The Economics of Iatroepidemics and Quakeries:
Physician Learning, Informational Cascades and Geographic Variation in Medical Practice[*]

Sushil Bikhchandani
UCLA
sushil.bikhchandani@anderson.ucla.edu

Amitabh Chandra
Dartmouth College
amitabh.chandra@dartmouth.edu

Dana P. Goldman
RAND
dgoldman@rand.org

Ivo Welch
Yale University
ivo.welch@yale.edu

First draft: September 3, 2000
This version: July 10, 2001

We formalize the role of physician learning and technology diffusion in the market for health care and offer a theoretical explanation for the well-documented phenomenon of variation in medical practice style. A simple Bayesian learning model demonstrates that physicians imitate one another because they learn information from the treatment choices of their colleagues, and closer colleagues provide more information. Therefore over time, the market will converge at the optimal dosage. However, our paper differs in one aspect, which may appear innocuous at first, but which has dramatic implications. In situations in which the physician's choice is discrete—i.e., in which the physician must select one from a select small number of choices—such learning can cease altogether. Therefore, we predict that discrete choice treatment methods suffer more from localized variation than continuous "dosage" treatment choices. Moreoever, it is quite possible that the medical community focuses on the wrong treatment choice (an iatroepidemic) and stays with it, even in the presence of legal repercussions and despite perfectly rational behavior by every physician. Finally, we argue that even small-scale medical studies can have a dramatic effect in overthrowing long-standing cascades and providing better treatment choices. We find empirical support for our model using data from the NHAMCS and the *Dartmouth Atlas of Cardiovascular Care.*

"It would appear from the available evidence that the extent of geographic variations [in medical practice use] can be explained only partially by standard economic phenomena such as price, income, and (using limited evidence) the distribution of illness across regions...Thus, none of the standard explanations for the cause of variations can be considered plausible...We thus turn our attention to the one remaining concept---differences in beliefs about the efficacy of treatment and decisions about which patients should receive treatment."

--Phelps and Mooney (1993), p. 153.

I.   Introduction

In 1938, Sir Allison Glover first presented a study which documented that tonsillectomy varies dramatically and systematically by locale within the United Kingdom [Glover (1938)]. Since then, an enormous body of literature in economics and medicine has thoroughly documented that similar variations in treatment are observed for many procedures and across many similar locales. In their pioneering paper in *Science*, Wennberg and Gittelson (1973) demonstrate that the chance of receiving a tonsillectomy varies from 7-70 percent across similar towns in Vermont. Wennberg, Freeman and Culp (1987) compare the use of medical procedures in Boston and New Haven (two cities that are very similar in terms of the presence of major academic medical centers, demographics, incomes, and health insurance coverage rates) and find that Boston residents spent almost 87 percent more per capita than New Haven residents on hospital care. In subsequent work, Wennberg (1990) examined variation in the use of surgical procedures at 16 university hospitals and established the existence of considerable variation in the practice of medicine at the nations' most elite hospitals. Traditional explanations—such sampling variation, differences in income, physician density, and underlying health status—cannot explain the large residual variation across very similar regions.[1] Such explanations also have an extremely difficult time rationalizing the existence of "quackeries" such as the use of leeches and bloodletting by physicians.[2]

---

[1] Phelps (1992, 1997, 1999) and Phelps and Mooney (1993) provide rigorous reviews of the immense empirical literature on geographic variation in physician practice style. In the light of their work, we do not review this literature in its entirety and refer the interested reader to their papers for an introduction to the small area variation phenomena. Briefly, this literature has found that the lowest rates of variation (as measured by the coefficient of variation (COV)) are found for procedures where the role of the intervention is well understood (e.g. hip fracture and acute myocardial infraction; COV typically less than 0.15). Other procedures such as treatments for back injury, diabetes, and hypertension have a COV in the range of around 0.5. Cardiac interventions tend to have the highest rates of variation in all the studies that we have reviewed.

[2] For a fascinating discussion of surgical fads (such as the use of elective hysterectomies, internal mammary ligation, and ileal bypass) and their deleterious consequences see Taylor (1979) and Robin (1974).

Most of this literature has focused on variation across geographic locales. However, similar questions arise over time rather than space when one considers the diffusion of medical technologies. While variation across time is not well-documented in medicine, it almost assuredly cannot be explained by traditional measures. Either way, the implications of the sustained presence of variation in physician practice style across time and space depend on the answer to three important questions:

First, is there one correct treatment? If so, the presence of variation implies considerable disagreement among physicians on the correct choice of treatment and raises troublesome implications for the presence of iatrogenic (physician induced) complications in the provision of healthcare. This is especially true if areas with high intervention rates provide "excessive" healthcare and hence have negative marginal productivity of healthcare.[3] An immediate corollary to this question asks whether one could derive a better treatment by taking the best of multiple approaches.

Second, what are the social costs of administering a treatment when another treatment is more efficient? If there is one optimal treatment choice given the costs of various procedures, the choice of suboptimal method is costly. This can be either because the less effective treatment is chosen, or because the more expensive treatment is no more effective.[4]

---

[3] A growing body of evidence from the medical literature supports the hypothesis that many aggressive interventions may actually result in deleterious outcomes for patients; hence our use of of the word "iatrogenic" to describe this phenomenon. For economists this is analogous to operating in a region of negative marginal product on the production function. Much of this literature has been published in *the New England Journal of Medicine,* the *Journal of the American College of Cardiology* and the *British Medical Journal.* See for example, the work of Bodel et al. (1998) who use a randomized trial to demonstrate that for patients with acute coronary disease the use coronary angiography (an invasive procedure) resulted in higher one year mortality than the group that received more conservative (ischemia driven management) therapy. Other references include Anderson et al. (1995), The TMI Study Group (1989), and SWIFT (1991). Indirect evidence comes from noting that during the medical malpractice debates of the 1970s, physicians in Los Angeles practiced a work slowdown (to force changes in California's malpractice law). During this time they only treatment emergency room cases. Surprisingly, the mortality rate actually fell during this period and Roemer and Schwartz (1979) suggest a causal interpretation for this result.

[4] Phelps and Parente (1990) and Dranove (1995) develop models to estimate the economic costs of iatroepidemics. Using welfare loss triangles and under the crucial assumption that the average rate of use approximates the "correct" rate, Phelps and Parente estimate that the annual welfare loss from cross regional variation alone exceeded $ 7 billion in 1990. It should be noted that their estimated is a lower bound for the size of the welfare loss as they ignore the costs of within regional variation. Dranove discuses the exact conditions under which this is a correct approximation and notes that the Phelps and Parente valuation could be seriously biased. In contrast to their work, our model predicts that for surgical (discrete) procedures there is no reason to believe that the average rate is close to the desirable rate.

Finally, is there a role for regulation in mitigating—or augmenting—localized treatment variations? [5] The answer to this last question depends on how treatment decisions are made, and perhaps upon the nature of the treatment. It is this issue that we address in this paper.

Despite the substantial empirical work documenting practice variation few explanations have been offered. Folland and Stano (1989, 1990), and Phelps and Mooney (1993) are the notable exceptions. As a result, there is little guidance as to what types of treatment variation might be expected to persist in the absence of—or despite—intervention. This paper argues that certain innate qualities of the treatment could lead to dramatic differences in the ways they diffuse. To demonstrate this point, we derive a information model of physician learning—building on the work of Phelps and Mooney (1993)— which provides a rich set of testable empirical predictions, and we confirm many of these predictions by using the findings of previous studies as well as our own research. The basic Bayesian learning model demonstrates that physicians imitate one another because they learn information from the treatment choices of their colleagues, and closer colleagues provide more information. Therefore over time, the market will converge at the optimal dosage. Our model differs in one aspect, which may appear innocuous at first, but which we will show to have dramatic implications. We demonstrate that such learning can cease altogether in situations in which the physician's/patient's choice is discrete, i.e., in which the physician must choose one treatment from a select small number of choices. This observation offers the empirical implication that discrete choice treatment methods suffer far more from localized or temporal variation than continuous ("dosage") treatment choices. Second, it argues that it is quite possible that the community focuses on the wrong treatment choice and stays with it, even in the presence of legal repercussions and despite perfectly rational behavior by every physician. And third, it argues that even small-scale medical studies can have a dramatic effect in overthrowing long-standing cascades and providing better treatment choices.

Our paper is organized as follows: Section II examines the empirical evidence put forth to explain the geographic variation phenomena. We demonstrate that most of the evidence favors a model of physician learning in the presence of uncertainty rather than one that is based on standard supply and demand explanations. To support our case, we present evidence from the Dartmouth Atlas of Health

---

[5] Cognizant of many of these implications, Congress established the Agency for Health Care Policy and Research (AHCPR) in 1989 with the purpose of determining "what works and to develop practice guidelines and standards to assess and assure quality of care." In 1999, President Clinton signed the Heath Care Research and Quality Act reauthorizing the AHCPR, and emphasizing its focus on quality by changing its name to the Agency for Healthcare Research and Quality (AHRQ).

Care (CECS, 1999a) and the Dartmouth Atlas of Cardiovascular Care (CECS, 1999b).[6] Next, we review the Phelps and Mooney (henceforth PM) model, and provide an intuitive motivation for our model and contrast our model with the PM model. We highlight empirical observations made in the literature that are rationalizable in our model, but not in PM. Specifically, our model is able to explain persistent variation across two contiguous hospitals. In contrast, the PM model predicts that over time all physicians learn the right technique and the market converges to the optimal dosage. In Section III we formalize a model that leads to the result that there is more variation in discrete treatment choices, derive a set of testable implications and discuss the robustness of the model. Section IV offers empirical support for our theoretical approach at three different levels, and Section V concludes.

II. Explaining Practice Variation

The basic facts that need to be explained are presented in Figure 1, which summarizes the variation in rates (on a log scale) at which 10 common surgical procedures are used relative to the US average (in 1996). Similar results have been documented for the variation in rates at which different diagnostic tests are utilized. Together the 10 procedures listed in Figure 1 comprised 42% of Medicare inpatient surgery and accounted for 44% of reimbursements for surgical care in 1995-96. When aggregated across all procedures and individuals these variations become even more pronounced, and therefore have enormous implications for the provision of healthcare as well as equity. Using the same data, Medicare payments for services reimbursed on a fee-for-service basis (including non-risk bearing health maintenance organizations) were $4,993 for each beneficiary. However, even after controlling for age, sex, race, illness patterns, and differences in regional prices, reimbursements per enrollee varied greatly: as noted in the Atlas, these ranged from $9,033 in the McAllen, Texas hospital referral region to $3,074 in Lynchburg, Virginia. This fact is illustrated in Figure 2.[7] The provocative nature of these

---

[6] The Atlases are produced at the Center for Evaluative Clinical Sciences at Dartmouth College and use 100% HCFA Claims data for the construction of utilization rates for procedures and diagnoses. The data are aggregated up to the Hospital Referral Region (HRR) level. There are approximately 300 such regions in the US. In Appendix B we provide a brief outline of how the Atlas data were constructed. We refer the reader who is interested in further details to the source documentation at: http://www.dartmouthatlas.org/

[7] Illness has been controlled for by using age-sex-race specific mortality and hospitalizations rates for five conditions: hip fracture, cancer of the colon or lung treated surgically, gastro-intestinal hemorrhage, acute myocardial infarction or stroke. These conditions were chosen because hospitalization for them is a proxy for the incidence of disease. The cost of living indices were computed by using non-medical regional price measures. Doing so avoids contaminating the analysis with physician workforce or hospital market conditions.

results has not gone unnoticed and several hypotheses have been put forward to explain these variations. Next, we briefly review the evidence in favor of each argument.[8]

Sampling. It is possible that much of the observed variation reflects random deviations from identical practice patterns across communities. However, work by McPherson et al. (1981) demonstrate this is not sufficient by assuming that hospitalization follows a Poisson process. Under this assumption, they recover the systematic component of COV and find that only 1-4 percent of total variation in Canada and 15 percent in the U.K. is attributable to "noise". Diehr et al. (1992) use simulation methods to formalize the extent to which "noise" could account for the findings of previous studies but find that many of them did indeed identify real variation.[9] Whereas this hypothesis may be a legitimate criticism for inferences made on smaller datasets, the Atlas' reliance on the enormous sample sizes derived from the 100% Medicare Claims data rejects its validity.

Patient Preferences. This class of explanations focuses on whether differences in patient preferences across regions are responsible for the observed variation in practice style. While simple and intuitive, a tastes based explanation is not useful for economists. The empirical evidence also rejects this hypothesis. In the highly publicized results of the SUPPORT trial [see Pritchard, et al (1998)] patients at five medical centers were asked for their preferences for end of life care (for example, dying at home (the preferred modality), or dying in the hospital). Despite the knowledge of these preferences it was found that the variation in care that patients actually received could not be explained by either their stated preferences or their clinical presentations. Instead it was found that actual care was best predicted by the regional use of that form of care.

Income and Price Variability. Expensive procedures can be more popular when they are cheaper or when the prospective recipient is wealthier. However, the evidence suggests that iatroepidemics occur in similarly wealthy regions, and in situations in which health insurance generally pays for the procedures and in which health insurance is equally prevalent. As documented by Phelps (1992, 1999) variations

---

[8] For the sake of clarity, we reiterate that we are not arguing that these factors do not matter. Indeed they do and are often precisely estimated in estimated regressions. At times they can explain a substantial portion of the observed raw variation. Our point is to note that even after controlling for these factors, a substantial portion of residual variation remains. This residual variation represents an upper bound on the size of the practice-style hypothesis.

[9] Almost every study in this literature uses the coefficient of variation to summarize the extent of the geographic variation phenomena. Diehr et al, also derive a useful result for testing whether the underlying rates of usage are the same across regions by noting that that the statistic $COV^2(k\text{-}1)Np/[(1\text{-}p)k] \sim \chi^2(k)$, where k = number of small areas, N = total population, and p = rate of hospitalization. We note that it is also possible to use empirical bayes methods to adjust for noise-driven variation.

have been found for Canada (where full coverage health insurance is universal) and the United Kingdom (where the provision of care is largely socialized). McPherson et al. (1982) find variation in utilization rates for New England, Norweigian counties and in districts of the West Midlands in the U.K.[10]

Illness Patterns. Some illnesses are more prevalent in some locales and there may be variation across areas in the severity of illnesses. While this is certainly the cause of variation in some cases, many illnesses show almost no variation across relatively homogeneous populations and are still subject to varying treatment choice. Figure 2 demonstrates that even when controls for observable dimensions of sickness are included there are substantial variations across regions. The degree to which patients are unobservably sicker would have to be immense to completely explain these differences. Fisher et al (1994) construct cohorts of Medicare beneficiaries on the basis of initial hospitalization for either AMI, stroke, GI bleeding, hip fracture or potentially curable surguery breat, colon or lung cancer. They find that there are substantial differences in the intensity with which beneficiaries were treated (as measured by readmission rates) even across similar teaching hospitals in the Boston area. Specifically, there is substantial variation across the readmission rates for Massachusetts General Hospital, Brigham and Women's Hospital, Beth Israel, and Boston University Medical Center. Most interestingly, there is no relationship between mortality (both 30 day and over the entire study period) and the intensity of hospitalization.[11]

---

[10] It is important to note that the level of aggregation used in these regressions plays an important role in the extent to which economic variables can explain the observed variation. The larger the level of aggregation (say the use of aggregate medical spending per capita instead of a procedure-specific measure) the larger will be the explanatory power of the regression. Phelps and Parente (1990) find that economic and demographic variables account for 45-75 percent of the observed variation in 134 separate diagnostic categories, and Escarce (1993) is able to explain only 43 percent of the observed variation in cataract surgery rates for Medicare beneficiaries through such controls. In an often overlooked paper, Carlisle et al. (1990) use data from Los Angeles county to examine the contributions of factors such as the severity of illness, distance to hospital, MD density, family income, and demographics on the rates of surgical procedures such as angio plasty, apppendectomy, CABG, endarterecctomy, hysterectomy (separately for the presence of uterine carcinoma and cervical carcinoma), mastectomy and pacemaker implantation. These variables are always individually significant, but only explain 9-43 percent of the observed variation in surgical rates. The use of mastectomies is an exception, where 67 percent of the observed variation can be explained, primarily through the severity of illness variable.

