THE EFFECT OF GENERALIST-SPECIALIST INTEGRATION ON PHYSICIAN BEHAVIOR, MEDICAL OUTCOMES, AND MEDICAL CARE SPENDING

Michael Punzalan

A DISSERTATION

in

Health Care Management & Economics

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

Robert J. Town
Gilbert and Shelley Harrison Professor, Professor of Health Care Management

Graduate Group Chairperson

Eric Bradlow, K.P. Chao Professor; Professor of Marketing, Statistics, and Education

Dissertation Committee:

Lawton R. Burns, James Joo-Jin Kim Professor, Professor of Health Care Management
Guy David, Associate Professor of Health Care Management
Mark V. Pauly, Bendheim Professor, Professor of Health Care Management, Professor of Business Economics and Public Policy
Robert J. Town, Gilbert and Shelley Harrison Professor, Professor of Health Care Management

THE EFFECT OF GENERALIST-SPECIALIST INTEGRATION ON PHYSICIAN BEHAVIOR, MEDICAL OUTCOMES, AND MEDICAL CARE SPENDING

To my parents Crispino and Maria, for providing the best for me

To my children Micah and Maila, for filling my life with joy

To my wife Carmi, for making my life complete

ACKNOWLEDGEMENT

I owe a large debt of gratitude to my department. I thank my committee members Rob Burns, Guy David, Mark Pauly, and Bob Town for their guidance and support throughout my project. I gratefully acknowledge the help of our departmental staff, particularly Joanne Levy, in helping me navigate the administration surrounding my degree.

I acknowledge my data providers for their role in making my research possible. I first thank Richard Creighton at SK&A for choosing to deal with me, a graduate student sharing a tiny grant for data acquisition. I also thank the Penn Workforce Working Group for allowing me access to the much larger extract of the SK&A data I analyzed for this dissertation. I thank the leadership and staff of the Health Care Cost Institute (HCCI) for granting me access to their claims database. I must also thank Adam Rifkind of the Office of Research Services at Penn for negotiating Penn's interest (and mine) in the HCCI data use agreement. Finally, I thank the Johns Hopkins ACG® System staff for granting me access to their severity scorer.

I am particularly indebted to a group of technologists at Wharton. I gratefully acknowledge the support of Wharton Computing's Research and Analytics team, particularly Hugh MacMullan. Without access to the Wharton Grid, this project was literally impossible. Hugh arranged a HPC Research Data Storage Grant for me when I lacked funding. In addition, Hugh, Gavin Burris, and Premal Vora helped me substantially increase the performance of my code during my project. I am extremely grateful for this team's support and apologize in advance to those I have mistakenly excluded from these acknowledgements.

I thank several family members and friends for their assistance. Merilee, Brian, Dan, Liz, Kris, Melodee, Melanie, and Adrian, thank you for reaching out to me and, in some cases, my family as we moved into, around, and out of Philadelphia. I thank my in-laws, Angelito, Esther, Margo, and Alister (and Kaylene) for helping Carmi and for watching our children when help was needed.

I am eternally grateful to my family for their unwavering belief in me. To my parents Maria and Crispino, I am thankful for a lifetime of support. Now that I have my own children, I understand how difficult I was to raise. Yet you always sought my best interests and I will do the same for your grandchildren. To my brothers Nicholas and Isaac, I am thankful for a lifetime of understanding. We tell the same stories every time we see each other, but somehow the stories become funnier every time we tell them. I thank my sister-in-law Phoebe for making sure we see each other and for laughing with, or more likely, at us. Both are fine by me.

To my daughters Micah and Maila, I thank you for making our family complete. Life was far less complicated and demanding before you arrived, but it is infinitely better now that you are here. To my wife Carmi, I thank you most of all. You made sure I saw this through to the end. You were my sounding board, my biggest cheerleader, and in the times I wanted to quit, the one who pushed me forward. I could not have done this without you.

ABSTRACT

THE EFFECT OF GENERALIST-SPECIALIST INTEGRATION ON PHYSICIAN BEHAVIOR, MEDICAL OUTCOMES, AND MEDICAL CARE SPENDING

Michael Punzalan

Robert J. Town

I study the effects of integration among generalist and specialist physicians on referral and treatment choices, inpatient admissions, and the cost of medical care. I first construct a novel data set that links the structure of physician organizations to physician behavior for two groups of individuals deemed most likely to benefit from generalist-specialist integration: individuals diagnosed with hypertension and diabetes mellitus. Due to differential selection across integrated and non-integrated practices, ordinary least squares regression is unlikely to identify the effects of interest. To surmount the identification problem, I simulate random assignment of beneficiaries to physician organizations in an instrumental variables framework. My estimates confirm that integration among generalist and specialists significantly affects referral decisions and treatment patterns. I fail to detect evidence that this kind of integration improves medical outcomes or reduces medical care spending. These findings strongly suggest that policies increasing generalist-specialist integration in piece-rate or fee-for-service environments are unlikely to achieve their goals. Finally, I consider the theoretical literature that might rationalize generalist-specialist integration patterns. I show that most potential explanations are inconsistent with my evidence, save for David and Neuman (2011), suggesting that task adhesion should be a focal point for future research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Introduction

Introduction

In this dissertation I estimate the effect of integration between generalist and specialist physicians on physician behavior, health care outcomes, and medical care spending. Estimates of these effects are important to health economics and health care policy.

First, physicians are arguably the most influential actors in the US health care system. In the hierarchy of health providers they rank highest, and lesser-credentialed providers such as nurses, nurse practicioners, and physician assistants, are expected to defer to their judgment. Due to massive differences in human capital, information asymmetry between physicians and patients causes patients to depend greatly on physician choices. Therefore, physician behavior is a first-order scholarly concern.

Second, there is great interest among health care policymakers in altering how physicians, as well as other providers, self organize. For example, the Patient Protection and Affordable Care Act of 2010 (ACA) provides significant financial incentives for provider integration through programs such as the Medicare Shared Savings Program.[1] Such interest exists due to the belief that integration, particularly of generalist and specialist physicians, solves important clinical coordination problems.

Lack of data has historically prevented researchers from empirically characterizing the effect of generalist-specialist integration. The appearance of new data on physician behavior and physician organizations, however, has made research like this dissertation possible.

[1] See `https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ACO/`.

The case for scope economies in medical practice

Clinical rationale

Proponents of provider integration argue that greater economies of scope in medical practices may increase the quality of "care coordination." Though the term has no precise definition, it generally refers to problems in health care production when there is more than one provider.[2]

Historically speaking, greater of economies of scope in medical practice can be traced back to the late 19th century, to the early history of the Mayo Clinic.[3] The Mayo Clinic was founded by William Worrall Mayo and his sons William James and Charles Horace. By the end of the late 1800s, their practice had become successful enough to warrant expansion. Instead of hiring more surgeons, which would strictly be an increase in scale, they hired more generalists. The Mayos tasked them with a filtering problem, to identify patients coming through the practice who would benefit from surgery.[4]

In the present day, concerns about care coordination are often raised for patients with chronic illnesses like diabetes and hypertension. These patients often require treatment from more than one provider, specifically a generalist as well as one or more specialists, and in this setting it can be difficult to ensure that a coherent medical care plan is implemented.[5] Low-quality care coordination is claimed to cause a wide-variety of adverse outcomes: lower quality care, higher health care costs, preventable hospital admissions and readmissions, unnecessary emergency room visits, medical errors, repeated or unnecessary tests, and poor outcomes.[6]

[2]See Chapter 2, McDonald et al. (2014).

[3]The information in this section is drawn from Starr (1982), an authoritative history.

[4]In today's terms we might call this activity "gatekeeping," e.g. see http://www.ncbi.nlm.nih.gov/pubmed/10359396.

[5]See Audet and Sinha (2012) and the cites therein.

[6]See https://ruralhealth.und.edu/projects/hwic/pdf/care-coordination.pdf.

Economic rationale

Arm's length dealing might also create significant financial disincentives to coordination. Consider a young female with Irritable Bowel Syndrome (IBS) who reports her chronic diarrhea and abdominal pain to her generalist. The generalist may obtain stool studies to rule out an infection and a blood test to rule out Celiac disease. He may also recommend dietary changes. If the patient's symptoms persist and the tests are negative for Celiac, the generalist may at that point diagnose her with IBS – if he has not already – and focus on relieving her symptoms with Immodium and anti-spasmodic drugs.

At any point during this process the generalist could refer the patient to a gastroenterologist. He may refer for clinical reasons. He may also retain the patient for financial reasons. If the generalist practices alone and performs tests, he expends some effort and the insurer pays him for his work. If he refers he expends no effort and is paid nothing. But if the generalist practices with a specialist and revenues are shared, the generalist's opportunity cost is lower. Prior research has shown that some health care providers integrate to change referral patterns.[7]

Integrating to change referral patterns also induces two moral hazard problems (see Holmstrom, 1982). Because the generalist in an integrated practice is paid whether he treats or not, he may refer too many patients to the specialist. For example, he may refer a patient with only mild symptoms. He may also shirk by providing insufficient effort, e.g. not ordering stool and blood studies before referring her to a specialist.

The prevalence of multispecialty practice

In Table 1 I show that multispecialty practice is economically significant and its prevalence has increased slightly during my sampling frame. In 2009, 1.8 percent of firms containing a generalist also contained a cardiologist, 0.9 percent contained an endocrinologist, and 13.7

[7] There is a collection of laws and regulations restricting physician self-referral, i.e. referral of a patient to certain entities in which the referring physician or a member of his or her family has a financial stake. These laws and regulations are commonly known as the "Stark law" or "Stark regulations."

percent contained any specialist. By 2012, integration rates increased to 2.1 percent for cardiology, 1.1 percent for endocrinology, and 15.6 percent for any specialty.

Table 1: Trends in physician organizations, 2009-2012

Panel A: Physician organizations

	2009	2010	2011	2012
Firms containing a generalist	73,440	69,411	76,111	76,860
Firms also containing a cardiologists	1,285	1,405	1,557	1,620
Firms also containing an endocrinologist	668	760	821	842
Firms also containing a specialist	10,042	10,617	11,524	12,003
Integration % for hypertension	1.75	2.02	2.05	2.11
Integration % for diabetes mellitus	0.91	1.09	1.08	1.10
Integration % for any disease	13.67	15.30	15.14	15.62

Panel B: Physicians

	2009	2010	2011	2012
Generalists	170,527	171,621	184,147	185,793
Integrated with cardiologists (count)	22,015	28,748	31,853	31,876
Integrated with endocrinologists (count)	18,857	26,350	29,256	28,845
Integrated with any specialist (count)	58,095	69,164	75,959	76,044
Integrated with cardiologists (percent)	12.91	16.75	17.30	17.16
Integrated with endocrinologists (percent)	11.06	15.35	15.89	15.53
Integrated with any specialist (percent)	34.07	40.30	41.25	40.93

Physician-level measures track these organizational trends. In 2009, 12.9 percent of generalists practiced with a cardiologist, 11.1 percent practiced with an endocrinologist, and 34.1 percent practiced with any specialist. These percentages increased to 17.2, 15.5, and 40.9 percent, respectively, by 2012.[8]

Literature review

Organizational form in medicine is tied to deep questions in organizational and health economics. In this section I describe what we currently know about the determinants of organizational form in medicine as well as the effects organizational form may have. Consistent with findings from an earlier review by Rebitzer and Votruba (2011), I find that

[8]The increases in integration rates from 2009 to 2010 are much sharper than increases later in the sample. And while the number of generalists increaes consistently from year to year, this is not true for the number of firms containing a generalist. The number of firms increases over the entire period, but there is a marked decrease between 2009 and 2010. It is difficult to know if this is a data quality issue or if these numbers actually reflect the truth, such as a strategic response to the ACA. However, these diverging trends do not drive my estimates, as patients in 2009 comprise only 6 percent of my diabetes mellitus sample and less than 1 percent of my hypertension sample.

the effects of integration among generalists and specialists are not well understood.

The determinants of organizational form in medicine

Organizational diversity is driven by technology and coordination

Rosen (1983) shows how economies of scope arising from skill complementarities can determine generalist-specialist integration. Suppose skill in completing tasks associated with generalist medicine reduces the cost of acquiring skills associated with specialist medicine. The stronger the spillover effect, the lesser the incentive for generalists to integrate with specialists, especially if tasks in specialist medicine are reimbursed at higher levels. However, it is unclear whether, if at all, these spillovers exist. Still, in the limiting case with no spillovers, David and Helmchen (2011) show that the cost of skill acquisition may still determine integration patterns. In their model they account for the fact that treatment of many diseases is treated according to a hierarchy of tasks, e.g. treatment algorithms taught in medical school or residency and fellowship, or embodied in standards and guidelines. For a given physician, the integration decision boils down to a choice about where in the task hierarchy he cedes control to another provider. He cedes control earlier in the task hierarchy when skill in the downstream task is more difficult to acquire or when the cost of coordination is lower.

Becker and Murphy (1992) demonstrate how coordination costs can limit organizational complexity. If coordination costs increase with firm size, firms cannot grow without limit. These coordination costs seem important in medicine. Newhouse (1973), Gaynor and Pauly (1990), Gaynor and Gertler (1995), and Rosenman and Friesner (2004) indicate that medical practices are subject to substantial diseconomies of scale.[9] These diseconomies arise due to revenue-sharing. Such sharing rules introduce free-riding incentives and are common in multi-physician medical practices.

[9] Though Reinhardt et al. (1979) and DeFelice and Bradford (1997), estimate increasing and constant returns to scale, respectively, in medical practices, these papers are not influential.

