Physician Practice Style and Healthcare Costs: Evidence from Emergency Departments.*

Gautam Gowrisankaran, Ph.D.† Keith A. Joiner, M.D. MPH‡

Pierre Thomas Léger, Ph.D.§

October 23, 2014

Abstract

We examine the impact of emergency department (ED) treatment on future healthcare costs and outcomes. We postulate that ED physicians may affect patient outcomes through acumen in diagnosis and appropriate disposition conditional on diagnosis. We make use of a unique dataset and quasi-experiment on patients who seek treatment at an ED in Montreal, Canada. Physicians there rotate across shifts between simple cases and difficult cases, implying that the assignment of patients will be quasi-random across physicians in an ED over time. We examine how the initial assignment of ED physician affects outcomes. We consider two serious, potentially life-threatening conditions, that present in the ED, angina and transient ischemic attacks, which are precursors to heart attacks and strokes respectively. We find that ED physicians who have high initial spending for angina and similar conditions also have high initial spending for transient ischemic attacks and similar conditions. ED physicians who spend more in the ED also have higher revisits to the EDs during the 0-5 days for both illness groups and higher costs of hospitalization in the 0-5 day period after the initial treatment for the angina and similar conditions category. We find no evidence that treatment by a high-spending ED physician leads to either better or worse outcomes, in terms of future hospital costs or ED revisits, over a 6-90 day period after the initial treatment. We also show that differences across patient outcomes across physicians are driven by both diagnostic and disposition differences although the variation in diagnosis skills is found to be quantitatively more important than variation in disposition skills.

Keywords: physician quality, emergency departments, quasi-random variation.

JEL Classification: I12

---

*Preliminary and incomplete. Do not quote without authors' permission. Léger acknowledges funding from HEC Montréal. Gowisankaran acknowledges funding from the Center for Management Innovations in Healthcare at the University of Arizona. The authors thank Michelle Houde and José Pérez at the Déapartement de santé publique de Montréal for data assistance and extraction and Andreas Krull, M.D. and Marie-Josée Ouimet, M.D., who provided many of institutional details about Montreal Emergency Departments.

†University of Arizona, HEC Montréal and NBER, gowrisankaran@eller.arizona.edu

‡University of Arizona, kjoiner@medadmin.arizona.edu

§HEC Montréal, CIRANO and CIRPEE, ptl@hec.ca

1 Introduction

As healthcare costs have continued to grow in the U.S. and other OECD countries, researchers and policy makers are increasingly interested in optimizing the production of healthcare, with the goal of increasing healthcare quality and access to care at the lowest possible cost. As pointed out by Porter (2010), of most importance is to increase the value of care, defined as outcomes/costs. Overall measures of population health, using metrics such as life expectancy, mortality and morbidity amenable to medical care, infant mortality, and more indicate the value of care delivered in the US falls far below that of most developed countries. With the advent of large, integrated health delivery systems, some assessment of value is possible for populations under their care. Hospital report cards are now commonplace, and are readily assessed through sites such as HealthGrades, Leapfrog, and Hospital Compare.

None of these capture the role of the individual physician in the production function for healthcare, which is the goal of this paper. In many medical settings, it is infeasible to evaluate the extent of variation in physician skill and practice style. This is particularly the case when medical decisions are made in the context of management of chronic disease. As such, they do not involve random pairing between physician and patient; they may involve care by a collection of care providers; and they cannot be adequately adjusted for underlying patient risk. One area of medicine where rankings have been prevalent is surgery.[1] Assessment of surgeon skill is facilitated by the fact that surgery occurs at known time point. Determining mortality and complication rates following surgical or invasive interventions are relatively straightforward, although patient risk adjustment is still needed to make comparisons across individual physicians. Although surgeon evaluations have been valuable, they cannot measure diagnostic skill or efficient resource utilization. Moreover, the large majority of current medical care consists of chronic disease management and not surgery. For chronic disease management, provider decisions include a nuanced evaluation of patient disease status, response to therapy (including complications from pharmaceuticals and other interventions), and coordination across multiple comorbidities, such as diabetes, hypertension, heart disease, and respiratory disease. Overall, disease management and diagnosis skills by physicians are important components of the healthcare production function, and yet we know little about how skills and practice styles vary across physicians.

[1] These report cards have been widely studied in the economics and public health literation; see, e.g., Dranove et al. (2003).

We present a new approach to measuring physician skill and resource utilization that circumvents some of the above limitations: we focus on the role of the physician in the emergency department (ED). Our study identifies the role of the ED physician by making use of a unique dataset that contains all health records for patients who visit an ED in Montreal, Quebec, Canada, during a 9 month period in 2006. For each ED encounter, we observe patient demographics, the diagnoses and procedures performed at the initial visit, and an identifier for the ED physician who is examining the patient. Importantly, our data link the ED visit with all future encounters of the patient with the health system during the 90 days following the initial visit, including office visits, hospital-based external clinic (referred to here as outpatient) visits, revisits to EDs, and hospitalizations. Our linked data allow us to find evidence on missed diagnoses, wrong diagnoses, inappropriate management decisions, and overuse or underuse of resources, all of which would be difficult to examine with commonly-used data sources such as hospital discharge data.

At most EDs in Montreal, ED physicians typically see patients with a full range of diagnoses and levels of complication. This procedure (discussed in detail below) leads to a quasi-random assignment of physicians to patients conditional on treatment at a particular ED. We exploit this quasi-random assignment to examine how the initial assignment of ED physician can contribute to the care received during the ED visit as well as the future pattern of health outcomes, healthcare consumption and healthcare costs.[2]

We focus on emergency departments (EDs) both because of the quasi-random assignment of patients but also because it is of significant interest to understand the role of ED physician practice style. EDs are a primary point of entry to the healthcare system for many patients, particularly patients who are underserved by primary care physicians. EDs are often seen as a particularly wasteful entry point to healthcare access, with researchers focusing on the (presumably negative) consequences that increased insurance can have on increasing ED usage Taubman et al. (2014). ED care is particularly complex, given the acute and time-sensitive nature of the diagnostic and therapeutic process, and hence are among the most common sites for misdiagnosis (Rothrock and Pagane (2000); Pope and Edlow (2012); Pope et al. (2000)). Finally, EDs are a prominent place of treatment for a number of widely-prevalent diseases such as ischemic heart disease and cerebrovascular disease (Friedmann et al. (2001)), which may be difficult to diagnose and where appropriate

[2] As discussed in detail below, we test the assumption of random assignment of patients based and eliminate certain hospitals and physicians as a consequence.

and timely care is particularly important.

We propose a simple conceptual model to highlight the role of the ED physician. In our model, the central role of the ED physician is the appropriate diagnosis and disposition of the patient. There are four possibilities to describe the consequences of the initial interaction between patients and providers in the ED. The physician either makes the correct diagnosis or does not. At that point, the physician either institutes appropriate treatment or does not. At both of these steps, decisions are made about resource utilization (laboratory tests, imaging studies, medications, specialty consultation, hospitalization, and more). In the optimal circumstance, the correct diagnosis is made, the disposition is appropriate, and only necessary resources are used. In contrast, overuse or underuse of resources, coupled with and sometimes connected to an incorrect diagnosis or inappropriate disposition, can result in worse outcomes and additional costs. Our model contains both potential horizontal and vertical dimensions to physician skill. On the vertical dimension, a physician who is better at diagnosis and the implementation of the appropriate treatment will be the best. On the horizontal dimension, some physicians may overtreat, leading to too high costs, while others may undertreat, leading to potential adverse health outcomes and costly future care.

The main empirical focus of our paper is to estimate an empirical counterpart to this conceptual model. More specifically, we estimate the impact of treatment by different physicians on the cost of physician services during the initial ED visit as well as healthcare resource use over a 90 day period following the initial ED visit. This allows us to examine the extent of variation in care patterns across ED physicians. We observe the physician related costs (based on fee-for-service payments to physicians) associated with the ED visit. Notably, we also observe multiple dimensions of outcomes, including the presence and costs of inpatient hospitalizations, external hospital clinic visits, physician office visits, and ED revisits. This allows us to understand the extent to which physicians' practice styles vary and correlate across different dimensions. Thus, we can uncover, for instance, whether physicians who spend relatively more in the ED also have relatively more or less future hospitalizations costs. Or, alternatively, whether those who exhibit more patient revisits to the ED within a given time window also exhibit greater patient hospital spending over the same time window.

Our analyses are based on two particular illnesses: angina and transient ischemic attacks (TIAs). Angina is a precursor to myocardial infarction (heart attack). It reflects partial blockage of one or more coronary arteries, which supply blood to the heart muscle. The initial presentation of angina is typically associated with exertion, which puts an increased load on the heart muscle. The classic symptom of angina is chest pressure or pain ("an elephant is sitting on my chest") but other presentations such as jaw or shoulder pain, indigestion, or nausea also occur. Diagnosis is considered more difficult in women, in which up to 60% may present without chest pressure/pain. Anginal symptoms usually resolve quickly, with cessation of exertion, because the relative limitation of blood flow to the heart is relieved. Confirmation of coronary artery disease depends upon prompt assessment using stress echocardiogram, coronary angiography, and/or nuclear medicine scans. The consequences of failing to consider or appropriately manage angina include myocardial infarction and sudden death.