[11] There are a number of studies that have found this result, primarily by controlling for the case-mix of the underlying population through regression analysis. Examples include Wennberg (1987), Roos and Roos (1982), and Carlisle et al. (1995). We believe that the definitive papers on this topic are Chassin et al. (1987) and Leape et al. (1990) because of their reliance of direct evidence. In the first paper the authors match patients to their detailed medical records. By using previously developed rating methods, they assign ratings of clinical appropriateness of each procedure for each patient. They find that the variation in appropriateness of procedure does not vary across regions of low and high use, suggesting that the patients are clinically similar. However, they find for example, that coronary angiography was used 2.3 times more often in the high-use site compared to the low use-site. Leape et al. (1990) adopt a similar approach by using medical records to study the appropriateness of coronary angiography,

Access to Medical Care. Procedures can be more often administered when there is space and expertise in the procedure available. There are two problems with this explanation: First, the question of variability can be rephrased as: why does medical space and expertise cluster in certain locales? In other words, capacity is an endogenous variable. Second, Wennberg (1987) demonstrates that hospitalization rates for hip-fractures have small variability, whereas those for fractured ankles and forearms exhibit much higher variation. The coefficient of variation in hospital admissions for acute myocardial infarction (AMI) is low, whereas that for angina pectoris exceeds 0.3. Because both procedures require a similar combination of inputs, (differences in say beds/surgeons are constant), variation patterns should be constant across multiple diseases requiring surgery.

Substitution across Procedures. If there are substitutes for a given procedure, different localities may use the procedures at different rates, yet stay on the same production isoquant. Phelps and Mooney (1993) provide systematic evidence to refute this conjecture. They find that the correlation between substitute procedures for low-back injury, coronary artery disease, cardiac arrhythmias, non-cancerous uterine disorders, cataracts, and the diagnosis of a stroke is surprisingly positive. They also found a positive correlation between the use of inpatient and outpatient treatment.[12] In Figure 3, we use data from the *Dartmouth Atlas of Cardiovascular Care* we illustrate the nature of this finding using another example. We regress the rates of imaging stress tests (echocardiography or nuclear studies) against rates for non-imaging (exercise) stress tests.[13] Both are used to evaluate patients with suspected coronary disease and we would expect there to be a negative relationship between the utilization rates of the two tests. However, the R^2 in the regression in Figure 3 is exactly zero. This fact may be interpreted to imply that the variation is attributable almost entirely to physicians *perceptions* of the relative efficacy of each test.

---

upper GI endoscopy and carotid endarterectomy. They find no relationship between the severity of illnesses and the use of a surgical intervention.

[12] They do find however, that in urban areas the correlation between the use of inpatient and outpatient treatments of knee surgery is negative, and that the use of intensive vs. nonintensive care beds for patients diagnosed with acute MI or angina pectoris was negative.

[13] We compare this relationship at both the Hospital Referral Region (HRR) and the Coronary Angiography Service Area (CASA). CASAs are analogous to HRRs and were constructed for the Dartmouth Atlas of Cardiovascular Care [see CECS(1999b, p.250) for exact details of the construction algorithm]. CASAs represent service areas according to where Medicare beneficiaries received their cardiac cauterizations. They were constructed in order to take into account the criticism that the HRRs were too big a service area for the study of cardiac catheterization. CASAs are not allowed to cross HRRs.

Legal Considerations. By adhering to the same local standard, physicians reduce the probability of experiencing legal repercussions in case of difficulties. Medical malpractice law defines "standard of care" in terms of local practice standards, not national ones. We do not believe however, that a substantial portion of the observed variation in practice style is attributable to this hypothesis; iatroepidemics occur as much in countries such as Great Britain, where the propensity to sue or discipline physicians is almost non-existent, and they are found to exist internationally across many different health care systems [see McPherson et al. 1981, 1982].

Physician Learning. PM and Phelps (1999) suggest that explanations based on the nature of physician learning are most likely to account for much of the empirically observed locality of treatment. In the PM model, physicians are Bayesian learners, who attempt to reach an optimal rate for the application of a particular treatment. Eventually, as physicians sample both their own and their colleagues experiences, the two will converge towards an optimal rate. This hypothesis suggests a number of implications: a physician's propensity to treat converges toward the community norm, and faster if the community is more informed and the doctor is less informed (e.g., younger). Specialists will pay more attention to other specialists and less attention to the general population. Among the implications of this theory is the hypothesis that the provision of more precise medical information in medical studies can enhance the learning of physicians, and thus offer dramatic social efficiency gains.

The PM model is powerful and intuitive. However, there are two important pieces of empirical evidence that contrast sharply with its central prediction that medical knowledge follows a gradual diffusion process:

(1) Even though Sir Allison Glover published his first study in 1938, and pediatric societies began to discourage the widespread use of tonsillectomy [see Phelps (1992)], Glover's own subsequent work a decade later found virtually identical variations [Gover (1948)], as did Phelps and Parente (1990) using New York data from the 1980s. Medical knowledge certainly diffuses over time, but there are instances when it ceases. A theoretical model should be able to reconcile *both* empirical features.

(2) One consequence of the PM model is that regions that are adjacent to one another should have similar practice styles (because geographic proximity aids diffusion). However, as the previously discussed evidence from Elliott et al (1994) notes, there is substantial variation in medical intensity for similar patients within city and within the class of teaching hospitals: Relative to Yale Hospital, Massachusetts General has a relative rate of readmission of 1.50

(95% CI: 1.30-1.73) but Boston University Medical Center's relative rate is 1.98 (95% CI: 1.64-2.39).[14]

(3)   In the PM model, an informational "shock" or intervention whereby physicians are informed about the risks and benefits of procedures that they are using should not drastically alter their practice style (because information ultimately diffuses through the economy). However, we provide evidence from the medical literature that such interventions do alter physician behavior. Furthermore, we also provide evidence that such interventions reduce mortality rates thereby suggesting that iatroepiedemics are a natural consequence of the nature of physician learning.

Our own paper agrees with the main premise offered in PM: physicians imitate one another because they learn information from the treatment choices of their colleagues, and closer colleagues provide more information. However, our paper distinguishes between discrete and continuous treatment choices. This distinction allows us to specify a theoretical model whose comparative statics are able to rationalize the three points above. We demonstrate that in the discrete case, Bayesian learning can cease altogether—a situation which is best illustrated by contrasting the Phelps-Mooney analysis with our own. (We will henceforth refer to a continuous choice as a "dosage," to a discrete choice as a "procedure.") The Phelps and Mooney analysis and our model apply equally well to the choice of the right dosage of tetracycline. The correct dose may not yet be known, and no doctor can *perfectly* infer success from past treatment outcomes. If doctor A were to have privately believed that a 100mg dose of alprazolam is appropriate, and he then observes that colleagues would have chosen a 200mg dose, he can adjust his dosage prescription to 150mg. His colleagues observe A's choice, and would reduce their own dosage (e.g., to 190mg), recognizing that A would have prescribed even less without A's assessment of community standards. Continuous experimentation and feedback between mutual observation of one another's dosage recommendations—or, analogously, from self-experimentation—would soon result in a local standard that efficiently aggregates the information of most local physicians. Similar communications among communities will eventually (at a slower rate) aggregate information across different medical communities, and local standards and national standard will converge. Some local

---

[14] The analysis controlled for age, sex, race, cause of admission, and coexisting conditions. For each hospital a cohort of Medicare beneficiaries was constructed on the basis of initial hospitalization for either AMI, stroke, GI bleeding, hip fracture or potentially curable surguery breat, colon or lung cancer [between October 1, 1987 and September 30, 1989]. The cohorts were followed for 35 months and the readmission rates were computed for this entire period.

treatment variation arises, because knowledge and observations of treatment successes by physicians will take a while to diffuse into their environment, and faster when physicians are closer.[15] Although physicians may administer a less effective dose for a while, eventually feedback between the choices of physicians would aggregate the information held by many physicians, and the treatment would converge towards the best treatment choice, given the aggregate information. Even without observing colleagues, an individual doctor could experiment with slightly different dosages and eventually calibrate on the optimal choice. It would take longer, and treatment patterns would not show localized variation, but eventually, patients would be well served.

But now consider a different choice, e.g., the choice of whether to perform an angiogram. The main choice is not "how much" angiogram, but whether to perform it or not. (A similar discrete choice may be whether to prescribe Alprazolam *at all*). A cardiologist observes that his three colleagues performed angiograms in a particular situation. Assuming this cardiologist suffers neither from hubris or humility (perhaps a rare quality for many cardiologists), he attributes the competencies of his colleagues to be of equal quality to his own. Counting three opinions pro, even if this cardiologist had privately believed that not doing an angiogram was the right choice, he would still find it appropriate to perform it—after all, three pros outweigh one con. The cardiologist is "herding" on the choices of his peers. Although he has private information and attempts to make the best choice given his entire information set, he acts as if he completely ignored his own information. This is an "informational cascade" (Bikhchandani, Hirshleifer, Welch [1992]).

The angiogram example differs dramatically from the alprazolam example. In the angiogram case, the cardiologic community observes the choice of an angiogram—but it does not learn whether our cardiologist privately believed that the angiogram was appropriate or inappropriate. This private information is lost to the community. In the alprazolam example, this was not the case. By using a little less treatment, doctor A's skepticism could have been inferred by his colleagues, and would have subtly influenced their own treatment choices. Eventually, the medical community would have converged on the correct dosage.

A similar story holds for the cardiologist who is learning from past experience (or self-experimentation) with similar patients. In the continuous regime, the cardiologist can test various doses

---

[15] A similar argument applies also to the treatment choice of a single physician. When the dosage can be varied, physicians can run "small" experiments to optimize the dosage.

and arrive at the optimal dosage. However, the cardiologist is less likely to experiment with surgical interventions.

Cascades are not the only problem preventing information aggregation. If physicians need to pay to find out more about alternatives to angiograms (e.g., spend the time to research health journal publications), they may not find it *ex-ante* worth their while to invest this time. After all, the likelihood of finding useful information may be too small for any individual.[16] However, we could have given the cardiologist information for free in our example and it would still be ignored. Once some bad news about an alternative procedure has emerged, none may continue research into such procedures. And even when faced with doubt, physicians cannot learn the correct choice by perturbing their treatment slightly in order to observe the outcome.[17] After all, the physician's choice is only whether or not to administer a treatment, not how much. In sum, it can become too expensive (or ethically inappropriate) for an individual physician to acquire information and/or experiment with alternative treatment methods; such a physician will then voluntarily copy the treatment choices of his peers, and the medical community will fail to receive any new information from this physician's perceptions and actions that the community could aggregate to come up with a better consensus in the future.

Although the intuition behind these two scenarios is fairly obvious, once explained, the intrinsic difference of discrete choice decisions can have enormous implications for public health policy. First, treatment choice is likely to converge towards a localized choice with incredible speed—all it takes is the majority of one opinion to sway everyone else. But this choice could be the wrong choice, and once every physician is in a cascade, there is no obvious mechanism to correct it and restart the aggregation of knowledge that experimentation and learning can provide. Second, it does not even take a majority for a community to settle on the wrong standard. Physicians choose procedures in sequence. If two cardiologists perform an angiogram, the third cardiologist would also perform it, even if he privately believed it was not the right treatment choice. (Because the first two physicians were informed, *on average* but not always, the cascade will focus on the right choice.) So would the doctor faced with the fourth, fifth, sixth—and one-hundredth—choice. Even if *only* the first two physicians happen to have privately believed that angiograms are appropriate, the entire community of thousands of physicians can end up administering them. It is the "tyranny" of the earlier choice—which in many cases is the status quo.

---

[16] These are "search" costs (Lippman and McCall, 1977).
[17] More applicable for penicillin and other threshold drugs. Perhaps worse for potentially fatal illnesses.

In sum, our paper adopts Phelps and Mooney's view that localized learning plays a major role in mutual imitation among doctors. It shares many of the implications of the Phelps and Mooney model— e.g., the presence of welfare losses due to slow learning, the need for micro-based (not macro-based) empirical studies of the transmission of information. But it is unique in at least three aspects:

(1) Every patient may be subjected to poor treatment choice, even over a long time span;

(2) Every member of the medical community may recognize that common opinion could potentially be based on the wrong choice, because it can be based on few independent assessments; in the extreme, solely on the first two doctors' choices. Still, no physician will find it in her interest to deviate and try another treatment choice. Aside from the legal repercussions if the alternative treatment were to fail (which serves to reinforce cascades), the information contained in the collective cascade is still more informed than the individual physician's information. After all, two opinions outweigh one opinion. Again, every physician is herding rationally, trying to make the best decision given the information at hand;

(3) The public release of an informational study can potentially overthrow a cascade, even if many physicians had followed it for a long time. (This is not the case in Phelps and Mooney, where physicians eventually aggregate their information and end up with the best consensus view. One study is unlikely to shift their perception.) Physicians realize that the fact that millions of angiograms are performed does not mean that thousands of physicians pooled their information to come up with a best treatment choice. Instead, it may have been the tyranny of the initial procedure, the first few physicians' information that sets every physician into a particular equilibrium. Therefore, just a little bit of extra information---worth as much as, say, three individual physicians' opinions---could potentially sway everyone to change treatment choice.

Each treatment has two outcomes – cured or not cured. (We could generalize this to multiple finite outcomes but that would only make the notation messy without yielding any additional insight.) Thus, angioplasty is one treatment; prescribing a certain dosage of blood thinner is another treatment, etc.

DISCRETE TREATMENTS

We model each discrete treatment choice **i** as a binomial distribution with **unknown** probability p_i of curing a patient and $(1 - p_i)$ of not curing her. The distribution of p_i is Beta (see De Groot pp. 40) with parameters α_i and β_i. The true value of p_i is unknown and is a point in the unit interval. The expected value of p_i is:

$$q_i = \frac{\alpha_i}{\alpha_i + \beta_i} \tag{1}$$

The (expected) probability that the first patient treated with i is cured is $\alpha_i/(\alpha_i + \beta_i)$. After observing the outcome on the first patient we revise our beliefs about the efficacy of i and proceed. The Beta distribution has a very simple updating rule (see De Groot pp. 160). Suppose that treatment i has been tried on $n = c + d$ patients, of which **c** patients were cured and **d** patients were not. Then the posterior distribution of p_i is Beta with parameters $\boldsymbol{\alpha_i' = \alpha_i + c}$ and $\boldsymbol{\beta_i' = \beta_i + d}$. Thus,

$$q_i(c,d) = (\alpha_i + c)/(\alpha_i + \beta_i + c + d) = (\alpha_i + c)/(\alpha_i + \beta_i + n) = \alpha_i'/(\alpha_i' + \beta_i') \tag{2}$$

is the expectation that the $(n+1)^{st}$ patient undergoing treatment i will be cured.

The variance of a Beta distribution with parameters α_i and β_i is

$$\text{Variance of } p_i = \frac{\alpha_i \beta_i}{(\alpha_i + \beta_i)^2 (\alpha_i + \beta_i + 1)} \tag{3}$$

Observe that with experience, the parameters increase. That is, $\alpha_i \leq \alpha_i + c = \alpha_i'$ and $\beta_i \leq \beta_i + d = \beta_i'$. Thus, as the number of trials n=c+d increases, the variance of p_i decreases. This is to the expected – the greater one's experience with a treatment the more one knows about its likely efficacy.

