that patient preferences play a relatively small role in explaining variations in care.

Because we are studying heart attack patients who were admitted through the ER, it is unlikely that patient demand is the main driver of either provider or procedure choice. We do not find any evidence of sorting on provider aggressiveness. We do find, however, evidence of a small amount of sorting within hospitals, which indicates that patients who are not good candidates for invasive procedures are more likely to be treated by more responsive cardiologists. In practical terms, this means that patients who are older and sicker tend to be matched with providers who are less likely to perform invasive surgery on them. As we will show, these patients subsequently end up with worse health outcomes. However, this is unlikely to be a quality of provider effect because the same teams are more likely to perform invasive procedures on highly appropriate patients, and achieve better than average outcomes for these patients. In order to deal with the possibility of patient selection, we control for patient appropriateness in all our models.

3. Data and Methods

Our analysis starts with hospital discharge data for all heart attack patients in Florida from 1992-2014.² Because we wish to focus on cases where there is actually a choice of procedures, we further restrict the sample to hospital-quarter-day of week combinations with non-zero invasive procedures and drop any hospital that did not have more than 35 invasive procedures per year. We also drop patients with diagnosis codes for which over 90% of patients received an invasive procedure or less than 10% of patients received an invasive procedure.

It is unfortunately impossible to know exactly who the decision maker is from the hospital record, so we assume that cardiologist had the most say. We focus on hospitals where

² The hospitalization data come from the Florida Agency for Health Care Administration (FL AHCA).

more than 50% of AMI patients had cardiologists listed on their hospital records, suggesting that these hospitals had cardiologists available for consultation. We also drop hospital-quarter-weekdays when there was apparently no cardiologist on call (because no patient appears with a cardiologist on the record). Within this sample, the share of patients with a cardiologist listed on their hospital records is 69%.³ If there is a cardiologist listed on the record, then we assume that she made the decision about procedure choice.

If no cardiologist is listed on the record, then we use the cardiologist who treated the greatest number of patients in that hospital-quarter-weekday cell. The cardiologist with the highest patient volume had the most tenure 83% of the time (where tenure is measured as the cumulative number of year-quarter-weekdays the physician had worked at that hospital up to the point that the patient arrived). Choosing the cardiologist with the highest patient volume also yields the same result as choosing the one with the most years of experience since residency 73% of the time. In what follows we will explore the relationship between our measures of practice style and physician characteristics in the sample with information about the cardiologist, and in the sample where we have imputed that information in order to judge the extent to which imputation error is likely to affect the estimated relationship between provider characteristics and practice style.

We focus on patients who were admitted through the ER (rather than AMI patients who arrive as referrals from other physicians or as transfers from other hospitals) since there is likely to be less capacity for physician selection in these cases. Finally, we restrict the sample to cardiologists who treated at least five AMI patients per three-year period, since it will be difficult to determine anything about the practice styles of physicians who see very few patients.

³ Overall, 60% of patients in our sample had a cardiologist listed on the record.

Appendix Table 2 shows the distribution of patients per cardiologist before and after this restriction is imposed.

Table 1 shows that these restrictions reduce the sample from 1,039,072 to 658,553 patients who were treated by 2,929 cardiologists. These restrictions also reduce the number of hospitals from 244 to 149, meaning that we are focusing only on hospitals that had the capability to perform invasive procedures. Of the patients in our sample, 59% have an invasive procedure, which is defined as receiving either cardiac catheterization or angioplasty.⁴

These data have information about all patient hospitalizations over the time period, including patient characteristics, admission sources, procedures, length-of-stays, charges, discharge outcomes, and physician license numbers which can be matched to Florida's physician license database to obtain additional information about the cardiologists in our sample. The available patient characteristics include gender, age, race, ethnicity, insurance, and up to ten ICD-9 diagnostic codes for each patient.

From the hospital discharge data we create several additional variables. First, we use the diagnosis codes to define patient comorbidities. Following the literature (e.g. Card et al., 2009) we pay special attention to the serious comorbidities included in the Charlson index (Charlson et al. 1987), which include cardiac arrhythmia, hypertension, congestive heart failure, peripheral vascular disease, dementia, cerebral vascular disease, coronary obstructive pulmonary disease, lupus, ulcers, liver disease, cancer, diabetes, kidney disease, and HIV. Patients with these conditions are likely to be poorer candidates for invasive procedures than healthier patients.

We measure health outcomes using information about the patient's disposition at the time of discharge. Many heart attack patients die in the hospital, which is obviously an extremely

⁴ Invasive procedures are defined by the following ICD-9 codes: 00.66, 36.0, 37.22, and 37.23.

negative outcome. Patients may also be discharged to home, to another hospital, or to another facility such as a skilled nursing facility or hospice. By focusing on being discharged home, we are implicitly assuming that this reflects a better health outcome than, for example, being transferred to a skilled nursing facility.

Additionally, we know whether the patient developed a hospital-acquired condition (HAC). HACs are defined by the Department of Health and Human Services and include infections like septicemia, clostridium difficile, pneumonia due to staphylococcus, catheter-associated urinary tract infection, vascular catheter infection, and surgical site infections following surgery.⁵ While current guidelines view HACs as avoidable and seek to eliminate them, it is undoubtedly the case that invasive procedures and longer hospital stays expose patients to greater risks of infection under current conditions. Thus, HACs can be viewed as one of the more serious potential costs of unnecessary surgical procedures.

One drawback to our visit-level hospital discharge data is that the data do not contain patient identifiers. Therefore, we can only measure health outcomes that occur on the index visit. We cannot measure health outcomes that occur outside of the hospital, in another hospital, or at the time of a follow-up visit. Ideally, a state agency or hospital administration that did have access to longitudinal data would probably wish to include these outcomes in the analysis. We do however know whether the patient had a previous AMI episode as this is recorded in the discharge data.

⁵ Other types of HACs include foreign body retained after surgery, air embolism, pressure ulcer stages III and IV, falls and trauma, manifestations of poor glycemic control, deep vein thrombosis and pulmonary embolism following orthopedic procedures, and lactogenic pneumothorax with venous catheterization. The HHS definitions come from a Fact Sheet report titled, "Hospital Acquired Infections (HAC) in Acute Inpatient Prospective Payment (IPPS) Hospitals," which was published in October 2012. To the HHS list we add septicemia, clostridium difficile, and pneumonia due to staphylococcus, three common infections that can be deadly and that plague patients in hospital environments.

Our second set of outcomes includes hospital costs and length-of-stay. The hospital discharge data contain hospital charges, which must be converted into hospital costs since hospitals generally only recover a fraction of the list price of a service. To convert charges to costs, we multiply the hospital charge by the hospital's cost-to-charge ratio (CCR) in the given year.⁶ Costs are further standardized to real 2009 dollars. We also separate hospital costs into categories such as pharmacy costs, laboratory costs, radiology costs, costs for medical devices, cardiology, operation rooms, and all other costs.

One drawback to using hospital discharge data is that the hospital costs are not the total costs for the patient's full episode of care. The hospital costs do not include physician fees or the costs of treating the patient at another facility. Since 7% of our sample gets transferred to another hospital, and all of our patients receive treatment from at least one physician on the initial visit, the differences between the hospital costs and the total costs for the episode of care are likely to be non-negligible. We use the length of stay and the hospital costs to index the amount of medical resources that were used to treat the patients.

To learn more about the characteristics of the cardiologists in our data, we match their medical license numbers to the Florida medical license database.⁷ We construct variables for cardiologist characteristics that include their experience, gender, whether they attended medical school in the United States, whether their medical school is ranked among the top-20 according to U.S. New and World Report research rankings⁸, and whether they speak Spanish. One potential contribution of this analysis is to show which, if any, observable characteristics are

⁶ We use the group cost-to-charge ratio because some individual hospital CCRs are missing. The group CCR assigns the same ratio to similar hospitals in the same geographic area.

⁷ Link to the Florida license database: <http://ww2.doh.state.fl.us/IRM00PRAES/PRASLIST.ASP>

⁸ In principal one might also wish to identify doctors who did prestigious residencies, but this information is not systematically collected in the licensing database.

correlated with practice style. It is of course possible that the decision is actually made by a team of physicians including the ER doctor (who is not observed in our sample of admitted patients) and to the extent that this is the case, the influence of any individual characteristics are likely to be diluted.

3.1 Identifying Good Candidates for Invasive Procedures

The first step in our analysis is to identify patients who appear, given their observable characteristics, to be good candidates for invasive procedures. We do this using a standard simple “machine learning” algorithm, which involves estimating a logit model for the use of the invasive cardiac procedure on all of the observable patient characteristics.⁹ Specifically, for each year of data we estimate the following model for visit i in quarter t :

$$(1) Pr(Invasive_{it} = 1) = F(\theta_{it,diag} + \Gamma X_{it} + \delta_{it,com} + \lambda_t + \varepsilon_{it})$$

Where $F(\cdot)$ is the logit function, $Pr(Invasive_{it} = 1)$ is the probability that the patient on visit i in quarter t receives an invasive procedure, $\theta_{it,diag}$ is a vector of 30 diagnosis codes for different types of acute myocardial infarction, X_{it} is the patient’s gender and a vector of age dummies (50-54, ... 65-69, 70, 71...89, 90+), $\delta_{it,com}$ is a vector of 14 comorbidities plus whether there was a previous AMI,¹⁰ and λ_t is a vector of quarter fixed effects.