TIAs are precursors to strokes, and are sometimes called mini-strokes. They result from a transient occlusion of a blood vessel in the brain. Unlike strokes, the signs and symptoms resolve quickly (usually within minutes), because the occlusion partially or fully resolves. The symptoms and signs, which vary enormously depending upon the part of the brain that is affected, include visual or speech changes, weakness, and numbness. Confirmation of a TIA depends upon additional tests, often done after discharge from the ED (carotid ultrasound, MRI or CT, echocardiogram). The consequences of failing to consider or appropriately manage a TIA include a nearly 5-fold increase in the incidence of stroke over the subsequent 90 days (Rothwell et al., 2007).

We consider angina and TIA in our analysis for two reasons. First, both conditions require a relatively high level of diagnostic acumen from the physician's perspective and as such are likely to capture physician diagnostic differences. That is, both are illnesses that can easily be misdiagnosed, and, misdiagnosis can lead to dramatic consequences. Second, for both conditions, encouraging the appropriate disposition of patients conditional on diagnosis may also be challenging.

In order to understand the skill of diagnosis and disposition for angina patients, our main analysis considers all patients who are diagnosed during an ED visit either with angina or an illness that could be indicative of a misdiagnosed angina. We refer to this broader sample of patients who may or may not suffer from angina as the angina+ sample. Similarly, we consider the set of patients who are diagnosed with TIA or an illness that could be indicative of a misdiagnosed TIA. We refer to this broader sample of patients who may or may not suffer from TIA as the TIA+ sample. Our unit of observation is one ED visit for one patient. Our main dependent variables include the patient's consumption of (physician related) healthcare during the initial ED visit as well as her consumption of medical services within 90 days following this visit. We separate consumption by type of facility, i.e., inpatient hospital, external hospital-based (outpatient) clinic, physician office, and ED. We then estimate each physician's average contribution (i.e., skill or practice style) across these different outcomes. Because the future use of medical care by patients (post ED visit) is likely a function of a proper diagnosis and disposition, the estimated skills or practice patterns capture both.

Once we have estimated each physician's average contribution to ED costs and future use of healthcare (like revisits to the ED or hospitalizations), we can compare them across the angina+ and TIA+ patient samples. Doing so will allow us to see whether physician practice styles or skills are more general (i.e., mostly invariant to the type of patient that is seen) or more illness-specific in nature. Again, in this context, skill measures based on patient outcomes (based on future use of medical care) capture both diagnostic and disposition elements

Next, we consider two particular subsets of the aforementioned samples. More specifically, we re-estimate each physician's average contribution to :(i) costs during the initial visits to the ED, and (ii) use and costs of future healthcare, while looking exclusively at patients in the TIA+ sample who were specifically diagnosed with TIA (or looking exclusively at patients in the angina+ sample who were specifically diagnosed with angina). Examining outcomes of patients based on a specific diagnosis of TIA or angina will allow us to examine disposition skills conditional on diagnosis. That is, will allow us to examine whether and to what extent patient outcomes vary across physicians once a particular (TIA or angina) diagnosis has been made. Furthermore, examining the correlation in skills across the wider and narrower illness categories (i.e., across the TIA+ and specific TIA samples, or, across the angina+ and specific angina samples), allows us to measure to what extent "skills" (or practice patterns) are driven by diagnostic vs good disposition.

Our paper relates to several studies. A recent literature has considered the quasi-random assignment of hospitals (Doyle et al., 2010, 2014; Epstein et al., 2010). Our study builds on this literature by examining in more depth the role of physicians as inputs into the healthcare production function. Other studies have evaluated misdiagnosis in the ED. These ED studies have typically restricted the sample to those subsequently identified to have an adverse event (e.g., complication, death, or hospitalization, sometimes associated with a medical liability suit) (Hastings et al. (2009); Kachalia et al. (2007); Abaluck and Agha (2014)). In our case, our unique data allows us to track all patient interactions with the health system for 90 days after the initial ED visit. This in turn allows us to measure costs and a variety of other well-defined outcomes.

Our results suggest important variation in physician practice styles and skills across the different populations. More specifically, we find that ED physicians vary in their average (physician-related) spending during ED visits. We also find that physicians' average spending during the ED (i.e., practice style) are highly correlated across illness groups (i.e., across the TIA+ and angina+ samples). Thus, a physician's average contribution to (physician-related) ED costs appears to be, in large part, general in nature rather than illness specific. We also find that physicians who are associated greater spending during the initial ED visit are also associated with: (i) more ED revisits within the first 5 days for both the TIA+ and angina+ samples, and (ii) more hospitalization costs for TIA+ sample. These results suggest that more spending in the ED need not reflect better care and may in fact reflect lower skills to begin with. There is weaker evidence of correlation in skills (based on a particular outcome such as hospitalization spending) across the TIA+ and angina+ groups - suggesting some illness-specific skills or practice styles.

Finally, we consider how physician practice styles or skills vary across the wider and narrower illness samples (i.e., across the TIA+ and TIA, and across the angina+ and angina samples). Because the former reflects both the physician's skill (or practice style) at diagnostic and disposition while the later captures skills (or practice style) conditional on a specific diagnosis, we can compare them to to say something about the relative role of diagnostic skills and disposition skills in achieving desired outcomes. We find that physician ED spending during the initial ED visit across the angina+ and angina population are positively correlated. This result suggest that practice styles are not fully explained by the actual diagnosis (with similar results for the TIA+ and TIA samples). We also find that physicians associated with more ED revisits during the first 90 days when treating the angina+ population are also associated with more ED revisits during the first 90 days when treating their angina population. The actual correlation coefficient suggests that both diagnostic and disposition skills are at play, but proportionally more driven by diagnostic differences than disposition. When examining the same correlation between hospital spending in the first 90 days for angina and angina+ samples, the relationship is positive but non-significant. This null result suggest that differences in outcomes (i.e., 90 day hospital spending) are mostly driven by diagnostic skills rather than different dispositions skills. Taken together, these results highlight the importance of proper diagnosis in avoiding potentially bad health outcomes and costly future ED and hospital care.

The remainder of the paper is organized as follows. Section 2 discusses the data and sample construction, Section 3 the model and estimation, and Section 4 the results. Finally, Section 5 concludes.

2 Data and Sample Construction

2.1 Data

Our study uses administrative data from *la Régie de l'assurance maladie du Québec* (RAMQ). The RAMQ data track all publicly-funded healthcare expenditures from the Canadian province of Quebec. More specifically, their database tracks all patients through time and across four types of care: EDs, office visits, hospital-based external clinics (outpatient) visits, and inpatient hospitalizations. RAMQ provides first-dollar coverage to all enrollees, which includes almost all residents of Quebec.[3]

We study residents of the Island of Montreal with an initial ED visit during the period April 1, 2006 to December 31, 2006. The Island of Montreal includes the city of Montreal and some suburban municipalities. Henceforth, we refer to this area simply as Montreal. Our study area includes about 1.9 million residents. We observe the entire universe of patients residing and ED physicians working in Montreal. Access to the data is provided through Montreal's public health department (*le Département de santé publique de Montréal*). Although we use the residents of Montreal and their ED visits as the initial point of interest, we track all future healthcare consumption that occurs within the provincial boundaries (not just those in Montreal). That is, we observe each patient's subsequent consumption of care (across EDs, office, external hospital clinic and inpatient care) as long as it is within the provincial boundaries and covered by the provincial health insurance program.

In Quebec (especially during the study period), physicians across locations are paid on a fee-for-service basis. Our data include the billed physician cost (i.e., the total fee-for-service payment) for office, outpatient and ED visits. Inpatient care cost data is recorded differently where a proxy for the total cost of each inpatient stay is constructed and called *Niveau d'intensité relative des resources utilisées* (NIRRU). We observe NIRRU and use it as our measure of hospital costs.[4]

[3]Although a privately funded sector exists in Quebec, it is very small and generally deals with non-covered services such as plastic surgery. This market remains insignificant as physicians who bill for any services privately must completely opt out of the public sector. An exception to the opt out rule is for imaging facilities.

[4]They do so because all non-physician related expenditures are covered by a "Global Budget" and thus not directly "billed" (associated with) to the patient.

One limitation of our data is that certain procedures are provided by hospitals and hence not directly billed to (associated with) the patient. These include complex imaging services, for instance.[5] Another limitation of our data is that NIRRU is only a proxy for costs, rather than reflecting an actual payment.