We do not differentiate between a doctor and/or the hospital/medical group/HMO (whatever we define our geographical unit to be) that the doctor is a part of. This decision-making group is referred to as a unit. A unit is defined by observability of outcomes – every doctor within a unit observes all outcomes of treatments undertaken at that unit. After n trials on treatment i, everyone at the unit observes the outcome and updates their beliefs accordingly. We assume that there is no experimentation – every trial (patient) is treated with the best available treatment.

Suppose that two distinct treatments, A and B, are available for a disease. For simplicity, assume that treatments are uncorrelated. Thus, experience with A is uninformative about the distribution of p_B. Initially, the two treatments are thought to be better than the conventional treatment, which has success probability r. That is,

$$r < q_i = \frac{\alpha_i}{\alpha_i + \beta_i} \qquad i = A, B. \tag{4}$$

Thus, we assume that r, probability of success of the conventional treatment, is known. This is reasonable as, presumably, there is extensive experience with the conventional treatment. (Recall from equation 3 that the variance of a Beta distribution goes to zero as the number of trials increases.) Equation (4) says that the new treatments A and B are expected to be better than the conventional treatment; if this were not the case then they would not be tried.

13

Before proceeding to a general analysis, a simple argument demonstrates how (i) a unit may settle on an inferior treatment choice and (ii) different units may choose different treatments. Throughout we assume that a unit selects the treatment that it believes to be the best at each stage. That is, it is considered unethical to experiment with a treatment which is expected (but not certain) to be inferior to other choices.

Suppose that the 1st patient at a unit is treated with A, i.e., $q_A \geq q_B$. Then if she is cured, we have $q_A(1,0) > q_B(0,0)$, and thus the 2nd patient will also be treated with A, and so on. If, at some point, it turns out that not many patients are cured after undergoing A, i.e., $c/(c+d)$ is sufficiently small that $q_A(c,d) < q_B(0,0)$, then and only then will B be tried. Ultimately, as $c+d \to \infty$ we have $q_A(c,d) \to p_A$. Thus, suppose that $q_A < p_A < p_B$ and the unit picked A first, because $q_B \leq q_A$. Then quite likely it would never (seriously) try B even though it is superior. Thus, when a unit under-estimates the efficacy of both new treatments, it may settle on an inferior treatment. And if some other unit picked B first (because it believed that $q_B \geq q_A$) it will stick with B and we will see variation in treatment choices across two different units. Moreover, if there are more than two new treatments, the variation across several units will be even greater.

First, we prove that a unit may converge on the wrong treatment. Let q_A', q_B' summarize the updated beliefs of the unit about the two treatments after some experience.[18] At any given time, a unit is in one of two cases depending on its history of experience, as reflected in beliefs about A and B. Of course, the unit may go from one case to the other and back, but ultimately, as it settles on a treatment choice, it must be in Case I.

CASE I:[19] $q_i' \geq q_j'$, $p_i > q_j'$. If $q_A' > q_B'$, say, then the unit picks A next. As long as posterior beliefs about A remain more favorable than beliefs about B then B is never tried and the unit settles on A. (A necessary condition is $p_A > q_B'$.) However, it may well be that $p_B > p_A$ and the unit settles on an inferior treatment.

It is possible that further experience with A is sufficiently negative that the new posterior belief about A (q_A^n) falls below q_B'. In this event, B will be tried and if $p_B > q_A^n$ and $p_A > p_B$ one may end up with the wrong choice again.

CASE II: $q_i' \geq q_j'$, $p_i < q_j'$. If $q_A' > q_B'$, say, then the unit picks A next. But as $p_A < q_B'$, sooner or later experience with A will be sufficiently negative ($q_A^n < q_B'$) that B will be selected. At that point, if $p_B > q_A^n$ then Case I applies. If, instead, $p_B < q_A^n$ then Case II applies. In the latter case, at a later point in time A will be used again as the posterior beliefs on B will become more pessimistic than q_A^n, the beliefs about A. Thus, in Case II the unit keeps switching between the two treatments until it goes back into Case I.

Variation in treatment choices by across units

Because a unit will not necessarily settle on the treatment choice with a higher probability of

success, two different units may select different treatments only one of which is the more effective one.

Two other factors contribute to greater variation across units. First, the initial assessments (q_A, q_B) of

---

[18] That is, if c_A, d_A is the experience with treatment A, then $q_A' = \alpha_A'/(\alpha_A' + \beta_B') = (\alpha_A + c_A)/(\alpha_A + c_A + \beta_A + d_A)$, etc.
[19] The unit knows the values of q_A, q_B only and not p_A or p_B. Therefore, the unit does not know whether it is in Case I or II.

two units may differ; one unit may believe that A is much more likely than B to be effective and the other unit's beliefs may be the opposite. This would lead one unit to experiment with A and the other with B. If these initial experiments are positive for the two units, they will stick to their initial choices. A second factor resulting in greater variation in treatment choices is that there may be more than two new treatments to consider, only one of which is the most effective. In this event, variations across different units would be that much greater.

The effect of observing other units' treatment choices

Suppose that unit 1 settles on treatment A after a period of trials with A and B (i.e., posterior beliefs satisfy $q_A^1 > q_B^1$). Furthermore, posterior beliefs are sufficiently sharply defined that the likelihood of switching back to B, based only on the unit's further experience with A, is negligible. At that point unit 1 observes unit 2's choice. How will this effect 1's beliefs and future actions?

If unit 2's choice is also A then this confirming evidence increases [decreases] unit 1's posterior beliefs about the efficacy of A [B]. Unit 1 will continue to choose A.

If, on the other hand, 2's choice is B then 1's posterior beliefs about A [B] decreases (increases). However, if q_A^1 is sufficiently greater than q_B^1 then A will continue to choose A.[20] Thus, differences in treatment choices may persist if all that unit 1 observes is what unit 2 is choosing after n trials (rather than the history of outcomes that 2 observed).

---

[20] An exact calculation using the Beta distribution makes this clear.

IV. Evidence[21]

In order to study the difference in the use of continuous versus discrete medical interventions, we limit both datasets to include only those patients who were diagnosed with having a cardiac incident (either hypertension, acute myocardial infraction, or coronary atherscherosis).[22] This clinical determination was made using standard 5-digit ICD-9-CM diagnostic codes. Our choice of studying those patients who presented themselves with a cardiac incident is not accidental: Heart attacks remain the leading cause of death in the United States, and the treatment of coronary artery disease proceeds along *both* surgical and well as pharmaceutical dimensions.[23] Surgical interventions include Coronary Artery Bypass Graft (CABG) where one or more coronary arteries are bypassed using grafts, and

[21] Measurement Issues: Most of the literature uses the Coefficient of Variation (CV) statistic to measure variation. Examining the standard deviations (or variance) across procedures is one way to proceed, however it is only meaningful if the two interventions are perfect substitutes (have linear isoquants) as potential treatments. This will only happen for a very small and therefore highly select group of patients. The CV is a better alternative because it is scale-free and hence allows for comparisons of variation across treatments with very different average rates of use. However we note that the CV is sensitive to the direction of the defined intervention: the CV for the use of ACE inhibitors is different from the CV for *not* using ACE inhibitors. This is a peculiar result given the fact that the standard deviation for the use of prescribing an ACE inhibitor is the same as the standard deviation of not prescribing the same medication. To overcome this limitation we propose the use of a new statistic, which we refer to as the Modified Coefficient of Variation (MCV). The MCV is defined as $\sigma/(\sqrt{\mu(1-\mu)})$ and is similar to the conventionally used coefficient of variation (notice that the MCV=CV. $\sqrt{\mu}/\sqrt{(1-\mu)}$. The advantage of our statistic is that it is directionally insensitive (unlike CV) and allows for "variation" to be determined more as a function of the standard deviation, than the mean rate of usage. Another limitation of CV is that (by construction) it tends to produce low estimates of variation for procedures with a high average rate of use. The MCV attaches more weight to the standard deviation, but retains the advantage of being scale free. Other researchers have proposed the use of the Extremal Quotient (EQ), or the ratio of the highest to the lowest rates of use [Chassin et al. (1986)]. However, is straightforward to show that the EQ has an expected value of infinity and therefore is not particularly meaningful for the purpose of formal inference. It is extremely difficult to derive the asymptotic distribution of the CV or MCV using asymptotic methods. The existing literature only reports the point estimate and ignores the sampling distribution. While it is tempting to derive these distributions by the direct application of the Slutsky theorem, this approach is incorrect because it ignores the fact that the standard deviation and mean are related. Using asymptotic theory we have found that: $\sqrt{n}(\hat{MCV}^2 - MCV^2) \xrightarrow{L} N\left(0, \sum \frac{w_i P_i(1-P_i)(1-2P_i)^2}{\mu^2(1-\mu)^2} + g(E(p-\mu)^2, E(p-\mu)^3)\right)$, where *g(.)* is a complicated function of the higher order moments of *p* (the average rate for each hospital). This intractable expression for the asymptotic variance can (potentially) be approximated by the application of bootstrap methods and this is an avenue for future research in the area. In this paper, we circumvent many of these problems by reporting kernel density estimates of the entire distribution of hospital fixed effects.

[22] This definition will excludes those with heart valve disorders, carditis, chest pain (excluding angina pectoris), pulmonary or cerebral embolisms, conduction , ventricular fibrillation, aneurysms and dysrhythmias). A detailed list of the diagnoses subsumed under AMI and coronary atherscherosis, as well as the programs to map these diagnoses to broader categories is available from the authors on request.

[23] It is not meaningful to compare variation in the use of a new surgical technique for appendectomy to a new drug for hypertension because the relevant clinical populations are entirely different. The treatment of coronary artery disease on the other hand, is uniquely suited for our analysis.

variants of Percutaneous Transluminal Coronary Angioplasty (PTCA) where a balloon or laser on the end of a catherter is used to clear the buildup of atherosclerotic plaque inside the arteries. Pharmaceutical interventions include the prescription of thrombolytics (which reduce the ability of blood to clot, and therefore improve blood flow), beta-blockers and ACE inhibitors which reduce the demand on the heart.

We present empirical evidence in support of our model at three levels: First, in section IV.A we present results from multiple panels of the National Hospital Ambulatory Medical Care Survey (NHAMCS, 1995-99). With these data it is possible to test the key prediction of our theoretical model: that there should be more variation in the use of surgical or discrete procedures than in the use of drugs or continuous procedures. If the model is correct, providers can experiment with the dosage and arrive at a consensus on the correct usage. Therefore, we expect to see a "preferred modality" in the use of continuous interventions, which should be less pronounced, or non-existent in the use of discrete interventions. To the extent that it is possible we also use the data from the *Dartmouth Atlas* to corroborate the results from the NHAMCS. The *Atlas* data have the advantage of being comprised of 306 HRRs but they are only representative for the Medicare population.

The NHAMCS data provide a rich source of information for the purpose of our analysis however, they are limited in that we cannot compare differences in the use of drugs and surgical interventions across space with these data. This is because the sampling frame for the NHAMCS only produces a nationally representative sample of hospitals; these hospitals are scattered all over the US and there is not enough information to study the behavior of adjacent hospitals. To study this part of our models predictions, in section IV.B we turn to the data from the Dartmouth Atlas of Cardiovascular Care in order to study the spatial implications of our model. If correct, we should expect that contiguous hospitals (or HRRs/CASAs) adopt similar levels of intensity in the usage of continuous interventions but not in the use of discrete interventions. We also use these data to reinforce the results obtained from the NHAMCS data but studying variation in the use of discrete and continuous diagnostic tests (with the NHAMCS we only studied procedures). The only limitation of the Atlas data is that it restricts the analysis to the Medicare (over 65) population, whereas the NHAMCS data samples all ambulatory care visits regardless of age. Finally, in IV.C we discuss evidence in favor of the public-policy implications of our model—that new information can break an existing cascade.

A. Evidence from the NHAMCS (1995-99).

The NHMACS is an annual survey and collects data on the provision of ambulatory care services in hospital emergency and outpatient departments, and is one of the only nationally representative surveys to collect data on the prescription behavior of hospitals. The survey is designed to be a national sample of visits to the emergency departments and outpatient departments of noninstitutional general and short-stay hospitals, excluding Federal, military, and Veterans Administration hospitals. Hospital staffs are instructed to complete Patient Record forms for a random sample of patient visits during a randomly assigned 4-week reporting period. The NHAMCS collects data on demographic characteristics of patients, expected sources of payment, patients' complaints, physicians' diagnoses, diagnostic/screening services, procedures, medication therapy, disposition, types of health care professionals seen, causes of injury where applicable, and certain characteristics of the hospital. For the purpose of our analysis the NHAMCS data includes a unique hospital identifier, allowing us to create hospital-specific rates for the use of different medications. We restrict our analysis to those patients who were admitted for a cardiovascular incident (as defined through the use of ICD-9-CM codes, [see Department of Health and Human Services (1980)]). Through the use of these codes we restrict our sample to those patients who have a diagnoses of hypertention.[24] The NHAMCS data are a rich source of data on the type of medications provided, and provide the names of up to six different prescription drugs. We standardize these drugs into broad classes of medications by using their National Drug Code (NDC) classifications.[25]

From 1995-1999 there were 12,509 cases of hypertension in the NHAMCS. However, most of these cases occurred in hospitals where there were less that 30 diagnoses of hypertension made during the survey period. To avoid problems of inference made with small sample sizes, we restricted ourselves to those only those cases were at least 30 diagnoses of hypertension were made for a given hospital. This restriction yielded a sample of 3,604 patients and 67 hospitals; approximately 40 percent of the sample

---

[24] In coding this variable we only relied on ICD-9-CM codes of 401-405 and excluded those with hypertention as a result of pregnancy and childbirth. Note that ischemic heart disease or AMIs are not included in our definition.

[25] The NDC System was originally established as part of an out-of-hospital drug reimbursement program under Medicare. The NDC serves as a universal product identifier for human drugs and is administered by the Food and Drug Administration (FDA). For more information on the mapping algorithm that converts drug names (both with tradenames and generic) to NDC codes see http://www.fda.gov/cder/ndc/. We note that the NHAMCS is available for 1993 and 1994 but we have not used the data from these years because the NDC codes used in the survey changed between 1994-95 rendering comparison with earlier years difficult. For example, in 1993 and 1994 there were no separate codes for calcium channel blockers, ACE inhibitors, or beta blockers. The NDC contains separate codes for whether Thrombolytics or Carbonic Anhydrase Inhibitors were administered but since these are non-standard interventions for hypertension we chose to drop them from our reported results.

were over the age of 65.[26] For each pharmaceutical intervention we estimate a fixed-effects regression (with hospital fixed-effects) using age, race, and gender as covariates. We allow for a non-parametric specification whereby the effect the race and gender were allowed to vary by age. We then recovered the hospital fixed-effects for our analysis. Specifically, for each patient i, and intervention j we estimate:

(1) $T_{ij} = X_i \beta + \delta_k + e_{ij}$

Here, T_{ij} is an indicator variable for whether or not a treatment j was administered to patient i, X_i is a vector of basic demographics for patient i, β is a vector of parameters giving the effect of each covariate on the probability of receiving treatment j. δ is a vector of hospital specific indicator variables (the fixed effects) and e_{ij} is the unexplained residual in the treatment equation. Note that by construction these fixed-effects will be mean zero (we have normed them to have a mean of one). Because of the extremely crude covariates that were used for underlying patient severity and demographics our estimates of the hospital fixed effects are biased towards being larger in absolute value than their true values because we believe that they will also be capturing the effect of many omitted variables such as income, health insurance status of the populations they serve and differences in illness patterns.