We estimate (1) using only patients in hospitals that have accredited teaching programs, which implicitly assumes that these hospitals define the standard of care in each year and try to follow the currently accepted best practices. We define teaching programs based on the

⁹ See An Introduction to Statistical Learning, by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani for a discussion of machine learning, with applications to medical decision making. The logistic model is one of the simplest (and robust) methods for modeling a binary decision. We will show that it can be used to construct an elegant two dimensional representation of provider choices.

¹⁰ The AMI diagnosis codes show whether the patient had a previous AMI.

Accreditation Council for Graduate Medical Education (ACGME)'s list (see: <https://apps.acgme.org/ads/Public/Programs/Search>) and use the 20 Florida hospitals with accredited internal medicine teaching programs. This includes institutions with national prestige such as the Mayo Clinic and the Cleveland Clinic, which have hospitals in Florida.

Note that if we construct an alternative patient appropriateness index using data from all hospitals, the correlation between this alternative index and the one that we use here is above 0.9. What this implies is that if called on to rank patients from least appropriate to most appropriate for surgery, most practitioners would rank them in the same way (even though they might well choose different cutoffs for deciding who would receive surgery).

We estimate the model year by year because if, for example, surgery becomes generally safer over time, then it may make sense to use aggressive procedures on more marginal patients over time. Thus, the standard of care is allowed to evolve over time; in fact we show that there is a general trend towards using more invasive procedures in older, sicker patients.

These estimates are shown for different years in Appendix Table 1. The estimates suggest that the weights that providers place on several important determinants of appropriateness for invasive procedures have evolved over time. For example, at the beginning of the sample invasive procedures were less likely to be performed on patients with dementia, but this is no longer the case. Providers have also become more likely to perform invasive procedures on patients with hypertension and on patients who are in their 80s, and less likely to perform these procedures on women. These trends show the importance of considering a provider's decision in the context of accepted practice for the time.

We use the estimated parameters from equation (1) to construct an index that ranks patients by their likelihood of receiving an invasive procedure. This index is a continuous

variable, which, in principle, can vary from negative infinity to positive infinity. Given this index, we can divide patients into terciles according to their appropriateness for invasive procedures. Tables 2, 3, and 4 show means for all patients, those in the lowest tercile of appropriateness for invasive procedures (low), and those in the highest tercile of appropriateness for invasive procedures (high).

Table 2 shows the mean characteristics of the patients themselves. We can see that patients who are good candidates for invasive procedures are less likely to be female, they are younger (and therefore less likely to have Medicare coverage and more likely to have private health insurance coverage), and they are less likely to have several other serious comorbid conditions including congestive heart failure or kidney disease.

Table 3 shows what happened to these patients. Overall, 60% have an invasive procedure, but as one might expect, this fraction rises to 86% among the high appropriateness patients and is only 28% among the least appropriate patients. The high appropriateness patients have lengths of stay that are a third shorter than the low appropriateness patients; however, their costs are 21% higher. In general outcomes appear somewhat better for this group in that they are much less likely to have hospital acquired conditions, more than almost twice as likely to be discharged home, and are only a quarter as likely to die in the hospital. This may however reflect their underlying relative youth and health status rather than their treatment.

Table 4 shows the characteristics of the cardiologists assigned to these patients. The unit of observation is still the patient. The table shows that there is little difference between cardiologists serving “high” and “low” appropriateness patients in terms of the types of provider characteristics that we can observe. They are similar in terms of experience, type of medical school attended, gender, and whether they speak Spanish.

3.2 Measuring Practice Style

The next step in our analysis involves the estimation of provider-specific regressions which show how the cardiologist's decision to use invasive cardiac procedures varies with the index we constructed above, which summarizes all of the information that we observe about the patient's characteristics and medical condition. The intercept in these regressions tell us whether the medical team in question is more or less likely to use invasive procedures relative to the state-wide standard defined in equation (1). This is the element of practice style that has been considered in Chandra and Staiger (2007), Epstein, Ketchum, and Nicholson (2010), and Epstein and Nicholson (2009). The innovation here is to also consider, given that mean level of procedure use, how likely the provider is to take the patient's characteristics into account when making a decision, and to target invasive procedures to those patients who appear to be the best candidates for them.

Given our long time period, we do not assume that a provider's behavior is constant over time. Instead we allow it to evolve with provider experience, measured as the number of years since the cardiologist assigned to the patient completed his or her residency. We create 3-year experience level bins, where $k = 1$ if the cardiologist has less than 3 years of experience, $k=2$ indicates 3-6 years of experience, and so on, up to more than 30 years of experience ($k=11$). Specifically, we estimate a logit model for each provider j in experience level k ,

$$Pr(Invasive_{ijk} = 1) = F(\alpha_{jk} + \beta_{jk} * Patient_Index_i + \varepsilon_{ijk}) \quad \text{for each } j * k = 1, \dots, JK \quad (2)$$

where again, $F()$ indicates the logit function. Here α_{jk} captures cardiologist j 's propensity to perform an invasive treatment on the median patient given experience level k , and β_{jk} captures the relative weight that cardiologist j places on the index summarizing the patient i 's

appropriateness for invasive treatment. Because we estimate separate equations for the same cardiologist at different levels of experience, α_{jk} and β_{jk} vary within cardiologists over time. This variation allows us to test whether cardiologists are more or less likely to consider observable patient characteristics as they gain experience. In what follows we will refer to these two dimensions of provider behavior as aggressiveness and responsiveness.

Theoretically the parameters α_{jk} and β_{jk} can take values from negative infinity to positive infinity, but there are some special cases that illustrate the intuition behind our model. When $\beta_{jk}=0$, providers ignore patient characteristics and have the same probability of performing an invasive treatment on all patients. One way to characterize this behavior is that the provider has a particular preference for invasive procedures that is independent of patient characteristics. When $\alpha_{jk}=0$, the provider's behavior depends only on the patient's appropriateness for invasive procedures. If in addition $\beta_{jk}=1$, then providers behave in exactly the way predicted by our equation (1) model. The coefficient α_{jk} can also be characterized as the probability that a provider will perform an invasive procedure on a patient with an index of patient appropriateness equal to zero, that is, on a patient of median appropriateness.

We are not the first to try to measure provider practice style. However, practice style is usually modeled as a provider-specific fixed effect. Instead, we allow there to be two dimensions to practice style. Moreover, we allow practice style to evolve over time. We will show that this richer model is useful in conceptualizing practice style.

Table 5a shows the distribution of the estimated parameters α and β in our sample. This table does not take the precision of the coefficient estimates into account. Therefore, Table 5b divides the data into cells according to whether the provider's behavior shows a statistically significant departure from the norm. That is, we ask whether α is significantly different from

zero, and whether β is significantly different than one. Since providers with few patients will have imprecisely estimated coefficients, this procedure takes into account the fact that it is difficult to judge practice style in a small practice. However, only 2.9% of patients see cardiologists in the bottom quartile of the distribution of 3-year practice size indicating (not surprisingly) that cardiologists with few patients account for few patients in our sample.

The largest cell in Table 5b is the middle cell where $\alpha=0$ and $\beta=1$. These are providers whose behavior is consistent with the state-wide norm established by the accredited teaching hospitals for that year. However, Table 5b shows that 13.8% of patients have a provider who is less responsive to patient characteristics than the norm (i.e. $\beta<1$) while very few (2.0%) have a provider that is more responsive to patient characteristics than the norm (i.e. have $\beta>1$). Turning to aggressiveness, 17.6% of patients have a provider who is less aggressive than the norm established in the teaching hospitals (i.e. $\alpha<0$) while 21.9% of patients have a provider who is significantly more aggressive than that norm (i.e. $\alpha>0$). Note that these determinations take the provider-specific sample size into account as providers with few patients have coefficients that are estimated with larger standard errors.

We can now return to the question of whether patients whose medical conditions make them good candidates for invasive procedures go to providers with systematically different practice styles than patients who are poor candidates for such surgery. Table 5c shows the results of taking the estimated α 's and β 's and regressing them on hospital*year effects as well as the patient's gender, age, comorbidities, and an indicator equal to one if the patient had a

previous AMI.¹¹ We do this separately for the low and high terciles of patient appropriateness for procedures.

We do not find any evidence of matching in the high appropriateness patients. Nor is there any evidence of matching on provider aggressiveness in either group of patients. The only significant matching effect we find is among low appropriateness patients, where a one standard deviation increase in appropriateness is associated with a 5.4% decline in the probability of having a low-beta provider. That is, within this group of generally less appropriate patients, patients who are better candidates for invasive procedures are less likely to have unresponsive providers. Hence, to the extent that there is a bias induced by matching, these estimates suggest that within the low-appropriateness group, those patients who are better candidates for invasive procedures will have more responsive providers, which suggests that they may have better outcomes.

Given this evidence of a small amount of matching in one group of patients, we include the continuous index of appropriateness in all of the models of outcomes and costs discussed below. Thus, we will be examining AMI patients admitted from the ER in the same hospital and year with the same measured appropriateness for invasive procedures but who, nevertheless, are treated by providers with different levels of responsiveness and aggressiveness.