Hierarchies may be limited due to physician preferences

Garicano and Hubbard (2007) explain why hierarchies may arise in organizations. Situating high-human capital individuals above low-human capital individuals can increase productivity by shielding high-human capital individuals from menial tasks. Analysis of confidential firm-level data from the Census supports their claims.[10]

The emergence of postgraduate training programs (e.g. internal medicine residency, cardiology fellowship) seems to fit this theory. However, hierarchies among non-trainee physicians are hardly common. One reason may be that most firms are too small to accomodate hierarchies. Another reason is that physicians have been known to fiercely guard their independence. Analyzing data from the 1997, 1998, and 2000 waves of the Community Tracking Study (CTS) Casalino et al. (2003) report that lack of physician cooperation and lack of physician leadership are two of most frequently cited barriers to large medical group practices (p. 1961). CTS Interview responses include:

> "Physicians don't think that administrative work is worth much" (p. 1961)

and

> "It's impossible to deal with doctors. I tried to put together a quality improvement program in our group, but it never happened. No physician was paid to run the organization. Meetings were a nightmare. I did so much for free, and all I got for it was a lot of flack" (pp. 1961-1962).[11]

[10] Another theory posits that "up-or-out" tournaments, implicitly defining another form of hierarchy, allows firms to control relationship-specific human capital amassed by its employees (Rebitzer and Taylor, 2007). This behavior is less relevant in my context.

[11] Notably, physicians – particularly younger ones – are increasingly joining larger groups. Liebhaber and Grossman (2007) report that the share of physicians in solo- or two-physician practice has declined steadily from 40.7 percent in the 1996-1997 wave of the Community Tracking Study to 32.5 percent in the 2004-2005 wave (see Figure 1). Furthermore, the percent of physicians with an ownership stake in their practice has declined from 61.6 percent in 1996-1997 to 54.4 percent in 2004-2005 (see Table 2).

Firms adjust boundaries to internalize referrals

Burns and Pauly (2002) ask hospital administrators their motives for acquiring generalist practices. A desire to retain and control referrals is commonly cited.

Several econometric studies document how referrals respond to organizational form. David and Neuman (2011) study gastroenterologists. In their data, gastroenterologists can perform procedures in either an ambulatory surgery center (ASC) or a hospital outpatient department (HOPD). More severe patients fare better in HOPDs. Some gastroenterologists have access to ASCs only while others may choose between ASCs and HOPDs. The authors find that higher-severity patients are more likely to be treated at a HOPD when their gastroenterologists have access to one. Gastroenterologists with access to ASCs alone refer only the highest-severity patients to HOPDs. Nakamura et al. (2007) document how referrals respond to changes in hospital firm boundaries. These authors look for changes in referral patterns after community hospitals are acquired by large tertiary care hospitals.[12]

In their data, thirty percent of acquisitions lead to a significant increase in referrals from community hospitals to the acquirer. Patients with more generous insurance were referred to the acquiring tertiary care hospital at higher rates. Patients undergoing more profitable procedures were also referred at higher rates. Hillman et al. (1990) and Mitchell and Sass (1995) are two examples from a sizable literature showing that physician ownership of an auxiliary facility (e.g. physical therapy) is associated with a greater propensity to refer to that facility.[13] Similarly, in an analysis of Medicare data, Barro et al. (2006) find evidence supporting the idea that cardiac specialty hospital formation is driven by the desire to capture referrals. David et al. (2011) study how integration of hospitals with home health

[12] In general, large tertiary care hospital is equipped for any condition a community hospital can treat as well as more severe conditions.

[13] This literature is generally referred to as the literature on "supplier-induced demand" (SID), a topic which shares considerable overlap with referral-motivated integration. The canonical reference on SID is Evans (1974). McGuire (2000) defines SID among physicians as "when the physician influences a patient's demand for care against the physician's interpretation of the best interests of the patient" (p. 504). Mitchell and Sass (1995) describe how SID may be tied to referrals. Consider a patient does not need but (probably) will not be hurt by an auxiliary service like such as physical therapy or lab work. A physician who owns an auxiliary facility that provides these services has a greater incentive to refer the patient for auxiliary services than a physician who does not.

agencies (HHAs) and skilled nursing facilities (SNFs) affects the site of patient recovery.[14] Hospital-based recovery is more costly than SNF- or HHA-based recovery. The authors show that patients in integrated hospitals recover more often in SNFs or with HHAs.

There are very few studies on referral-motivated integration outside of health care. The rare exception is Garicano and Hubbard (2009), who find that lawyers specializing in fields where referrals are more prevalent are more likely to integrate with other lawyers in that field.

Bargaining power

Bargaining is a central feature of health care markets and the extent to which firms in this market do so distinguishes health care as an industry. The prevalence of Health Management Organizations (HMOs) and Preferred Provider Organizations (PPOs) as well as the increase of insurer market power in a growing number of geographic markets have provided a strong incentive for providers to bolster their bargaining power,[15] and the strength of these incentives has been clearly demonstrated in recent years (Dafny, 2010).[16]

There is an extremely small literature on the effect of hospital-physician integration on bargaining power. Gal-Or (1999) models these interactions and generates an extremely rich set of results. In short, the effect depends on the comparative degrees of competitiveness among hospitals and physician firms, respectively. Interestingly, exclusive dealing between integrated parties may increase barganing power among excluded parties. The empirical literature consists of two papers. Ciliberto and Dranove (2006) study hospital-physician merg-

[14] As above, hospitals specialize in acute care while HHAs and SNFs specialize in post-treatment recovery.

[15] For examplle, Town and Vistnes (2001) finds that that "hospital mergers, even in urban areas with many nearby hospitals, can lead to significant price increases" (p. 733).

[16] Blair and Herndon (2004) notes that physicians "have pursued numerous legislative initiatives at both the federal and state levels to grant self-employed physicians, who are currently prohibited from unionizing, an antitrust exemption to bargain collectively with health plans" (p. 989). Bordonaro (2013) provides a recent example. In the 2000-2001 round of the Community Tracking Study (CTS), bargaining leverage with health plans was the most frequently cited benefit (81 percent of interviewees) of practicing in a large medical group. In comparison, only 20 percent cited better lifestyle and 15 percent cited improved quality (Casalino et al., 2003, p. 1961). The American Medical Association provides a substantial guide to evaluating and negotiating contracts with insurers (American Medical Association, 2012).

ers in California from 1994 through 2001. In most markets they study, vertical integration does not affect hospital prices. In rural areas, prices decline. Cuellar and Gertler (2006) study hospital-physician mergers in Arizona, Florida, and Wisconsin from 1994 through 1998. They find that integration increases prices without improving efficiency.

Foreclosure

Foreclosure occurs when access by a party in a vertical relation is restricted to the other (upstream or downstream) party. It may have strong effects in referral-intensive industries. Personal conversations and the trade literature strongly indicate that many medical specialists highly value referral sources. Consider a cardiology practice that acquires a primary care practice and internalizes its referrals. Clearly, the other practices are worse off. Patients may also be worse off due to the restriction of choice, potentially reduced match quality between patients and physicians, and price increases due to increased bargaining power. I am not aware of studies on foreclosure in health care, particularly on foreclosure of referrals. Evidence from non-medical contexts is mixed (Lafontaine and Slade, 2007, see p. 672).

Quality

Quality (perceived or actual) is an extremely salient concern in health care (see the comprehensive survey of Kolstad and Chernew, 2009). The premium paid by insurers to "superstar" hospitals is well-documented (Ho, 2009). It is possible that physician firms may expand firm boundaries if doing so generates a prestige-driven premium. Firms may also try to exploit their prestige by acquiring other firms and rebranding them. Physician firms may also attempt to generate or exploit a prestige-driven premium by strategically integrating. I am unaware of an empirical literature on this topic.[17]

[17] Physicians also self-regulate referrals on a quality basis. Both Navathe and David (2009) and Johnson (2011) show that physicians divert referrals from worse to better performers.

Integration, generally speaking, does not seem to increase welfare

The evidence is mixed at best, negative at worst.

David and Neuman (2011), David et al. (2011), and Afendulis and Kessler (2011) find that integration is welfare-improving. Nakamura et al. (2007) as well as Hillman et al. (1990), Mitchell and Sass (1995), and other papers in the SID literature find that integration is welfare-reducing. The results from Barro et al. (2006) are mixed. Huckman (2006) estimates economically small effects.

In two exhaustive reviews, both Burns and Pauly (2002) and Burns and Muller (2008) conclude that integration reduces welfare. Burns and Pauly (2002) write,

> "During the 1990s many hospitals pursued twin strategies of vertical and horizontal integration.... While the forms of integration varied across hospitals and markets, their economic performance, after a decade of experience, was generally uniform: Nothing worked" (p. 128).

Burns and Muller (2008) do not find evidence that hospital-physician integration reduces costs, increases quality, increases access to care, or increases clinical integration. Madison (2004) analyzes data on Medicare heart attack patients and does not find that integration improves treatment choices, expenditures, or outcomes.

However, integration encouraging greater specialization improves welfare

In a very careful study of obstetricians, Epstein et al. (2010) document how larger practices encourage physicians to develop expertise in treating certain types of high-risk pregnancies. The authors argue that this specialization leads to better matching between patients and physicians. They estimate that such specialization averted almost 11,000 hospital days and 5,500 complications for women in Florida and New York during their sample period.

The link between firm size, specialization, and productivity is potentially very important. I am unaware of other empirical studies that estimate the effect of firm size on labor specialization, at least for extremely high-skilled labor. However, one hopes the effect is positive. Donohoe (1998) reviews papers from 1966 through 1997 that compare generalist and specialist care, concluding,

> "Evidence is strongest that the knowledge base and quality of care provided by specialists exceeds those of generalists for certain conditions such as myocardial infarction, depression, and acquired immunodeficiency syndrome." (p. 1596).

In a similarly-themed review, Harrold et al. (1999) report,

> "specialists were reported to be more knowledgeable about conditions encompassed within their specialty.... [S]pecialists practicing in their area of expertise were more likely to use medications associated with improved survival and to comply with routine health maintenance screening guideliens; they used more resources including diagnostic tests, procedures, and longer hospital stays" (p. 499).

Afendulis and Kessler (2007) find that anti-specialization among cardiac physicians increases costs without improving outcomes.[18] Prior to the invention of percutaneous transluminal coronary angioplasty (PTCA), cardiologists provided only medical services while cardiac surgeons provided only surgical services. In the 1970s, the emergence of PTCA led to the creation of a new medical specialty, interventional cardiology, whose practitioners were trained in both medical and surgical (i.e. specifically PTCA) services. The authors show that these anti-specialized physicians provided care at higher cost without a discernable improvement in outcomes.

[18] The authors frame their paper as a study of integration. I interpret their work as a study of anti-specialization. In the authors' terminology, an "integrated" physician provides both medical and surgical services. However, integration of physicians in the authors' sense is not analogous to integration among generalists and specialists. Consider a firm that contains only cardiologists. If the firm integrates by adding cardiac surgeons, the firm's endowment of man-hours increases. Yet a physician's endowment of man-hours is fixed. Therefore the authors' definition of "integration" corresponds to the case in which all the physicians

That said, the gains from specialization depend on the condition. In another review, Harrold et al. (1999) find that specialists outperformed generalists in the treatment of myocardial infarction, stroke, and asthma but generalists outperformed specialists when treating hypertension, diabetes, chronic obstructive pulmonary disease, unstable angina, and low back pain.

New evidence is needed

There are very few settled questions in the broader literature on health care organizations. The most convincing result is that scope economies are limited by coordination costs, particularly the free-riding problem. It also appears that integration cannot generally be expected to increase welfare, though the case of specialization-increasing welfare is one very important exception.

The effect of generalist-specialist integration is unknown. This is unfortunate given the central role physicians play in the health care system and the demonstrated interest policymakers have in changing patterns of organization. In the remainder of this document I present new contributions from this project.

in a practice learn new skills, e.g. cardiologists learn PTCA or become interventional cardiologists.

CHAPTER 2 : Data

Sources of data

The first source of data for this project is a health insurance claims database from the Health Care Cost Institute (HCCI). The HCCI is an independent non-profit that warehouses health insurance claims from large, private market insurers such as Aetna, Humana, and UnitedHealthcare.[1] My data set contains 3.5 billion rows detailing 3.1 billion claims. These claims are associated with over 1 million providers.[2] In each year over 36 million unique beneficieries appear in the data. Each claim contains billing (e.g. allowed amount, coinsurance, copay, and deductible), care (e.g. diagnosis codes, procedure codes, NPI, place of service code), product (e.g. market segment, product segment such as HMO and PPO), and demographic (e.g. age band, gender, and state) data. My extract spans the years 2008-2012.

My second data set is an annual survey of physician firms conducted by SK&A, a health care marketing research firm. The unit of observation is the office. For each office, SK&A records street address, other contact information (e.g. phone numbers), and affiliations (e.g. firm identifier aggregating offices into firms, hospital affiliation, health system membership). For each physician observed in the office, SK&A reports his or her name, degree (MD or DO), medical specialty or specialties, and NPI. While these data are rich, physician and firm identifiers require substantial imputation. I describe the imputation in Chapter F of the Supplementary Material. After imputation my extract spans 2009-2012.

My third source of data is generated by a severity scorer developed by health services researchers at Johns Hopkins University. The ACG® System software computes a severity score for any patient with a year's worth of claims.[3] Higher scores indicate higher severity. I use the software to address a missing data problem, that physicians observe patient health status but I do not. I choose to generate severity scores with the ACG® System software

[1] See `http://www.healthcostinstitute.org/data-contributors`. Page last accessed February 18, 2016.
[2] Both physician and non-physician providers are identified.

so that I may avoid ad hoc procedures in defining my own severity measures. In work on referrals preceding mine, health services researchers have also relied on the ACG® System software to address this same missing data problem. (see Forrest et al., 2002, 2006).

I use the data in the following way. I observe the specialty composition of each firm in the SK&A data. Since I observe NPIs in both the HCCI and SK&A data, I can link physician behavior to firm composition.[4] My sampling frame is determined by the overlap of my data extracts: 2009-2012.

Data set construction

Patient data

Patient sampling

I analyze two groups of patients: those with a primary care physician claim for hypertension and those with a primary care physician claim for diabetes mellitus. An individual may appear in both groups. These conditions have high rates of incidence and prevalence, lead to potentially severe complications, and both generalists and specialists may treat them.