In Table 1 we present summary statistics for the NHAMCS data. Note that while there is variation in the use of the different drugs, the variation is not as substantial as what was observed for the surgical interventions previously discussed. In Figure 4 we display the entire distribution of these fixed effects and note that without exception the usage rates lie within -.1 to .1 times the average rate (of zero). We superimpose a normal distribution to facilitate comparison. The variation is smallest in the use of antianginal agents, antihypertensives and alphablockers, and is larger in the use of Diuretics, Calcium Channel Blockers and ACE Inhibitors.

Note to co-authors: are results are actually a little stronger that you can see. They are contaminated by sampling "noise" because of the inclusion of some very small hospitals. Im working of using Bayes Shrinkage methods to fix this problem.

B. Evidence from the HCUP

---

[26] Our results are robust to relaxing this sample size criteria to 20 observations per hospital but are unsurprisingly not robust to sample sizes. We have deliberately not calculated prevalence rates of hypertension as the sample is comprised of patient visits and that is not the appropriate denominator to use. We have also limited out sample to those hospitals that saw at least 30 patients over the age of 65 for hypertension. Our results were identical to those obtained for all ages.

Whereas the NHAMCS provides a rich set of data on the prescribing behavior of physicians, information on the use of discrete surgical procedures is extremely scarce primarily because the NHAMCS is a survey of ambulatory care. Therefore, the use of relatively common surgical procedures such as CABG or angioplasty is not captured in these data. To obtain data on such procedures we used the 1996 National Inpatient Sample (NIS) from the Heathcare Cost and Use Project (HCUP). The NIS-HCUP contains a sample of hospitals selected from the 19 states in the U.S., and includes the population of discharges (about 7 million) at each selected hospital. The sample of hospitals in the NIS (about 900-1,000 each year) is selected to reflect the national characteristics of community hospitals. Each patient record on the NIS-HCUP includes information on basic demographics of the patient, expected primary and secondary payers, total charges, principal and secondary diagnoses, and principal and secondary procedures.

Kernel Density estimates of the distribution of fixed effects from the HCUP data are presented in Figures 5A (patients with AMI) and 5B (patients diagnosed with coronary atherslerosis). The results are consistent with our model: for both conditions the variation in the use of the two surgical interventions (CABG and PTCA) are much greater than those obtained for the use of pharmaceutical intervention in the NHAMCS data. Furthermore, we see that the variation in CABG treatment rates is less than that of PTCA. This is consistent with our hypothesis that with CABG a surgeon can experiment with the number of grafts.

C. Evidence from the Dartmouth Atlas of Cardiovascular Care

In this section we discuss the spatial implication of our learning model. We expect adjacent regions to have similar practice styles in the use of continuous interventions and not necessarily in the use of discrete interventions. It is only possible to test these implications of our models using the Atlas data at the HRR level as the NHAMCS does not sample adjacent hospitals. In Figure 4a we illustrate the enormous geographic variation in a relatively "standard" procedure—percutaneous coronary interventions. In 1996 over 200,000 of these procedures were conducted with an average rate of 7.5 per 1000 Medicare enrollees. As in previous figures the data have been standardized for demographics and illness patterns and are reported at the HRR level. Note how in Texas, Pennsylvania and California the ratio of rates (to the US average) can vary drastically even across adjacent HRRs. In fact, adjusted rates differ by a factor of eight. In Figure 4b we study the rates of Coronary Artery Bypass Grafting (CABG) relative to the US average (of 6.5/1000 Medicare beneficiaries). There was less variation in the usage of

CABG relative to PTCA but it was still substantial. On the basis of our theatrical model we would a priori expect less variation in the use of CABG because the physician can choose (experiment) the number of grafts—1,2,3 or 4. Therefore, while CABG remains a surgical intervention, it does have characteristics that make it a quasi-continuous variable. In Figure 4c we study the rates of aortic valve replacement—a procedure that is designed to overcome stenosis (a narrowing of the valves resulting in pressure overload) or insufficiency (a valve that does not close completely, thereby causing excessive pressure to buildup on the heart because the blood backwashes). The variation in aortic valve replacement is substantial—controlling for age, race, sex and illness there is still a four fold different across HRR regions. We believe that the evidence summarized in Figures 4a-4c is consistent with the spatial implications of our model for discrete interventions. We now turn to variations for continuous interventions.

Figures 5a-5c present data on the use of drugs that are dispensed for the treatment of AMI.[27] In Figure 5a we note how there is a lot less variation relative to surgical interventions in the use of aspirin at hospital discharge (there was even more agreement in the use of aspirin therapy during hospitalization). More importantly, we also note that there is a clustering effect in the degree to which adjacent HRRs use aspirin therapy. This prediction is true in both the PM model as well as ours. Similar (although not identical results) are obtained for the use of ACE inhibitors (Figure 5b) and Beta Blockers (Figure 5c).

C. Can Information Shatter a Cascade?

Empirical evidence is consistent with this conjecture; physicians have been found to dramatically alter their practice styles after the release of information that alerts them to the extent of variation. Release of information has reduced the use rates for hysterectomies (Dyck et al., 1977), tonsillectomies (Wennberg et al, 1977) and pediatric admissions, prostatic surgeries, hysterectomies and back surgeries (American Medical Association, 1986). None of these studies incorporates a formal experimental design and we alert the reader to this deficiency. Our discussion of the evidence is only meant to suggest that the data are generally (not formally) supportive of the claims of our model.[28]

---

[27] Each figure presents the fraction of patients who were hospitalized for AMI in 1994-95 and were "ideal" candidates for the treatments considered. The determination of which patients were ideally suited for a particular intervention was made by reviewing clinical data. The methods used for this determination are discussed at length in O'Conner et al (1999).

[28] A comprehensive review of the literature on the effects of feedback information in clinical practice is Mugford et al (1991).

In the classic Wennberg study, feedback data were provided by the Vermont Medical Society to physicians performing tonsillectomies in 13 hospital service areas in Vermont. After clinicians were shown the result of the existing variation in tonsillectomy rates across the area it was found that there was a 32 percent decline in tonsillectomy rates relative to the US "control" group. Before the exchange of information there was a 13 fold difference in per capita rates between the highest and lowest HSAs. After the intervention, it fell to 4.5. It is also interesting to note that the reduction took place rapidly— most of the change occurred in the year immediately following the feedback. The Wennberg study only demonstrates that feedback can change behavior—there is no direct evidence that the intervention simultaneously caused better care.

Dyck et al (1977) study hysterectomy rates in seven hospitals through a prospective committee review. They find that after the review committee made its finding known the overall rates of hysterectomies fell. Additonally, the number done for "unacceptable" reasons fell as well.

Evidence in favor of the hypothesis that physicians may actually be practicing in the realm of iatroepidemics comes from a study conducted by the Northern New England Cardiovascular Disease Study Group (NNECVDSG) based at Dartmouth Medical School. The NNECVDSG designed an intervention whereby all 23 cardiothorasic surgeons practicing CABG surgery in Maine, Vermont and New Hampshire were given feedback of outcome data, training in continuous quality management and site visits to other medical centers. In comparing (post intervention) observed and expected hospital mortality rates they found a 24% reduction in hospital mortality as a result of the informational intervention. In Figure 6 we illustrate the key findings from this study. The mortality results are actually understated as there is evidence that the postintervention group (n=6,488) were actually sicker (as measured by age, and the rates of diabetes, chronic obstructive pulmonary disease and peripheral vascular disease) than the preintervention group (n=6,638). The NNECVDSG study is supportive of our prediction that information can change existing behaviors because it provides an insight into the "black-box" of physician decision making. The intervention was comprised of 1) clinicians receiving anonymous risk-adjusted outcome data for himself, his affiliated medical center and the region 2) a series of site visits in which a cardiac surgeon and a perfusionist from each center undertook a round-robin tous of the other centers. The teams observed the entire treatment process at each site, from the initial cardiac catherterization conference to surgery and post-operative care. Most importantly, they were encouraged to think about similarities and dissimilaries with the procedures followed at home.

The three studies discussed above all provide evidence in favor of our general hypothesis that 1) information can break cascades and 2) that physicians may have locally converged at the incorrect choice of treatment. However, we should also observe that (1) is not true in the case of continuous interventions. This evidence is more difficult to find, in part by construction. Consider the following thought experiment: assume that our model is correct. If so, then physicians have converged to the right dosage. If this is true it would be unnecessary to conduct an intervention that seeks to change prescribing behavior. Therefore, it is perhaps unsurprising that we found it difficult to find evidence of such interventions. In fact, we have been able to uncover only one example of an intervention designed to test the effect of new information on prescribing behavior. Consistent with our model, this study did not find a change in prescribing behavior in the post-intervention period. In particular, Schaffer et al (1983) found that neither mailing physicians a brochure on "correct" prescription behavior nor visits by drug educators affected prescribing behavior for antibiotics.

V. Conclusions

In this paper we have formalized the role of technology diffusion in the market for health care and offer a theoretical explanation for the well-documented phenomenon of geographic variation in medical practice. A simple Bayesian learning model can demonstrate that physicians imitate one another because they learn information from the treatment choices of their colleagues, and closer colleagues provide more information. Therefore over time, the market will converge at the optimal dosage. However, in our paper we demonstrate that such learning can cease altogether in situations in which the physician's choice is discrete, i.e., in which the physician must select one from a select small number of choices. Our model offers a rich set of theoretical predictions, which we believe, will motivate a large empirical research program. We predict that discrete choice treatment methods suffer far more from localized variation than continuous "dosage" treatment choices. Second, we argue that it is quite possible that the community focuses on the wrong treatment choice and stays with it, even in the absence of legal repercussions and despite perfectly rational behavior by every physician. And third, it argues that even small-scale medical studies can have a dramatic effect in overthrowing long-standing cascades and providing better treatment choices.

While it is beyond the scope of this paper (and current data!) to test all the predictions of our model, we find broad empirical support for our model using data from the NHAMCS and

HCUP data. Specifically, we do find that for patients who experienced a cardiac incident there wa s much greater variation in the use of surgical interventions than in the use of medications.

The implications of our model for public-policy in affecting (or explicitly improving) the production of health care are immense. Our model argues that in the case of discrete surgical treatments, or at the extensive margin of prescribing a new drug, tremendous cost savings can be realized through the large information dissemination programs that educate physicians about the appropriateness of the a given indication. Therefore, institutions such as the AHRQ or specialty organizations like the American College of Obstetricians and Gynecologists (ACOG) can dramatically improve the quality of medical care through the use of briefs that discuss the empirical evidence in favor of a particular technique. An example of such an informational intervention is the issuance of guidelines by the American College of Cardiology and the American Heart Association that recommends that non-imaging stress tests (as opposed to coronary angiography) should be the preferred modality in the diagnoses of coronary artery disease.

Appendix A: Definitions and Terminology[29]

ACE inhibitor: A group of antihypertensive medications that work by inhibiting an enzyme (angiotensin converting enzyme) that is important in the regulation of blood pressure. Studies have also indicated that it may help prevent or slow the progression of kidney disease in patients with diabetes. Examples include: captopril, ramipril, enalapril, losartan potassium, bepridil and lisinopril.

Angioplasty, balloon: Use of a balloon catheter for dilatation of an occluded artery. It is used in treatment of arterial occlusive diseases, including renal artery stenosis and arterial occlusions in the leg. For the specific technique of balloon dilatation in coronary arteries, angioplasty, transluminal, percutaneous coronary is available.

Angioplasty, laser: A technique utilizing a laser coupled to a catheter which is used in the dilatation of occluded blood vessels. This includes laser thermal angioplasty where the laser energy heats up a metal tip, and direct laser angioplasty where the laser energy directly ablates the occlusion. One form of the latter approach uses an excimer laser which creates microscopically precise cuts without thermal injury. When laser angioplasty is performed in combination with balloon angioplasty it is called laser-assisted balloon angioplasty.

Antihypertensive: An agent that reduces high blood pressure.

Atherosclerosis: The progressive narrowing and hardening of the arteries over time because of the buildup of cholesterol and fatty material within a blood vessel.

Bypass graft: An alternative blood vessel that is created by a surgeon to reroute blood flow. Grafts may be synthetic (dacryon) or autologous (a vein from the patients own leg used as a substitute for the diseased vessel).

Beta-blocker: A large group of medications that act to block specific receptors in the nervous system. The effect of beta-blockade results in slowing of the heart rate, reduction in blood pressure and reduced anxiety. Beta-blockers are used in the treatment of angina, heart arrhythmias, high blood pressure, mitral valve prolapse and other conditions.

CABG: Coronary Artery Bypass Grafting Surgery. Surgery in which a vein is harvested from the leg, or an artery is harvested from the internal mammary artery to bypass the coronary artery which has narrowed because of the buildup of atherosclerotic plaque.

Calcium Channel blocker: A drug that blocks the entry of calcium into cells, thereby preventing cell death and loss of function caused by excess calcium. Calcium channel blockers are used primarily in the treatment of certain heart conditions and stroke, but are being studied as potential treatments for Alzheimer's disease.

---

[29] We have defined these terms in using information from consultations with physician colleagues and the online medical dictionary available at http://medical-dictionary.com

Echocardiogram: A test which uses high-frequency sound waves to image the heart and surrounding tissues.

PTCA: percutaneous coronary transluminal angioplasty: Dilatation of an occluded coronary artery (or arteries) by means of a balloon catheter to restore myocardial blood supply.

Stress Echocardiogram: An echocardiogram that is performed after a period of physical exertion. Chemical stimulation of the heart (to mimic exertion) is used in some cases where physical activity is not possible. In some cases, exertion may manifest a cardiac abnormality not obvious during echocardiography in the resting heart.

Thrombolytic: Medications that dissolve blot clots (for example streptokinase, tissue plasminogen activator or TPA and urokinase).

Hospital Service Area: Hospital Service Areas (HSAs) represent local health care markets for community-based inpatient care. The definitions of HSAs used in the 1996 edition of the Atlas were retained in the 1999 edition. HSAs were originally defined in three steps using 1993 provider files and 1992-93 utilization data. First, all acute care hospitals in the 50 states and the District of Columbia were identified from the American Hospital Association Annual Survey of Hospitals and the Medicare Provider of Services files and assigned to a location within a town or city. The list of towns or cities with at least one acute care hospital (N=3,953) defined the maximum number of possible HSAs. Second, all 1992 and 1993 acute care hospitalizations of the Medicare population were analyzed according to ZIP Code to determine the proportion of residents' hospital stays that occurred in each of the 3,953 candidate HSAs. ZIP Codes were initially assigned to the HSA where the greatest proportion (plurality) of residents were hospitalized. Approximately 500 of the candidate HSAs did not qualify as independent HSAs because the plurality of patients resident in those HSAs were hospitalized in other HSAs. The third step required visual examination of the ZIP Codes used to define each HSA. Maps of ZIP Code boundaries were made using files obtained from Geographic Data Technologies (GDT) and each HSA's component ZIP Codes were examined. In order to achieve contiguity of the component ZIP Codes for each HSA, "island" ZIP Codes were reassigned to the enclosing HSA, and/or HSAs were grouped into larger HSAs. This process resulted in the identification of 3,436 HSAs, ranging in total 1996 population from 604 (Turtle Lake, North Dakota) to 3,067,356 (Houston) in the 1999 edition of the Atlas. Intuitively, one may think of HSAs as representing the geographic level at which "front end" services such as diagnoses are received.

Hospital Referral Region: Hospital service areas make clear the patterns of use of local hospitals. A significant proportion of care, however, is provided by referral hospitals that serve a larger region. Hospital referral regions were defined in the Atlas by documenting where patients were referred for major cardiovascular surgical procedures and for neurosurgery. Each hospital service area was examined to determine where most of its residents went for these services. The result was the aggregation of the 3,436 hospital service areas into 306 hospital referral regions. Each hospital referral region had at least one city where both major cardiovascular surgical procedures and neurosurgery were performed. Maps were used to make sure that the small number of "orphan" hospital service areas - those surrounded by hospital service areas allocated to a different hospital referral region - were reassigned, in almost all cases, to ensure geographic contiguity. Hospital referral regions were pooled with neighbors if their populations were less than 120,000 or if less than 65% of their residents' hospitalizations occurred within the region. Hospital referral regions were named for the hospital service area containing the referral hospital or hospitals most often used by residents of the region. The regions sometimes cross state boundaries. Intuitively, one may think of HRRs as representing the geographic level at which "back end" services such as invasive surgery are received.