3.3. Models of Costs, Patient Outcomes, and Correlates of Provider Practice Style

Given the data we have constructed, we can now ask how the variations in provider treatment style affect costs and patient outcomes, and also what characteristics of providers are

¹¹ As discussed above, we cannot follow patients over time, but the diagnosis for each AMI patient indicates whether there has been a previous AMI.

associated with differences in practice style. We have defined a standard of care using data from teaching hospitals. This standard of care has two dimensions which can be thought of as the average level of use of invasive procedures (α) and the extent to which the provider responds to the observable condition of the patient (given largely by their exact diagnoses, age, and comorbidities) when deciding whether to perform an invasive procedure (β). If the standard is useful then we will find that patients of providers who deviate from the standard will have poorer outcomes.

In order to measure the effect of provider characteristics on outcomes, we estimate models of the form:

$$Y_{ijkht} = \phi_1 * LowResponsiveness_{ijkt} + \phi_2 * LowAggressiveness_{ijkt} + \phi_3 * HighAggressiveness_{ijkt} + \Pi Z_i + \Omega X_i + \rho Appropriateness_i + \delta_{ht} + \varepsilon_{jkt}, \quad (3)$$

where *LowResponsiveness* corresponds to an estimated provider β that is significantly less than one; *LowAggressiveness* corresponds to an estimated provider α that is significantly less than zero, and *HighAggressiveness* corresponds to an estimated provider α that is significantly greater than zero. For the time being, we ignore the possibility that β is significantly greater than one since it is so rare in our data. However, in Appendix Table 4 (discussed further below) we present models that also include an indicator for this possibility. These measures are specific to patient i of provider j with experience level k in hospital h in year-quarter t .

The outcomes vector Y_{ijkht} includes health outcomes, total hospital costs, itemized costs, length-of-stay, and the probability of receiving an invasive procedure. Health outcomes include the probability of developing a hospital-acquired condition (HAC), the probability of being discharged home (Home), and the probability of dying in the hospital (Died). The subscript h

indicates that outcomes may also vary at the hospital level. The cost variables are in logs, the length of stay variable is in levels, and the invasive procedure variable is binary.

The vector Z_i includes other observable characteristics of providers, including experience, gender, whether they attended medical school in the US or abroad, and whether they attended top-20 medical schools. The vector X_i includes gender, age, whether there was a previous heart attack, and the patient's comorbidities. As discussed above, we also include a measure of each patient's appropriateness for invasive procedures, $Appropriateness_i$. We estimate alternative models adding in patient race, ethnicity, and type of health insurance, which may have independent effects on treatment choices and on outcomes. We control for hospital*year fixed effects (δ_{ht}) in order to be sure that we are capturing provider-level differences as well as to capture things like technological improvements that may improve survival over time.

Finally, we examine the way that the estimated parameters of provider practice style vary with provider characteristics, and with their experience. In order to conduct this examination we estimate:

$$\hat{\alpha}_{jkt} = \phi_1 * \hat{\alpha}_{jkt-1} + \Pi Z_{it} + \delta_{ht} + \varepsilon_{jkt}, \quad (4a)$$

$$\hat{\beta}_{jkt} = \phi_2 * \hat{\beta}_{jkt-1} + \Pi Z_{it} + \delta_{ht} + \varepsilon_{jkt}. \quad (4b)$$

where the estimated parameters at time t depend on provider characteristics including experience, lagged values of the dependent variables to capture persistence, hospital*year fixed effects and the patient level variables included in the other regression models.

4. Estimation Results

Table 6 shows the estimated coefficients on the practice style measures included in equation (3). As discussed above, these models control for measured patient appropriateness in

addition to diagnoses and other patient characteristics. For the sake of brevity these coefficients are not shown, but they generally have the expected sign; that is, older, sicker patients tend to have worse health outcomes other things being equal. With regard to the aggressiveness of the providers, Table 6 suggests that *all* patients do better with more aggressive physicians and worse with less aggressive providers. As discussed above, these models control for measured patient appropriateness for invasive procedures, which is of course an important determinant of outcomes. Patients with aggressive physicians are more likely to be discharged home, less likely to get hospital acquired infections, and less likely to die in the hospital and the advantages are actually greater among the “low appropriateness” patients who on average, have only a 28% probability of receiving invasive procedures. Among these patients, having an aggressive provider is associated with a 2 pp reduction in the probability of dying in hospital, on a baseline of 17%. This result suggests that aggressive procedures may actually be under rather than over-utilized, and that in particular, that some of the older, sicker patients who tend to be deemed inappropriate for aggressive procedures would benefit from them.

Turning to the measure of responsiveness, Table 6 suggests that among the patients who are highly appropriate for procedures and conditional on the aggressiveness of the providers, providers who are not responsive to observable patient characteristics (here patient diagnosis, previous AMI, age, and comorbidities) tend to have worse patient outcomes. Among these high appropriateness patients, low provider responsiveness predicts more hospital-acquired conditions, a higher probability of dying in the hospital, and a lower probability of being discharged home.

Among low appropriateness patients, however, low provider responsiveness predicts better outcomes. Recall that in our framework, being more responsive means that a provider

with a given rate of doing invasive procedures is more likely to allocate procedures to the high appropriateness patients. So we can think of an increase in responsiveness in these models as a reallocation of a given number of procedures away from low appropriateness towards high appropriateness patients. If in fact, all patients would benefit from procedures, then this means that responsiveness will benefit high appropriateness patients and harm low appropriateness patients, just as we see here.

Although we did detect some evidence of patient selection in the low appropriateness patients, it seems unlikely that the findings in Table 6 could be driven by selection. As we showed in Table 5c, among low appropriateness patients, those who are better candidates for invasive procedures (i.e. younger and less sick) tend to have more responsive physicians, which would tend to bias estimates towards finding favorable effects of responsiveness.

The cost data in Table 7 are consistent with this interpretation. Not surprisingly, for all patients, high aggressiveness is associated with higher costs, a higher probability of invasive procedures, and a longer length of stay, and vice versa. Among high appropriateness patients, having a low responsiveness provider is associated with lower probabilities of receiving invasive procedures and correspondingly lower costs. Among low appropriateness patients, we see the reverse pattern with low responsiveness providers being more likely to do invasive procedures, which in turn generate higher costs and longer lengths of stay.

Table 8 shows that the same patterns that apply to total costs are reflected in virtually every category of costs. More aggressive providers have higher costs, especially for operating room and cardiology, while less aggressive providers have lower costs. Since low responsiveness is associated with more invasive procedures in low appropriateness patients, and with fewer invasive procedures in high appropriateness patients, costs show patterns consistent

with this behavior. As discussed above, we only observe hospital costs. In most cases, there will also be additional costs associated with both the treating providers and anesthesiologists. Since these costs are likely to rise with invasive procedures, the comparisons above are likely to understate the extent to which additional procedures drive up costs.

Since our cost data is incomplete, we hesitate to use our estimates to calculate a cost per life saved from doing invasive procedures (versus not doing these invasive procedures). We can however examine the percent change in probability of death that accompanies a percent change in cost. Under the assumption that unobserved costs are proportional to hospital costs, this may be a valid comparison. What we find is that for the high appropriateness patients, a 7% decrease in costs is associated with a 23% increase in the probability of death, for an elasticity of -3.3. For the low appropriateness patients, a 5% increase in costs is associated with a 6.5% reduction in the probability of death, for an elasticity of -1.3. Thus, while it is more cost effective to allocate invasive procedures to highly appropriate patients than to less appropriate patients, there may still be considerable potential gains to changing practice style in a way that increases the use of invasive procedures among the less appropriate patients.

4.1 Correlates of Practice Style

Table 9 shows how the two dimensions of practice style that we have identified (responsiveness and aggressiveness) vary with other provider characteristics. Because we allow practice style to evolve over time, we can also ask how it varies with experience (measured as years since residency). The first and fourth columns show models without lagged practice style measures while the remaining columns do control for lagged practice style. The estimates on the lags suggest that practice style (and especially aggressiveness) is quite persistent over time,

although it does evolve, a result that is consistent with Epstein and Nicholson's (2009) work on obstetricians. This is an interesting observation given past work asking whether providers are "punished" for aggressive practice styles (Dranove et al., 2011).

Conditional on lagged practice style, more experienced providers are actually less responsive and less aggressive than newer graduates: We find that providers with 6-9 years of experience are more responsive than the omitted category (those with 3-6 years). After that slight peak, responsiveness is relatively flat until about 20 years of experience, after which it declines over time. Aggressiveness declines more continuously with experience. We also find that U.S. trained providers tend to be less responsive and more aggressive except for those trained at top 20 medical schools. Those trained at top schools are both more responsive and more aggressive than other U.S. trained physicians. Female providers have higher responsiveness, and lower aggressiveness, while the reverse is true for Spanish-speaking providers.