To generate my samples, I begin by searching the physician claims data for any claim I can link to a generalist in the SK&A data. This results in lists of patients who received service for hypertension or diabetes mellitus from a generalist. Because treatment of non-adults likely differs from treatment of adults, I exclude individuals less than 18 years old. I also exclude individuals 65 and older due to the possibility of multiple insurance providers. I then randomly sample patient identifiers within each list and draw claims for those individuals. The sampling rate for hypertension patients is 25 percent. For diabetes mellitus patients it is 40 percent. I sample to reduce computational burden. There are millions of individuals

[3] See The Johns Hopkins ACG System (2016) for further technical details.

[4] In order to protect patient confidentiality, patient and provider identifiers in the HCCI data are encrypted. HCCI also coordinates encryption of the SK&A identifiers so that I may link observations across data sets without compromising patient confidentiality. While this is an important practical matter, the details of this procedure are unimportant for understanding my results.

on each list.

Severity scores and treatment windows

After random sampling there is still sample attrition. First, I need to generate a severity score for each patient. The severity scorer requires one year's worth of claims. For example, if the first generalist encounter for hypertension is January 1, 2011, that individual must be in the data starting January 1, 2010.[5] Second, I need to observe the patient throughout what I call the treatment window, the number of days after the first generalist encounter for hypertension or diabetes for which I study behavior and outcomes. For example, if the first generalist encounter for hypertension is January 1, 2011 and the window is 360 days, then that individual must be in the data in December 2011. The length of the treatment window is arbitrary. As a practical matter, longer windows induce more attrition but give more time for effects to appear. Since I study two chronic illnesses, I set relatively long windows of 360 and 720 days. For some acute conditions, shorter windows would have been more appropriate.

Sample attrition is severe – more than half of the sample is lost due to scoring and coverage cessation – perhaps leading me to oversample individuals with longer periods of continuous coverage. However, the coverage of the HCCI data set is very broad, covering tens of millions of individuals across the United States. This suggests a minimal role for sample selection bias.

It is also important to note that I do not observe patient age in years in the HCCI data. Instead I observe patient age in 5-year bins. Because the scorer requires age, I randomly assign patient age within the bin and retain that age for regression analysis.

[5] Since I only observe month and year of membership, I set a conservative requirement of thirteen months of continuous coverage.

Severity score computation

I generate severity scores with medical claims and a severity scorer, the ACG® System Software. The software maps claims and demographics to a single index of severity. Ideally I would score all individuals simultaneously, but since my data set is so large I implement a modified procedure. For each condition I draw 1,000 random samples of 50,000 patients and score the patients in each sample. I then compute the modal score for each patient across all samples in which he or she appears and set that as the patient's severity score.[6] The scorer generates several risk measures. I utilize the total cost risk score as it is forward looking and generates wide variation in the severity measure.

Claims extraction

Each of my outcomes is computed from an extract of physician services, outpatient services, inpatient services, or pharmacy claims.[7] Due to the size of the underlying data and the subtleties of medical coding, filtering these claims is non-trivial.

For physician services claims, I search for claims with primary diagnosis code (ICD-9) indicating hypertension or diabetes mellitus alone, or one of these conditions with its complications. Hypertension claims have primary diagnosis codes beginning with 401, 402, 403, and 404. Diabetes mellitus claims have primary diagnosis codes beginning with 250.[8] My filters are consistent with several medical coding sources.[9]

It is possible providers code differently in practice. For example, instead of coding a

---

[5] See `http://acg.jhsph.org/` for more information.

[6] I experimented with computing means, but in practice this made little difference. It also may not have been necessary to score random samples, but without knowing the internals of the scorer this seemed to be the most reasonable approach.

[7] The claims were provided by HCCI in separate files according to this classification.

[8] To illustrate, codes beginning with 402 denotes hypertensive heart disease, i.e. heart disease with presumed causal link to hypertension. Codes for diabetes mellitus without complications begin with 2500, while a code beginning with 2504 denotes an opthalmic condition with presumed causal link to diabetes. I argue these filters are conservative, but reasonable.

[9] See, for example HMSA Provider Resource Center (2014), `http://www.aafp.org/fpm/2008/0600/p35.html`, and `http://www.fortherecordmag.com/archives/042312p28.shtml`.

diabetes-caused retinopathy as 2504X, where X is either empty or another digit, they might choose a diagnosis code in the 360XX-379XX range. However, there is strong guidance from payers and a small industry of consultants, educators, and publications that concern themselves with optimized medical claims coding. This further suggests my filters are reasonable.

For inpatient and outpatient services I extract claims more aggressively. Especially for inpatient claims, providers are much less likely to have documentation that a given condition is caused by hypertension or diabetes mellitus. I certainly introduce error in that I am collecting too many claims, but it is unlikely that this error varies with integration status.

For pharmacy claims, I only collect claims for hypertension patients. Ideally I would have claims for both. I am able to collect claims for hypertension patients because HCCI provides a variable that allows me to filter claims for medications plausibly linked to hypertension, i.e. those designated as "Cardiovascular Drugs." In contrast, while there is a designation for "Hormones and Synthetic Substitutes," conversations with clinicians indicate that drugs outside this class might also be used in treating diabetes mellitus or its complications.[10]

Physician, firm, and area data

Physician specialty and firm integration status

Physician specialty is observed directly in the SK&A data.[11] With physician specialties and firm identifiers I can compute integration status at the firm-year level. For example, in my diabetes analysis the primary specialist of relevance is the endocrinologist. Therefore, finding firms integrated for diabetes mellitus is a matter of flagging firms containing both

[10] I attempted to utilize all prescription drug data for diabetes mellitus patients, but the estimates are uninformative. Also, the resulting file sizes were almost prohibitively large.

[11] Before computing any other variables, it was necessary to impute physician NPIs and group identifiers. Without imputation, I cannot link physicians across years in the SK&A data due to how the data set was constructed. The data set is effectively a repeated cross section of US physicians. Though some national identifiers are in the data, there is no variable that links physicians across all years in my extract. There is also potentially severe misclassifications of firms due to data collection issues. I describe my imputation procedure in the Appendix. That said, missing NPIs do not appear to be a major concern in the most recent years of the data, though I am not aware of the quality the firm identifiers for the newer data.

generalists and endocrinologists. Cardiology is the relevant specialty for hypertension.

Firm and area characteristics

I compute firm and office size using reported size and my office and group identifiers. Using physician specialty data, firm and office identifiers, and address data I compute physician counts, the number of relevant specialists in the ZIP code, and the number of offices and firms in the ZIP code that contain a relevant specialist.[12]

Finally, for each generalist in the data I compute the distance to the nearest specialist. I geocode each office in the data based on 5-digit ZIP code, separate the offices containing generalists from the offices containing relevant specialists, and then for each office compute the set of specialist distances.[13]

Differential distances

I rely on this variable for my instrumental variables regression. For each patient I set the differential distance equal to the distance in miles to the nearest integrated generalist minus the distance in miles to the nearest non-integrated generalist. I begin by geocoding patient locations.[14] Patient locations are provided to me at the 5-digit ZIP code level every month the patient is insured by one of HCCI's data contributors. I select the ZIP code corresponding to the month and year of the first generalist visit for the condition. I then geocode physician locations using office ZIP codes in the SK&A data. I then compute all great-circle distances between patient and generalist office ZIP codes, divide offices by integration status, and compute minimums.

[12] ZIP codes for most locations are actually available at the 9-digit level. However, the opportunity cost of geocoding each location at the 9-digit level instead of the 5-digit level was high, as there are substantially more 9-digit ZIP codes than 5-digit ZIP codes and the Texas A&M geocoding service (see next footnote) does impose rate limits. Furthermore, the loss of precision in defining local markets by 5-digit ZIP codes is likely low. Therefore I truncate ZIP codes to 5 digits.

[13] I gratefully acknowledge Texas A&M Geoservices for providing me free access to their geocoding platform. See `http://geoservices.tamu.edu/Services/Geocode/` for more information about their service.

[14] I, again, gratefully acknowledge Texas A&M Geoservices.

Subsample construction

In addition to my full-sample analysis, I re-estimate my models on two sets of subsamples. These subsamples divide patients into groups more or less likely to be affected by integration.

Low- and high-access subsamples

The first set divides patients by geographic access to specialists. Integration should have a greater effect in the low-access sample. I initially attempted to divide individuals among urban and rural locations by assigning Rural-Urban Commuting Codes to ZIP codes.[15] However, the rural patient samples were too small to make any inferences.[16] I instead subdivide my sample by distance to the nearest specialist. For each patient I have the distance between his or her generalist and the nearest relevant specialist. I compute the 33rd and 67th percentile for this distance. My high-access subsample contains all individuals with nearest specialist distance below the lower quantile (nearest specialist is closer). My low-access subsample contains all individuals above the higher quantile (nearest specialist is further away). There are more patients in the high-access group because the distribution of nearest specialist distance is negatively skewed.

Hospital affiliation subsamples

The SK&A data contains a flag that indicates whether a provider has a hospital affiliation. Integration should have a stronger effect in the no-affiliation sample. The values for this flag are non-empty and empty.[17] Providers in a firm with a non-empty value for this variable are included in the hospital-affiliated subsample. All others are included in the not-affiliated subsample.

[15] A link to the underlying data is available at United States Department of Agriculture, Economic Research Service (2016).

[16] This may be an artifact of HCCI exclusion rules, as patients in rural areas face higher risk of being identified in case of a data breach.

[17] Unfortunately an empty value is consistent with both no affiliation and unknown. Since this flag is relevant only to this subgroup analysis and the cost of imputation is high, I do not attempt imputation for this variable.

CHAPTER 3 : Descriptive analysis

Descriptive analysis

Introduction

I first introduce notation. The notation is used throughout the rest of this document. I then describe my basic descriptive approach and define the outcomes that will be analyzed. I conclude with a brief discussion of the results. The estimates strongly indicate that integration is associated with provider choices and financial outcomes.

Notation

Let i index patients and let j denote physicians, where j is i's generalist physician. Let c denote the patient's condition. Let t_0 be the date of the initial diagnosis. Let I_{j,t_0} be an indicator that equals 1 if j's firm is integrated at t_0 for condition d. Let Δ denote the treatment window, the total length of time in days that I collect claims for i pertaining to d. As specified in the previous chapter, I study outcomes at 360 and 720 days.

Then let $o_{i,j,t_0}^{c,\Delta}$ be the outcome for patient i given all the treatment he has received in the window Δ. For example, the outcome could be the total paid for i's care or the number of new prescriptions he has filled within the window. Since the values of c and Δ will be made clear in the text and tables, I suppress these superscripts for ease of reading.

Descriptive analysis

In my descriptive analysis I compute means for the group of patients with an integrated generalist as well as the group of patients with a non-integrated generalist and then compare them. In symbols the difference is

$$E\left[o_{i,j,t_0}\left|I_{j,t_0}=1\right.\right]-E\left[o_{i,j,t_0}\left|I_{j,t_0}=0\right.\right]. \tag{3.1}$$

Outcomes

My analysis studies a wide range of outcomes. They may be divided into two categories: behavioral responses and clinical and cost measures. The behavioral responses are:

Primary specialist visit: An indicator that equals 1 if I observe a physician services claim for the relevant specialist within the claims window

Primary specialist RVUs: For patients with a primary specialist claim, this is the sum of primary specialist RVUs observed in the claims window[1]

Generalist RVUs: The sum of generalist RVUs observed in the claims window

Outpatient services visit: An indicator that equals 1 if I observe an outpatient services claim within the claims window

Outpatient RVUs: For patients with an outpatient services claim, this is the sum of outpatient RVUs observed in the claims window

New cardiovascular prescription, generalists: An indicator that equals 1 if I observe a pharmacy claim for a new cardiovascular prescription if the prescribing physician is the patient's generalist. Hypertension patients only.[2]

New cardiovascular prescription, specialists: An indicator that equals 1 if I observe a pharmacy claim for a new cardiovascular prescription if the prescribing physician is a relevant specialist. Hypertension patients only.[3]

The first variable allows me to determine if firm boundaries affect referral. Though I do not measure referral directly, it is reasonable to assume that generalists will refer primarily to the specialists in the group. The remaining variables allow me to observe whether firm

[1] HCCI provides procedure codes. Using these procedure codes I merge in RVUs from the PFS Relative Value files provided by the Centers for Medicare & Medicaid Services, available at `https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/PFS-Relative-Value-Files.html`.

[2] The reasons driving the restriction are detailed in Section 2.2.1.

[3] Ibid.

boundaries affect the total amount of work physicians do as well as whether firm boundaries cause the substitution of one kind of work for another (e.g. outpatient tests versus physician appointments). Though simple models of referrals (see Garicano and Santos, 2004) make the reasonable assumption that effort is unidimensional, I study a large number of behavioral measures because physicians may respond to integration along multiple margins.

The remaining regressands are used to measure clinical and financial effects of integration:

Inpatient admission: An indicator that equals 1 if I observe an inpatient admission in the claims window

Physician cost: The sum of allowed amounts in physician services claims within the claims window

Outpatient cost: The sum of allowed amounts in outpatient services claims within the claims window

Pharmacy cost: The sum of allowed amounts in pharmacy claims within the claims window. Hypertension patients only.[4]

Total non-inpatient cost: The sum of all costs above. This excludes pharmacy costs for diabetes mellitus patients.

Total cost: The above measure plus the cost of inpatient care.

Estimates of the effect of integration on these outcomes, particularly inpatient admissions and total costs, may allow us to make judgments about the integration-inducing policies.

Results

Table 2 shows that outcomes respond to the structure of the generalist's firm. However, these differences are quantitatively small. For example, patients of non-integrated practices

[4] Ibid.

are more likely to visit a specialist than patients of non-integrated practices.[5] In the hypertension sample the differences are 3 and 4 percentage points at 360 and 720 days, respectively. In the diabetes mellitus sample, the differences are 4 and 3 percentage points at 360 and 720-days, respectively. Patients of non-integrated practices are very slightly less likely to have any outpatient services utilization than their integrated counterparts. For hypertension patients, there is no difference at 360 days and the difference is 1 percentage point at 720 days. For diabetes patients, there is no difference in either window. Differences in costs are also extremely small. The largest difference is in physician costs. The differences for hypertension patients measure 26.4 dollars at 360 days and 33.2 dollars at 720 days, while the differences for diabetes mellitus patients are larger, respectively 61.9 dollars and 96.1 dollars.