---

[30] We have duplicated the definitions used in CECS (1999a, 1999b). For further details on the construction methods see http://www.dartmouthatlas.org/99US/toc8.php.

REFERENCES

American Medical Association (1986). *Confronting Regional Variations: The Maine Approach*, Department of Health Care Review.

American Medical Association (1990). *Current Procedural Terminology (CPT)*, 4th ed., Chicago: The American Medical Association.

Banerjee, A.V. (1992), "A Simple Model of Herd Behavior", *Quarterly Journal of Economics*. 107(3): 797-817.

Bikhchandani S., D Hirshleifer and I. Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Exchange as International Cascades", *Journal of Political Economy* 100(5):992-1026.

Center for Evaluative Clinical Sciences, Dartmouth Medical School (1999a), *The Dartmouth Atlas of Health Care*. Hanover, NH

Center for Evaluative Clinical Sciences, Dartmouth Medical School (1999b), *The Dartmouth Atlas of Cardiovascular Care*. Hanover, NH

Department of Health and Human Services (1980), *International Classification of Diseases (ICD-9-CM)*, 2nd ed., Washington, DC: U.S. Government Printing Office (PHS)-80-1260.

Detsky, A.S. (1989), "Are Clinical Trials a Cost-Effective Investment?", *Journal of the American Medical Association* 262(13):1795-1800.

Detsky, A.S. (1990), "Using Cost-Effectiveness Analysis to Improve the Efficiency of Allocating Funds to Clinical Trials", *Statistics in Medicine* 9:173-183.

Diehr, P., K.C. Cain, W. Kreuter, and S. Rosenkranz (1992), "Can Small Area Analysis Detect Variation in Surgery Rates? The Power of Small Area Variations Analysis", *Medical Care* 30(6):484-502.

Dranove, David (1995), "A Problem with Consumer Surplus Measures of the Cost of Practice Variations," *Journal of Health Economics* 14: 243-51.

Dyck, F.J., F.A. Murphy, J.K. Murphy, et al. (1977), "Effect of Surveillance on the number of Hysterectomies in the Province of Saskatchewan," *New England Journal of Medicine* (296): 1326.

Fisher, Elliott S., Wennberg, John E., Stukel, Therese A., Sharp, Sandra (1994), "Hospital Readmission Rates for Cohorts of Medicare Beneficiaries in Boston and New Haven" *New England Journal of Medicine* 331: 989-95.

Fleiss, J. (1984), *Statistical Methods for Rates and Proportions*, New York: Wiley and Sons, 1984, 2nd ed.

Glover, A.F. (1938), "The Incidence of Tonsillectomy in School Children", *Proceedings of the Royal Society of Medicine* 31:1219-1236.

Leape, L.L., R.E. Park, D.H. Solomon et al. (1990), "Does Inappropriate Use Explain Small Area Variations in the Use of Health Care Services?", *Journal of the American Medical Association* 1990 263(5):669-672.

Lippman, S.A. and J.J. McCall (1979) eds. *Studies in the Economics of Search*, Amsterdam: North Holland.

Lippman, SA, McCall J.J. (1976), "The Economics of Job Search: A Survey, Part 1", *Economic Inquiry* 14(2):155-189

McPherson K., P.M. Strong, A. Epstein and L. Jones, (1981), "Regional Variations in the Use of Common Surgical Procedures: Within and Between England and Wales, Canada, and the United States," *Social Science in Medicine* 15A:273-288.

McPherson K., J.E. Wennberg, O.B. Hovind and P. Clifford (1982), "Small-Area Variations in the Use of Common Surgical Procedures: An International Comparison of New England, England, and Norway," *New England Journal of Medicine* 307(21): 1310-1314.

Mugford, Miranda, Banfield, Philip, O'Hanlon Moira (1991), "Effects of Feedback of Information on Clinical Practice," *British Medical Journal* 303: 398-402.

O'Conner, Gerald, Quinton, Hebe B., Traven Neal D., Ramunno, Lawrence, Dodds, T. Andrew, Marciniak, Thomas A., and Wennberg, John E. (1999). "Geographic Variation in the Treatment of Acute Myocardial Infraction: The Cooperative Cardiovascular Project," *Journal of the American Medical Association* 281: 627-33.

Phelps, C.E. (1997), *Health Economics*, Reading, MA: Addison Wesley Longman, 1997.

Phelps, C.E. and C. Mooney (1993) "Variations in Medical Practice Use: Causes and Consequences", in *Competitive Approaches to Health Care Reform*, Richard J. Arnauld, Robert F. Rich and William White, eds. Washington, DC: The Urban Institute Press.

Phelps, C.E. and S.T. Parente (1990), "Priority Setting in Medical Technology and Medical Practice Assessment", *Medical Care* 29(8):703-723.

Phelps, C.E., C. Mooney, A.I. Mushlin and N.A.K. Perkins (1992), "Doctors Have Styles, and They Matter!" Department of Community and Preventive Medicine, University of Rochester, Rochester, NY.

Pritchard, RS, Fisher ES, Teno JM (1998), "Influence of Patient Preferences and Local Health System Characteristics on the Place of Death," *Journal of the American Geriatric Society* 46: 1242-50.

Robin, Eugene D. (1984). *Matters of Life and Death: Risks vs. Benefits of Medical Care.* (New York: Freeman).

Roos, N.P., L.L. Roos, and P.D. Henteleff (1977), "Elective Surgical Rates -- Do High Rates Mean Lower Standards? Tonsillectomy and Adenoidectomy in Manitoba", *New England Journal of Medicine* 297:360-365.

Roos, N.P. and L.L. Roos (1982), "Surgical Rate Variations: Do They Reflect the Health or Socioeconomic Characteristics of the Population?", *Medical Care* 20(9):945-958.

Schaffer, W., Ray WA, Ferderspiel CF, Miller WO, (1983), "Improving Antibiotic Prescribing in Office Practice: A Controlled Trial of Three Education Methods," *Journal of the American Medical Association* 250: 1728-32.

Stano, M., and S. Folland (1988), "Variations in the Use of Physician Services by Medicare Beneficiaries", *Health Care Financing Review* 9(3):51-57.

Taylor, Richard (1979). *Medicine out of Control: The Anatomy of a Malignant Technology.* (Melbourne: Sun Books).

Wennberg, J.E., L. Blowers, R. Parket, et al. (1977). "Changes in Tonsillectomy Rates associated with Feedback and Review," *Pediatrics* 59: 821-26.

Wennberg, J.E., J.L. Freeman and W.J. Culp (1987). "Are Hospital Services Rationed in New Haven or Over-utilised in Boston?", *Lancet* 1, 1185-1188.

Wennberg, J.E., J.L. Freeman, R.M. Shelton and T.A. Bubolzt (1989) "Hospital use and mortality among Medicare beneficiaries in Boston and New Haven", *New England Journal of Medicine* 321:1168-73.

Wennberg J. and A. Gittelsohn (1973), "Small Area Variations in Health Care Delivery. *Science* 182:1102-1108.

Wennberg J. and A. Gittlesohn (1982), "Variations in Medical Care Among Small Areas", *Scientific American* 246(4):120-134

Woodward, R.S. and F. Warren-Boulton (1984), "Considering the Effect of Financial Incentives and Professional Ethics on 'Appropriate' Medical Care," *Journal of Health Economics* 3(3) 223-237.

Young P.(1993), "The Evolution of Conventions", *Econometrica* 61(1):57-84.

Figure 1: Surgical Variation for Ten Common Procedures
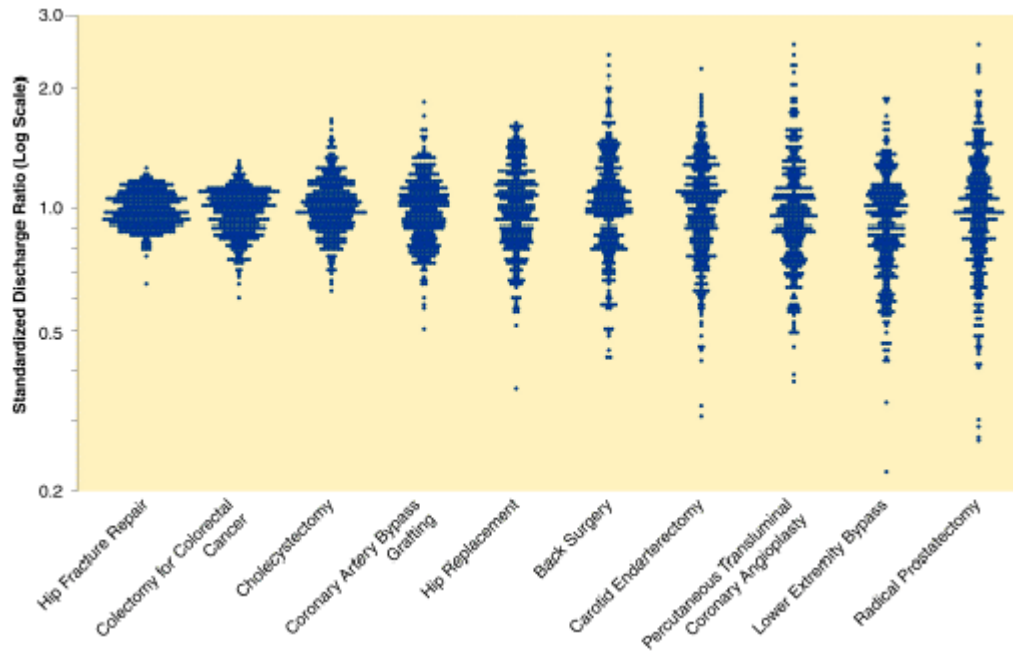
Source: Figure 5.1 in the *Dartmouth Atlas of Health Care*. Each data point represents an observation for a Hospital Referral Region relative to the US average standardized for age-gender-race and illness. See CECS (1999a) and text, for construction details.

Figure 2: Differences in Health and Regional Differences in Prices Cannot Explain Variations in Medicare Spending
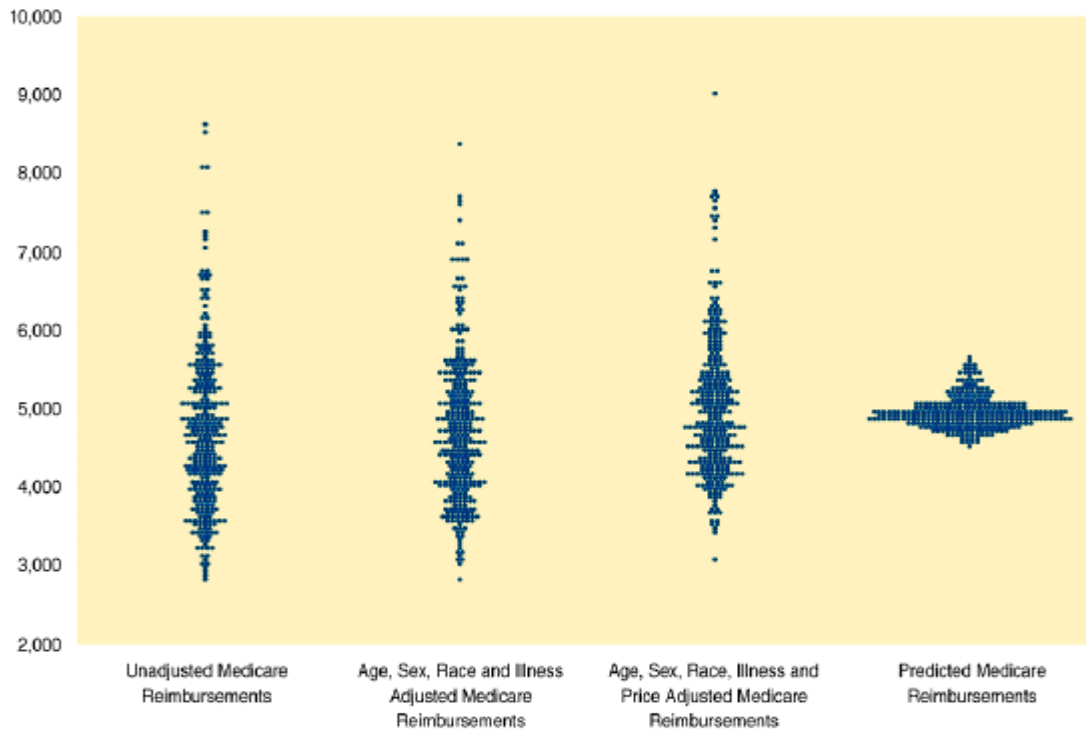

Source: Figure 1.6 in the *Dartmouth Atlas of Health Care*. Each data point represents an observation for a Hospital Referral Region. See CECS (1999a) and text, for construction details.