As discussed above, there are two key limitations of our data that could be expected to affect the estimates in Table 9. First, we do not actually observe the decision maker, but assume that it is the cardiologist. Suppose, for example, that it was the ER doctor who generally decided on the procedure and then called in the cardiologist. In this case, one would not expect the measured characteristics of the cardiologist to have much effect on procedure choice, unless these characteristics were correlated with those of the actual decision maker. Hence, the fact that we see very systematic relationships between the cardiologist's characteristics and practice style indicates that at the very least, the identity of the cardiologist is a reasonable proxy for the provider or team of providers that is actually making the decision.

A second limitation is that we do not actually know the cardiologist for a significant fraction of our sample but we impute it using the busiest (and generally most senior) cardiologist on duty. Table 10 shows the results of estimating models similar to those of columns (3) and (6) of Table 9 for the subsample in which we know the cardiologist, and the subsample where the cardiologist is imputed. If each cardiologist was a completely autonomous decision maker, then we might expect to see a relationship between cardiologist characteristics in the cardiologist-known sample, but not in the cardiologist-imputed sample. On the other hand, if it is in fact the most active cardiologist on duty who sets the tone for the unit, then one might actually see a stronger relationship between provider characteristics and practice style in the cardiologist-imputed sample.

Table 10 shows that in fact the qualitative relationships between provider characteristics and practice style are quite similar in the two sub-samples: Aggressiveness and responsiveness are both lower for more experienced physicians, especially for those with more than 20 years of experience. We find a weaker effect of attending a U.S. medical school in the cardiologist-imputed sample, but the impact of attending a top 20 medical school is greater. The impact of gender is also greater in the cardiology-imputed sample and the impact of Spanish language is similar. Overall, the R-squared is higher in the imputed-cardiologist sample, though since these are non-overlapping samples we hesitate to make too much of this result and conclude that the relationship between provider characteristics and practice style is fairly similar in the two sub-samples.

4.2 Robustness

Appendix Table 3 shows estimates of a model that controls for combinations of aggressiveness and responsiveness. Table 5b suggested that very few patients saw providers with an estimated β significantly greater than one. Moreover, relatively few patients had providers with β significantly less than one and α significantly greater than one. Therefore we collapse this latter category together with β less than one and α equal to zero and consider only four alternatives to the baseline case in which $\beta=1$ and $\alpha=0$. These alternatives are: $\alpha<0$ and $\beta\geq 1$, $\alpha>0$ and $\beta\geq 1$, $\alpha<0$ and $\beta<1$, and $\alpha\geq 0$ and $\beta<1$.

Appendix Table 3 shows that given typical responsiveness, more aggressive providers have better patient outcomes in both groups of patients. The next two rows show that low responsiveness in the provider is associated with worse patient outcomes in patients who are highly appropriate for invasive procedures. In patients who have low appropriateness for invasive procedures, responsiveness has no effect if the physician is less aggressive than average – in these physicians who perform small numbers of procedures, there may not be much scope for reallocating procedures away from the low appropriateness patients. However, if the physician is typically aggressive, then low responsiveness is harmful to the low appropriateness patients, as in Table 6.

Appendix Table 4 shows models similar to those in Table 6 except that they also include the patients' race, ethnicity, and type of insurance. These are variables that might be predictive of patient outcomes, and might also be correlated with the type of treatment they receive. However, including these variables has virtually no effect on the estimated coefficients of practice style.

5. Conclusions

In this paper we contribute to the growing literature on provider practice style by developing a framework that considers not only the provider's overall propensity to perform an invasive procedure, but also the way that the provider chooses to allocate procedures across patients. To do so, we use the behavior of providers in accredited teaching hospitals to predict which heart attack patients arriving in any ER are more or less appropriate candidates for invasive procedures. This prediction is made using the patient's diagnosis, demographic characteristics, previous AMI history, and co-morbidities. Given these predictions we can identify providers who are consistently more or less aggressive than the average, or whose behavior in terms of matching patients to procedures (which we call responsiveness) deviates significantly from the norm defined in the teaching hospitals.

Our results suggest that all patient types benefit from aggressive treatment of heart attacks using invasive procedures. Thus, being assigned a provider who is careful to allocate a given number of invasive procedures to the most appropriate patients is good for these highly appropriate patients, but bad for the lower appropriateness patients who are consequently less likely to obtain these procedures. These results may be surprising given the extensive public discussion of the possible overuse of aggressive, invasive procedures in patients who may not need them. In our data, attempts to target procedures to the patients who are most likely to benefit look less like reducing unnecessary procedure use, and more like rationing of a procedure that is highly beneficial to most patients.

It is worth reiterating that our models focus on differences in treatment patterns between cardiologists working in the same hospital and year. We also restrict our attention to hospitals that frequently perform invasive cardiac procedures. Hence, the differences in outcomes that we

find are not likely to be a function of the many possible differences between hospitals or even within hospitals across years.¹²

These results extend the literature on provider practice style, showing that systematic differences can be identified using simple machine learning protocols applied to routine hospital discharge data. In addition, we show that these dimensions of practice style can be predicted using observable provider characteristics. Both aggressiveness and responsiveness are persistent over time, although providers tend to become both less responsive and less aggressive with experience. We also show that U.S. trained cardiologists tend to be both less responsive and more aggressive than those trained abroad. Providers from top 20 medical schools are an exception in that they are both more responsive and much more aggressive than other physicians. Lastly female cardiologists are significantly more responsive and less aggressive than their male counterparts, while the reverse is true among Spanish-speaking physicians in Florida.

Our results suggest that simple machine learning algorithms can indeed be used to characterize practice style and identify those styles associated with the best patient outcomes. In principal, this information could in turn be used to evaluate and improve protocols affecting patient treatment. In our example, reallocating procedures away from less appropriate towards more appropriate patients hurts the former and helps the latter. This implies that invasive cardiac procedures are underutilized and that more intensive use of these procedures would likely be associated with better health outcomes in both types of patients.

¹² Factors that could explain differences between hospitals or within hospitals over time include, for example, differences in the patient pool, capacity constraints, hospital-level protocols allocating procedures to patients, or differences in the quality of nursing care.

References

- Baicker, K., E. S. Fisher, and A. Chandra (2007, May-June). Malpractice liability costs and the practice of medicine in the medicare program. Health Affairs 26 (3), 841–852.
- Card, David, Carlos Dobkin and Nicole Maestas. Does Medicare Save Lives? Quarterly Journal of Economics, 124(2): 597-636, 2009.
- Chandra, A., D. Cutler, and Z. Song (2012). Chapter six - who ordered that? The economics of treatment choices in medical care. In T. G. M. Mark V. Pauly and P. P. Barros (Eds.), Handbook of Health Economics, volume 2 of Handbook of Health Economics, pp. 397 – 432. Elsevier.
- Chandra, A. and D. O. Staiger (2007). Productivity spillovers in health care: Evidence from the treatment of heart attacks. Journal of Political Economy 115 (1), pp. 103–140.
- Charlson, Mary E.; Pompei, Peter; Ales, Kathy L.; MacKenzie, C. Ronald. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. Journal of Chronic Diseases 40 (5): 373–83, 1987.
- Currie, J. and W. B. MacLeod (2008, May). First do no harm? Tort reform and birth outcomes. Quarterly Journal of Economics 123 (2), 795–830.
- Cutler, D., J. Skinner, A. D. Stern, and D. Wennberg (2013, August). Provider beliefs and patient preferences: A new look at regional variation in health care spending. Technical Report 19320, NBER.
- Doyle, Joseph J., Steven M. Ewer, Todd H. Wagner (2010). Returns to provider human capital: Evidence from patients randomized to provider teams, Journal of Health Economics, 29, 866-882.
- Dranove, D., S. Ramanarayanan, and A. Sfekas (2011). Does the market punish aggressive experts? Evidence from cesarean sections. B E Journal of Economic Analysis & Policy 11 (2).
- Epstein, A.J., Ketchum, and S. Nicholson (2010, Winter). Specialization and matching in professional service firms. Rand Journal of Economics, 41 (4), pp. 811-834.
- Epstein, A. J. and S. Nicholson (2009). The formation and evolution of provider treatment styles: An application to cesarean sections. Journal of Health Economics 28, 1126–1140.
- Finkelstein, A., M. Gentzkow, and H. Williams (2014, July). Sources of geographic variation in health care: Evidence from patient migration. Technical report, University of Chicago.
- Garber, A. M. and J. Skinner (2008). Is American Health Care Uniquely Inefficient? The Journal of Economic Perspectives 22 (4), pp. 27–50.

Gaynor, M., J. B. Rebitzer, and L. J. Taylor (2004, Aug). Provider incentives in health maintenance organizations. Journal of Political Economy 112 (4), 915–931.

Gruber, J., J. Kim, and D. Mayzlin (1999, AUG). Provider fees and procedure intensity: the case of cesarean delivery. Journal of Health Economics 18 (4), 473–490.

Gruber, J. and M. Owings (1996). Provider financial incentives and cesarean section delivery. The RAND Journal of Economics 27 (1), pp. 99–123.

James, Gareth, Witten, Daniela, Hastie, Trevor, Tibshirani, Robert. An Introduction to Statistical Learning with Applications in R (New York: Springer Science) 2013.

Kessler, D. and M. McClellan (1996). Do providers practice defensive medicine? Quarterly Journal of Economics 111 (2), 353–90.