For each patient/ED observation, our data include the patient's gender, age group (one of 18), 3-digit postal code as well as two measures of socio-economic status, as constructed for us by the public health department. The first of these measures, known as the *Indice de défavorisation matérielle* (material deprivation index), is a score of 1 to 5 which seeks to reflect the individual's material (i.e., economic) wellbeing. It reflects mostly variations in education, employment and income variations. The second known as the *Indice de d'éfavorisation social* (social deprivation index) seeks to reflect the individual's social and family support and wellbeing (also a score from 1 to 5). It reflects mostly variations in family structure and marital status. Both of these are constructed using a variety of sources and are based in large part on geographical (i.e., postal code) location.[6]

Our data contain all visits to a Montreal ED by a Montreal resident from April 1, 2006 through December 31, 2006. Each ED visit constitutes one observation. For each observation, the data also include the ED's unique identifier, the date of service, the physician's unique identifier, 4 digit ICD-9 diagnosis codes, procedure codes as well as an aggregate of the fees paid to physicians for services provided.[7][8] The data also link the initial ED visit with future visits for 90 days (which may be in the same or another ED or hospital, or in an office or external hospital clinic setting). Thus our data terminate on March 31, 2007. For each future visit, the data track all subsequent healthcare consumption, irrespective of location of consumption as long as it is within the provincial boundaries.

2.2 Sample Construction

From this universe of patients in Montreal, we exclude initial visits at 3 of the 20 EDs: two EDs serve less than 1000 patients per year, and one serves only children. We also exclude all ED physicians

[5]Additionally, imaging services are one market for which there exists a robust private market that will not be included in our data.

[6]These are known as the *Pampalon* indices (Pampalon), Robert et Guy Raymond, Indice de défavorisation matérielle: son application au sector de santé et du bien-être.

[7]An ED visit can include consultation with several MDs as well as several diagnoses and acts. The data include unique MD identifiers for each diagnosis provided and act administered to the patient.

[8]Although the date of service was used by the *le Département de santé publique de Montréal* to construct the sample and track patients for 90 days, the data we received does not include the actual date.

who are not present in all months studied as well as those who saw less than 200 cases during the 9-month period, as these physicians are less likely to have an equitable mix of cases. Our base sample includes 280 ED physicians who practice in 17 Emergency Departments. We observe 321,256 ED visits, each of which constitutes one observation. These visits are made by 199,442 distinct patients. 26% of the observations occur on the weekend.

Our analysis uses patients with diagnoses of angina+ and TIA+. Our definition of angina+ is angina and related diagnoses which might be confounded with angina in an ED setting. Our goal is to choose a set that is sufficiently broad that all ED physicians would diagnose angina patients with one of these diagnoses. But, we also would like the set to be sufficiently narrow that dimensions of physician practice style are similar for all patients within these classes.

Table 1: ICD-9 diagnosis codes for angina+ and TIA+

Angina+		TIA+	
Code	Description	Code	Description
786	Chest pain	346	Migraine
789.6, 530.8	Gastro-esophageal reflux	345	Epilepsy
530.5	Esophageal dysmotility	780.2	Syncope
530.1	Esophagitis	432.1	Subdural hematoma
535.5	Gastritis	431	Intra-cerebral hemmorage
733.6	Costochondrititis	721.1	Compressive myelopathy
307.8	Psychosematic/psychogenic		
420	Pericarditis		
422	Myocarditis		
441	Acute aortic syndroms		
415.1	Pulmonary embolism		
486	Pneumonia		
489	Asthma		

Table 1 indicates the ICD-9 diagnosis codes that we use for each of these diagnoses. The diagnoses for angina+ include the most common non-cardiac causes for chest pain (786), including gastro-esophageal reflux (789.6, 530.8) , esophageal dysmotility (530.5), esophagitis (530.1) gastritis (535.5), costochondritis (733.6) and psycosomatic/psychogenic (307.8). This diagnosis for TIA+ includes the most common diagnoses which present with symptoms that overlap with TIA, including migraine (346), epilepsy (345) and syncope (780.2). For both angina+ and TIA+ there are other serious and less serious conditions that can mimic angina and TIA, which we have excluded from our analysis because of their infrequency. Of relevance to our analysis, it is more straightforward to determine if a patient has coronary artery disease (which leads to angina) than to determine with certainty that a patient has angina.

We now discuss assignment of patients to physicians in the ED. EDs in Montreal are staffed with 1 or more physicians. When 2 or more ED physicians are present, physicians are assigned to either the heavy cases (most often patients who arrive by ambulance) or the light cases (most often patients who enter through the front door) for the duration of their shift. If two physicians are assigned to the same shift type, then the allocation of patients to physicians is random in the sense that it is based uniquely on the triage order (which provides how much time the patient can wait before seeing a physician, or equivalently, who should be seen next). The triage order is done by a triage nurse. Finally, shift allocations (i.e., defined by its time and shift type) are done several months in advance in an equitable manner. As ED physicians are paid fee-for-service, where payments are invariant to any physician characteristics such as experience or tenure, most ED physicians are expected to work all shift types in similar proportions. Exceptions may nonetheless exist, especially among older physicians who may only work part-time or may be given only one type of patients (generally milder cases).

Although in a given shift the allocation of patients to physicians is conditionally deterministic, the quasi-random allocation of ED physicians to shift-types over time should lead all ED physicians to see very similar pools of patients over the long run. There is at least one exception to this assignment rule: in at least one ED, physicians were not assigned to heavy or mild cases but rather the allocation of patients was based uniquely on the triage order (and therefore, purely random in nature from the onset).

In order to test and deal with the possibility of non-random assignment of patients, no matter how small, we examine whether physicians within an ED treat patients with observably different patient characteristics. We use two observable patient characteristics here: age group and gender.[9] Although patient pools could still differ in unobservable manners, we believe that this test is informative in identifying the non-random assignment of patients. We proceed by regressing the patient characteristic of interest (gender or age group) on physician dummies, separately for each ED. We then test whether the MD dummy variables are statistically different from each other.

We present results from F-tests for gender and age group, in the first two columns of Table 2. Of the seventeen EDs, we find evidence that 9 do not assign their patients purely randomly with

[9]Other patient characteristics such as the probability of a certain diagnosis might be assigned in different ways by different ED physicians.

Table 2: F-tests for randomness

	Gender	Age	Gender(restricted)	Age(restricted)
1	1.87	21.34*	1.50	1.88
2	2.17	25.54*	2.63	10.25*
3	2.18*	85.36*	2.28*	3.69*
4	1.43	0.29	1.55	0.26
5	1.91*	2.63*	1.72	2.38*
6	4.66*	6.02*	5.04*	3.83*
7	12.70*	2.34	2.86	2.18
8	2.35*	7.62*	1.86	2.41
9	1.48	1.14	1.48	1.14
10	6.39*	3.12*	2.25	1.64
11	6.81*	19.38*	7.78*	3.67*
12	1.80	7.14*	2.00	5.34*
13	2.34*	41.54*	1.81	3.54*
14	2.20	9.51*	2.39	8.28*
15	1.96	4.50*	1.80	1.91
16	4.54*	6.73*	2.30*	4.41*
17	1.03	2.17*	1.06	1.73

respect to their gender (EDs 3, 5, 6, 7, 8, 10, 11, 13, 16) while we find evidence against random assignment with respect to their age for most hospitals (except for EDs 4, 7 and 9). In order to get a sense of whether differences uncovered above are meaningful, we examine the differences between gender mean and age mean patient assignments for an individual physician and the ED averages at the ED at which the physician works. These distributions (omitted for compactness) suggest that F-test results may be driven in part by a few atypical physicians. The presence of apparent outliers is consistent with the fact that certain older physicians (especially those without advanced emergency-medicine training) may be assigned, on average, less-severely ill patients.

In light of this possibility, we eliminate ED physicians whose gender mean or age mean are significantly different than their own ED peers. More specifically, we drop physicians whose gender mean is more than 5% from the mean for their ED, and whose age group mean is more than 0.5 of an age group around the mean for their ED. Once we exclude these outliers we rerun the same two F-tests as before. Results from these F-tests are presented in the last two columns of Table 2. The F-test results suggest that a total of eight EDs (i.e., EDs 1, 4, 7, 8, 9, 11, 15, 17) now do not violate the randomness assumption on both observable patent characteristics. From these eight EDs, we focus on five EDs in particular: EDs 1, 4, 9, 15 and 17. We use this narrower set of five EDs as only a very few atypical physicians were excluded (i.e., a total of 5 physicians dropped across the 5 chosen EDs).

Table 3 presents summary statistics on the entire sample of EDs and physicians, while Table 4 presents summary statistics on our estimation sample. By comparing these tables, one can see that we retain 71 ED physicians out of a total of 280 for the angina+ and TIA+ samples. Our final sample is similar to the full set of EDs in terms of most variables.

In our final sample, mean physician-related spending during ED visits is about \$49 for angina+ and \$64 for TIA+. The expected number of ED revisits over a 90-day period is high. More than ED visits, the costs of hospitalization for both samples are high, and higher for TIA+ than for the angina+ sample. More specifically, the cost of hospitalization for the first 90 days are considerably higher for the angina sample compared to the angina+ sample (with over \$9,690 for the first 90 days for the angina sample compared to \$4398 for angina+ sample).[10] The number of physicians in the angina and TIA samples are fewer than angina+ and TIA+, respectively. This is in part because both are relatively rare (especially TIA) but, most importantly, both are very difficult to diagnose (and we exclude physicians who do not examine someone with TIA or angina).