Figure 3: Rates for Non Imaging and Imaging Stress Tests Compared

Panel A: Data at the Hospital Referral Region (HRR) Level
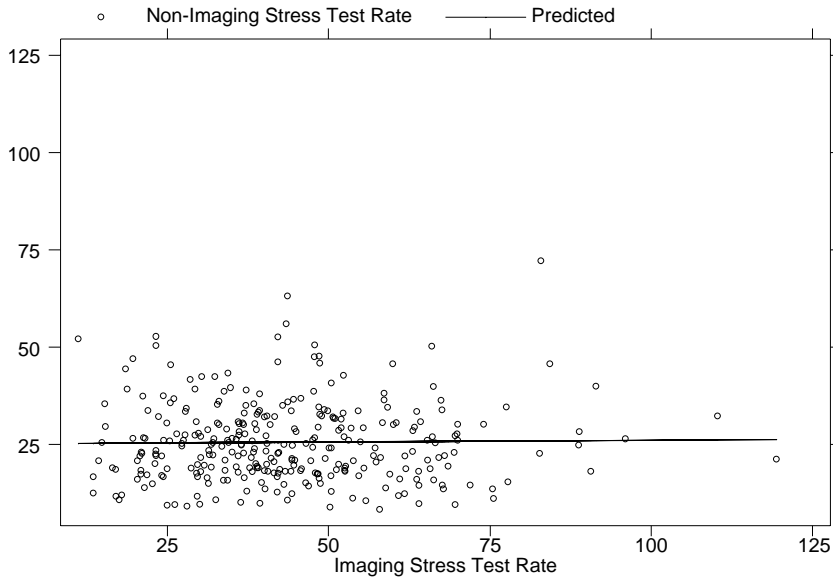
Panel B: Data at the Coronary Angiography Service Area (CASA) Level
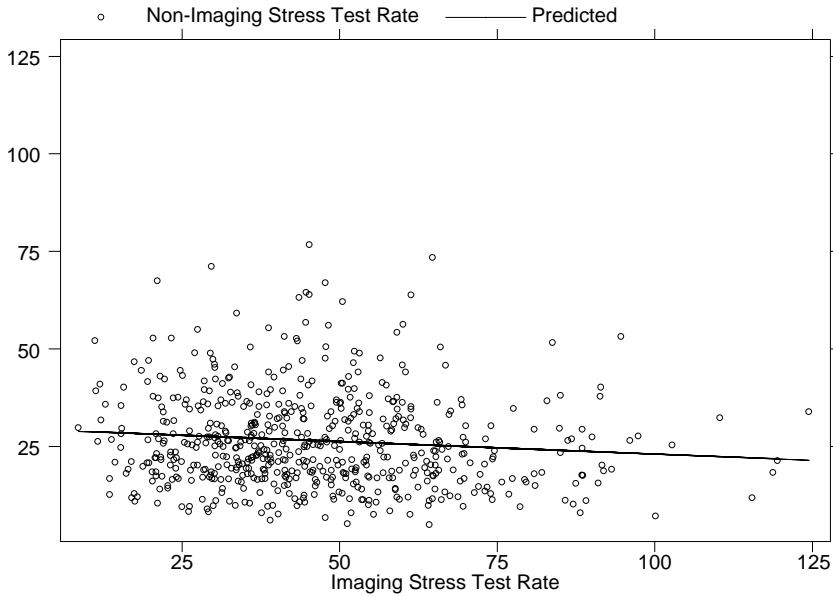
Notes: Each rate was constructed per one thousand Medicare beneficiaries. In Panel A, n=306 with an R^2 of 0.0002. In Panel B, n=589 with an R^2 of 0.0095. Regressions were weighted using the number of Medicare beneficiaries in each HRR or CASA as weights. The results were unchanged if the number of admissions for AMI were used as weights. See text and Appendix for details of the HRR and CASA construction.
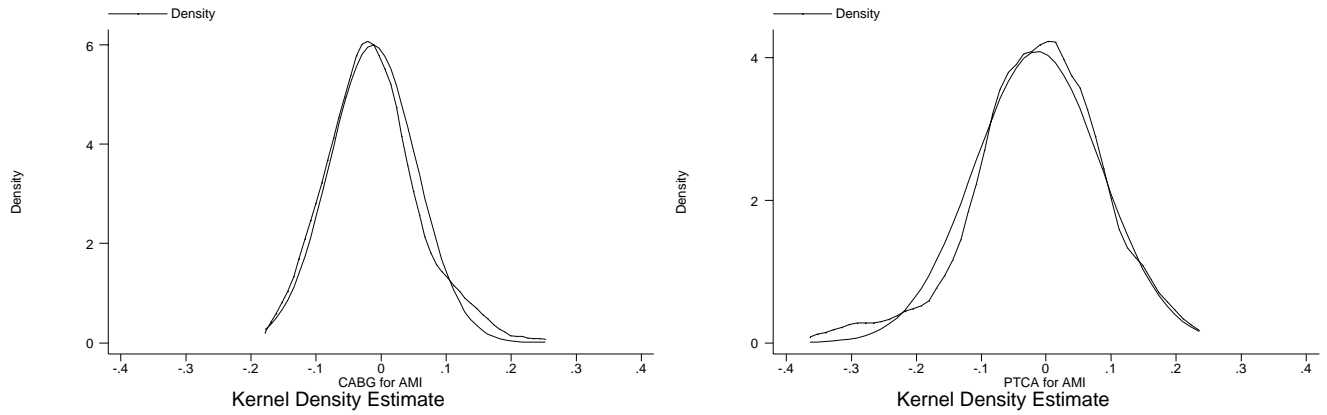
Figure 4: Kernel Density Estimates of the Distribution of Hospital Fixed Effects for the Treatment of Hypertension , NHAMCS Data 1995-99
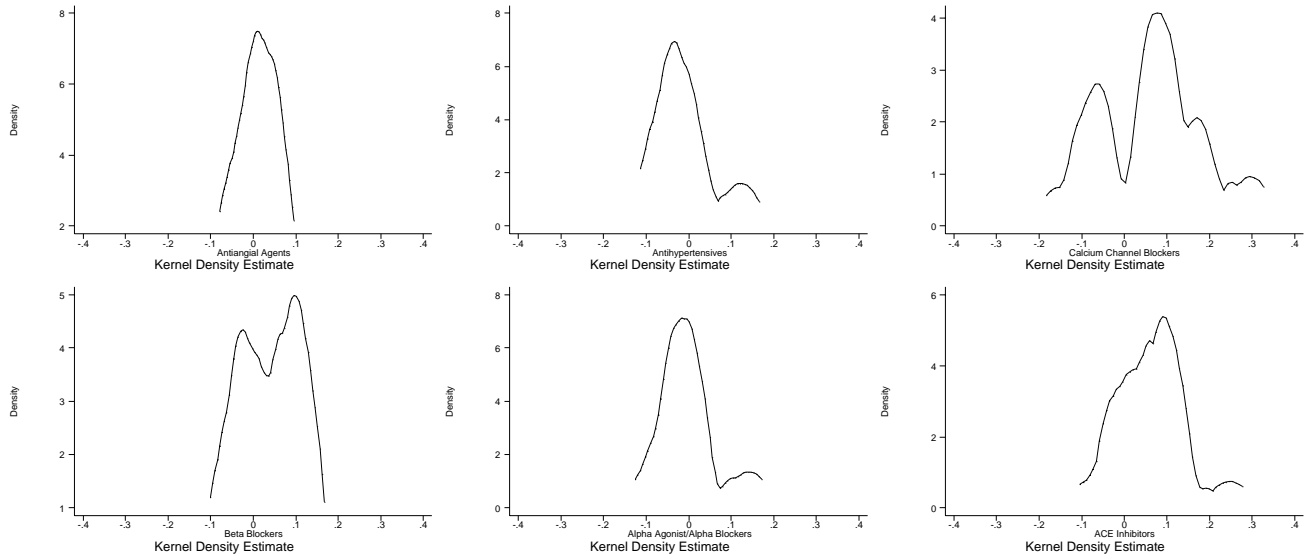
Notes: Table reports usage rates of pharmaceutical interventions for the treatment of hypertension. The rates are computed as the hospital fixed-effects from a patient level regression that standardizes the data for race, gender and age. There are 67 hospital in the sample, each with at least 30 cases of hypertension.

Panel A: CABG and PTCA for AMI

Kernel Density Estimate — CABG for AMI

Kernel Density Estimate — PTCA for AMI
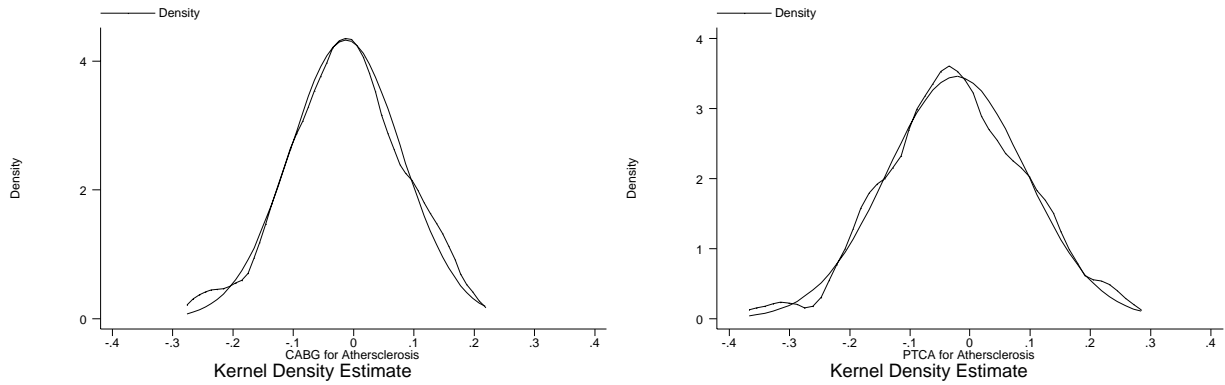
Notes: Table reports usage rates of surgical interventions for the treatment of AMI. The rates are computed as the hospital fixed-effects from a patient level regression that standardizes the data for race, gender and age. There are 120 hospital in the sample, each with at least 30 cases of AMI.

Figure 5B: Kernel Density Estimates of the Distribution of Hospital Fixed Effects for the Treatment of Coronary Atherschlerosis , HCUP Data 1996

Panel B: CABG and PTCA for Atherslerosis

Kernel Density Estimate

Kernel Density Estimate

Notes: Table reports usage rates of surgical interventions for the treatment of coronary atherschlerosis. The rates are computed as the hospital fixed-effects from a patient level regression that standardizes the data for race, gender and age. There are 120 hospital in the sample, each with at least 30 cases of atherschlerosis.
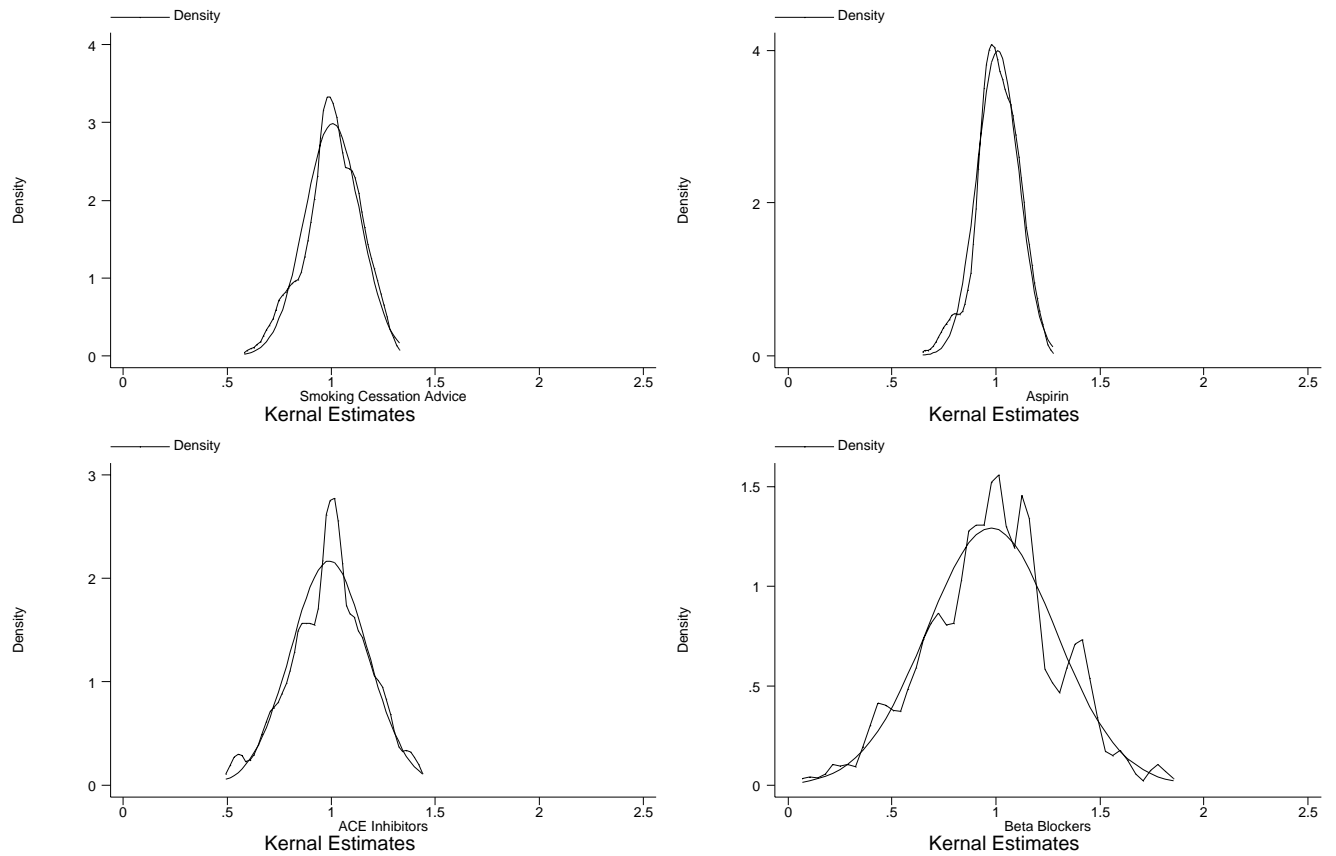
Figure 6: Kernel Density Estimates of the Distribution of HRR level Treatment Rates for Continuous Interventions, Dartmouth Atlas Data

Notes: Table reports HRR treatment rates relative to the US average for 301 HRRs using data from the Dartmouth Atlas of Cardiovascular Care [CECS (1999b)]. Rates are standardized for age-race-gender and illness.

Figure 7: Kernel Density Estimates of the Distribution of HRR level Treatment Rates for Discrete Interventions, Dartmouth Atlas Data
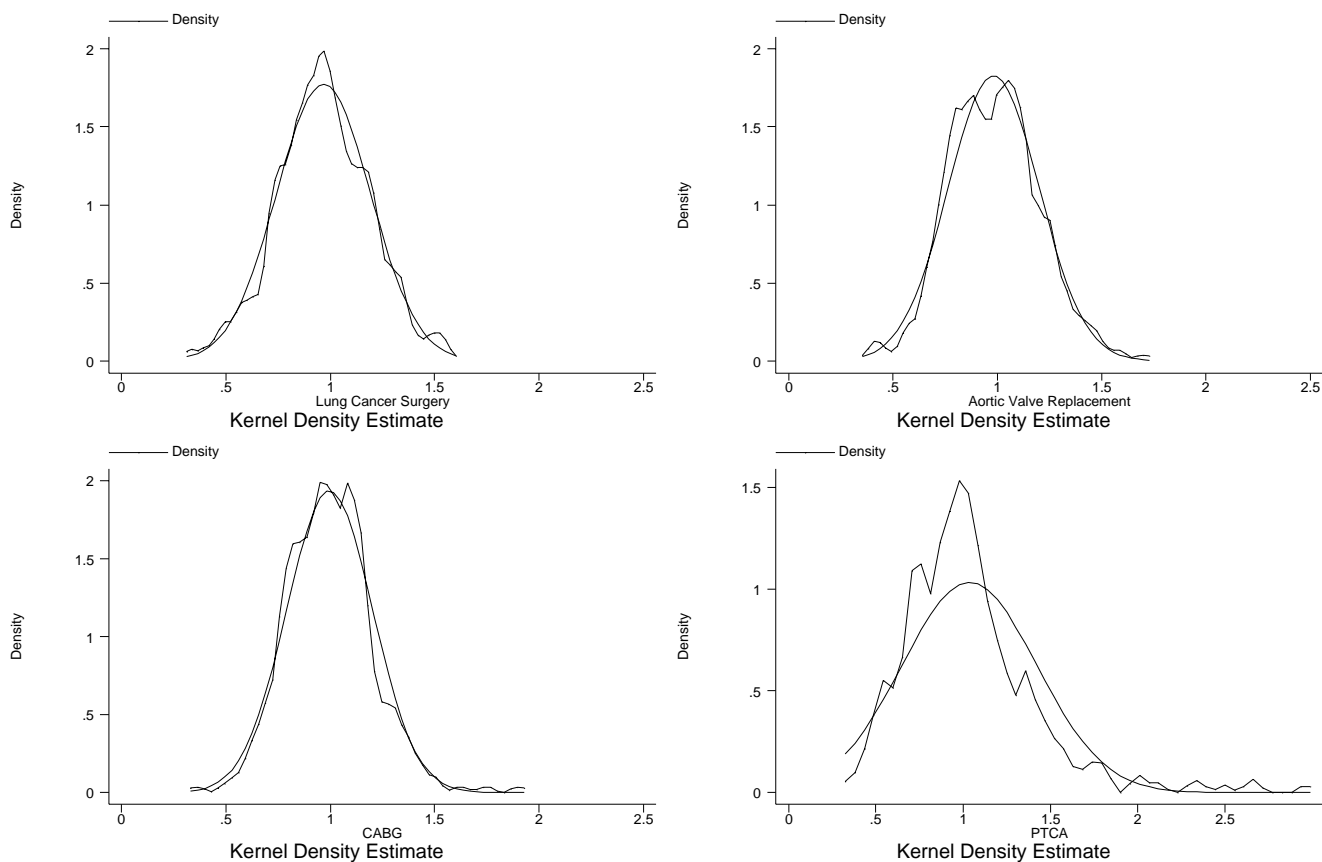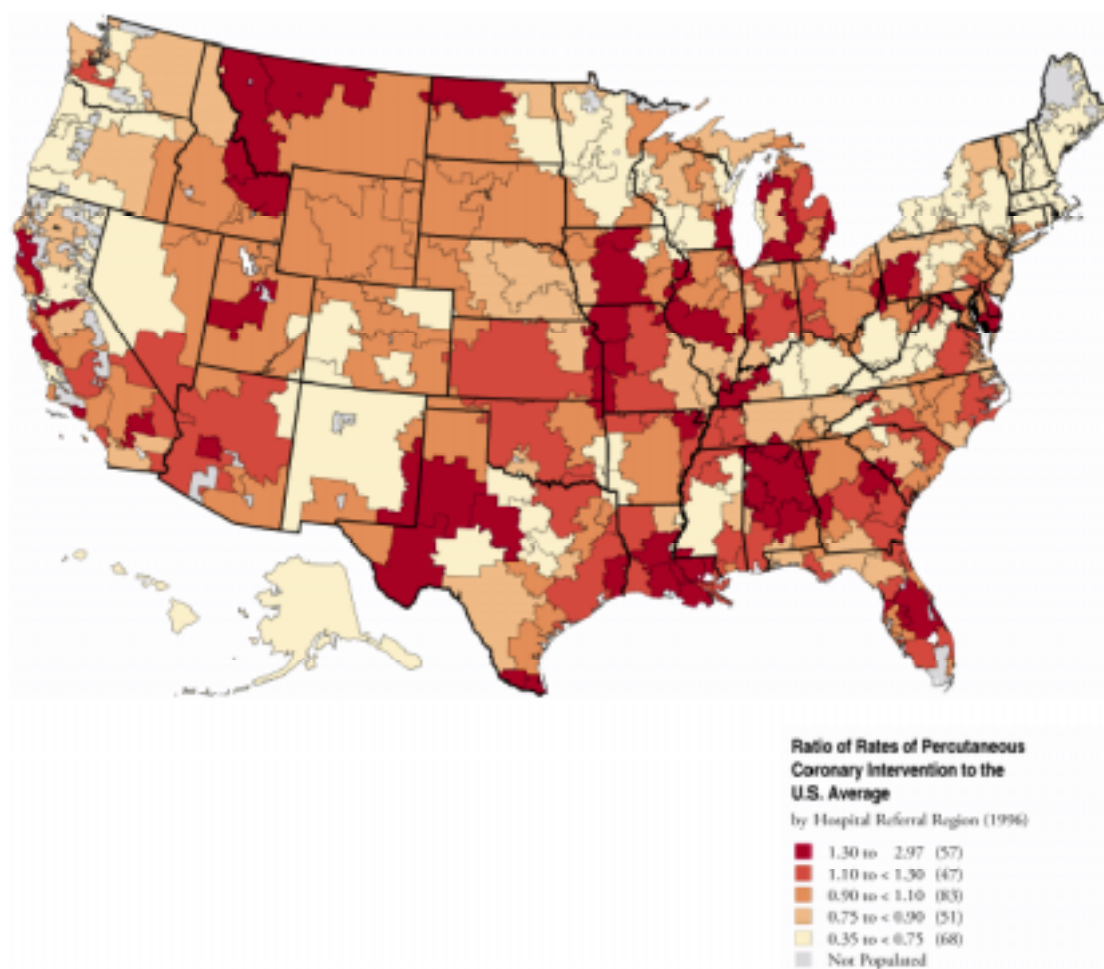