In my regression analysis I will show that this simple descriptive analysis paints a misleading picture, and that integration has large, economically significant effects along several important margins.

[5] My dissertation proposal contained the opposite finding. However, in that document I defined integration very broadly, calling a practice integrated if it contained a generalist and at least one of the top ten specialists for the condition. Because these lists contained specialties like emergency medicine I opted for the most conservative definition of integration.

Table 2: Differences in mean outcomes for patients of non-integrated and integrated generalists

Panel A: Hypertension

	360 days					720 days				
	Non-integrated	Integrated	Difference	t-test value	p-value	Non-integrated	Integrated	Difference	t-test value	p-value
Primary specialist visit	0.17	0.14	-0.03	-14.93	0.000	0.23	0.20	-0.04	-12.70	0.000
Primary specialist RVUs	4.43	3.69	-0.74	-3.87	0.000	4.97	4.26	-0.72	-3.53	0.000
Generalist RVUs	3.92	3.31	-0.61	-35.31	0.000	4.96	4.19	-0.77	-26.60	0.000
Any outpatient service	0.03	0.04	0.00	2.74	0.006	0.05	0.05	0.01	3.27	0.001
Outpatient RVUs	2.37	2.03	-0.34	-1.94	0.053	2.47	2.10	-0.37	-1.96	0.050
New prescription, generalist	0.40	0.38	-0.02	-5.96	0.000	0.41	0.39	-0.02	-4.52	0.000
New prescription, specialist	0.92	0.93	0.01	0.72	0.471	0.94	0.96	0.02	1.52	0.129
Inpatient admission	0.01	0.01	0.00	-1.76	0.078	0.01	0.01	0.00	-1.80	0.072
Physician costs	191.16	217.52	26.35	20.92	0.000	248.82	282.06	33.24	16.05	0.000
Outpatient costs	19.77	17.06	-2.70	-1.73	0.083	28.86	25.82	-3.05	-1.25	0.212
Pharmacy costs	443.90	435.42	-8.47	-1.10	0.270	831.73	819.70	-12.03	-0.67	0.504
Physician + outpatient + pharmacy costs	654.82	670.00	15.18	1.90	0.057	1109.42	1127.58	18.16	0.99	0.324

Panel B: Diabetes mellitus

	360 days					720 days				
	Non-integrated	Integrated	Difference	t-test value	p-value	Non-integrated	Integrated	Difference	t-test value	p-value
Primary specialist visit	0.28	0.24	-0.04	-15.83	0.000	0.37	0.34	-0.03	-10.09	0.000
Primary specialist RVUs	7.73	6.64	-1.10	-10.57	0.000	10.92	9.79	-1.13	-8.09	0.000
Generalist RVUs	4.55	4.02	-0.52	-25.40	0.000	6.44	5.73	-0.71	-19.19	0.000
Any outpatient service	0.06	0.06	0.00	2.67	0.008	0.08	0.09	0.00	2.20	0.028
Outpatient RVUs	2.27	2.02	-0.25	-0.87	0.382	2.43	1.97	-0.46	-2.36	0.018
Inpatient admission	0.02	0.02	0.00	-2.59	0.010	0.04	0.04	0.00	-1.46	0.143
Physician costs	274.31	336.25	61.94	21.58	0.000	440.31	536.41	96.10	17.45	0.000
Outpatient costs	36.93	33.93	-3.00	-0.79	0.430	54.04	50.01	-4.03	-0.74	0.461
Physician + outpatient costs	311.24	370.18	58.94	12.25	0.000	494.34	586.42	92.07	11.64	0.000

CHAPTER 4 : Regression analysis

Model specification

I take the following model to the data:

$$o_{i,j,t_0} = \beta_0 + I_{j,t_0}\beta_1 + s_{i,t_0}\beta_2 + x'_{i,t_0}\beta_4 + w'_{j,t_0}\beta_5 + z'_{j,t_0}\beta_6 + D_{l(j)} + D_{t_0} + \epsilon_{i,j,t_0} \tag{4.1}$$

In words, I regress each of my outcomes on the dummy for integration, patient severity score (s_{i,t_0}), patient sex, age, and age squared (x_{i,t_0}), firm and office characteristics (w_{j,t_0}), area characteristics (z_{j,t_0}), and state and year fixed effects ($D_{l(j)}$ and D_{t_0}, respectively). The idea is to control for as many confounders as possible in my estimation.

Patient characteristics should affect physician behavior. Physicians analyze each patient's history and physical, notes from prior appointments, and information gleaned from the patient during the current appointment with knowledge gained in medical school, postgraduate training (if any), external information sources like clinical guidelines, and learning-by-doing to formulate treatment plans. Though I do not observe this patient data, for each patient I have a severity score, a forward-looking single index of severity that summarizes the patient's health status at the time of first diagnosis. My approach assumes more severe patients should receive more medical care.[1] Provided the severity scorer accurately measures severity differences among patients, the computed scores should provide a reasonable substitute for the clinical information I do not observe. There are a few good reasons to believe this strategy is credible. As mentioned in Chapter 2, Section 2.1, researchers in health services have already demonstrated confidence in the severity scorer used for my project. Second, there are significant financial incentives from providers and insurers that likely drive the development of accurate severity scorers.[2] Third, I find that the effect of

[1] The exception is those who would benefit from palliative care for life-ending illnesses. Because severity is highly negatively skewed, these cases should have minimal influence on my results. For an authoritative reference on palliative medicine, see .

[2] Tangentially speaking, there is a small, but growing amount of evidence from the field that large amounts of data plus machine learning or statistical modeling can lead to large clinical gains. In recent tests, IBM's

severity is positive and economically significant in all my regressions, which provides some reassurance from the data that the scorer does a reasonably good job. Because physicians may react in a non-linear fashion to patient severity, I include both the linear and squared terms for the severity score.[3] Because I have access to only two other patient observables, gender and age, I include them in the patient characteristics vector as well.

Firm and office characteristics include distance to the nearest relevant specialist and its square, office size, and firm size. The most salient factor is the distance to the nearest relevant specialist. Distance is known to strongly affect provider choice and medical care consumption, (see McClellan et al., 1994; Kessler and McClellan, 2000, and cites therein). Therefore the farther away is the nearest relevant specialist, the less likely it is a patient will see him or her. The empirical literature on diseconomies of scale in physician practices strongly indicate I should include the size measures that I have, though in practice the coefficients on these measures are economically and statistically insignificant far more often than not.

Area characteristics are computed at the ZIP code level. They include number of relevant specialists, number of offices containing a relevant specialist, and number of firms containing a relevant specialist. Because distance to the nearest specialist imperfectly captures local specialist availability, I compute measures that account for specialist density within the ZIP code.[4]

Finally, standard errors for all my regressions are clustered by state.

Watson – a computing system designed to answer open-ended questions – reportedly outperformed physicians in diagnosing lung cancer.

[3] I experimented with the addition of higher-order terms. For all outcomes in my test cases, only the linear term mattered. Only after I was able to process the pharmacy claims data did it become clear that a cubic term would improve fit for my new prescription regressions. However, after reestimating the pharmacy regressions, I was able to confirm that the inclusion of this term does not significantly affect the nature of my results.

[4] Ideally I would also include county-level measures of specialist density. However, the ZIP-to-county crosswalks available to me led to substantial data loss.

Descriptive statistics for explanatory variables

Descriptive statistics for my explanatory variables are presented in 3. Panel A contains descriptives for the hypertension sample and Panel B contains descriptives for the diabetes mellitus sample. Within each panel I present descriptives for both the 360- and 720-day windows. For each variable I compute the number of observations, mean, standard deviation, and 25th, 50th, and 75th percentiles. Because the 360- and 720-day statistics are highly similar, I focus on the 360-day estimates and the differences between the hypertension and diabetes mellitus samples.

Hypertension

After sampling and filtering, there are 183,920 patients in the 360-day sample and 114,806 patients in the 720-day sample. I lose about 1.6 percent of my samples whenever I implement the IV estimator, due to the inability of the geocoder to assign coordinates to those patients' ZIP codes. This loss is minimal.

Patient demographics compare favorably to nationally representive surveys. Mean values for age and sex are broadly consistent with estimates from the National Health and Nutrition Examination Survey (NHANES) for 2011-2012. In the NHANES 2011-2012, prevalence strongly increases with age and is slightly higher among men. In my hypertension sample, mean age is 50.0, the 25th percentile for age is 44.0, and 51.0 percent of individuals are male.[5] The histograms in Figures 1 and 2 show that the distribution of severity is highly negatively skewed. Mean severity in the 360-day sample is 1.01 while values at the 50th and 75th percentiles are 0.59 and 1.13, respectively.

Only twenty percent of patients are treated at integrated practices. More than fifty percent of patients are treated in a ZIP code containing a relevant specialist. Between twenty five and fifty percent of patients live in a ZIP code where the nearest integrated and non-

[5] Recall, patient age ranges from 18 to 64 in both samples.

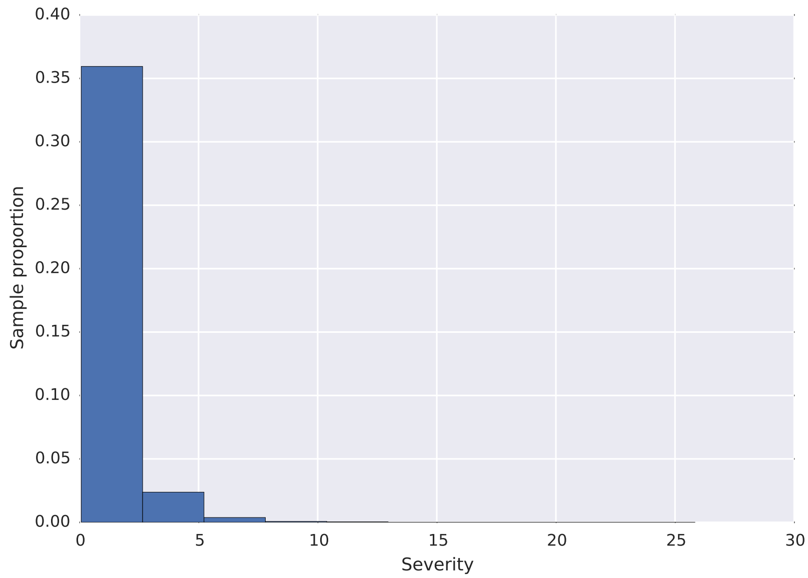

Figure 1: Distribution of severity in the 360-day hypertension sample

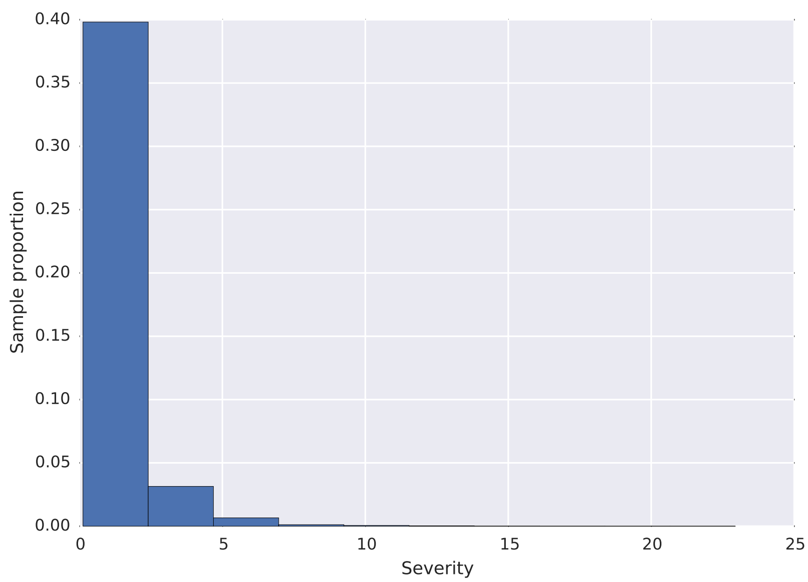

Figure 2: Distribution of severity in the 720-day hypertension sample

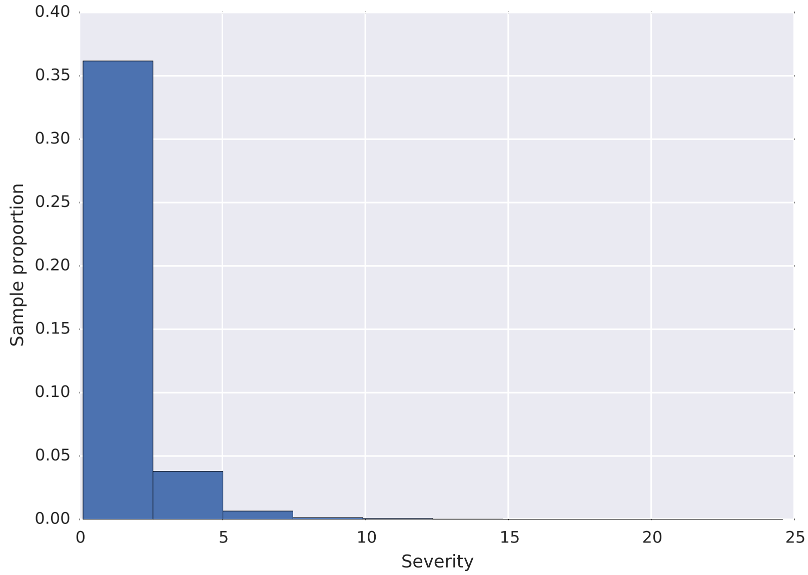

Figure 3: Distribution of severity in the 360-day diabetes sample

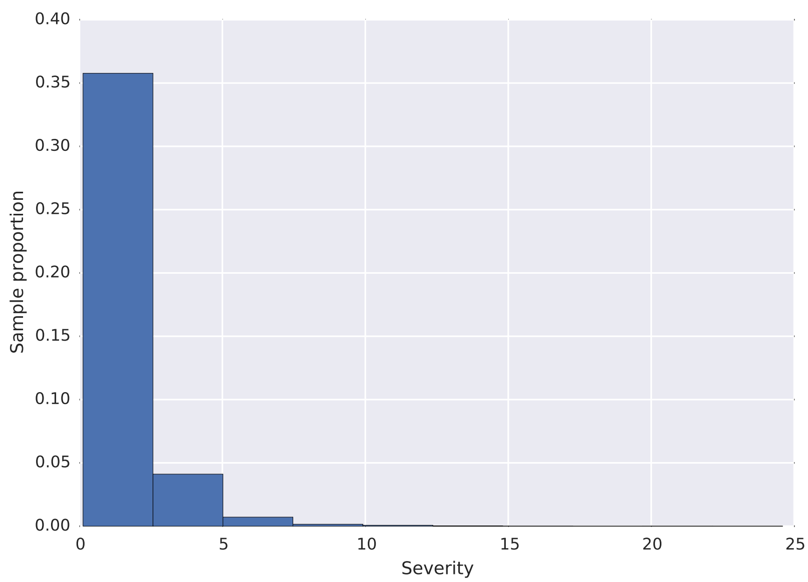

Figure 4: Distribution of severity in the 720-day diabetes sample

integrated practices are equally distant. For the rest, the nearest integrated practice is at least 3.3 miles further than the nearest non-integrated practice. More than half of the sample is treated in practices of five physicians or less, is treated by physicians without a hospital affiliation, and lives in ZIP codes where there are no more than two relevant specialists.