3 Model and Estimation

3.1 Model

We postulate a simple model of ED patient treatment, health outcomes, and healthcare expenditures. The basic premise of our model is that the most important function of the ED is the appropriate diagnosis and disposition of the patient. An ED that does not adequately recognize signs and symptoms characteristic of life-threatening diseases will send patients home without adequate follow-up treatment, with the risk of further complications. An ED that is overly aggressive in its treatments will result in extra resources being used without delivering any health benefit. Appropriate disposition of the patient depends on the ability to accurately diagnose signs and symptoms and the knowledge of appropriate treatments. Since many diagnostic and therapeutic interventions occur only after discharge from the ED, disposition also depends on the ability to convince patients

[10] That is, at \$3,602 during the first 5 days and \$2,014 during the subsequent 85 days for the TIA+ sample compared to \$2,411 and \$1,987 for angina+ sample, respectively.

Table 3: Summary statistics for all EDs and ED physicians in Montreal

Statistic	Angina	Angina+	TIA	TIA+
Number of EDs	17	17	17	17
Number of ED physicians	217	280	132	280
Mean age	14.08	12.56	15.44	13.02
Mean female	.51	.53	.55	.56
Number of patients	1784	38556	431	8485
Spending during ED visit	66.30	56.28	60.12	69.03
Number of ED revisits 0-5 days	.32	.30	.36	.37
Number of ED revisits 6-90 days	.83	.74	.47	.82
Number of office visits 0-30 days	.72	.64	.83	.61
Number of external clinic visits 0-30 days	1.06	.86	1.61	.99
Number of hospitalizations 0-5 days	.32	.19	.27	.23
Number of hospitalizations 6-90 days	.20	.16	.18	.18
Costs of hospitalization 0-5 days ($)	5068.41	2662.93	3192.13	3786.46
Costs of hospitalizations 6-90 days ($)	2806.74	2086.77	2026.87	2261.83

Note: patient statistics treat each patient/ED encounter as a unique observation.

Table 4: Summary statistics for estimation sample

Statistic	Angina	Angina+	TIA	TIA+
Number of EDs	5	5	5	5
Number of ED physicians	55	71	37	71
Mean age	15.11	12.66	15.48	13.01
Mean female	.46	.54	.48	.59
Number of patients	343	8406	104	1644
Spending during ED visit	51.82	49.21	59.11	63.57
Number of ED revisits 0-5 days	.36	.27	.44	.34
Number of ED revisits 6-90 days	.80	.71	.48	.77
Number of office visits 0-30 days	.66	.60	.62	.61
Number of external clinic visits 0-30 days	1.34	.88	1.64	1.03
Number of hospitalizations 0-5 days	.39	.16	.30	.22
Number of hospitalizations 6-90 days	.22	.17	.21	.19
Costs of hospitalization 0-5 days ($)	6,602	2411	3,115	3,602
Costs of hospitalizations 6-90 days ($)	3,088	1,987	2,190	2,014

Note: patient statistics treat each patient/ED encounter as a unique observation.

to obtain the appropriate treatments.

Our model is meant to be fairly general, and can apply to any case where appropriate diagnosis and disposition of the patient are difficult skills. As discussed in the introduction, our empirical work considers two potentially life-threatening conditions, angina and TIA. For both these conditions, diagnosis might be challenging and confirmation of the diagnosis depends on tests that are often done after release from the ED. This section illustrates the model for TIA, but the same model applies equally well to angina.

Consider a patient i who experiences an illness shock θ_i and then presents at the ED with symptoms that may or may not be indicative of a TIA (such as visual or speech changes, weakness, or numbness). We assume that there are two possible illness shocks: $\theta = \theta^{TIA}$ if the patient has experienced TIA, or $\theta{=}\theta^N$ if the patient has experienced the null disease, which is the absence of a TIA, but may exhibit similar symptoms.[11]

We do not directly specify the information that the ED physician observes, regarding the probability that patient i has θ^{TIA} or θ^N. We allow the possibility that these shocks are fully or partially observable to the ED physician. In our model, the physician, and not the patient, make all decisions, and hence the information known to the patient is not relevant.

We further assume that there are two treatments T. The appropriate treatment for a TIA is given by T^{TIA} and the appropriate treatment for illness θ^N is given by T^N. T^{TIA} is more expensive than T^N, involving referrals for scans for instance.

We now exposit the patient's health production function. The health stock is a latent value H_i^* which we assume is additive in the patient's baseline health endowment, illness shock, and treatment.

The baseline health endowment $\overline{H}_i$ is the endowment prior to the onset of the illness shock. The illness shock θ is given by $\overline{v}(\theta_i) + \varepsilon_i$. The mean component of the health shock is $\overline{v}(\theta_i)$ and $\overline{v}(\theta^{TIA}) < \overline{v}(\theta^N)$ so that TIA confers a worse health shock on average. The ε_i is the deviation of the health shock from the mean level. Denote the mean value of a particular treatment, which is assumed to be illness-dependent, as $v(\theta_i, T_i)$. We assume that: (i) $v(\theta^N, T^{TIA}) = v(\theta^N, T^N)$, so that there is no extra expected value in receiving the intensive treatment for the null illness θ^N, and (ii) $v(\theta^{TIA}, T^{TIA}) > v(\theta^{TIA}, T^N)$, so that there is a positive value in getting an intensive treatment

[11] As noted in Section 2, disease θ^N groups together all conditions which are in TIA+ but not in TIA. Importantly, we assume that the follow-up care that is useful for a TIA is not useful for these other conditions.

given illness θ^{TIA}.

Taken together, the expression for the post-treatment latent (i.e., unobservable) health stock is given by:

$$H_i^* = \overline{H}_i + \overline{v}(\theta_i) + v(\theta_i, T_i) + \varepsilon_i. \tag{1}$$

Although the health stock is unobservable to the econometrician, the consumption of medical services is observable. More specifically, denote H_i as an observable healthcare event, occurring after the initial illness shock and care at the ED, for example a hospitalization. In such an example, the observable hospitalization occurs if $H_i = 1\{H_i^* < 0\}$, that is, the patient is hospitalized if his or her (latent) health falls below a given threshold. Although in this example, the outcome variable is binary in nature (i.e., hospitalized or not), it could as well be a count variable (i.e., the number of hospitalizations) or a continuous variable (i.e., the costs associated with hospitalizations).

Corresponding to our health production function, we also define a dollar expenditure function. Expenditures for patient i are a function of the illness shock, treatment, and a residual term:

$$D_i = e(\theta_i, T_i) + u_i, \tag{2}$$

where: (i) the e function provides the deterministic part in the relationship between the illness-treatment pair at the ED, and costs at the ED, and (ii) D is measured in dollars, and hence observable (unlike the health stock). Unlike with health shocks, we take no stand as to the costs resulting from the different treatments. While T^{TIA} will cost more than T^N in the current ED visit, it is possible that T^N will cost more in the long-run, due to adverse outcomes such as strokes.

Considering the implications of our model, there are three things that can occur to patient i from a health point of view. She might have had the null disease, in which case the treatment does not affect her health status. Or, she might have had a TIA, in which case the T^{TIA} treatment will result in a higher expected final latent health and a lower probability of a stroke (the potentially avoidable adverse outcome) than would an inappropriately low treatment T^N. From an expenditure point of view, her medical expenditures will be higher if she gets T^{TIA} regardless of her underlying health status.

We seek to use our framework and data to evaluate ED physician contribution to desired outcomes (i.e., "skill" or "quality"). In our model, the "quality" (i.e., contribution to desired outcomes) of an ED physician will depend on (1) her ability to distinguish θ^{TIA} from θ^N; and (2) her ability to get patients to receive the appropriate follow-up care, T^{TIA} and T^{N} respectively. While this statement suggests a vertical dimension to quality, quality will also be partly horizontal. In particular, some physicians may overtreat patients with the null disease, resulting in too many procedures and high expenditures, but good health outcomes, while other physicians may undertreat patients with TIA, resulting in low expenditures in the ED but potentially adverse health outcomes down the road.

To understand the role of physician skill, define $P_j(T|\theta)$ to be the probability that physician j picks treatment T when faced with health state θ, and define $P_i(\theta)$ to be patient i's probability of having illness θ. Let $c(i)$ denote the assignment of patient to some physician j.

Patient i's expected (latent) health-stock will then be equal to:

$$\begin{aligned} E[H_i^*] = \overline{H}_i + P_i(\theta^N)(v(\theta^N, T^N) + \overline{v}(\theta^N)) + P_i(\theta^{TIA})(v(\theta^{TIA}, T^N) + \overline{v}(\theta^{TIA})) + \varepsilon_i \\ + P_i(\theta^{TIA})P_{c(i)}(T^{TIA}|\theta^{TIA})(v(\theta^{TIA}, T^{TIA}) - v(\theta^{TIA}, T^N)) \end{aligned} \quad (3)$$

where the expectation here is over the probability of TIA and the physician's treatment action. In (3), the first line is the underlying health status of patient i and the second line is the impact of being treated by a particular physician (i.e., which in turn depends on his or her skills (related to both diagnostic and disposition)). Note that physician skill here only matters to the extent that the patient has TIA, as treatments do not affect health outcomes given that the patient has the null disease (i.e., $v(\theta^N, T^N) = v(\theta^N, T^{TIA})$).