Notes: Table reports HRR treatment rates relative to the US average for 301 HRRs using data from the Dartmouth Atlas of Cardiovascular Care [CECS (1999b)]. Rates are standardized for age-race-gender and illness.

Figure 8a: Discrete Intervention: Rates of Percutaneous Coronary Interventions Relative to the US average in 1996.

Ratio of Rates of Percutaneous
Coronary Intervention to the
U.S. Average

by Hospital Referral Region (1996)

- 1.30 to 2.97 (57)
- 1.10 to < 1.30 (47)
- 0.90 to < 1.10 (83)
- 0.75 to < 0.90 (51)
- 0.35 to < 0.75 (68)
- Not Populated

Source: Dartmouth Atlas of Cardiovascular Care [CECS (1999b)]. The data are disaggregated at the HRR level using 100 percent Medicare claims data for 1996. See text for construction details.

Figure 8b: Discrete Intervention: Rates of Coronary Artery Bypass Graft (CABG) Surgery Relative to the US average in 1996
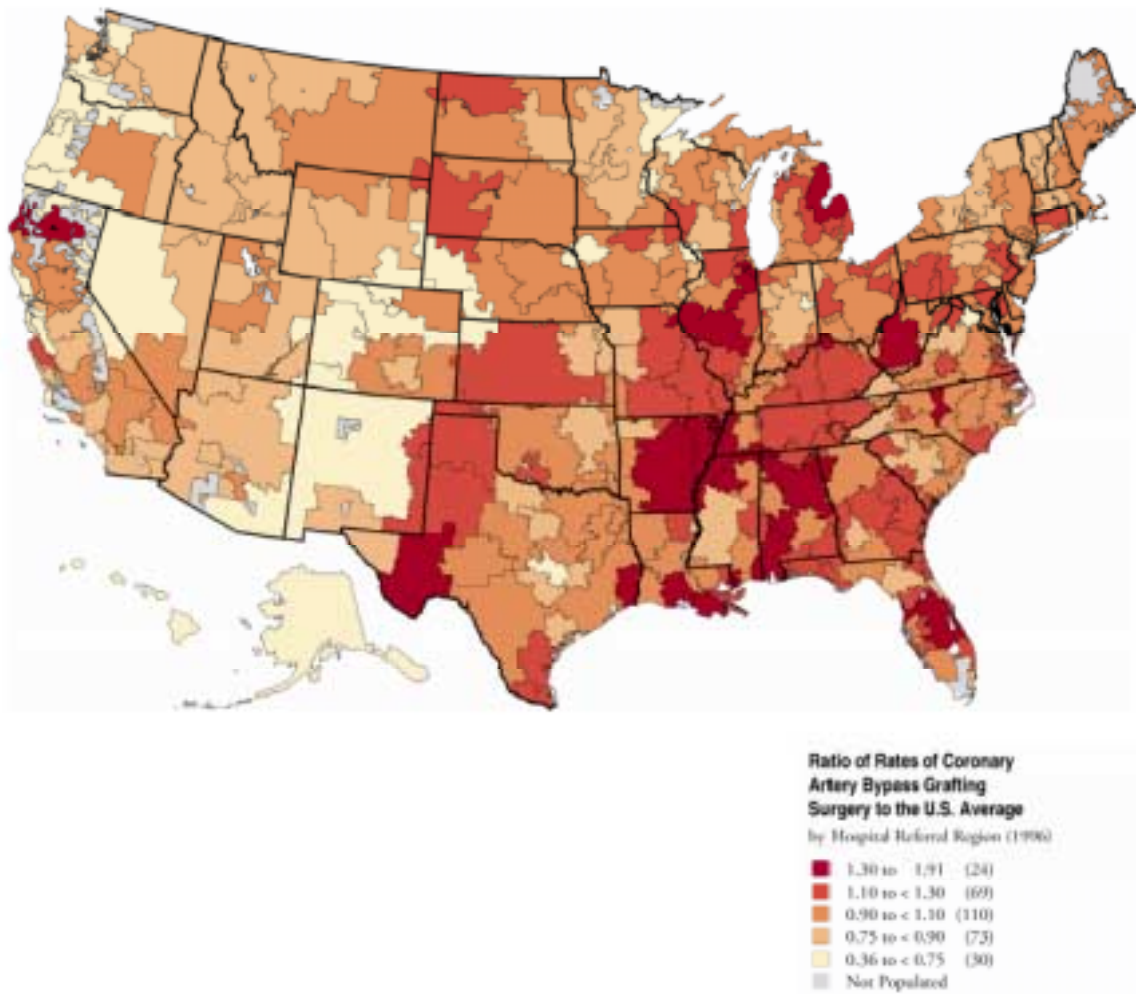
Ratio of Rates of Coronary
Artery Bypass Grafting
Surgery to the U.S. Average
by Hospital Referral Region (1996)

- 1.30 to 1.91 (24)
- 1.10 to < 1.30 (69)
- 0.90 to < 1.10 (110)
- 0.75 to < 0.90 (73)
- 0.36 to < 0.75 (30)
- Not Populated

Source: Dartmouth Atlas of Cardiovascular Care [CECS (1999b)]. The data are disaggregated at the HRR level using 100 percent Medicare claims data for 1996. See text for construction details.

Figure 8c: Discrete Intervention: Rates of Aortic Valve Replacement Relative to the US average in 1996.

Ratio of Rates of Aortic Valve
Replacement to the U.S. Average
by Hospital Referral Region (1996-97)

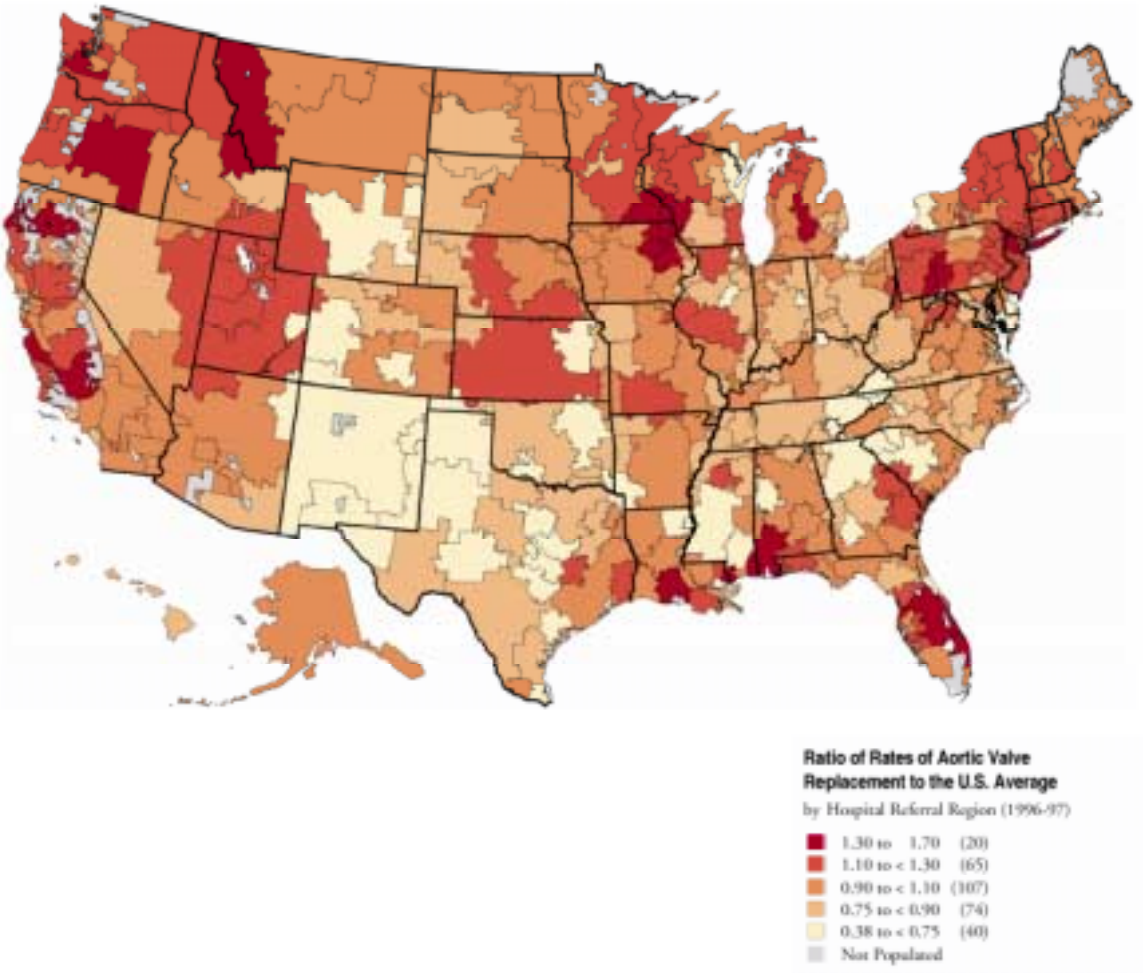|		
■	1.30 to 1.70	(20)
■	1.10 to < 1.30	(65)
■	0.90 to < 1.10	(107)
■	0.75 to < 0.90	(74)
□	0.38 to < 0.75	(40)
▨	Not Populated	

Source: Dartmouth Atlas of Cardiovascular Care [CECS (1999b)]. The data are disaggregated at the HRR level using 100 percent Medicare claims data for 1996. See text for construction details.

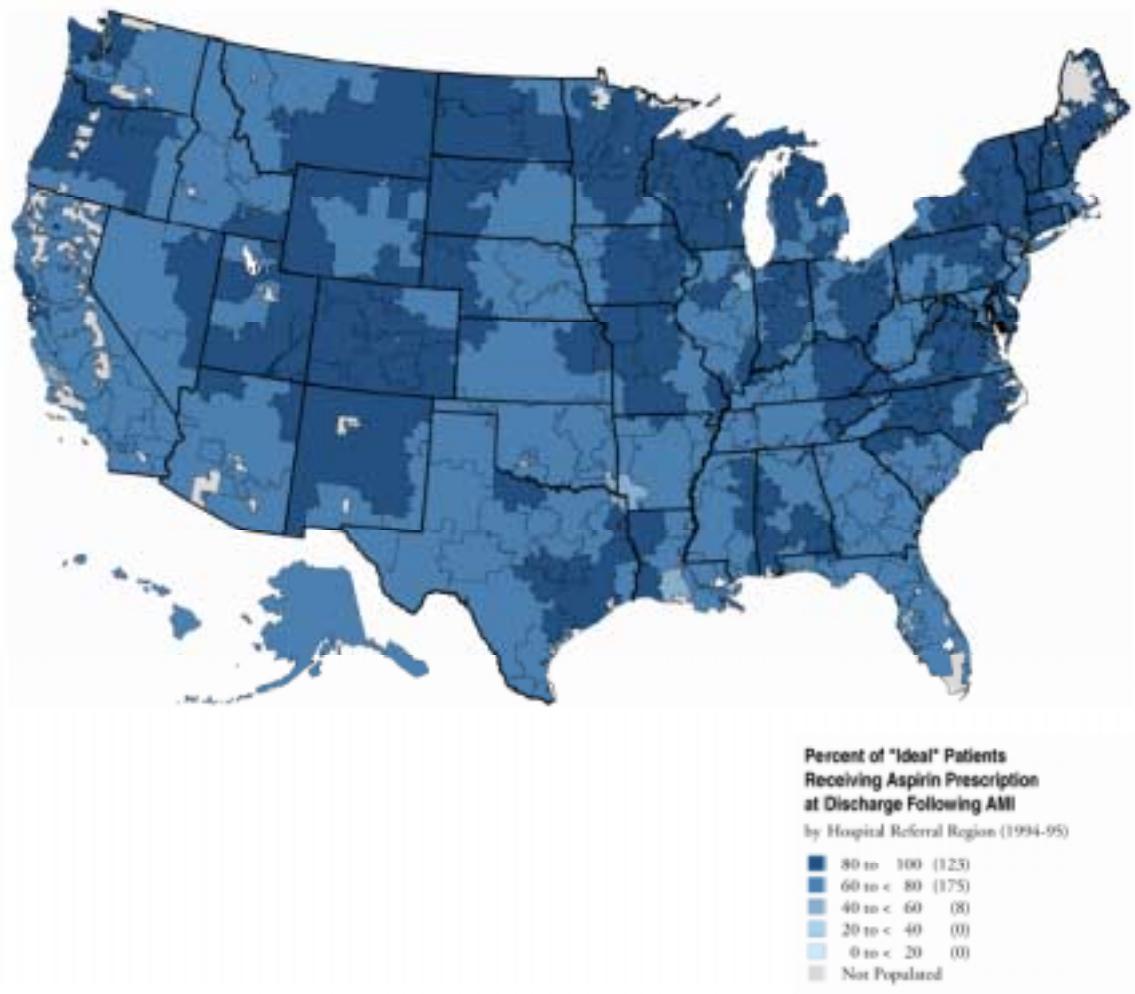**Figure 9a: Continuous Intervention: Use of Aspirin at Hospital Discharge for patients with AMI in 1996**

Source: Dartmouth Atlas of Cardiovascular Care [CECS (1999b)]. The data are disaggregated at the HRR level using 100 percent Medicare claims data for 1996. See text for construction details.