Diabetes mellitus

After sampling and filtering, there are 191,956 patients in the 360-day sample and 134,411 patients in the 720-day sample. Attrition is lower in this sample compared to the hypertension sample. It is unlikely to be driven by health status, as severity in this sample is comparatively greater. Similar to the hypertension sample, I lose a very similar amount of patients for IV estimation: 1.6 percent in the 360-day window and 1.7 percent in the 720-day window.

In the diabetes mellitus sample, mean age is 51.2 and 54 percent of patients are male in the 360-day window, making this sample slightly older and more male than the corresponding hypertension sample. Though mean age in my sample is lower than mean age at diagnosis in the population, I have dropped all individuals aged 65 and older from my analysis, suggesting my sample is representative of the working population. And the greater amount of males is consistent with sex-specific diabetes prevalence rates in my sample period. The distribution of severity for diabetes patients is also highly negatively skewed (see Figures 3 and 4).

Only seventeen percent of patients are treated in integrated practices. This is three percentage points less than in the hypertension sample. Less than 50 percent of patients are treated in a ZIP code containing the relevant specialist and the median distance between the treating generalist . The median distance from the treating generalist to the nearest specialist is 3.1 miles. Furthermore, more than 75 percent of patients live in a ZIP code where the nearest integrated practice is farther than the nearest non-integrated practice.

Endocrinologists are comparatively less accessible than cardiologists, as more than fifty percent of patients live in a ZIP code where there are no endocrinologists.

Threats to identification

There are two primary threats to identification: sorting by physicians among practices and sorting by patients among practices. Below I describe the nature of these threats and my strategies for dealing with them. In the case of patient sorting I argue that I have an instrument that simulates random assignment, strongly mitigating the threat of confounding from this source. In the case of physician sorting, I have a comparatively weaker response. However, there are reasons to believe bias due to physician unobserved heterogeneity is comparatively small.

Physician sorting

It is not difficult to rationalize physician sorting among integrated and non-integrated practices. A generalist might choose to work in an integrated firm because he or she values formal connections with more specialized physicians. A generalists might also be willing to sell his or her group in order to monetize equity or improve work-life balance. If preferences that induce this behavior are correlated with clinically relevant characteristics like skill or risk aversion, then OLS is confounded. The most natural response is to exploit the physician identifiers in the data add physician fixed effects. While I am able to add these intercepts to my original specification, estimation fails when I add fixed effects in IV regression (the solution to my patient sorting problem) because my instrument becomes dangerously weak. Because I lack data on physician characteristics, the next best alternative is to reason about the bias.[6]

I argue that the bias can be signed with confidence for two outcomes. In the regressions for

[6] The obvious third-party data source that might help me is the American Medical Association Physician Masterfile (see `http://www.ama-assn.org/ama/pub/about-ama/physician-data-resources/physician-masterfile.page`). However, incorporating this data at this stage in my project is infeasible. In future work my first task would be to purchase this data and add demographic data for the physicians in the HCCI and SK&A data who also appear in the Masterfile.

Table 3: Descriptive statistics for explanatory variables

Panel A: Hypertension

	360 days						720 days					
	N	Mean	Std. Dev.	25th percentile	50th percentile	75th percentile	N	Mean	Std. Dev.	25th percentile	50th percentile	75th percentile
INT minus NON	180,990	7.52	13.71	0.00	3.31	8.89	112,961	7.73	12.89	0.00	3.43	9.21
Severity	183,920	1.01	1.13	0.36	0.59	1.13	114,806	1.03	1.16	0.39	0.62	1.15
Age (in years)	183,920	49.98	9.88	44.00	51.00	58.00	114,806	50.59	9.58	45.00	52.00	58.00
Male	183,920	0.51	0.50	0.00	1.00	1.00	114,806	0.50	0.50	0.00	1.00	1.00
Integrated	183,920	0.20	0.40	0.00	0.00	0.00	114,806	0.20	0.40	0.00	0.00	0.00
Firm size (in physicians)	183,920	95.96	283.07	2.00	5.00	50.00	114,806	97.99	293.98	2.00	5.00	50.00
Firm size (in offices)	183,920	16.21	44.13	1.00	1.00	10.00	114,806	16.29	43.68	1.00	1.00	10.00
Office size (in physicians)	183,920	7.17	17.57	1.00	3.00	6.00	114,806	7.21	17.54	1.00	3.00	6.00
Hospital affiliation at firm level	183,920	0.29	0.46	0.00	0.00	1.00	114,806	0.30	0.46	0.00	0.00	1.00
Distance to nearest specialist	183,920	3.07	8.62	0.00	0.00	3.46	114,806	3.08	7.81	0.00	0.00	3.48
Relevant specialists in ZIP	183,920	6.91	13.55	0.00	2.00	9.00	114,806	6.92	13.51	0.00	2.00	9.00
Relevant offices in ZIP	183,920	1.81	2.86	0.00	1.00	2.00	114,806	1.79	2.81	0.00	1.00	2.00
Relevant firms in ZIP	183,920	1.74	2.72	0.00	1.00	2.00	114,806	1.71	2.67	0.00	1.00	2.00

Panel B: Diabetes mellitus

	360 days						720 days					
	N	Mean	Std. Dev.	25th percentile	50th percentile	75th percentile	N	Mean	Std. Dev.	25th percentile	50th percentile	75th percentile
INT minus NON	188,934	10.85	19.53	0.93	4.43	11.58	132,185	11.10	19.04	1.25	4.59	11.93
Severity	191,956	1.21	1.35	0.47	0.70	1.50	134,411	1.27	1.40	0.49	0.71	1.60
Age (in years)	191,956	51.22	9.32	46.00	53.00	59.00	134,411	51.77	9.04	46.00	53.00	59.00
Male	191,956	0.54	0.50	0.00	1.00	1.00	134,411	0.54	0.50	0.00	1.00	1.00
Integrated	191,956	0.17	0.38	0.00	0.00	0.00	134,411	0.17	0.38	0.00	0.00	0.00
Firm size (in physicians)	191,956	84.43	263.01	1.00	4.00	36.00	134,411	84.63	264.40	1.00	4.00	35.00
Firm size (in offices)	191,956	14.99	42.08	1.00	1.00	8.00	134,411	14.80	41.34	1.00	1.00	7.00
Office size (in physicians)	191,956	6.46	15.23	1.00	3.00	6.00	134,411	6.57	15.73	1.00	3.00	6.00
Hospital affiliation at firm level	191,956	0.26	0.44	0.00	0.00	1.00	134,411	0.27	0.44	0.00	0.00	1.00
Distance to nearest specialist	191,956	7.70	14.13	0.00	3.13	8.00	134,411	7.73	13.89	0.00	3.13	8.13
Relevant specialists in ZIP	191,956	1.08	2.52	0.00	0.00	1.00	134,411	1.09	2.54	0.00	0.00	1.00
Relevant offices in ZIP	191,956	0.58	1.08	0.00	0.00	1.00	134,411	0.58	1.06	0.00	0.00	1.00
Relevant firms in ZIP	191,956	0.57	1.05	0.00	0.00	1.00	134,411	0.57	1.03	0.00	0.00	1.00

the probability of a primary specialist visit, the bias is positive. Generalists with greater propensity to refer, ceteris paribus, might be more willing to join integrated practices. The estimate for integration in the primary specialist RVUs regressions is biased downward (toward negative infinity). If generalists with greater propensity to refer differentially sort into integrated practices, it is likely the marginal patient for these generalists has lower severity than the marginal patient for generalists with lesser propensity to refer. Because the marginal patients for integrated generalists have lower severity, there is less work for the relevant specialist to do when he or she sees them.

For the remaining outcomes, factors that should drive selection are likely to work in opposition to each other. This indeterminacy suggests, perhaps strongly, that the quantitative impact of physician selection is not large. In the case of generalist RVUs, generalists who select into integrated firms might do so because they can use referrals as a mechanism for shifting work onto other providers. This would induce downard bias in the integration estimate. On the other hand, generalists who select into integrated firms might also do so because they are relatively risk averse and value the "insurance" integration with a more specialized provider can offer. Selection of this nature would induce upward bias in the integration estimate, as risk averse physicians are more likely to overutilize. In the case of outpatient RVUs, if the desire to shirk induces selection into integrated firms, generalists in integrated firms might choose to refer before ordering the full set of relevant tests, biasing the integration estimate downward. On the other hand, if risk aversion induces selection, one might expect generalists in these firms to overutilize outpatient services, biasing the integration upward. For the same reasons shirking and risk aversion produce countervailing effects on my spending measures. In the case of prescription drug utilization, it is not at all obvious why physician selection should induce bias one way or the other. And with respect to inpatient admission, it is unclear whether integrated practices are attracting better or worse generalists (or specialists, for that matter).[7]

[7]If demographics for generalists in integrated and non-integrated practice were available, one might reason from those. For example, there is evidence that age is negatively correlated with physician quality (see Choudhry et al., 2005). Knowledge of average age in integrated and non-integrated practice would allow

Though I cannot control for physician sorting in my IV regressions, I still refer to non-IV panel (OLS FE) estimates for heuristic purposes. The more important is physician sorting, the further apart will be the OLS and OLS FE estimates.[8] Alternatively, when OLS and OLS FE estimates are close, I have suggestive evidence that physician sorting does not contaminate my IV estimate. This is at best a heuristic procedure, but compared to the alternative of not reporting OLS FE estimates at all, I argue my approach is more transparent and conservative.[9]

Patient sorting

Patients may sort among integrated and non-integrated practices in ways that affect outcomes. For example, if younger patients differentially select into integrated practices and severity is negatively correlated with age, then patient sorting induces downward bias in the integration estimate. Alternatively, if more severe patients differentially select into integrated practice because they believe it will be easier to secure specialist referrals, patient sorting could induce upward bias in the integration estimate. Even though the little we know about the nature of patient selection among generalists suggests that patients are not sorting on clinically relevant features (see Kolstad and Chernew, 2009, especially p. 41S), it seems extreme to assume that sorting does not affect outcomes. Therefore, because I have an arguably powerful instrumental variable, I take a conservative approach and also estimate model parameters with IV. In the following paragraphs I reiterate the instrument's definition from Chapter 2. explain how this instrument simulates random assignment, provide a preliminary test for randomization, discuss how physician sorting affects my IV estimates, and provide some history on the use of differential distances in health economics

me to take a stand on the bias in the inpatient admission regression. However, to the best of my knowledge, demographic information of this nature is not available.

[8]Though this immediately suggests a Hausman test, the test is not valid because I am still confounded by patient selection. For completeness, I implement the test in my full-sample regressions and the test generally rejects random effects estimation. Note that the alternative hypothesis is not differential selection among practices, just that physicians have idiosyncratic tendencies.

[9]I explored the possibility that the OLS and OLS FE estimates might be used to assess the nature of physician selection, but when the effect of patient selection is large relative to physician selection, some distinct patterns of patient and physician selection are observationally equivalent.

and outcomes research.

The differential distance is defined as the distance in miles to the nearest integrated generalist minus the distance in miles to the nearest non-integrated generalist. The greater the magnitude of this distance, the greater is the patient's inducement towards choosing the nearer provider.

I now explain why this instrument is valid, relevant, and simulates random assignment. First, residential choice is orthogonal to generalist locations and organizational forms. Introspection and anecdote suggest the desire to sort among residental locations seems much more driven by public goods as well as residential and cultural amenities than access to medical care. For example, commuting time and quality are probably highly salient to working-age renters and homebuyers. Checklists for first-time homebuyers often include the quality of the public school system. Though the extant empirical literature is small, it would be highly surprising to find that crime does not reduce location-specific demand for housing. In contrast, physician accessibility and the structure of their organizations seem much less likely to affect residential choices. The cost of acquiring the information to sort on these features seems large. To sort on integration, individuals would have to know the locations of the generalists in their area and the specialty composition of their practices. The time cost of collecting and organizing this data is not trivial. Consumers are already known to have difficulties navigating the US health care system.[10] It is hard to believe that an economically significant amount of individuals in my data would have collected the information necessary for sorting. And even if I underestimate willingness to collect this data, individuals sorting in my sample also requires patients to be highly forward-looking. Patients appear in my sample only if they receive a diagnosis of hypertension or diabetes after one year of not being diagnosed with these conditions. Though information on these conditions, their complications, the effect all these may have on quality of life and mortality, and how strongly lifestyle modifications can reduce the incidence of this disease, the

[10] See, for example `https://www.academyhealth.org/files/issues/NavigatingHealthCare.pdf`.

rising rate – an important risk factor for both hypertension and diabetes – provides casual evidence that individuals are not forward-looking enough to induce sorting on physician locations and organizations. Second, distance is a demand shifter. Distance is known to affect hospital demand (, see). If distance affects provider choice when the stakes are high, it must certainly affect demand when stakes are much lower. Because regular appointments with primary care providers are relatively low-stakes, it is not hard to imagine that travel times are highly salient to potential patients. Furthermore, there is little evidence that patients prioritize clinically relevant features when choosing a primary care physician and it is unclear whether consumers are even good judges of quality in the first place (see, especially, p.41S Kolstad and Chernew, 2009). Since primary care is an experience good – even for the potentially irrelevant features like waiting room amenities – and travel times are salient, I argue that distance is highly likely to shift demand.