Define $S^H_{ic(i)} \equiv P_i(\theta^{TIA})P_{c(i)}(T^{TIA}|\theta^{TIA})\left(v(\theta^{TIA}, T^{TIA}) - v(\theta^{TIA}, T^N)\right)$. $S^H_{ic(i)}$ is the treatment effect, in terms of expected incremental health benefit, from physician assignment $c(i)$. Hence, it is a measure of the impact of the skill of assignment $c(i)$ (i.e., to physician j) on i's health. Using this definition, we can rewrite (3) as:

$$E[H_i^*] = S^H_{ic(i)} + \overline{H}_i + P_i(\theta^N)\left(v(\theta^N, T^N) + \overline{v}(\theta^N)\right) + P_i(\theta^{TIA})\left(v(\theta^{TIA}, T^N) + \overline{v}(\theta^N)\right) + \varepsilon_i. \quad (4)$$

We can also exposit the expectation of D_i, in a similar fashion to $E[H_i^*]$, where:

$$E[D_i] = S^D_{ic(i)} + P_i(\theta^N)e(\theta^N, T^N) + P_i(\theta^{TIA})e(\theta^{TIA}, T^N) + u_i. \quad (5)$$

Note that the equation underlying $S^D_{ic(i)}$ in (5) is different from $S^H_{ic(i)}$ in (4) because expenditures can be affected by the treatment regardless of the underlying health state. $S^D_{ic(i)}$ is the treatment effect of assignment $c(i)$ in terms of lowering expenditures to patient i.

3.2 Estimation

Using our conceptual model as a basis, we seek to estimate the ED-physician component of different outcome measures. Our idea is to uncover the physician treatment effects across different dimensions, notably costs and health outcomes. We then examine the extent to which these different treatment effects for a given physician correlate with each other. We estimate the fixed effects as empirical counterparts to versions of (4) and (5).

We now discuss our basic regression analyses. Focusing on health status, we parametrize the expected value of the part of health status that is not a function of assignment to ED physician as:

$$\overline{H}_i + P_i(\theta^N)(v(\theta^N, T^N) + \overline{v}(\theta^N)) + P_i(\theta^{TIA})(v(\theta^{TIA}, T^N) + \overline{v}(\theta^N)) = x_i\beta^H, \tag{6}$$

where x_i are observables and β^H is a vector of parameters to estimate. In our data, we observe information on the socio-economic status of the patient, gender, and age, as well as the patient's choice of hospital. These variables all enter into x_i.

We further assume that the physician effects are homogenous across a set of patients $\mathcal{I}$, so that $S^H_{ic(i)} = S^H_{ic(i')}$ for all patients i and i' in $\mathcal{I}$. The interpretation of $\overline{S}^H_j$ is the physician's j contribution or "skill" in producing health outcome, where skill is in quotes because many outcomes, such as (costly) hospitalizations, are generally not associated with a good health outcome.

Substituting (6) into (4), replacing $\overline{S}^H_j$ for $\overline{S}^H_{ic(i)}$, and solving for the realization H^*_i instead of the expectation, we obtain:

$$H^*_i = \overline{S}^H_j + x_i\beta^H + \varepsilon_i, \tag{7}$$

where ε_i captures the random component of health outcomes. Analogously, we can obtain an estimating equation for cost outcomes $C^*_i = \overline{S}^C_j + x_i\beta^C + u_i$, which has a similar derivation and identical form to (7).

Equation (7) forms the basis of our main estimating equation. This specification expresses physician "skill" as inputs into the patient's health production function. We will focus on the physician fixed effects $\overline{S}_j^H$ and $\overline{S}_j^C$ that stem from these regressions.

In general, there is a central complication that makes estimating the role of ED physicians difficult: the underlying health status of the patient may vary and may be systematically correlated with the assignment to ED physician assignment $c(i)$. Patients who present at the ED with symptoms of a particular disease may be assigned to a physician with more expertise in that disease, for instance. If patients are even partially informed regarding their illness, this will then create an endogeneity bias in evaluating the impact of physician quality. We address this endogeneity concern using the fact that the assignment to patients at many EDs in Montreal is quasi-random. The quasi-random nature of the assignment implies that the assignment of physician *within an ED* will not be systematically related to ε_i or u_i. That is, with random assignment of physicians, $\overline{S}_j^H$ will not be correlated with ε_i while $\overline{S}_j^C$ will not be correlated to u_i

We now discuss specifics of our empirical implementation. All our analysis is based on linear regressions. All specifications include fixed effects for each ED physician and the same patient demographics in x_i. Focusing on TIA, our model considers patients who present at the ED with symptoms that could plausibly be indicative of TIA. As in Section 2, denote this set of patients as TIA+. Accordingly, most of our TIA regressions define the sample $\mathcal{I}$ to be the set of TIA+ patients. The underlying assumption is that the set of diagnoses in TIA+ (and equivalently for angina+) is sufficiently broad that a diagnosis of TIA+ would be similar across all ED physicians in a given ED. We estimate multiple versions of (7). The dependent variables vary across specifications and include: (i) the patient's number of revisits to the ED during within 5 days or between 6 and 90 days; (ii) the patient's number of physician office visits as well as external (outpatient) clinic visits during the first 30 days; and (iii) the patient's number of hospitalizations as well as their costs in the first 5 days or between 6 and 90 days. These allow us to estimate each physician's average contribution, or skill/practice-style, related to future healthcare consumption (which serve as proxy to health). We also consider the specification where the dependent variable is (physician-related) spending during the initial ED visit. This allows us to estimate each physician's average contribution (or cost-control skill or practice style) to physician-related costs during the ED visit.

Note that patients may self-select into different EDs based on unobservable attributes of health status, needs and demands. Different areas of the region may also have people with differential likelihoods of TIA given similar symptoms, implying that even if patients select the ED closest to their home, we might see differences across EDs in ε_i. Thus, our data only contain quasi-random assignment of patients conditional on admission to a given ED. Thus, we can only identify the level of the skill relative to the mean level at that ED: the interpretation of $\overline{S}_j^H$ is of the skill level relative to other physicians at the same ED.

Using the TIA+ sample for $\mathcal{I}$, the interpretation of the fixed effects is as the skill of the physician in terms of *both diagnosis and disposition of patients*. In some cases, such as revisits to an ED, we view a positive $\overline{S}_j^H$ fixed effects as unambiguously negative. In other cases, such as office visits and (outpatient) hospital-based external clinic visits, a positive coefficient is likely positive, as an office or outpatient visit might constitute a useful follow-up that would prevent future ED visits or hospitalizations. Coefficients on future costs regressions are more ambiguous, as costs within a relatively short term might forestall future costs.

We also estimate regressions where $\mathcal{I}$ includes just patients who are diagnosed with TIA (as opposed to TIA+) in the initial ED visit. The interpretation of these coefficients is as the skill of the ED physician for *enacting the appropriate disposition for the patient given the diagnosis*. We note that the variation for these regressions is less plausibly exogenous than for the TIA+ sample: physicians who are better at diagnosing TIA from the TIA+ patient pool may be faced with a different underlying mix of patients with the TIA diagnosis than physicians who are worse at diagnosing TIA.

Finally, we can obtain a measure of the importance of diagnosis skills relative to disposition skills related to a particular outcome, by examining the gap between (1) the TIA+ fixed effect (which indicates diagnosis and disposition together) and (2) the TIA fixed effect (which indicates disposition conditional on diagnosis).[12]

4 Results

We now consider the impact of ED physicians across different measures and diseases by reporting statistics on $\overline{S}_j^H$ and $\overline{S}_j^C$. First, we consider the mean physician effects for the angina+ sample. Each regression includes fixed effects for each physician in the sample and controls for demographics. Because of the number of coefficients and lack of comparability across hospitals, we do not report

[12]We can also calculate the diagnosis probability as the probability of a TIA diagnosis conditional on TIA+. However, this measure is neither unambiguously positive or negative, as variation might indicate overdiagnosis of TIA, underdiagnosis of TIA, or misdiagnosis of TIA. Nonetheless, it is informative as to the variation in diagnosis ability across physicians.

the coefficient values.

Figure 1 presents the probability density function of physician fixed effects (i.e., demeaned average contribution) associated with spending during the initial ED visit for the angina+ sample. Figure 2 presents the probability density function of physician fixed effects associated hospital spending in the first 5 days, also for the angina+ sample. Figures 3 and 4 do present the same probability density functions but for the TIA+ sample of patients.