McClellan, Mark(2011). Reforming Payments to Healthcare Providers: The Key to Slowing Healthcare Cost Growth While Improving Quality? The Journal of Economic Perspectives, 25 (2), 69-92.

Sutherland, Jason, Fisher, Elliott, Skinner, Jonathan (2009). Getting Past Denial – The High Cost of Health Care in the United States, New England Journal of Medicine, 361, Sept. 24, 1227-1230.

Tsai CL, Magid DJ, Sullivan AF, Gordon JA, Kaushal R, Michael Ho P, et al. Quality of care for acute myocardial infarction in 58 U.S. emergency departments. Academic Emergency Medicine. Sep 2010;17(9):940-50.

Zafari, A. Maziar, Yang, Eric. “Myocardial Infarction Treatment and Management,” Dec. 1, 2014. <http://emedicine.medscape.com/article/155919-treatment#aw2aab6b6b3>

Table 1: Derivation of Analysis Sample

Diagnosis Name & ICD-9-CM Code	# Patients	# Cardiologists	# Hospitals
All Patients Admitted from the ER			
AMI Codes beginning with "410"	1,039,072	-	244
Invasive procedure	532,202	-	233
Restricting to hospital-quarters with non-zero invasive procedures*			
All Patients	757,958	-	149
Invasive procedure	404,466	-	149
Restricting to patients at hospitals where >50% of patients receive cardiologists**			
All Patients	658,553	2,929	149
Invasive procedure	389,492	2,830	149

*(1) Drop hospital-quarters with 0 invasive procedures. (2) Drop hospital-quarter-days-of-week with 0 invasive procedures. (3) Drop hospital-years with < 36 invasive procedures. (4) Drop AMI diagnosis-years where > 90% or < 10% of patients receive an invasive procedure.

** (1) We restrict the sample to cardiologists whose license numbers match the license numbers in Florida's medical practitioner database. (2) We restrict the sample to cardiologists who treat at least 5 AMI patients per 3-year period. (3) We restrict the sample to patients who either have a cardiologist listed or who are admitted while a cardiologist is on call. (4) We restrict to patients who visit hospitals where more than 50% of AMI patients receive cardiologist assignments.

Table 2: Mean Patient Characteristics

Appropriateness for Surgery:	All	Low	High
Female	0.40	0.53	0.27
Age	69.91	80.69	59.65
White	0.79	0.83	0.76
Black	0.08	0.07	0.10
Hispanic	0.10	0.08	0.11
Medicaid	0.04	0.02	0.06
Medicare	0.66	0.88	0.38
Private Insurance	0.21	0.07	0.39
Self Pay or Other	0.09	0.03	0.17
Morbidity Index	0.45	-1.33	2.02
Subsequent AMI	0.05	0.12	0.003
#Diagnoses	8.20	8.98	7.16
Arrhythmia	0.26	0.32	0.20
Hypertension	0.43	0.33	0.56
Congestive Heart Failure	0.32	0.51	0.11
Peripheral Vascular Disease	0.05	0.05	0.04
Dementia	0.03	0.09	0.00
Cerebral Vascular Disease	0.07	0.14	0.01
COPD	0.16	0.20	0.09
Lupus	0.02	0.03	0.01
Ulcer	0.01	0.01	0.00
Liver Disease	0.02	0.03	0.00
Cancer	0.06	0.10	0.02
Diabetes	0.21	0.18	0.22
Kidney Disease	0.15	0.28	0.03
HIV	0.003	0.004	0.002
N	658,553	217,323	223,853

Table 3: Mean Procedure and Outcome Rates

Appropriateness for Surgery:	All	Low	High
Any invasive procedure	0.59	0.28	0.86
Length of Stay	6.81	7.68	5.18
Total Hospital Costs	19380	16601	20099
Pharmacy Costs	2932	2722	2674
Laboratory Costs	2514	2657	2043
Radiology Costs	1059	1252	704
Medical Devices Costs	2702	1819	3466
Cardiology Costs	3617	1754	5453
Operating Room Costs	1025	687	1135
Other Costs	5532	5710	4625
Hospital-Acquired Conditions	0.14	0.22	0.06
Discharged to Home	0.65	0.47	0.81
Died in the Hospital	0.10	0.17	0.04
N	658,553	217,323	223,853

Notes:

Costs converted to 2009 dollars using hospital CPI.

See <https://research.stlouisfed.org/fred2/series/DHSPRG3A086NBEA#>

Table 4: Provider Characteristics

Appropriateness for Surgery:	All	Low	High
Experience (yrs since residency)	13.48	13.62	13.37
US Medical School	0.59	0.59	0.59
Top-20 Medical School	0.12	0.12	0.12
Female Physician	0.06	0.05	0.06
Spanish-Speaking Physician	0.22	0.23	0.22
N	658,553	217,323	223,853

Table 5a: Distributions of Provider Responsiveness (Beta) and Aggressiveness (Alpha) Across Patients

Patient Percentile	Responsiveness (Beta)	Aggressiveness (Alpha)
1%	-0.216	-2.386
5%	0.284	-1.278
10%	0.482	-0.818
25%	0.729	-0.354
50%	0.926	0.054
75%	1.106	0.464
90%	1.319	0.946
95%	1.500	1.412
99%	2.073	2.621
N	658,553	658,553

Note: For each provider, we estimate a model for every three years of our sample period and obtain an estimate of alpha and beta.

Table 5b: Fraction of Estimated Provider Coefficients that are Significantly Different than Beta=1 and Alpha=0

	Beta<1	Beta=1	Beta>1	Total
Alpha<0	0.028	0.138	0.010	0.176
Alpha=0	0.069	0.527	0.0096	0.606
Alpha>0	0.041	0.177	0.0007	0.219
Total	0.138	0.842	0.020	

N= 658,553 patients.

Table 5c: Provider Practice Style Conditional on Hospital, Year and Patient Gender, Age, Comorbidities, and Previous AMI

	(1)	(2)	(3)	(4)
	Low	High	Low	High
	Beta < 1	Beta < 1	Beta =1	Beta =1
Z-Patient Appropriate	-0.0074* (0.0032)	0.0016 (0.0030)	0.0083* (0.0033)	-0.0018 (0.0033)
Hospital*Year FE	Y	Y	Y	Y
Patient Age Categorie	Y	Y	Y	Y
Previous AMI	Y	Y	Y	Y
Patient Comorbidities	Y	Y	Y	Y
Physician Characterist	Y	Y	Y	Y
N	217,323	223,853	217,323	223,853
R ²	0.01	0.00	0.01	0.01
	Alpha < 0	Alpha < 0	Alpha > 0	Alpha > 0
Z-Patient Appropriate	-0.0031 (0.0035)	-0.0051 (0.0034)	0.0025 (0.0035)	-0.0004 (0.0044)
Hospital*Year FE	Y	Y	Y	Y
Patient Age Categorie	Y	Y	Y	Y
Previous AMI	Y	Y	Y	Y
Patient Comorbidities	Y	Y	Y	Y
Physician Characterist	Y	Y	Y	Y
N	217,323	223,853	217,323	223,853
R ²	0.01	0.01	0.00	0.00

Notes: A one-standard deviation increase in the patient appropriateness index means that low-propensity patients are 0.74 pp (5.4% = 0.74/13.8) less likely to have a low-beta cardiologist. Standard errors clustered at the provider level and appear in parentheses. * indicates $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 6: Outcomes and Practice Style Among Patients with High and Low Appropriateness

	(1)	(2)	(3)	(4)	(5)	(6)
Appropriateness for Invasive Procedure:	High	High	High	Low	Low	Low
Outcome:	Hosp. Acquired Infection	Died in Hospital	Discharged to Home	Hosp. Acquired Infection	Died in Hospital	Discharged to Home
Low Responsiveness (Beta<1)	0.007*** (0.002)	0.009*** (0.001)	-0.025*** (0.003)	-0.010*** (0.003)	-0.011*** (0.003)	0.008* (0.004)
Low Aggressiveness (Alpha<0)	0.010*** (0.002)	0.009*** (0.001)	-0.019*** (0.003)	0.014*** (0.003)	0.013*** (0.002)	-0.024*** (0.003)
High Aggressiveness (Alpha>0)	-0.003* (0.001)	-0.005*** (0.001)	0.013*** (0.002)	-0.011*** (0.003)	-0.019*** (0.002)	0.021*** (0.003)
Hospital*Year FE	Y	Y	Y	Y	Y	Y
Patient Appropriateness Index	Y	Y	Y	Y	Y	Y
Patient Age Categories & Gender	Y	Y	Y	Y	Y	Y
Previous AMI	Y	Y	Y	Y	Y	Y
Patient Comorbidities	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
N	223853	223853	223853	217323	217323	217323
R ²	0.05	0.06	0.29	0.08	0.08	0.12

Notes: Standard errors are clustered at the provider level and shown in parentheses. * indicates p<0.05, ** indicates p< 0.01, *** indicates p<0.001. Alphas and Betas vary with each 3 years of physician experience.

"Low appropriateness" indicates patient is below the 34th percentile of our index of appropriateness for invasive procedures. "High appropriateness" indicates patient is above the 66th percentile.