Because differential distance is a demand shifter and residential choice is orthogonal to provider locations and characteristics, these differences effectively serve as randomization devices. In the next paragraph I discuss prior use of this randomization in outcomes research and health economics. In this paragraph I discuss how randomization at the patient level interacts with generalist sorting among integrated and non-integrated practices. For sake of argument, consider a clinical trial designed to test the efficacy of some treatment regimen. At Site A the treatment is provided and at Site B subjects receive a placebo. The investigators randomize subjects between A and B. If there are no systematic differences between the two sites, then the difference in mean outcomes between A and B is indeed the effect of the regimen. However, if the staff at A are somehow superior to the staff at B and staff quality impacts outcomes, then the difference in mean outcomes between A and B reflects both the effectiveness of the regimen and the differential quality in staff. Physician selection that affects behavior or outcomes is analogous to differences in staff quality in my example, implying that random assignment of patients to integrated and non-integrated practices does not bypass confounding due to physician sorting. But as discussed earlier, the bias for only two of my outcomes is determinate while selection may induce countervailing effects for the rest.

To my knowledge, the earliest – if not the first – paper utilizing differential distances as an instrument is McClellan et al. (1994). In this paper, the authors attempt to estimate the effect of treatment intensity for acute myocardial infarction (AMI) on mortality for elderly individuals from non-experimental data. Their identification problem is similar to mine, as patients in the data may systematically select among the available treatments. The authors make the argument that patients who live closer to a hospital outfitted for cardiac catheterization and revascularization (more intensive treatments) than hospitals not outfitted for the procedure are disproprotionately more likely to receive cardiac catheterization if they are admitted to the hospital after an AMI.[11] This instrument is well-known in the health economics literature and has appeared in top general-interest and health economics outlets in the last several years (e.g. see Chandra and Staiger, 2007; Cutler, 2007).

Finally, in Table 4 I compare mean values of patient age, gender, and severity for the group of patients with negative differential distances against the means for the group of patients with positive differential distances. If my instrument induces random assignment, the differences in patient observables should be small. The results are similar across windows, so I discuss results only for the 360-day window. While differences are statistically significant, all differences I detect are small. In the hypertension sample, the difference in mean age between the positive and negative groups is less than a year and the proportion of males differs by only 2 percentage points. The difference in mean severity is 0.05 which is small compared to mean severity in the full sample (1.01). In the diabetes sample, the difference in mean age between the two groups is also less than a year, the difference in the proportion of males is also 2 percentage points. The difference in mean severity is 0.09 which is small compared to mean severity in the full sample (1.21). My sample is not comparable to the

[11]For completeness, note that the mechanism McClellan et al. (1994) rely on is slightly different than mine. In their setting, differential distances affect assignment through their effect on the set of ambulance transit times among hospitals. Because AMI is a highly serious condition, ambulance drivers are more likely to choose the hospital and they are more likely to choose closer hospitals as a clinical matter. In my setting, choice is much more strongly driven by patient preferences.

Medicare sample analyzed by McClellan et al. (1994), but it is worth reviewing the differences in age and gender reported in Table 4 of their paper. Those authors assign patients into one of two groups. The low-distance group contains all patients with a differential distance less than or equal to 2.5 miles. The high-distance group contains all the rest. Mean age is equal between the two groups, but patients in the high-distance group are 1.8 percentage points more likely to be male. I argue the differences in age and gender I observe in my data compare favorably to these results and suggest I should have high confidence in the quality of my instrument.

Table 4: Comparison of observables for differential distance subgroups

Panel A: Hypertension

	360 days					720 days				
	Distance > 0	Distance < 0	Difference	t-test value	p-value	Non-integrated	Integrated	Difference	t-test value	p-value
Age (in years)	50.69	49.95	-0.74	-5.97	0.000	51.25	50.57	-0.68	-4.44	0.000
Male	0.53	0.51	-0.02	-2.98	0.003	0.53	0.50	-0.03	-3.79	0.000
Severity	1.06	1.00	-0.05	-3.51	0.000	1.08	1.03	-0.06	-2.94	0.003

Panel B: Diabetes mellitus

	360 days					720 days				
	Non-integrated	Integrated	Difference	t-test value	p-value	Non-integrated	Integrated	Difference	t-test value	p-value
Age (in years)	52.07	51.20	-0.87	-7.18	0.000	52.54	51.75	-0.79	-5.56	0.000
Male	0.56	0.54	-0.02	-2.25	0.025	0.55	0.54	-0.01	-1.55	0.120
Severity	1.30	1.21	-0.09	-4.72	0.000	1.37	1.27	-0.10	-4.10	0.000

CHAPTER 5 : Hypertension: Evidence from the full sample

Introduction

Hypertension is a condition in which the "long-term force of the blood against [one's] artery walls is high enough that it may eventually cause health problems, such as heart disease."[1]

Lifestyle changes such as improving diet and exercise as well as quitting smoking are commonly recommended. It is also common to prescribe antihypertensives. These strategies are not mutually exclusive. Common complications of uncontrolled hypertension include heart attack; stroke; aneurysm; heart failure; weakened and narrowed blood vessels in the kidneys; thickened, narrowed, or torn blood vessels in the eyes; metabolic syndrome; and trouble with memory or understanding.[2] Many patients can be successfully treated by generalists, but many are unlikely to respond to treatment, making them strong candidates for referral.[3]

Full-sample results

Cardiologist visit

I present regression estimates in Table 5. In each table I present estimates from the 360-day window in the left panel and estimates from the 720-day window in the right panel. Within each panel I present estimates from OLS, OLS FE, and IV estimation. I include coefficient estimates for the integration dummy, severity and its square, the maledummy, age and its square, and distance to the nearest specialist and its square. Standard errors are clustered by state. I exclude the remaining coefficients because they much more often than not are economically and statistically insignificant. I provide a full complement of tables at the end of this document for interested readers.

[1] http://www.mayoclinic.org/diseases-conditions/high-blood-pressure/basics/definition/con-20019580

[2] http://www.mayoclinic.org/diseases-conditions/high-blood-pressure/basics/complications/con-20019580

[3] See http://www.medscape.com/viewarticle/582072 and http://www.ash-us.org/HTN-Specialist/Certified-HTN-Centers.aspx.

Table 5: Effect of integration on the probability of a cardiologist visit (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.013	0.000	0.115*	-0.012	0.017	0.099
	-1.5578	0.0166	1.8587	-1.0207	1.0209	1.1674
Severity	0.021***	0.021***	0.020***	0.032***	0.034***	0.032***
	7.4166	6.4125	7.4929	9.2295	9.0450	9.3042
Severity squared	-0.001*	-0.001**	-0.001*	-0.001***	-0.002***	-0.001***
	-1.8903	-2.4353	-1.6450	-3.0610	-4.5398	-2.8362
Male	-0.001	0.003	-0.001	0.001	0.006*	0.001
	-0.5669	1.0745	-0.4055	0.4932	1.6732	0.5823
Age	-0.001	-0.001	-0.001	0.001	0.001	0.001
	-1.1130	-0.9580	-1.0891	0.8630	0.7970	0.9325
Age squared	0.000	0.000	0.000	-0.000	-0.000	-0.000
	1.5159	1.5007	1.4427	-0.0862	-0.0457	-0.1857
Distance to nearest specialist	-0.000	0.000	0.000	-0.001	0.001	-0.000
	-1.2506	0.1392	0.2255	-1.4083	0.7829	-0.2306
Distance to nearest specialist squared	-0.000	0.000	-0.000	0.000	-0.000*	-0.000
	-0.3386	0.6409	-1.2106	0.0109	-1.9288	-0.5674
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

OLS estimates are negative and statistically insignificant. They indicate integration decreases the probability of a specialist visit by 1.3 percentage points at 360 days and 1.2 percent at 720 days, respectively. The OLS FE estimates are small and positive but also statistically insignificant, indicating no change at 360 days and a 1.7 percentage point increase at 720 days.

In contrast, IV estimates are positive, and larger by an order of magnitude. The 360-day estimate is statistically significant at the 90 percent level, indicating a 11.5 percentage point increase in the probability of a specialist visit within this time frame. Though the 720-day estimate is not statistically significant, it is also large, indicating a 9.9 percentage point increase within this time frame. Lack of significance in the longer window might be explained by sample attrition, as the 720-day sample is only 61 percent of the size of the 360-day sample.

Table 6: Effect of integration on generalist RVUs (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.516***	0.044	1.314***	-0.620***	0.225*	1.784***
	-7.8311	0.6890	3.3685	-6.4062	1.7899	3.2531
Severity	0.059***	0.015	0.054***	0.083***	0.087***	0.080***
	3.8077	0.9491	2.9755	3.6206	2.6296	3.0253
Severity squared	-0.002	0.001	-0.001	-0.005	-0.007	-0.004
	-0.8323	0.3507	-0.2939	-1.6084	-1.2850	-1.2847
Male	-0.015	-0.001	-0.015	0.037	0.074**	0.036
	-0.9242	-0.0595	-0.9312	1.1524	2.2553	1.0701
Age	-0.010*	-0.003	-0.010*	0.018	0.023*	0.016
	-1.7958	-0.3781	-1.8547	1.6083	1.9037	1.5259
Age squared	0.000	0.000	0.000	-0.000	-0.000	-0.000
	1.3709	0.6776	1.3534	-0.8954	-0.8623	-0.8046
Distance to nearest specialist	-0.026***	-0.006	-0.019***	-0.030***	-0.012	-0.020***
	-9.9631	-0.8612	-5.2159	-7.4358	-0.7582	-3.9925
Distance to nearest specialist squared	0.000***	0.000	0.000***	0.000***	-0.000	0.000
	5.1907	0.8308	2.7761	3.2357	-0.5231	1.4219
F	–	–	52.6	–	–	48.3
N	183,467	183,467	180,545	114,544	114,544	112,706

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Though IV is potentially biased upward due to physician selection, the actual bias may indeed be small. The distance between the OLS and OLS FE estimates appears small, especially compared to the distance between the OLS and IV estimates. This suggests the role of physician selection is smaller than the role of patient selection. Furthermore, the effect of severity is positive, economically significant, and both terms are statistically significant, lending further credibility to my result. The effect of the other reported observables – male, age and its square, and distance to the nearest specialist and its square – are extremely small. In fact, these coefficients are economically insignificant in almost every regression. Therefore in subsequent sections I will only discuss these coefficients when they are comparable in importance to integration or severity.

Table 7: Effect of integration on cardiologist RVUs (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.115	-0.768**	0.709	0.181	-0.715	1.862
	0.8114	-2.1075	0.5028	1.4184	-1.3990	1.3508
Severity	0.446***	0.461***	0.430***	0.521***	0.435***	0.495***
	6.3418	4.4826	6.0344	6.2632	2.8906	5.8232
Severity squared	-0.027***	-0.035***	-0.026***	-0.026***	-0.018	-0.023***
	-4.4724	-3.4253	-3.9487	-3.4478	-0.9929	-3.0215
Male	0.217***	0.264***	0.221***	0.304***	0.377***	0.316***
	3.6843	2.6841	3.7571	5.7829	4.8433	5.8003
Age	-0.092***	-0.031	-0.088***	-0.079***	-0.059	-0.076***
	-4.1687	-0.8577	-4.2878	-2.8973	-1.4092	-2.8630
Age squared	0.001***	0.000	0.001***	0.001***	0.001*	0.001***
	4.4263	0.9720	4.6497	3.2204	1.6830	3.1856
Distance to nearest specialist	-0.015	-0.058	-0.013	-0.020**	-0.006	-0.014*
	-1.5499	-1.4009	-1.2538	-2.1081	-0.1676	-1.9000
Distance to nearest specialist squared	0.000**	0.002	0.000**	0.000**	0.000	0.000*
	2.5302	1.2188	2.2715	2.1914	0.3147	1.7199
F	–	–	67.6	–	–	51.6
N	29,297	29,297	28,914	25,291	25,291	24,945

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Physician utilization

Estimates for the generalist RVU regressions are presented in Table 6. OLS estimates are negative and statistically significant at the 99 percent level, indicating reductions in generalist utilization by 0.52 RVUs at 360 days and 0.62 RVUs at 720 days. In contrast, OLS FE estimates are positive in both windows, indicating smaller increases in generalist utilization of 0.04 RVUs at 360 days and 0.22 RVUs at 720 days. The latter estimate is significant at the 90 percent level. The IV estimates are positive and larger still, indicating that generalist utilization increases by 1.31 RVUs at 360 days and 1.78 RVUs at 720 days. Both IV estimates are statistically significant at the 99 percent level.

I present estimates for the cardiologist regressions in Table 7. OLS estimates are positive, small, and statistically insignificant, indicating an increase of 0.12 cardiologist RVUs at 360 days and 0.18 cardiologist RVUs at 720 days. In contrast, OLS FE estimates are negative and noticeably larger. The 360-day estimate is statistically significant at the 95 percent level, indicating a decrease of 0.77 cardiologist RVUs at 360 days. The 720-day estimate is also negative, but smaller in magnitude and statistically insignificant, indicating a decrease of 0.72 RVUs within this time period.

I do not detect statistically significant effects when I implement the IV estimator. However, the estimates are economically significant and positive. If the IV estimator is indeed biased downward, as I argue in Chapter 4 Section 4.3.1, then my IV estimates provide strongly suggestive evidence that integration induces greater cardiology utilization. Treating the IV coefficients as lower bounds for the effect of integration, my estimates indicate increases of at least 0.71 RVUs within 360 days and 1.86 RVUs within 720 days. And even if physician selection biases my estimates, the effect of integration on cardiologist RVUs dominates the effect of severity for most patients. Within 360 days, the effect of integration is greater than the effect of severity for patients with a severity score of 1.86 or lower. Within 720 days, the threshold value is even higher, a score of 4.86. Both values are far into the right tail of the window-specific severity distributions.