Figure 9b: Continuous Intervention: Use of ACE Inhibitors at Discharge for patients with AMI in 1996

Percent of "Ideal" Patients
Receiving ACE Inhibitors at
Discharge Following AMI

by Hospital Referral Region (1994-95)

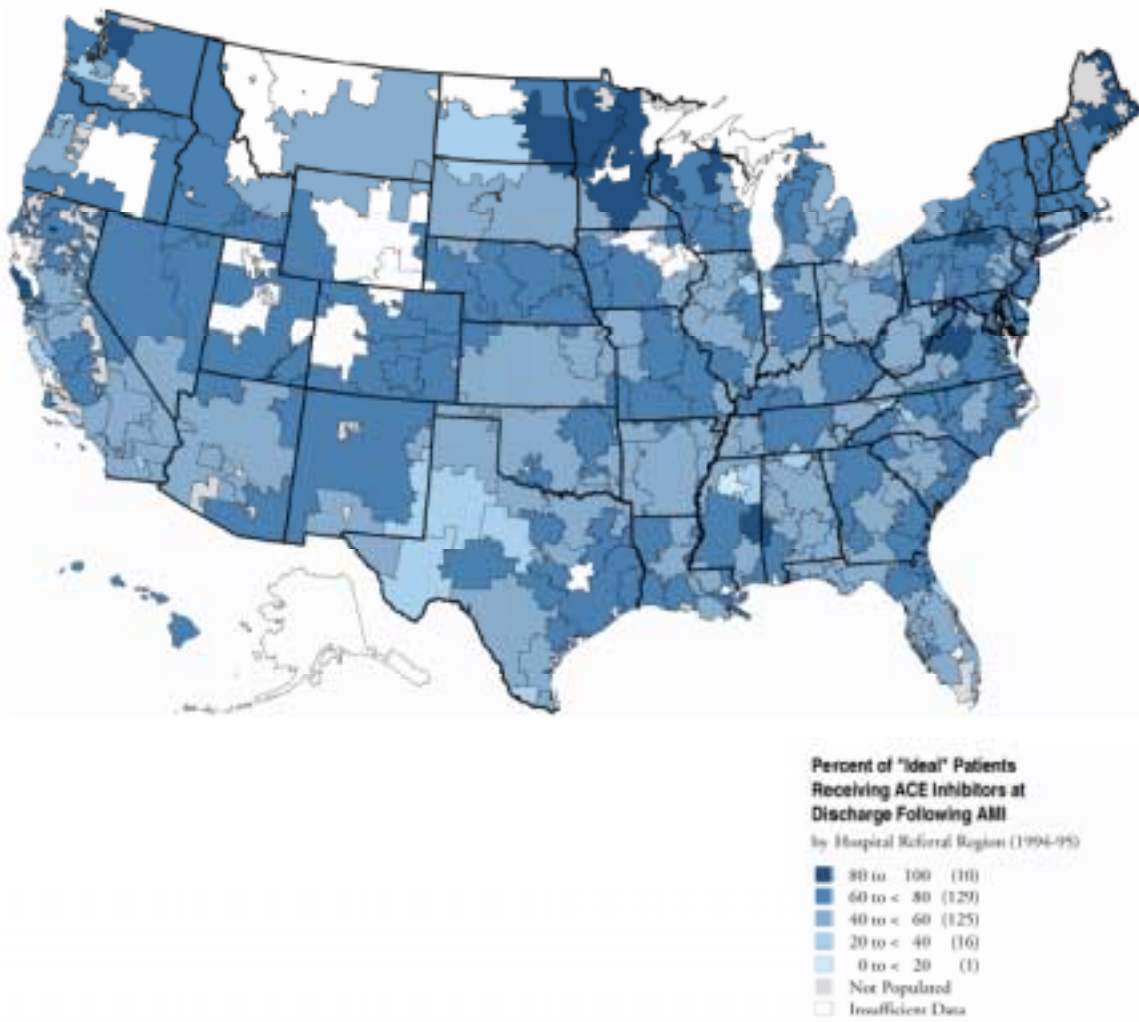|		
■	80 to 100	(10)
■	60 to < 80	(129)
■	40 to < 60	(125)
■	20 to < 40	(16)
□	0 to < 20	(1)
▨	Not Populated	
□	Insufficient Data	

Source: Dartmouth Atlas of Cardiovascular Care [CECS (1999b)]. The data are disaggregated at the HRR level using 100 percent Medicare claims data for 1996. See text for construction details.

Figure 9c: Continuous Intervention: Use of Beta Blockers at Discharge for patients with AMI in 1996

Percent of "Ideal" Patients
Receiving Beta Blocker Prescription
at Discharge Following AMI
by Hospital Referral Region (1994-95)

- 80 to 100 (7)
- 60 to < 80 (53)
- 40 to < 60 (139)
- 20 to < 40 (62)
- 0 to < 20 (7)
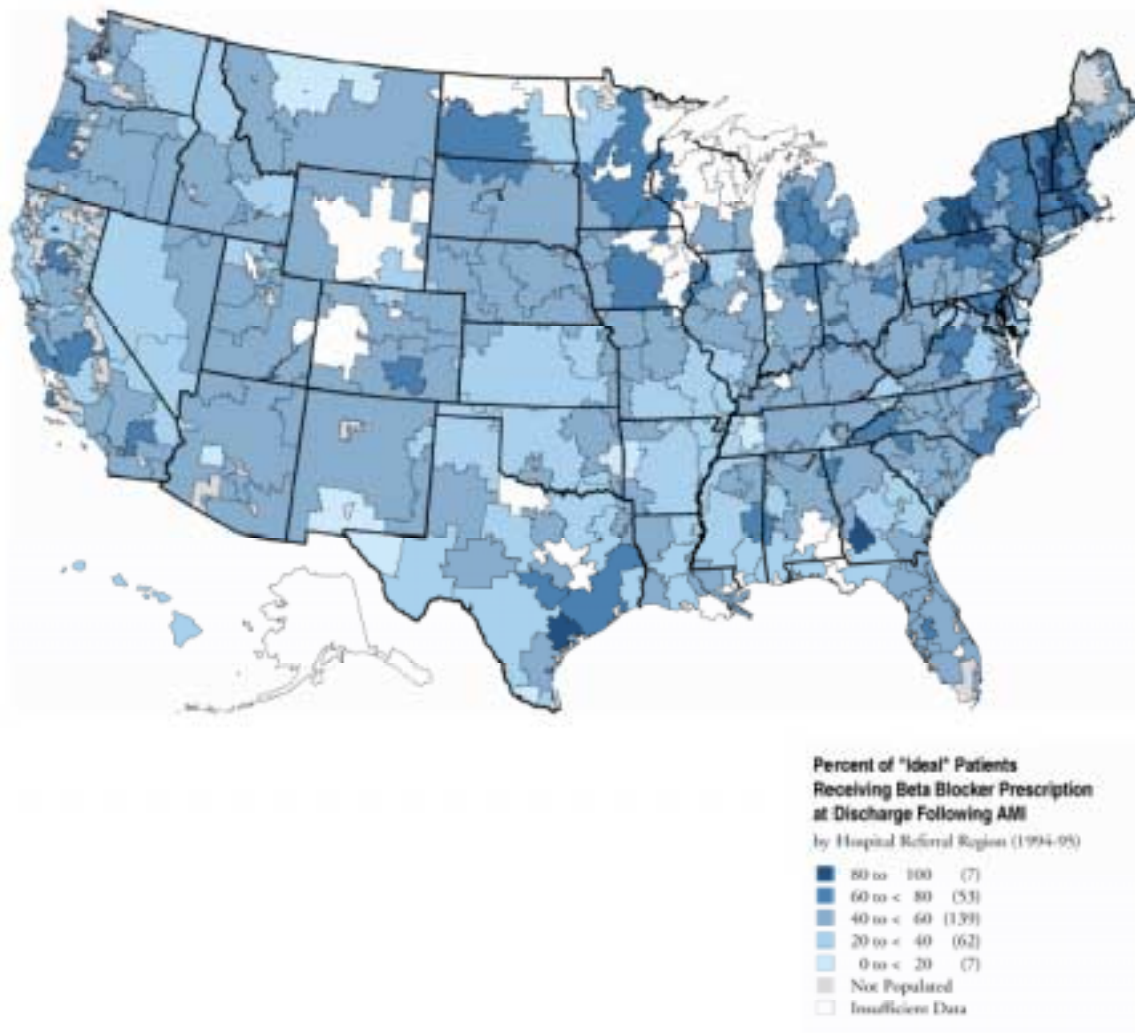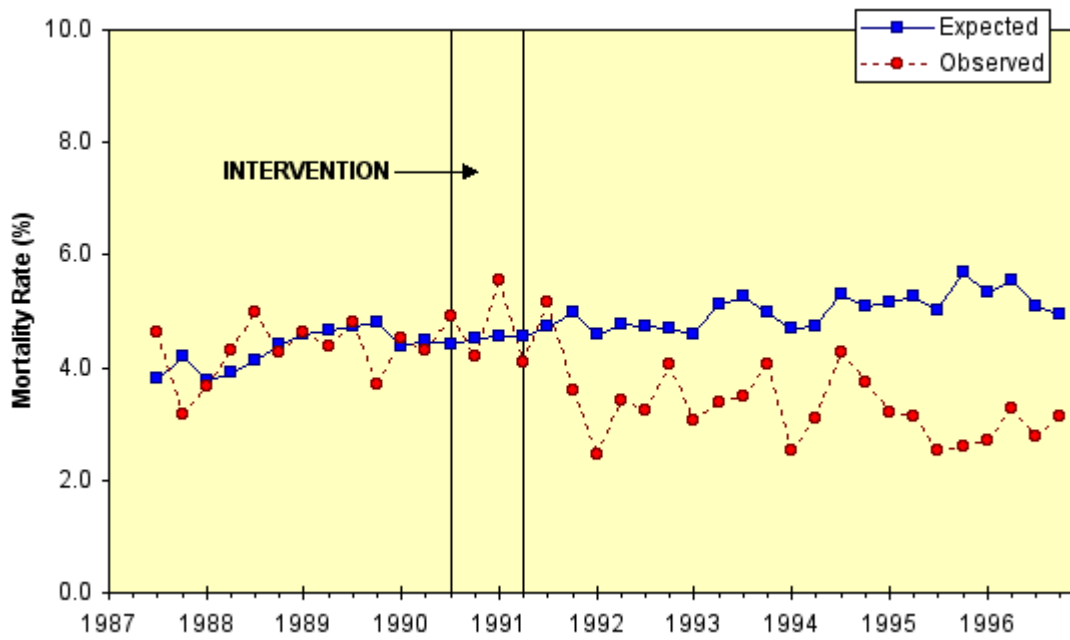- Not Populated
- Insufficient Data

Source: Dartmouth Atlas of Cardiovascular Care [CECS (1999b)]. The data are disaggregated at the HRR level using 100 percent Medicare claims data for 1996. See text for construction details.

Figure 10: Iatroepidemics, Information Releases and Hospital Mortality Rates, Evidence from the NNECDSG Intervention

Source: Figure 5.14 in the Dartmouth Atlas of Health Care [CECS(1999a)]. The NNECDSG group [see O'Conner et al (1996)] found that techniques that promote continuous quality mangement reduced hospital mortality rates associated with CABG surgery by 24%. See text for details.

Table 1: Hospital Level Fixed Effects for the Treatment of Hypertension, NHAMCS Data 1995-1999

Panel A: Fixed Effects and Standard Deviations

Pharmaceutical Treatment	Mean	Std. Dev.	Min	Max
Antiangial Agents	1	0.0500	0.9317	1.1961
Antihypertensives	1	0.0745	0.9074	1.2151
Diurectics	1	0.1610	0.7403	1.3995
Calcium Channel Blockers	1	0.1357	0.7024	1.2652
Beta Blockers	1	0.0941	0.8415	1.3296
Alpha Agonist/ Alpha Blockers	1	0.0667	0.9085	1.2307
ACE Inhibitors	1	0.1329	0.7224	1.3654

Panel B: Correlation Coefficients for Hypertension Treatments

	Antiangial Agents	Anti-Hypertensives	Diurectics	Calcium Channel Blockers	Beta Blockers	Alpha Agonist/ Alpha Blockers	ACE Inhibitors
Antiangial Agents	1.0000						
Anti Hypertensives	-0.0133	1.0000					
Diurectics	-0.0371	0.0786*	1.0000				
Calcium Channel Blockers	0.0934*	0.0176	0.3262*	1.0000			
Beta Blockers	0.3129*	-0.1485*	0.1352*	0.2136*	1.0000		
Alpha Agonist/ Alpha Blockers	0.1630*	-0.0085	0.2161*	0.2147*	0.2177*	1.0000	
ACE Inhibitors	0.1245*	0.0549*	0.4485*	0.4158*	0.4352*	0.2332*	1.0000

Notes: Table reports usage rates of pharmaceutical interventions for the treatment of hypertension relative to the US average. The rates are computed as the hospital fixed-effects from a patient level regression that standardizes the data for race, gender and age. There are 67 hospital in the sample, each with at least 30 cases of hypertension. In Panel B, an asterix indicates significance of the pairwise correlation coefficient at the 5-percent significance level.

Table 2: Hospital Referral Region Summary Statistics, Data from Dartmouth Atlas of Cardiovascular Care

Panel A: Treatment Means and Standard Deviations

Treatment	Mean	Std. Dev.	Min	Max
Aspirin at Discharge	0.9995	0.0989	0.6705	1.2345
Smoking Cessation Advice	1.0017	0.1329	0.6098	1.2939
ACE Inhibitors	1.0013	0.1887	0.5127	1.4289
Beta Blockers at Discharge	0.9988	0.3155	0.1006	1.8634
Reperfusion (Thrombolytics)	0.9997	0.2766	0.1750	1.9484
CABG	0.9977	0.2064	0.3614	1.9030
Surgery for Lung Cancer	1.0313	0.2395	0.3637	1.6775
PTCA	1.0008	0.3732	0.3423	2.8612

Panel B: Correlation Coefficients for AMI Treatments

	Aspirin at Discharge	Beta Blockers at Discharge	Reperfusion	ACE Inhibitors	Smoking Cessation Advice	CABG	PTCA
Aspirin at Discharge	1.0000						
Beta Blockers at Discharge	0.3508*	1.0000					
Reperfusion (Thrombolytics)	0.2994*	0.1405*	1.0000				
ACE Inhibitors	0.2456*	0.2676*	0.1699*	1.0000			
Smoking Cessation Advice	0.0995	0.04680	0.0099	-0.0044	1.0000		
CABG	-0.1956*	-0.1819*	-0.2771*	-0.1618*	0.02670	1.0000	
PTCA	-0.1773*	-0.2628*	-0.2093*	-0.3837*	0.00290	0.3447*	1.0000

Notes: Table reports HRR treatment rates relative to the US average for 301 HRRs using data from the Dartmouth Atlas of Cardiovascular Care [CECS (1999b)]. In Panel B, an asterix indicates significance of the pairwise correlation coefficient at the 5-percent significance level.