Table 7: Practice Style, Procedure Use, Log(Costs), and Length of Stay for Patients with High and Low Appropriateness for Invasive Procedures

	(1)	(2)	(3)	(4)	(5)	(6)
Appropriateness for Invasive Procedure:	High	High	High	Low	Low	Low
Outcome:	Any Invasive Procedure	Total Costs	Length of Stay	Any Invasive Procedure	Total Costs	Length of Stay
Low Responsiveness (Beta<1)	-0.08*** (0.00)	-0.07*** (0.01)	-0.10 (0.07)	0.08*** (0.00)	0.05*** (0.01)	0.16* (0.07)
Low Aggressiveness (Alpha<0)	-0.09*** (0.01)	-0.11*** (0.01)	-0.30*** (0.07)	-0.11*** (0.00)	-0.08*** (0.01)	-0.01 (0.09)
High Aggressiveness (Alpha>0)	0.05*** (0.00)	0.09*** (0.01)	0.22** (0.08)	0.17*** (0.00)	0.13*** (0.01)	0.29*** (0.09)
Hospital*Year FE	Y	Y	Y	Y	Y	Y
Patient Appropriateness Index	Y	Y	Y	Y	Y	Y
Patient Age Categories & Gender	Y	Y	Y	Y	Y	Y
Previous AMI	Y	Y	Y	Y	Y	Y
Patient Comorbidities	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
N	223853	223853	223853	217323	217321	217323
R ²	0.14	0.38	0.11	0.16	0.38	0.11

Notes: Standard errors are clustered at the physician level and shown in parentheses. * indicates p<0.05, ** indicates p< 0.01, *** indicates p<0.001. Alphas and Betas vary with each 3 years of physician experience.

"Low appropriateness" indicates patient is below the 34th percentile of our index of appropriateness for invasive procedures. "High appropriateness" indicates patient is above the 66th percentile.

Table 8: Practice Style, and Detailed Log(Costs) for Patients with High and Low Appropriateness for Invasive Procedures

Appropriateness for Invasive Procedure:	(1) High	(2) High	(3) High	(4) Low	(5) Low	(6) Low
Type of Cost:	Pharmacy	Laboratory	Radiology	Pharmacy	Laboratory	Radiology
Low Responsiveness (Beta<1)	-0.05*** (0.01)	-0.02 (0.01)	0.02 (0.02)	0.04*** (0.01)	0.02* (0.01)	0.00 (0.01)
Low Aggressiveness (Alpha<0)	-0.10*** (0.02)	-0.06*** (0.01)	-0.04* (0.02)	-0.05*** (0.01)	-0.01* (0.01)	0.03** (0.01)
High Aggressiveness (Alpha>0)	0.09*** (0.02)	0.06*** (0.01)	0.05** (0.02)	0.10*** (0.01)	0.04*** (0.01)	-0.00 (0.01)
R ²	0.28	0.43	0.22	0.29	0.48	0.26
Type of Cost:	Medical Devices	Cardiology	Operating Room	Medical Devices	Cardiology	Operating Room
Low Responsiveness (Beta<1)	-0.14*** (0.02)	-0.11*** (0.01)	-0.21*** (0.06)	0.09*** (0.02)	0.18*** (0.01)	-0.04 (0.04)
Low Aggressiveness (Alpha<0)	-0.17*** (0.02)	-0.14*** (0.01)	-0.53*** (0.05)	-0.18*** (0.02)	-0.29*** (0.01)	-0.28*** (0.03)
High Aggressiveness (Alpha>0)	0.16*** (0.02)	0.07*** (0.01)	0.28*** (0.05)	0.27*** (0.02)	0.39*** (0.01)	0.27*** (0.04)
R ²	0.29	0.31	0.38	0.23	0.23	0.27
Hospital*Year FE	Y	Y	Y	Y	Y	Y
Patient Appropriateness Index	Y	Y	Y	Y	Y	Y
Patient Age Categories & Gend	Y	Y	Y	Y	Y	Y
Previous AMI	Y	Y	Y	Y	Y	Y
Patient Comorbidities	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y

Notes: Standard errors are clustered at the physician level and shown in parentheses. * indicates p<0.05, ** indicates p< 0.01, *** indicates p<0.001. Alphas and Betas vary with each 3 years of physician experience.

"Low appropriateness" indicates patient is below the 34th percentile of our index of appropriateness for invasive procedures. "High appropriateness" indicates patient is above the 66th percentile.

Table 9: Relationship between Provider Characteristics, Aggressiveness and Responsiveness

	(1)	(2)	(3)	(4)	(5)	(6)
Physician Characteristics	Responsiveness	Responsiveness	Responsiveness	Aggressiveness	Aggressiveness	Aggressiveness
Responsiveness (t-1)		0.1640*** (0.0025)	0.1655*** (0.0025)			0.0387*** (0.0018)
Aggressiveness (t-1)			0.0167*** (0.0020)		0.4755*** (0.0018)	0.4791*** (0.0018)
3-6 years experience	-0.0532*** (0.0057)			0.0086 (0.0054)		
6-9 years experience	-0.0377*** (0.0059)	0.0161** (0.0051)	0.0165** (0.0051)	0.0380*** (0.0053)	0.0406*** (0.0045)	0.0438*** (0.0045)
9-12 years experience	-0.0501*** (0.0058)	-0.0123* (0.0051)	-0.0120* (0.0051)	-0.0209*** (0.0053)	-0.0160*** (0.0045)	-0.0160*** (0.0045)
12-15 years experience	-0.0311*** (0.0059)	0.0119* (0.0052)	0.0135** (0.0052)	-0.0566*** (0.0054)	-0.0170*** (0.0045)	-0.0169*** (0.0045)
15-18 years experience	-0.0427*** (0.0060)	-0.0079 (0.0052)	-0.0058 (0.0052)	-0.1180*** (0.0054)	-0.0704*** (0.0046)	-0.0706*** (0.0046)
18-21 years experience	-0.0467*** (0.0062)	-0.0164** (0.0055)	-0.0129* (0.0055)	-0.1346*** (0.0057)	-0.0301*** (0.0048)	-0.0290*** (0.0048)
21-24 years experience	-0.0682*** (0.0067)	-0.0283*** (0.0059)	-0.0244*** (0.0059)	-0.1728*** (0.0064)	-0.0581*** (0.0053)	-0.0554*** (0.0053)
24-27 years experience	-0.1609*** (0.0079)	-0.1170*** (0.0074)	-0.1128*** (0.0074)	-0.1651*** (0.0075)	-0.0642*** (0.0063)	-0.0610*** (0.0063)
27-30 years experience	-0.1546*** (0.0089)	-0.1052*** (0.0084)	-0.1018*** (0.0084)	-0.2448*** (0.0085)	-0.1588*** (0.0071)	-0.1562*** (0.0071)
>30 years experience	-0.1873*** (0.0121)	-0.1275*** (0.0127)	-0.1218*** (0.0127)	-0.2890*** (0.0091)	-0.1232*** (0.0079)	-0.1150*** (0.0079)
US Medical School	-0.0121*** (0.0031)	-0.0184*** (0.0033)	-0.0192*** (0.0033)	0.0226*** (0.0029)	0.0096*** (0.0027)	0.0081** (0.0027)
Top-20 Medical School	0.0478*** (0.0040)	0.0601*** (0.0041)	0.0603*** (0.0041)	0.0286*** (0.0039)	0.0186*** (0.0035)	0.0189*** (0.0035)
Female Physician	0.0398*** (0.0060)	0.0335*** (0.0063)	0.0353*** (0.0063)	-0.0806*** (0.0059)	-0.0513*** (0.0057)	-0.0528*** (0.0057)
Spanish-Speaking Physician	-0.0404*** (0.0037)	-0.0310*** (0.0040)	-0.0318*** (0.0040)	0.0508*** (0.0034)	0.0113*** (0.0032)	0.0136*** (0.0032)
Hospital*Year FE	Y	Y	Y	Y	Y	Y
Patient Appropriateness Ind	Y	Y	Y	Y	Y	Y
Patient Characteristics	Y	Y	Y	Y	Y	Y
Previous AMI	Y	Y	Y	Y	Y	Y
Patient Comorbidities	Y	Y	Y	Y	Y	Y
N	658553	543192	543192	658553	543192	543192
R ²	0.14	0.18	0.18	0.24	0.43	0.43

Notes: Robust standard errors are shown in parentheses.

< 3 years of experience is the omitted category in columns (1) and (4)

< 3 years and 3-6 years of experience are the omitted categories in columns (2), (3), (5), and (6).