In recent iterations of RBRVS, an office visit for an established patient can generate 0.6 to 1.2 RVUs. The former value is for a 10-minute visit, the latter value is for a 15-minute visit.[4] For simplicity, let 1 RVU represent the average value of an office visit. Therefore one way to size these results is to think of integration increasing utilization by two routine office visits within 360 days (0.71 RVUs for cardiologists and 1.31 RVUs for generalists) and 3.5 routine office visits within 720 days (1.86 RVUs for cardiologists and 1.78 RVUs for generalists). These modest per-patient effects can lead to large aggregate effects. It is estimated that about 1 in 3 adults in the United States, approximately 70 million individuals, were affected by hypertension in 2011.[5] According to 2012 data from the National Ambulatory Medical Care Survey (NAMCS), there were 929 million physician office visits. About 80 percent

[4]See `https://www.medicalhomeportal.org/link/301` for coding guidelines and the RVU15A Physician Fee Schedule data file from CMS for the RVU values.

[5]Notably, this is higher than the 1 in 4 rate in 1999-2000. See `http://circ.ahajournals.org/content/112/11/1651.full`.

Table 8: Effect of integration on the probability of outpatient service (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.003	-0.007	-0.081***	-0.002	-0.008	-0.131***
	-0.6830	-0.8194	-3.8415	-0.3023	-0.9343	-4.1249
Severity	0.002*	0.002*	0.002*	0.005***	0.006***	0.005***
	1.8294	1.9593	1.8970	4.4204	4.8593	4.6121
Severity squared	0.000	0.000	0.000	-0.000	-0.000*	-0.000
	1.0219	0.0480	0.9182	-0.0344	-1.8132	-0.2186
Male	-0.002**	-0.001	-0.002***	-0.002	-0.001	-0.003*
	-2.4416	-1.2185	-2.7345	-1.6235	-0.7781	-1.7595
Age	-0.001**	-0.000	-0.001*	-0.001	-0.000	-0.001
	-2.0116	-0.9616	-1.9109	-1.4796	-0.2272	-1.4555
Age squared	0.000*	0.000	0.000	0.000	0.000	0.000
	1.7180	0.7068	1.6116	1.3623	0.1633	1.3638
Distance to nearest specialist	0.001***	0.000	0.000	0.001***	-0.000	0.001
	3.5183	0.6550	1.6261	3.3481	-0.0839	1.4355
Distance to nearest specialist squared	-0.000	0.000	0.000	-0.000	0.000	0.000
	-0.3998	1.2718	0.4728	-0.5513	1.2047	0.3302
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

of patients in my sample are treated in non-integrated practices. For sake of argument, if all patients in non-integrated practices were moved to integrated practices – ignoring factors such as behavioral responses and capacity constraints – it would induce an increase in physician work equivalent to 13.0 million office visits, or 1.4 percent of total office visits in 2012.

Outpatient utilization

There is strong evidence that integration reduces outpatient utilization on the extensive margin and suggestive evidence that integration increases outpatient utilization on the intensive margin.

I present results from the any outpatient claim regression in Table 8. Estimates from OLS, OLS FE, and IV regressions are all negative. OLS estimates are small and statistically insignificant, indicating a reduction of 0.3 percentage points at 360 days and 0.2 percentage

Table 9: Effect of integration on outpatient RVUs (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.47**	-5.88	3.30	-0.46**	1.29	3.62
	-2.350	-0.853	1.338	-2.551	0.941	1.637
Severity	0.389***	0.195	0.260**	0.351**	0.193	0.236*
	3.0762	0.6553	1.9935	2.3561	0.5030	1.7876
Severity squared	0.000	0.040	0.022	0.006	0.031	0.027
	0.0208	0.7505	1.1152	0.2562	0.4491	1.3921
Male	-0.150	-0.068	-0.162	-0.050	0.025	-0.048
	-1.1087	-0.2575	-1.1707	-0.3435	0.0962	-0.3225
Age	-0.040	-0.176	-0.023	-0.069	-0.103	-0.072
	-0.6453	-1.5062	-0.3255	-1.0135	-0.5926	-1.0021
Age squared	0.000	0.002	-0.000	0.000	0.001	0.000
	0.2980	1.4783	-0.0188	0.6473	0.5641	0.6121
Distance to nearest specialist	-0.048***	0.090	-0.036**	-0.047***	0.274	-0.033*
	-3.0347	0.2420	-2.2132	-3.0628	0.9829	-1.9028
Distance to nearest specialist squared	0.000	0.000	0.000	0.000	-0.004	0.000
	1.4733	0.0356	0.8518	1.3824	-1.0220	0.5653
F	–	–	28.5	–	–	27.5
N	6,048	6,048	5,874	5,539	5,539	5,373

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

points at 720 days. OLS FE estimates are not much larger and are also statistically insignificant. They indicate reductions of 0.7 percentage points and 0.8 percentage points in the 360- and 720-day windows, respectively. In contrast, IV estimates are larger by an order of magnitude and statistically significant at the 99 percent level. They indicate reductions in the probability of an outpatient claim by 8.1 percentage points at 360 days and 13.1 percentage points at 720 days. These extensive margin effects I estimate are quantitatively large. In the 360-day window, only 3 percent of patients have an outpatient claim for hypertension.

Estimates for the outpatient RVU regressions are presented in Table 8. OLS estimates are modest, significant at the 95 percent level, and almost identical across windows. These estimates indicate reductions of 0.47 and 0.46 RVUs within 360 and 720 days, respectively. Estimates from fixed effect regressions are statistically insignificant and far different from each other. The 360-day estimate is negative and large, indicating a reduction of 5.88 RVUs within the window. The 720-day estimate is positive and smaller in magnitude, indicating an increase of 1.29 RVUs within the window. It is possible the observed differences are a result of small sample sizes. Because I estimate the model only for individuals with an outpatient claim, and outpatient service rates are small, sample sizes are merely $6,048$ at 360 days and $5,373$ at 720 days. In one sense my IV results are consistent with this fact, as both estimates are statistically insignificant. However, in contrast to the OLS FE results, they are same-signed. This set of estimates indicate outpatient RVU increases of 3.30 and 3.62 at 360 and 720 days, respectively.

Like the extensive margin estimates, these intensive margin estimates are quantitatively large. Mean outpatient RVUs in the full sample are 2.24 in the 360-day window and 2.33 in the 720-day window. The effects are also large when compared to the severity effects. In both windows the effect of severity is positive and economically significant, but less than the effect of integration for all but the largest values of the severity score.

Prescription drug utilization

In Tables 10 and 11 I present estimates from the regression of the new cardiovascular prescription indicators on my regressors.[6] Table 10 contains the result for generalists, that is among all patients, the regressand equals 1 for patients who obtain a new prescription from his or her generalist and 0 otherwise. The OLS and OLS FE estimates are small, positive, and statistically insignificant. Within the 360-day window, the estimates are identical. Both models indicate a 0.3 percentage point increase in the probability that a generalist issues (and the patient fills) a new cardiovascular drug prescription.[7] Within the 720-day

[6]Though I estimated models in which the dependent variable was the count of new cardiovascular prescriptions or the count of all cardiovascular prescriptions, I do not present those results. First, the estimates for the effect of integration in those regressions statistically insignificant. More importantly, the count of prescriptions can be a misleading measure of drug utilization. A higher count can be a result of greater willingness to adjust prescriptions if hypertension control is not being achieved. Higher counts are also consistent with prescribing activity inconsistent with more recent sets of clinical guidelines. If those guidelines describe more effective care, then those closer to the standard might not need to adjust medications as often as those further from the standard.

[7]I do not observe non-compliant patients. Therefore there is downward bias in my new prescription indicator. However, because my instrument simulates random assignment of patients between integrated and non-integrated practices, and mean observables in the integrated and non-integrated groups are highly

Table 10: Effect of integration on the probability of a new generalist prescription (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.003	0.003	0.191*	0.003	0.010	0.149
	0.3487	0.3037	1.9476	0.3835	0.5957	1.2263
Severity	0.250***	0.264***	0.250***	0.254***	0.277***	0.254***
	20.7245	20.1813	20.5839	29.1518	29.8085	29.0214
Severity squared	-0.022***	-0.024***	-0.022***	-0.022***	-0.025***	-0.021***
	-10.8831	-10.8842	-10.7604	-15.3139	-16.9706	-15.0855
Male	0.036***	0.035***	0.036***	0.036***	0.037***	0.036***
	8.3540	6.4379	8.0643	6.8944	5.3762	6.6959
Age	0.008***	0.006***	0.008***	0.007***	0.006***	0.006***
	6.4583	5.8946	6.6942	4.6668	5.3543	4.5236
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	-8.1335	-7.6811	-8.4191	-6.6150	-7.5880	-6.4511
Distance to nearest specialist	0.000	0.003**	0.001*	-0.000	0.001	0.001
	0.1922	2.3444	1.7707	-0.2026	0.5264	0.7894
Distance to nearest specialist squared	-0.000	-0.000***	-0.000**	-0.000	0.000	-0.000*
	-1.3569	-2.6949	-2.3528	-0.8971	0.0816	-1.8610
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

window, the OLS estimate indicates a 0.3 percentage point increase while the OLS FE estimate indicates a 1.0 percentage point increase. None of these estimates is statistically significant. It is encouraging that the estimates are similar.

IV estimates are significantly larger. At 360 days, integration increases the probability of a generalist prescription by 19.1 percentage points. This effect is statistically significant at the 90 percent level. At 720 days, the estimate is similar, though statistically insignificant, indicating a 14.9 percentage point increase in the probability of a new cardiovascular drug prescription.

The effect of integration is large in absolute terms. However, the effect of severity is larger than the effect of integration for an economically significant range of severity scores. For patients with a severity score less than 0.824, the effect of integration is larger. For patients with a severity score in between 0.824 and 10.540, basically all but the most severe patients in my sample, the effect of severity is far larger.[8] The difference is large. For nearly all individuals in this latter group, the effect of severity exceeds the effect of integration by at least 1.0 percentage points.

Table 11 contains the regression for specialists, that is, among those patients with a cardiologist visit, the regressand equals 1 for those who obtain a new prescription from a cardiologist and 0 otherwise. There is little evidence in my data that cardiologists change their prescribing behavior in response to integration. If there is an effect it would be very

similar, it is unlikely that non-compliant patients are non-randomly distributed among practice types.

[8]It is possible my fit would be better if it included a cubic term in severity. Unfortunately, the outcomes I used to set my initial specification did not include the new prescription regressand due to the cost of processing the pharmacy files. These files were significantly larger than the physician, outpatient, and inpatient files.

That said, I argue it is important to impose a consistent functional form in these regressions for sake of consistency and comparability. On these grounds it is more difficult to argue for the inclusion of a cubic term when the quadratic terms are almost always economically insignificant, as well as statistically insignificant, in almost all other regressions. That said, my pharmacy results are likely to provide a useful first approximation at the very least, as the model predicts worst for the highest-severity patients, the smallest group of patients in my sample.

When I reestimate the pharmacy model with a cubic term in severity, the results are highly similar. The effect of integration increases (360 days: +1.0 percentage points, 720 days: XXX), but the cubic term is positive as well as economically and statistically significant.

Table 11: Effect of integration on the probability of a new cardiologist prescription (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.008	0.013	0.031	0.007	-0.005	-0.023
	1.1363	0.7345	0.4621	1.2027	-0.2385	-0.2906
Severity	0.045***	0.039***	0.044***	0.061***	0.062***	0.060***
	8.0071	4.9654	7.9371	11.5580	14.5721	11.8846
Severity squared	-0.003***	-0.003**	-0.003***	-0.004***	-0.004***	-0.004***
	-3.4169	-2.1672	-3.3634	-5.8719	-7.3658	-5.9398
Male	0.016***	0.014***	0.016***	0.020***	0.017***	0.020***
	5.1870	3.7692	5.2936	5.8089	3.1559	6.0583
Age	0.001	-0.001	0.002*	0.000	-0.001	0.000
	1.4950	-0.2911	1.7152	0.1388	-0.3260	0.0756
Age squared	-0.000	0.000	-0.000*	0.000	0.000	0.000
	-1.4330	0.3332	-1.6750	0.0573	0.5940	0.1301
Distance to nearest specialist	-0.000	-0.000	-0.000	-0.001	-0.003	-0.001
	-0.1221	-0.1486	-0.1561	-0.8822	-1.3300	-0.8988
Distance to nearest specialist squared	0.000	0.000	0.000	0.000	0.000	0.000
	0.4393	0.7447	0.5369	1.0042	1.1406	0.7325
F	–	–	8.5	–	–	17.8
N	30,046	30,046	29,654	25,924	25,924	25,569

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

small when compared to the effect of integration on generalist prescribing. OLS estimates are positive but small and statistically insignificant, indicating increases in the probability of a new cardiologist prescription of 0.8 and 0.7 percentage points within the 360- and 720-day windows, respectively. The 360-day OLS FE estimate is slightly larger, indicating a 1.3 percentage point increase in the probability of a cardiologist prescription, while the 720-day estimate is negative but smaller in magnitude, indicating a 0.5 percentage point reduction in the window. IV estimates are like the OLS FE estimates but greater in size. Within 360 days, the probability of a new cardiologist prescription increases by 3.1 percentage points. Within 720 days, the probability decreases by 2.3 percentage points. It is worth noting that the F statistic for the 360-day regression takes on a low value of 8.5, even though I do not make any strong claims based on this particular estimate.

Notably, there appears to be a much tighter relationship between cardiologist prescribing behavior and severity. The effect of severity in both windows is positive and economically significant. Furthermore, the t-statistics on the linear term – the quantitatively more important one – are extremely large for all estimators and in both windows. The t-statistics on the coefficient for the linear severity term when estimated with OLS, OLS FE, and IV are, respectively, 8.01, 4.97, and 7.94 in the 360-day window. In the 720-day window, t-statistics are, respectively, 11.56, 14.57, and 11.88.