Figures 1 and 3 suggest that physicians vary in their average contribution (i.e., their fixed effect) in terms of (physician-related) costs during the initial ED visit for both the angina+ and TIA+ samples, respectively. That is, when presenting with symptoms that may or may not be indicative of a TIA (or angina), physician's vary in the average contribution to ED costs during the initial ED visit (relative to mean contribution at the same ED). Furthermore, the variation is greater in the TIA+ sample than in the angina+ sample. This may be driven by the underlying differences in diagnosing and treating TIA+ patients relative to angina+ patients (which we discuss below when examining hospitalizations). Similarly, figures 2 and 4, suggest that physicians vary in terms of their average contribution (i.e., their fixed effect) to hospital spending in the first 5 days following the initial ED visit. Although each of these density functions exhibit considerable "lumping" of physicians' "skills" at the centre, the tails nonetheless suggest potentially important variation. This suggests that spending during the initial ED cost and subsequent hospital spending during the first 5 days would likely vary considerably if patients were to be treated by either ends of the "skills/practice-style" of the distribution.

Table 5 reports the mean difference between the 75th and 25th percentile of physician effects across each hospital. Each row of Table 5 provides statistics on one regression.

We find that the 75th percentile physician bills about $9 more than the 25th percentile physician for the angina+ sample of patients. Note that the billing data for EDs reflects only physician-related costs (i.e., fee-for-services for procedures performed by physicians) in the ED. Thus, for instance, if a ED physician orders a CT scan, this would likely be done by the hospital and not billed (or associated with) with the patient .[13] There is also a substantial interquartile range in the costs of hospitalization across ED physicians, of $859 in the 0-5 day window and $959 for the 6-90 day window. There is a mean of 0.093 in the interquartile range for the 0-5 day mean ED revisits, a value that grows to 0.206 for the 6-90 day window. Similarly, there are substantial differences in

[13] In fact, all non-physician related costs are generally covered by the hospital's global budget.

Figure 1:

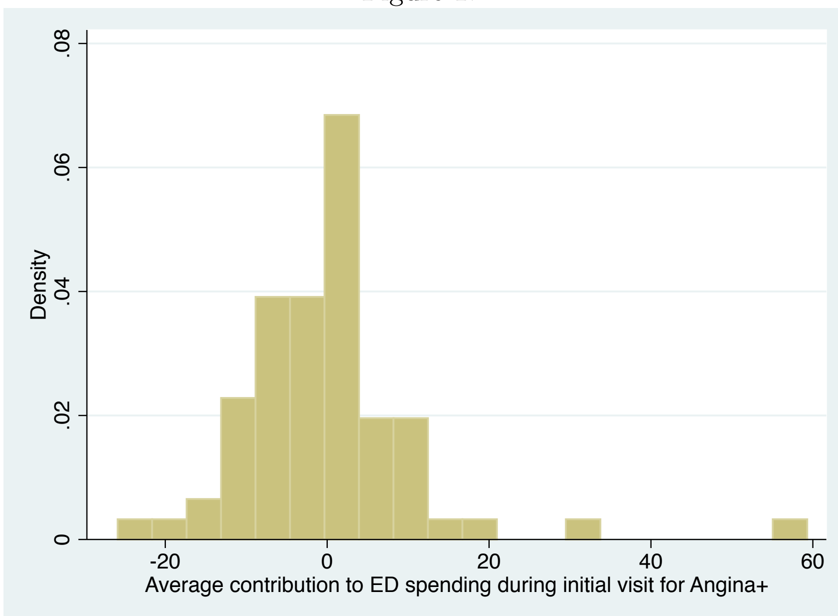


Figure 2:

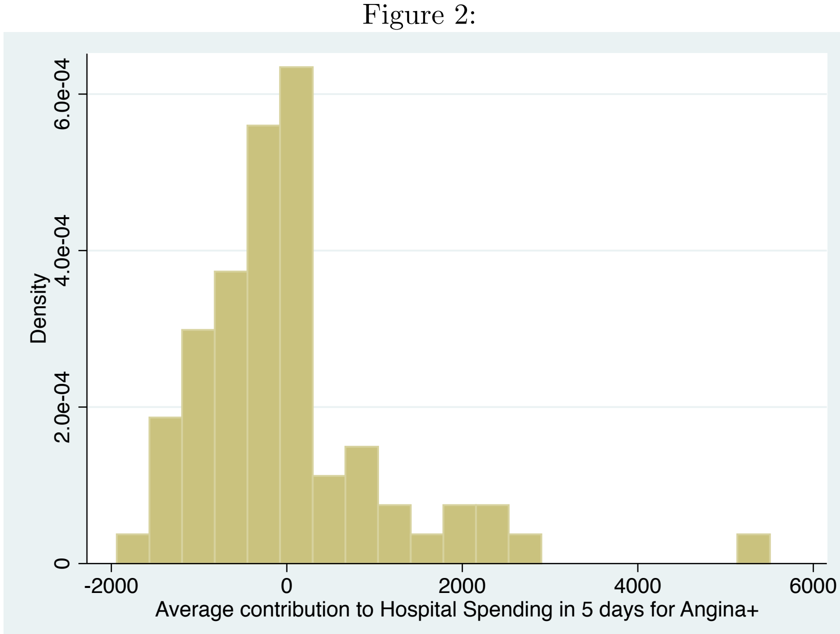


Figure 3:

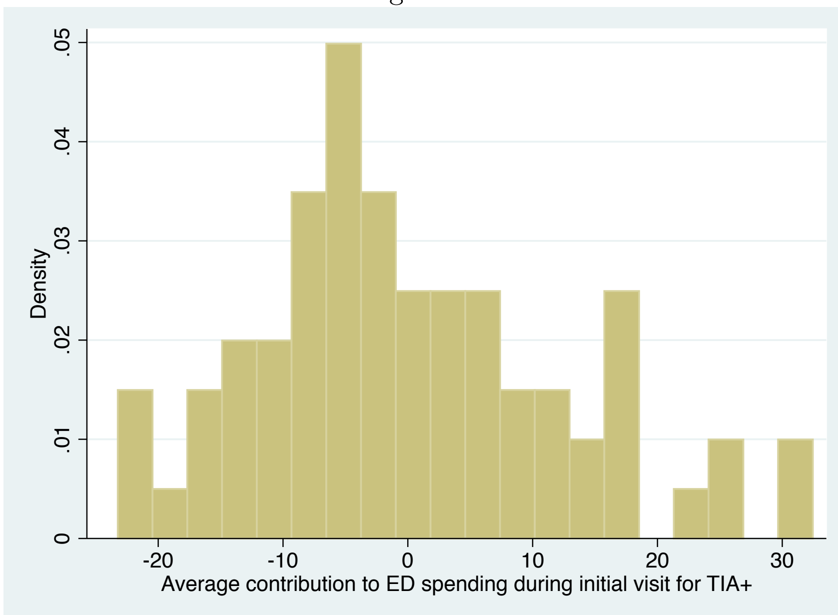
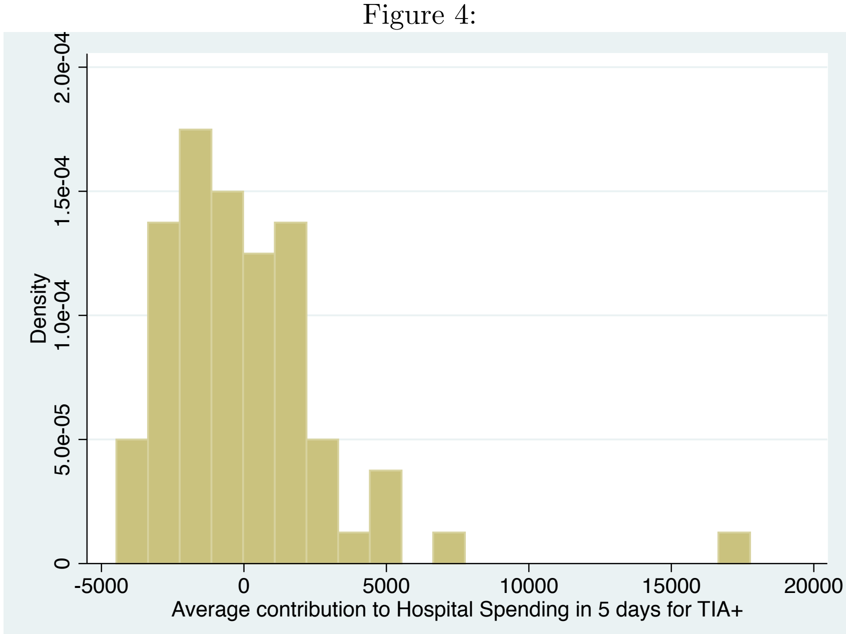


Figure 4:



the number of office visits, external clinic visits, and hospitalizations between the 75th and 25th quartiles.[14]

Table 5: Difference between 75th and 25th percentile physician for angina+

Dependent variable	Mean difference between 75th & 25th percentile
Spending during ED visit ($)	9.11
Number of ED revisits 0-5 days	0.093
Number of ED revisits 6-90 days	0.206
Number of office visits 0-30 days	0.200
Number of external clinic visits 0-30 days	0.146
Number of hospitalizations 0-5 days	0.062
Number of hospitalizations 6-90 days	0.063
Costs of hospitalization 0-5 days ($)	859
Costs of hospitalizations 6-90 days ($)	959

Note: each row provides the mean, across hospitals, of the difference between the 75th and 25th percentile of the given statistic, and of the standard deviation of the difference.