We lose < 3 years of experience in columns (2), (3), (5), and (6) because we control for the previous 3-year lag in alpha/beta

Table 10: Practice Style and Characteristics of All, Observed Cardiologist, and Imputed Cardiologist Samples

	(1)	(2)	(3)	(4)	(5)	(6)
Cardiologist Imputation?:	All	Observed	Imputed	All	Observed	Imputed
Physician Characteristics	Responsiveness	Responsiveness	Responsiveness	Aggressiveness	Aggressiveness	Aggressiveness
Responsiveness (t-1)	0.1655*** (0.0025)	0.1657*** (0.0030)	0.1544*** (0.0037)	0.0387*** (0.0018)	0.0386*** (0.0021)	0.0409*** (0.0030)
Aggressiveness (t-1)	0.0167*** (0.0020)	0.0111*** (0.0023)	0.0358*** (0.0034)	0.4791*** (0.0018)	0.4997*** (0.0020)	0.3992*** (0.0033)
6-9 years experience	0.0165** (0.0051)	0.0167** (0.0061)	0.0338*** (0.0092)	0.0438*** (0.0045)	0.0218*** (0.0053)	0.0889*** (0.0080)
9-12 years experience	-0.0120* (0.0051)	-0.0087 (0.0060)	-0.0134 (0.0093)	-0.0160*** (0.0045)	-0.0281*** (0.0053)	0.0085 (0.0080)
12-15 years experience	0.0135** (0.0052)	0.0142* (0.0062)	0.0125 (0.0091)	-0.0169*** (0.0045)	-0.0269*** (0.0054)	-0.0027 (0.0081)
15-18 years experience	-0.0058 (0.0052)	0.0007 (0.0062)	-0.0258** (0.0095)	-0.0706*** (0.0046)	-0.0878*** (0.0055)	-0.0197* (0.0082)
18-21 years experience	-0.0129* (0.0055)	-0.0070 (0.0066)	-0.0250** (0.0096)	-0.0290*** (0.0048)	-0.0502*** (0.0057)	0.0094 (0.0085)
21-24 years experience	-0.0244*** (0.0059)	-0.0025 (0.0070)	-0.0708*** (0.0113)	-0.0554*** (0.0053)	-0.0689*** (0.0062)	-0.0581*** (0.0097)
24-27 years experience	-0.1128*** (0.0074)	-0.1006*** (0.0086)	-0.1233*** (0.0136)	-0.0610*** (0.0063)	-0.0717*** (0.0073)	-0.0671*** (0.0121)
27-30 years experience	-0.1018*** (0.0084)	-0.0759*** (0.0099)	-0.1470*** (0.0151)	-0.1562*** (0.0071)	-0.1527*** (0.0083)	-0.1958*** (0.0137)
>30 years experience	-0.1218*** (0.0127)	-0.1039*** (0.0142)	-0.0923*** (0.0278)	-0.1150*** (0.0079)	-0.1085*** (0.0090)	-0.2086*** (0.0164)
US Medical School	-0.0192*** (0.0033)	-0.0201*** (0.0039)	0.0050 (0.0061)	0.0081** (0.0027)	0.0068* (0.0032)	-0.0017 (0.0051)
Top-20 Medical School	0.0603*** (0.0041)	0.0534*** (0.0048)	0.0757*** (0.0077)	0.0189*** (0.0035)	0.0140*** (0.0040)	0.0373*** (0.0064)
Female Physician	0.0353*** (0.0063)	0.0238** (0.0073)	0.0717*** (0.0116)	-0.0528*** (0.0057)	-0.0397*** (0.0067)	-0.1148*** (0.0101)
Spanish-Speaking Physician	-0.0318*** (0.0040)	-0.0328*** (0.0048)	-0.0326*** (0.0072)	0.0136*** (0.0032)	0.0114** (0.0038)	0.0200*** (0.0057)
Hospital*Year FE	Y	Y	Y	Y	Y	Y
Patient Appropriateness Index	Y	Y	Y	Y	Y	Y
Patient Characteristics	Y	Y	Y	Y	Y	Y
Previous AMI	Y	Y	Y	Y	Y	Y
Patient Comorbidities	Y	Y	Y	Y	Y	Y
N	543192	372985	170207	543192	372985	170207
R ²	0.18	0.17	0.27	0.43	0.43	0.48

Notes: See Table 9.

Appendix Table 1: Modeling the Probability of Invasive Procedures

	(1)	(2)	(3)	(4)	(5)
Year:	1992	1997	2002	2007	2012
female	-0.158 (-1.23)	-0.127 (-1.57)	-0.198* (-2.56)	-0.301*** (-4.45)	-0.306*** (-4.73)
<u>Co-morbidities</u>					
arrhythmia	-0.229 (-1.71)	-0.0480 (-0.57)	-0.171* (-2.08)	0.168* (2.00)	0.370*** (3.77)
hypertension	0.171 (1.38)	0.0530 (0.66)	0.237** (3.07)	0.709*** (9.52)	0.871*** (12.09)
congestive heart failure	-0.346** (-2.75)	-0.603*** (-7.25)	-0.464*** (-5.63)	-0.361*** (-4.85)	-0.158* (-2.00)
peripheral disease	-0.213 (-0.80)	0.149 (0.84)	-0.201 (-1.37)	1.124** (3.23)	0.0969 (0.42)
dementia	-2.028*** (-3.45)	-1.572*** (-5.33)	-1.668*** (-7.64)	-0.875** (-2.61)	-0.519 (-1.86)
cere disease	-0.547* (-2.43)	-0.560*** (-4.01)	-0.944*** (-7.13)	-1.245*** (-8.68)	-0.769*** (-6.23)
chronic obstructive pulmonary disease	-0.444** (-2.86)	-0.205* (-2.04)	-0.453*** (-4.88)	-0.180 (-1.89)	-0.195 (-1.61)
lupus	-0.261 (-0.50)	-0.684** (-2.67)	-0.312 (-1.40)	-0.423 (-1.70)	-0.338 (-1.27)
ulcer	0.268 (0.44)	0.0453 (0.09)	-0.939* (-2.26)	-0.564 (-1.24)	0.490 (0.92)
liver disease	-1.991** (-3.17)	-1.151*** (-3.38)	-0.685** (-2.74)	-0.732** (-2.72)	-0.932*** (-4.94)
cancer	-0.518 (-1.88)	-0.245 (-1.79)	-0.832*** (-6.30)	-1.002*** (-6.23)	-0.483** (-3.08)
diabetes	-0.306* (-2.15)	-0.0833 (-0.91)	-0.257** (-3.09)	0.574*** (5.26)	0.337*** (3.43)
kidney disease	-1.116*** (-4.33)	-0.706*** (-5.16)	-1.023*** (-9.10)	-0.853*** (-10.60)	-0.734*** (-9.82)
hiv	-1.818 (-1.64)	-1.405** (-2.58)	-1.656*** (-5.40)	-1.067** (-2.85)	-1.164** (-3.14)
Age Group					
age 50-54	0.212 (0.74)	0.115 (0.46)	0.283 (1.25)	0.269 (1.42)	0.251 (1.30)
age 55-59	0.219 (0.74)	-0.0738 (-0.33)	0.253 (1.20)	0.0209 (0.12)	0.160 (0.88)
age 60-64	-0.352 (-1.32)	0.133 (0.60)	-0.0408 (-0.20)	0.163 (0.97)	-0.146 (-0.86)

age 65-69	-0.388 (-1.50)	-0.218 (-1.01)	-0.185 (-0.93)	-0.300 (-1.83)	-0.0701 (-0.43)
age 70	-0.316 (-1.25)	-0.368 (-1.81)	-0.0764 (-0.39)	-0.200 (-1.23)	-0.295 (-1.83)
age 71	-0.213 (-0.57)	-0.486 (-1.76)	-0.766** (-2.79)	-0.296 (-1.18)	-0.742** (-3.19)
age 72	-0.477 (-1.15)	-0.591* (-2.14)	-0.0772 (-0.28)	-0.227 (-0.84)	-0.620** (-2.61)
age73	-0.608 (-1.53)	-0.509 (-1.90)	-0.244 (-0.89)	-0.495 (-1.90)	-0.443 (-1.79)
age74	-0.976* (-2.49)	-0.800** (-2.78)	-0.770** (-3.02)	-0.177 (-0.72)	-0.459* (-1.98)
age75	-0.978* (-2.30)	-0.827** (-3.01)	-0.588* (-2.27)	-0.234 (-0.94)	-0.236 (-0.96)
age76	-1.521** (-3.28)	-0.861** (-3.10)	-0.625* (-2.33)	-0.988*** (-4.47)	-0.234 (-0.87)
age77	-0.528 (-1.35)	-1.012*** (-3.66)	-0.177 (-0.61)	-0.747** (-3.16)	-0.910*** (-3.92)
age78	-0.920* (-2.14)	-1.047*** (-3.64)	-0.659* (-2.43)	-0.486* (-1.97)	-0.833*** (-3.37)
age79	-1.146** (-2.68)	-1.126*** (-4.05)	-1.169*** (-4.66)	-0.610** (-2.72)	-1.069*** (-4.50)
age 80	-0.887 (-1.77)	-1.223*** (-4.43)	-0.775** (-2.81)	-1.062*** (-4.95)	-0.878*** (-3.74)
age 81	-1.835*** (-3.62)	-1.444*** (-5.26)	-0.839** (-2.96)	-0.954*** (-4.11)	-1.014*** (-4.44)
age 82	-2.379*** (-4.33)	-0.986** (-3.25)	-0.746* (-2.46)	-0.808*** (-3.32)	-1.275*** (-5.62)
age 83	-2.314*** (-3.79)	-1.579*** (-4.86)	-0.892** (-3.24)	-0.889*** (-3.82)	-1.016*** (-4.33)
age 84	-3.404*** (-4.66)	-1.870*** (-6.07)	-1.036*** (-3.68)	-1.157*** (-4.95)	-1.261*** (-5.17)
age 85	-2.557** (-3.22)	-2.272*** (-6.87)	-1.481*** (-5.40)	-1.518*** (-6.44)	-1.619*** (-6.89)
age 86	-3.141** (-3.19)	-2.233*** (-6.05)	-1.724*** (-4.98)	-2.023*** (-7.97)	-1.368*** (-5.36)
age 87		-1.989*** (-4.87)	-1.910*** (-5.66)	-2.115*** (-7.75)	-1.942*** (-7.09)
age 88	-3.193** (-2.89)	-2.591*** (-6.49)	-1.767*** (-4.77)	-1.948*** (-5.92)	-1.955*** (-6.96)
age 89		-2.555*** (-5.25)	-2.426*** (-6.50)	-2.298*** (-7.45)	-1.932*** (-6.23)