Next, we consider the mean physician effects for the TIA+ sample. Table 6 presents analogous results to Table 5. Overall, the differences across physicians for TIA+ are quantitatively larger than for angina+. We find that the 75th percentile physician bills about $15 more than the 25th percentile physician for the TIA+ sample. Here, the interquartile range in the costs of hospitalization across ED physicians is also larger, $3,065 in the 0-5 day window and $1,929 for the 6-90 day window. The interquartile range in ED revisits across these physicians is also quite large, at 0.155 for 0-5 days and 0.160 for 6-90 days. Similarly, there is a large interquartile range in the number of office visits, external clinic visits, and hospitalizations.

Note that the differences in interquartile ranges for physicians fixed effects for TIA+ is much larger than for angina+. For instance, Table 4, shows that the mean cost of hospitalization within 90 days for the angina+ sample is $4,398 (i.e., $2,411+$1987) while the mean cost of hospitalization within 90 days for the TIA+ sample is $5,616 (i.e., $3602+$2014). This is not surprising given that identifying the medical causes behind symptoms in TIA+ is typically a more uncertain and ambiguous endeavor then identifying the medical causes behind symptoms in angina+. This is because angina+ typically concerns symptoms referable to organs amenable to direct evaluation

[14] Because of the likely role of the tails of the distribution, we will also consider the 90th and 10th percentile difference. Furthermore, the mean of the standard deviation of differences for both the 75th vs the 25th percentile and for the 90th and 10th percentile will be added.

(heart, esophagus, stomach, chest wall) while TIA+ invokes a variety of symptoms pertaining to the entire brain and nervous system, both much more complex. This complexity is likely driving some of the interquartile differences in physician spending during the initial ED visit.

Table 6: Difference between 75th and 25th percentile physician for TIA+

Dependent variable	Mean difference between 75th & 25th percentile
Spending during ED visit ($)	15.189
Number of ED revisits 0-5 days	0.181
Number of ED revisits 6-90 days	0.500
Number of external clinic visits 0-30 days	0.443
Number of office visits 0-30 days	0.274
Number of hospitalization 0-5 days	0.155
Number of hospitalization 6-90 days	0.160
Cost of hospitalization 0-5 days ($)	3,065
Cost of hospitalization 6-90 days ($)	1,929

Note: each row provides the mean, across hospitals, of the difference between the 75th and 25th percentile of the given statistic, and of the standard deviation of the difference.

Having considered the means of the physician effects, we now turn to examining the relation between the different effects across a physician. Here, we correlate different measures of the physician fixed effects $\overline{S}_j^H$ and $\overline{S}_j^C$ with each other. Because we cannot compare these effects across hospitals, we difference each physician's effect from the mean level at the ED at which she works. The correlations that we report are then performed on the means.

Table 7 reports the correlation across different physician fixed effects for the angina+ sample, while Table 8 reports the correlations for the TIA+ sample. We see a number of patterns in common across these two diseases. First, physicians who have higher (physician related) spending in the initial ED visit also have higher number of revisits to the ED within the subsequent 5 days. In particular, the correlation between the initial ED spending and the number of ED revisits within 5 days is 0.273 for the angina+ sample and 0.299 for the TIA+ sample and statistically significant in both cases. There is also a correlation between ED physicians who spend more in the initial ED visit and the costs of inpatient hospitalization for admissions made during the first 5 days after the initial ED visit. For the TIA+ sample, this correlation is 0.198 and statistically significant while for angina+, it is 0.133 but not statistically significant.

Taken together, these results suggest that physicians who have higher spending during the initial visit also exhibit greater costs (i.e., more visits to the EDs and greater hospitalization costs) within the first 5 days after the ED consultation. These results is in line with Doyle et al. (2010) who finds, examining two teams of residents in a particular hospital, that more spending by physicians is not necessarily associated with better outcomes and, that spending may substitute, in part, for diagnostic acumen. Interestingly, the correlation effects seem to play out in the 5 day period following the initial ED visit. That is, there is no significant correlation between physicians' average ED (physician-related) spending during the initial visit and average contribution to revisits or hospitalizations during the 6 to 90 day period.

Table 7: Correlation in physician fixed effects for angina+ patients

	ED revisits 0-5 days	ED revisits 6-90 days	Office visits 0-30 days	External clinic visits 0-30 days	Inpatient costs 0-5 days ($)	Inpatient costs 6-90 days ($)
Spending during ED visit ($)	0.273**	−0.107	0.0402	0.153	0.133	0.0576
ED revisits 0-5 days		0.120	−0.216*	0.209*	−0.0714	0.0711
ED revisits 6-90 days			0.184	−0.0720	0.0549	0.0913
Office vists 0-30 days				−0.265**	−0.0980	−0.180
External clinic visits 0-30 days					0.123	0.0758
Inpatient costs 0-5 days ($)						0.207*

Note: *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Another pattern that appears from Tables 7 and 8 is a negative relation between ED revisits within the first 5 days and office visits within the first 30 days. In both cases, physicians who succeed in having patients visit physician offices see far fewer ED revisits among their patients. The correlations here are −0.216 for angina+ and −0.277 for TIA+; both are statistically significant. Thus, we find that there is differential efficacy across ED physicians in getting patients to avoid ED revisits, and that this skill correlates with the ability to get patients to obtain appointments with physicians for office visits. Interestingly, external clinic visits appear to work in the opposite direction from physician visits. In particular, for both angina+ and TIA+, physicians with more external clinic visits also have more ED revisits within 0-5 days. Because we do not observe who may have referred the patient to the external hospital-based clinic (the physician during the initial

Table 8: Correlation in physician fixed effects for TIA+ patients

	ED revisits 0-5 days	ED revisits 6-90 days	Office visits 0-30 days	External clinic visits 0-30 days	Inpatient costs 0-5 days ($)	Inpatient costs 6-90 days ($)
Spending during ED visit ($)	0.299**	0.111	-0.134	0.161	0.198*	-0.0155
ED revisits 0-5 days		-0.0398	-0.277**	0.254**	0.168	-0.00677
ED revisits 6-90 days			0.0714	0.0178	-0.0783	0.292**
Office vists 0-30 days				0.0449	-0.121	-0.212*
External clinic visits 0-30 days					0.358***	0.113
Inpatient costs 0-5 days ($)						0.233**

Note: *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

visit or during the revisit), it is difficult to uncover the exact reason for this. In addition, physicians with greater hospitalizations costs in the first 5 days are also associated with more external clinic visits in the first 30 days for TIA+. The positive correlation between the two suggest that they may act as substitutes to each other in terms of outcomes. External clinics are located at hospitals but do not involve a hospital admission.

We postulate that there are multiple types of physicians and that physicians who spend more on the initial ED visit are more conservative in their diagnosis and management and tend to hospitalize patients more in conjunction with the initial visit. They may also encourage patients to return to the ED if symptoms reoccur. These physicians may also preferentially direct patients to specialists in the external clinic rather than back to a primary care provider. However, it is also possible that external clinic visits are hard to schedule. ED physicians who routinely suggest follow up in an external clinic may end up with many patients who are unable to schedule such visits in a timely manner and hence end up being hospitalized or revisiting the ED.

Having analyzed the correlation of physician effects for the two sample of angina+ and TIA+, we next seek to understand the relation between physician effects across these two illness samples. Table 9 reports correlations in physician effects across the two samples. We choose here a smaller set of correlations to report than in Tables 7 and 8 due to space constraints.

The most striking finding here is the strong correlation of cost types across the two samples.

Table 9: Correlation in physician fixed effects across Angina+ and TIA+ patients

	Spending during ED visit TIA+	ED revisits 0-90 days angina+	ED revisits 0-90 days TIA+	Hospital spending 0-90 days angina+	Hospital spending 0-90 days TIA+
Spending during ED visit angina+ ($)	0.609***	-0.00564	-0.0634	0.129	-0.288**
Spending during ED visit TIA+ ($)		0.191	0.194	0.128	0.163
ED revisits 0-90 days angina+			-0.193	0.0769	0.316***
ED revisits 0-90 days TIA+				-0.113	0.0787
Hospital spending 0-90 days angina+ ($)					-0.0255

Note: *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Physicians who spend more in the initial ED visit for the angina+ sample also spend more for the TIA+ sample. The correlation here is 0.609 and statistically significant. This would suggest that physician practice styles (measured by spending on physician services) are correlated across patients seen, where some physicians are generally more "conservative" in their approach than others.

There is much weaker evidence of correlation in outcomes: the correlation between the physician effect for ED revisits for angina+ and for hospital spending for TIA+ is 0.316 and statistically significant. As both of these are considered "bad" outcomes in general, this positive relationship may be suggestive that some skills (or lack thereof) are general and common across patient populations. Other coefficients are not statistically significant and in some cases negative. For instance, the 0-90 day hospital spending fixed effects between the two conditions has a negative correlation of -0.113, albeit one that is not statistically significant.

We next turn to the question of the correlations between the angina and angina+ samples and between the TIA and TIA+ sample. Recall that the "plus" samples capture skills based on diagnosis and disposition, while the angina and TIA samples capture skills at disposition conditional on a diagnosis, with the assumption that the distribution of unobservables ε_i (or u_i) among the diagnosed angina (or TIA) population is orthogonal across ED physicians within an ED. Tables 10 and 11 report these correlations for angina/aingina+ and TIA/TIA+ respectively.

Examining Table 10, we find that physician ED spending (during the initial ED visit) across the angina+ population and angina population are positively correlated. This would suggest that

Table 10: Correlation in physician fixed effects across angina and angina+ patients

	Spending during ED visits angina+	ED revisits 0-90 days angina	ED revisits 0-90 days angina+	Hospital spending 0-90 days angina	Hospital spending 0-90 days angina+
Spending during ED visit angina ($)	0.382***	0.382***	0.0883	-0.0478	-0.111
Spending during ED visit angina+ ($)		0.0206	-0.00564	-0.0326	0.129
ED revisits 0-90 days angina			0.310**	0.0815	0.138
ED revisits 0-90 days angina+				-0.151	0.0769
Hospital spending 0-90 days angina ($)					0.290**

Note: *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 11: Correlation in physician fixed effects across TIA and TIA+ patients

	Spending during ED visit TIA+	ED revisits 0-90 days TIA	ED revisits 0-90 days TIA+	Hospital spending 0-90 days TIA	Hospital spending 0-90 days TIA+
Spending during ED visit TIA ($)	0.405**	0.0826	-0.257	-0.159	-0.271
Spending during ED visit TIA+ ($)		-0.195	0.194	-0.109	0.163
ED revisits 0-90 days TIA			-0.00607	0.277*	-0.185
ED revisits 0-90 days TIA+				-0.294*	0.0787
Hospital spending 0-90 days TIA ($)					0.334**

Note: *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

physicians who are associated with more ED spending when considering all patients who may or may not have angina are also associated with more ED spending when considering patients who are diagnosed with angina. This result in return suggests that differences are not uniquely driven by the actual diagnosis but also, in part, the physician's practice style. We find the same positive relationship when examining the TIA+ and TIA samples which are presented in Table 11).

Next, consider the correlation of skills associated with a particular outcome (in terms of future consumption of medical services) across the angina+ and angina, and across the TIA+ and TIA sample, respectively. In such a setting, an absence of correlation between a physician's average contribution to a particular outcome (like ED revisits within 5 days) for the TIA+ sample and his or her average contribution to the same outcome but for the TIA sample would suggest that the actual diagnosis of the TIA is at the root of the difference. That is, the absence of a correlation would suggest that physicians who exhibit relatively more ED revisits within 5 days for the TIA+ population no longer do so once the TIA diagnosis has been made. On the other hand, a correlation of one would suggest the physician's contribution to a particular outcome (like ED revisits within 5 days) is invariant to the actual specific diagnosis of TIA. This in turn would suggest that differences in skills (based on a particular outcome) are not driven by differences in diagnosis skills but rather differences in disposition skills.[15]

We find that physicians associated with more ED revisits during the first 90 days when treating the angina+ population are also associated with more ED revisits during the first 90 days when treating the angina population (with a correlation of 0.310 which is statistically significant). The fact that the correlation is neither null nor 1 is suggestive that both diagnostic and disposition are at play. When examining the same correlation but for the TIA and TIA+ sample, we find no statistically significant effect. This would suggest that physician's contribution to outcomes is driven almost uniquely by diagnosis. We also find a positive correlation between hospital spending in 0-90 days for angina and angina+ of 0.310 (which is statically significant). Similarly, we find a positive correlation between the two 0.334 (which is statistically significant) when considering the TIA+ and TIA samples.

The takeaway from these results is that there is a lot of correlation between the skills of diag-

[15]We also find a positive correlation between a physician's average contribution to ED revisits within 90 days for the TIA+ sample and the hospital costs for the TIA sample. This is not surprising as bad health outcomes are likely to lead to more revisits to the ED for the TIA+ sample (those with and without TIA) and, consequently, more hospitalizations and costs for those suffering from TIA.

nosis and disposition together (the "plus" samples) and disposition conditional on diagnosis (the angina/TIA samples). Given that the correlation between these fixed effects are, in all cases, less than 0.5, suggests that there is also a substantial ED physician component to diagnosis. In particular, the fact that the correlations between (1) disposition and (2) diagnosis and disposition are less than 0.5 implies that variation in diagnosis skills may be quantitatively more important than variation in disposition skills.

5 Conclusion

In this paper, we have used detailed data from emergency departments in Montreal, Quebec, Canada in order to understand how ED physicians contribute to healthcare costs and outcomes. A central advantage of our data is that we observe the assignment of ED physician and observe future interactions with the provincial medical system within 90 days of the initial ED visit. We use the fact that the assignment of ED physicians in Montreal, conditional on choice of ED, is close to random. This allows us to control for unobservable health status of patients by including physician effects. We develop this point fully in a theoretical model.

Using our data, we compute the physician effect on various outcomes for two samples of patients, those who are diagnosed with conditions that might indicate angina (angina+) and those who are diagnosed with conditions that might indicate a transient ischemic attack (TIA+). For both conditions, we find that physicians who have high spending in the initial ED visit also have high future interactions with the medical system in terms of ED revisits and costs of hospitalization.

We also find that ED physicians whose patients are likely to have office visits for physicians following the initial ED visit also have fewer ED revisits. This suggests that the skill of helping patients obtain referrals to physicians in an office setting may help lower the prevalence of ED treatment. We also find that physicians who provide costly care for angina also are likely to provide costly care for TIA, but that the physician effects on the outcomes do not correlate across conditions. Finally, we seek to separate out the skills of patient diagnosis and patient disposition, by examining the physician effects for the patients with angina and TIA against the effects for the broader samples of angina+ and TIA+. We find evidence that both diagnosis and disposition are important, with suggestive evidence that diagnosis is more important for both diseases.

References

Abaluck, J. and Agha, L. (2014). Negative tests and the efficiency of medical care: What determines heterogeneity in imaging behavior? Working Paper, Yale.

Doyle, J., Graves, J., Gruber, J., and Kleiner, S. (2014). Measuring returns to hospital care: Evidence from ambulance referral patterns. *Journal of Political Economy*, Forthcoming.

Doyle, J., Wagner, T., and Ewer, S. (2010). Returns to physician human capital: Analzying patients randomized to physician teams. *Journal of Health Economics*, 29:866–882.

Dranove, D., Kessler, D., McClellan, M., and Satterthwaite, M. (2003). Is more information better? the effects of "report cards" on health care providers. *Journal of Political Economy*, 111(3):555–588.

Epstein, A., Ketcham, J., and Nicholson, S. (2010). Specialization and matching in professional services firms. *Rand Journal of Economics*, 41:811–834.

Friedmann, P., Jin, L., Karrison, T., Hayley, D., Mulliken, R., and Chin, J. W. M. (2001). Early revisit, hospitalization or death among older persons discharged from the ed. *The American Journal of Emergency Medicine*, 19:125–129.

Hastings, S., Whitson, H., Purser, J., Sloane, R., and Johnson, K. (2009). Emergency department discharege diagnoses and adverse health outcomes among older adults. *Journal of the American Medical Association*, 57:1856–1861.

Kachalia, A., Gandhi, T., Puopolo, A., Yoon, C., Thomas, E., Griffey, R., Brennan, T., and Studdert, D. (2007). issed and delayed diagnoses in the emergency department: a study of closed malpractice claims from 4 liability insurers. *Annals of Emergency Medicine*, 49:196–205.

Pope, J., Aufderheide, T. P., Ruthazer, R., Woolard, R. H., Feldman, J. A., Beshansky, J. R., Griffith, J. L., and Selker, H. P. (2000). Missed diagnoses of acute cardiac ischemia in the emergency department. *New England Journal of Medicine*, 342:1163–1170.

Pope, J. and Edlow, J. (2012). Avoiding misdiagnosis in patients with neurological emergencies. *Emergency Medicine International*, 2012.

Porter, M. E. (2010). What is value in health care. *New England Journal of Medicine*, 363:2477–2481.

Rothrock, S. and Pagane, J. (2000). Acute appendicitis in children: emergency department diagnosis and management. *Annals of Emergency Medicine*, 36:39–51.

Rothwell, P., Giles, M., Chandratheva, A., Marquardt, L., Geraghty, O., Redgrave, J., Lovelock, C., Binney, L., Bull, L., Cuthbertson, F., Welch, S., Bosch, S., Carasco-Alexander, F., Silver, L., Gutnikov, S., and Mehta, Z. (2007). Effect of urgent treatment of transient ischemic attack and minor stroke on early recurrent stroke (express study): a prospective population-based sequential comparison. *The Lancet*, 370:1432–1442.

Taubman, S. L., Allen, H. L., Wright, B. J., Baicker, K., and Finkelstein, A. N. (2014). Medicaid increases emergency-department use: Evidence from oregon's health insurance experiment. *Science*, 17(17):263–268.