age 90		-2.508*** (-4.71)	-2.648*** (-5.81)	-2.355*** (-7.36)	-2.113*** (-6.74)
age > 90		-3.453*** (-8.70)	-3.446*** (-11.31)	-3.238*** (-14.06)	-2.849*** (-12.88)
<u>Heart diagnoses (ICD codes)</u>					
410.01	1.128 (1.77)	0.801 (1.67)	2.014* (2.54)	2.482** (3.08)	
410.02			-1.950 (-1.56)	-1.034 (-0.78)	
410.10		-0.159 (-0.09)		-0.135 (-0.10)	
410.11	1.486* (2.39)	1.145* (2.48)	2.173** (2.77)	2.092** (2.77)	
410.12	0.00773 (0.01)	-1.153 (-1.73)	-1.861* (-2.08)	-0.355 (-0.40)	-1.473 (-1.60)
410.21	2.104** (2.85)	1.386** (2.76)	2.235* (2.54)	2.098* (2.55)	
410.22	1.094 (0.95)		-0.963 (-0.63)		1.795 (0.87)
410.31	2.211* (2.28)	0.991 (1.88)	2.341** (2.66)		
410.41	1.430* (2.34)	1.138* (2.48)	2.333** (2.95)	2.542*** (3.31)	
410.42	-0.398 (-0.52)	-0.109 (-0.18)	-0.945 (-1.11)	-1.646 (-1.91)	-1.045 (-1.08)
410.51	1.518* (2.33)	0.755 (1.45)	1.382 (1.66)	2.023* (2.27)	
410.52	1.543 (1.34)		-0.656 (-0.49)		
410.61	0.344 (0.42)	1.607 (1.67)	0.538 (0.56)		3.005** (2.69)
410.62			1.043 (0.89)		
410.70		0.651 (0.66)		0.318 (0.27)	1.972* (2.31)
410.71	1.108 (1.79)	0.628 (1.38)	1.284 (1.67)	0.673 (0.91)	1.437* (2.18)
410.72	-0.634 (-0.85)	-1.786** (-3.11)	-1.122 (-1.54)	-1.543* (-2.14)	-1.078 (-1.66)
410.80			-1.472 (-1.15)		
410.81	0.699	0.528	0.712	0.796	1.607*

	(1.11)	(1.12)	(0.90)	(1.02)	(2.23)
410.82	-0.584	0.0963	0.0321	-1.606	0.166
	(-0.50)	(0.12)	(0.02)	(-1.13)	(0.14)
410.90	0.127	-1.397*		-1.152	0.969
	(0.12)	(-1.97)		(-0.97)	(0.93)
410.91	0.0360	-0.733	0.207	0.0297	1.118
	(0.06)	(-1.56)	(0.27)	(0.04)	(1.69)
410.92	-0.464			-2.040*	-1.151
	(-0.64)			(-2.49)	(-1.56)
Quarter 1	0.0235	-0.0834	-0.240*	0.0118	0.190*
	(0.15)	(-0.79)	(-2.35)	(0.13)	(2.15)
Quarter 2	-0.0827	-0.0493	-0.126	-0.0693	0.0685
	(-0.51)	(-0.46)	(-1.19)	(-0.74)	(0.78)
Quarter 3	0.0875	0.0512	0.0575	0.0503	0.0992
	(0.51)	(0.44)	(0.53)	(0.53)	(1.10)
Constant	0.144	0.859	0.751	0.618	-0.252
	(0.23)	(1.73)	(0.95)	(0.82)	(-0.37)
N	1,595	3,832	4,972	6,263	5,658
Pseudo R ²	0.2134	0.2313	0.2773	0.2874	0.2003

Notes: These models are estimated using only patients in programs accredited to teach internal medicine. These are coefficients from the logit model (not marginal effects).

Diagnoses codes 410.00, 410.20, 410.30, 410.32, 410.40, 410.50 and 410.60 were collinear with the other variables included in the model and hence dropped from the regressions.

Appendix Table 2: Distribution of Patients Per Physician Before and After Sample Restriction

Percentile	Per 3-years	Per 3-years
	Before	After
1%	3	7
5%	11	14
10%	20	23
25%	48	50
50%	105	107
75%	223	226
90%	389	389
95%	540	549
99%	1,012	1,012
N	668,616	658,553

(1) We restrict the sample to cardiologists whose license numbers match the license numbers in Florida's medical practitioner database. (2) We restrict the sample to cardiologists who treat at least 5 AMI patients per 3-year period.

Appendix Table 3: Robustness of main results to an alternative categorization of alpha and beta.

	(1)	(2)	(3)	(4)	(5)	(6)
Appropriateness for Invasive Procedure: Outcome:	High	High	High	Low	Low	Low
	Hospital Aquired Infection	Died in Hospital	Discharged Home	Hospital Aquired Infection	Died in Hospital	Discharged Home
Low Aggressiveness, Typical Responsiveness (Alpha<0, Beta>=1)	0.010*** (0.002)	0.009*** (0.002)	-0.020*** (0.003)	0.014*** (0.003)	0.012*** (0.003)	-0.024*** (0.004)
High Aggressiveness, Typical Responsiveness (Alpha>0, Beta>=1)	-0.002 (0.002)	-0.004*** (0.001)	0.011*** (0.002)	-0.013*** (0.003)	-0.019*** (0.002)	0.022*** (0.004)
Low Aggressiveness, Low Responsiveness (Alpha<0, Beta<1)	0.019*** (0.005)	0.021*** (0.003)	-0.037*** (0.008)	0.002 (0.005)	0.001 (0.005)	-0.013 (0.007)
Typical Aggressiveness, Low Responsiveness (Alpha>=0, Beta<1)	0.005** (0.002)	0.007*** (0.002)	-0.021*** (0.003)	-0.013*** (0.003)	-0.018*** (0.003)	0.015*** (0.004)
Hospital*Year FE	Y	Y	Y	Y	Y	Y
Patient Appropriateness Ind	Y	Y	Y	Y	Y	Y
Patient Age & Gender	Y	Y	Y	Y	Y	Y
Previous AMI	Y	Y	Y	Y	Y	Y
Patient Comorbidities	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
N	223853	223853	223853	217323	217323	217323
R ²	0.05	0.06	0.29	0.08	0.08	0.12

Notes: See Table 5.

Appendix Table 4: Robustness of Main Results to Inclusion of Race, Ethnicity, and Insurance Status

	(1)	(2)	(3)	(4)	(5)	(6)
Appropriateness for Invasive Procedure:	High	High	High	Low	Low	Low
Outcome:	Hosp. Acquired Infection	Died in Hospital	Discharged to Home	Hosp. Acquired Infection	Died in Hospital	Discharged to Home
Low Responsiveness (Beta<1)	0.007*** (0.002)	0.009*** (0.001)	-0.024*** (0.003)	-0.009*** (0.003)	-0.011*** (0.003)	0.008* (0.004)
Low Aggressiveness (Alpha<0)	0.010*** (0.002)	0.009*** (0.001)	-0.018*** (0.003)	0.014*** (0.003)	0.013*** (0.002)	-0.024*** (0.003)
High Aggressiveness (Alpha>0)	-0.003 (0.001)	-0.005*** (0.001)	0.012*** (0.002)	-0.011*** (0.003)	-0.019*** (0.002)	0.020*** (0.003)
Quarter since sample start	Y	Y	Y	Y	Y	Y
Hospital FE	Y	Y	Y	Y	Y	Y
Patient Appropriateness Index	Y	Y	Y	Y	Y	Y
Patient Age & Gender	Y	Y	Y	Y	Y	Y
Previous AMI	Y	Y	Y	Y	Y	Y
Patient Comorbidities	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
Race, Ethnicity, Insurance	Y	Y	Y	Y	Y	Y
N	223853	223853	223853	217323	217323	217323
R ²	0.06	0.06	0.29	0.08	0.08	0.13

Notes: Standard errors are clustered at the physician level and shown in parentheses. * indicates p<0.05, ** indicates p< 0.01, *** indicates p<0.001. Alphas and Betas vary with each 3 years of physician experience. "Low appropriateness" indicates patient is below the 34th percentile of our index of appropriateness for invasive procedures. "High appropriateness" indicates patient is above the 66th percentile.