Inpatient admission

In Table 12 I present estimates of the effect of integration on the probability of inpatient admission. OLS estimates are negative and statistically significant, indicating 0.1 percentage point and 0.2 percentage point decreases in the probability of an inpatient admission within, respectively, the 360- and 720-day windows. In contrast, OLS FE estimates are positive though statistically insignificant, indicating 0.1 and 0.2 percentage point increases in the probability of an inpatient admission. IV estimates are positive and, in fact, larger than the OLS FE estimates, indicating a 0.3 percentage point increase in the probability of an inpatient admission at 360 days and a 1.7 percentage point increase in the probability of an

Table 12: Effect of integration on the probability of inpatient admission (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.001*	0.001	0.003	-0.002*	0.002	0.017*
	-1.8231	0.3018	0.2956	-1.7608	0.5926	1.6749
Severity	0.005***	0.004***	0.005***	0.006***	0.006***	0.006***
	5.7489	6.7284	5.5657	4.4685	4.2897	4.3987
Severity squared	0.000	-0.000	0.000	0.001**	0.000	0.001**
	0.1595	-0.2038	0.2452	2.0910	1.0081	2.2255
Male	0.002***	0.003***	0.002***	0.004***	0.004***	0.004***
	4.4119	4.8486	4.3467	5.1932	4.5535	5.1437
Age	-0.000	-0.000**	-0.000	-0.000***	-0.001***	-0.000***
	-1.3047	-2.4722	-1.2163	-2.6117	-3.5466	-2.6110
Age squared	0.000**	0.000***	0.000**	0.000***	0.000***	0.000***
	2.2893	3.3922	2.1866	3.9681	5.1348	3.9297
Distance to nearest specialist	-0.000**	0.000	-0.000	-0.000	-0.000	-0.000
	-2.0515	0.0982	-1.3565	-1.5380	-1.3444	-0.3278
Distance to nearest specialist squared	0.000	-0.000	0.000	0.000	0.000	-0.000
	1.0040	-0.7686	0.8083	0.5961	1.1014	-0.0083
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

inpatient admission at 720 days. While the 360-day estimate is not statistically significant, the 720-day estimate is statistically significant at the 90 percent level. Severity has a positive impact on inpatient admissions, both severity coefficients are statistically significant, and the effect of severity is not dominated by the effect of integration.

The result is surprising. Unless one believes medical care is generally flat-of-the-curve, one would expect greater utilization to lead to lower rates of inpatient admission. It is possible that the positive effect I observe is an artifact of non-emergency utilization, specifically higher rates of elective inpatient utilization within integrated practices. I have tried to study this issue by using admission type classifiers in the data – e.g. emergency, urgent, or emergent admission – but my estimates lacked precision.

Table 13: Effect of integration on physician services spending (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	2.81	6.25	63.87	6.6	29.1***	97.0*
	0.703	1.449	1.335	1.31	3.30	1.77
Severity	5.98***	4.94***	5.77***	9.48***	9.84***	9.34***
	5.555	4.331	5.488	6.587	5.739	6.429
Severity squared	-0.248**	-0.299*	-0.206**	-0.402**	-0.420*	-0.365*
	-2.3494	-1.8981	-1.9732	-2.0880	-1.7479	-1.8213
Male	1.89*	2.95**	1.89*	7.38***	10.66***	7.32***
	1.654	2.000	1.655	3.230	3.629	3.145
Age	0.581*	0.937**	0.593*	3.54***	3.51***	3.48***
	1.8983	2.3698	1.9281	4.505	4.319	4.523
Age squared	-0.011***	-0.012***	-0.011***	-0.038***	-0.034***	-0.037***
	-3.3082	-2.9189	-3.3028	-4.4163	-3.9291	-4.4337
Distance to nearest specialist	-1.138***	-0.957	-0.866***	-1.48***	-0.70	-1.09***
	-5.1959	-1.5463	-3.6119	-5.431	-0.634	-3.654
Distance to nearest specialist squared	0.012***	0.015	0.010**	0.015***	-0.007	0.012**
	3.6066	1.0094	2.5324	3.2576	-0.5412	2.4116
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Medical spending

In Tables 13 through 17 I present my estimates for the effect of integration on medical spending.

Generalist and cardiologist spending

In Table 13, the explanatory variable is the sum of spending attributable to the patient's generalist and cardiologist physicians. OLS estimates are small and insignificant, indicating spending increases of 2.8 dollars at 360 days and 6.6 dollars at 720 days. OLS FE estimates are larger, showing increases of 6.3 dollars at 360 days and 29.1 dollars at 720 days. The 360-day estimate is statistically insignificant while the 720-day estimate is significant at the 99 percent level. IV estimates are substantially larger in both windows. In the 360-day window, being treated by an integrated generalist causes an increase in physician services spending of 63.9 dollars. In the 720-day window, the increase is 97.0 dollars. While the

Table 14: Effect of integration on outpatient services spending (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-4.37*	1.38	-43.46***	-5.52	0.52	-64.02***
	-1.838	0.351	-2.697	-1.522	0.100	-3.017
Severity	0.078	3.756**	-0.139	0.294	7.961**	-0.111
	0.0201	2.0711	-0.0367	0.0510	2.1646	-0.0198
Severity squared	1.13	0.24	1.14	1.69	0.20	1.73
	1.547	0.703	1.589	1.548	0.367	1.609
Male	-1.69	-1.56	-2.19*	-1.25	-0.96	-1.95
	-1.412	-1.376	-1.873	-0.757	-0.485	-1.246
Age	0.453	0.678	0.547	0.511	1.180	0.578
	0.9325	1.2024	1.2227	0.7276	1.6419	0.8866
Age squared	-0.007	-0.009	-0.008*	-0.008	-0.014*	-0.009
	-1.3008	-1.4288	-1.6688	-1.0454	-1.8230	-1.3037
Distance to nearest specialist	0.403**	0.941	0.194	0.289	-0.161	0.007
	2.1828	0.8963	0.9109	1.0234	-0.0948	0.0237
Distance to nearest specialist squared	0.000	0.016	0.001	0.003	0.009	0.004
	0.0667	0.4863	0.3258	0.6927	0.5296	1.0377
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

360-day estimate is statistically insignificant, the 720-day estimate is significant at the 90 percent level. Though the increases appear modest, the effect of integration dominates the (positive) effect of severity for all values of the severity score. The effect of the other observables are non-zero, but are still dominated by the effect of integration.

Outpatient spending

Table 13 the estimates from the outpatient services spending regressions. OLS estimates are negative and small in both windows. Within the 360-day window, the integration coefficient indicates a 4.4 dollar decrease in outpatient services spending. In the 720-day window the estimated effect is a 5.5 dollar decrease in outpatient services spending. The 360-day estimate is statistically significant at the 90 percent level while the 720-day estimate is statistically insignificant. Both OLS FE estimates are small and statistically insignificant, indicating spending increases of 1.4 and 0.5 dollars at 360 and 720 days, respectively. My

Table 15: Effect of integration on cardiovascular drug spending (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	17.6	9.6	224.0**	35.3	75.8	381.6
	1.40	0.43	2.39	1.11	0.81	1.33
Severity	616.8***	637.0***	616.0***	1136.8***	1181.0***	1137.7***
	23.32	24.27	23.38	29.58	36.18	29.57
Severity squared	-40.9***	-43.8***	-40.7***	-74.8***	-82.2***	-74.6***
	-11.84	-11.62	-11.79	-15.25	-18.01	-15.04
Male	137.2***	136.6***	138.0***	246.7***	233.4***	246.7***
	18.49	15.83	19.50	13.66	10.69	13.80
Age	13.7***	13.4***	13.5***	27.7***	31.2***	27.6***
	4.78	4.59	4.89	5.16	5.22	4.99
Age squared	-0.138***	-0.136***	-0.136***	-0.281***	-0.329***	-0.281***
	-4.0847	-3.9423	-4.1383	-4.4553	-5.0510	-4.2601
Distance to nearest specialist	-1.67	5.81***	-0.40	-3.59	8.97**	-1.48
	-1.480	2.633	-0.341	-1.451	2.215	-0.640
Distance to nearest specialist squared	0.009	-0.095**	-0.004	0.005	-0.164**	-0.018
	0.4843	-2.3517	-0.1845	0.1870	-2.1938	-0.6537
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

IV estimates are far larger, negative, and statistically significant at the 99 percent level. They indicate that integration reduces outpatient spending by 43.5 dollars within 360 days and 64.0 dollars within 720 days.

These estimates indicate that spending decreases due to reductions along the extensive margin more than offset the spending increases due to greater intensive margin utilization. Also, the effect of integration is large in the sense that it is greater in magnitude than the effect of severity for most values of the severity score in my sample. For individuals with scores less than or equal to 6.24 – i.e. most of the sample – the effect of integration is greater.

Cardiovascular drug spending

Table 15 contains the estimates from the regression for total cardiovascular drug spending. OLS estimates are positive but statistically insignificant, indicating spending increases of

17.6 dollars at 360 days and 35.3 dollars at 720 days. OLS FE estimates are also positive and statistically insignificant. The 360-day estimate indicates a 9.6 dollar increase in cardiovascular drug spending. Within 720 days, the increase is 75.8 dollars.

IV estimates are positive and far larger. They indicate increases in cardiovascular drug spending by 224.0 and 381.6 dollars, respectively, in the 360- and 720-day windows. Or, on a monthly basis, integration increases spending, respectively, by 18.7 and 15.9 dollars per month. Given mean spending per month of about 35 dollars, the effects I observe are quantitatively significant. That said, the effect of integration is small compared to the effect of severity for a wide range of individuals. For individuals with very low severity – a score less than 0.373 in the 360-day sample and and 0.343 in the 720-day sample – the effect is of integration is greater. Beyond these thresholds the effect of severity is greater. For example, at the mean values of severity, respectively 1.01 and 1.03, spending increases through the severity channel by 48.4 and 45.5 dollars per month.

Total spending

In Table 16 I set my outcome as the sum of physician, outpatient, and cardiovascular drug spending. This measure of costs represents the spending attributable to outpatient care. Both OLS and OLS FE estimates are positive and statistically insignificant. OLS estimates indicate small spending increases of 16.0 and 36.4 dollars at, respectively, 360 and 720 days. OLS FE estimates indicate larger increases of 17.2 and 105.4 dollars within the same windows. IV estimates are far larger. Within 360 days, the increase in spending is 244.4 dollars. This estimate is statistically significant at the 95 percent level. Within 720 days, the increase in spending is 414.6 dollars, though the estimate is not statistically significant.

In Table 17 I add inpatient spending to my cost measure. While office-based physicians do not directly affect inpatient spending – as in they might not have ordered the corresponding inpatient services – they have some connection to it through their pre-admission care. OLS estimates are small and statistically insignificant, indicating a decrease in spending of 13.7

Table 16: Effect of integration on all (non-inpatient) medical spending, (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	16.0	17.2	244.4**	36.4	105.4	414.6
	1.21	0.69	2.12	1.13	1.08	1.35
Severity	622.9***	645.7***	621.7***	1146.6***	1198.8***	1147.0***
	24.58	24.98	24.60	29.77	36.42	29.68
Severity squared	-40.0***	-43.8***	-39.7***	-73.5***	-82.5***	-73.2***
	-12.09	-11.88	-12.03	-14.45	-18.77	-14.22
Male	137.4***	138.0***	137.7***	252.8***	243.2***	252.1***
	19.14	15.94	20.04	13.98	11.28	13.90
Age	14.7***	15.0***	14.6***	31.7***	35.9***	31.6***
	4.49	4.43	4.60	5.36	5.73	5.20
Age squared	-0.156***	-0.157***	-0.155***	-0.327***	-0.377***	-0.327***
	-4.0747	-3.9383	-4.1411	-4.7825	-5.5586	-4.5938
Distance to nearest specialist	-2.41**	5.80***	-1.07	-4.77**	8.11*	-2.56
	-2.228	2.755	-0.989	-1.979	1.723	-1.194
Distance to nearest specialist squared	0.021	-0.065	0.007	0.024	-0.162**	-0.003
	1.1706	-1.4752	0.3532	0.8539	-2.0812	-0.0977
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 17: Effect of integration on all medical spending (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-13.7	-7.0	538.9*	31.1	137.0	1060.2*
	-0.35	-0.09	1.83	0.44	0.56	1.78
Severity	717.1***	754.9***	713.5***	1127.2***	1265.0***	1125.4***
	15.05	24.03	14.88	8.65	10.83	8.70
Severity squared	-36.2***	-48.6***	-35.5***	-28.3	-58.6***	-27.3
	-3.56	-10.93	-3.46	-1.12	-2.84	-1.08
Male	190.5***	190.1***	188.9***	380.2***	373.4***	382.1***
	10.85	10.17	11.00	11.51	8.25	11.30
Age	8.40	4.74	9.57	32.0***	22.6*	32.5***
	0.990	0.470	1.138	2.62	1.67	2.68
Age squared	-0.050	-0.009	-0.062	-0.249*	-0.149	-0.258*
	-0.5331	-0.0798	-0.6643	-1.8378	-1.0309	-1.8975
Distance to nearest specialist	-6.41***	-7.68	-3.82*	-12.19***	5.55	-6.87*
	-2.873	-1.252	-1.655	-2.637	0.387	-1.652
Distance to nearest specialist squared	0.086**	0.082	0.063	0.112*	-0.148	0.060
	2.4861	0.8645	1.5580	1.9303	-0.8487	1.0948
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

dollars at 360 days and an increase in spending of 31.1 dollars at 720 days. OLS FE estimates are also statistically insignificant, indicating a decrease in spending of 7.0 dollars at 360 days and an increase in spending of 137.0 dollars at 720 days. IV estimates are far larger and are both statistically significant (90 percent level). I observe a 538.9 dollar increase at 360 days and a 1,060.2 dollar increase at 720 days. Because the hospitalization rate in my sample is about 1 percent, inpatient spending is necessarily negatively skewed, and right-tail values must be driving this result.

Nevertheless, these effects are large on a per-person and population basis. The average deductible for an individual with employer-sponsored health insurance (single coverage) in 2012 was 1,097 dollars. Therefore, the estimated one-year effect of integration on medical care spending represents 49.1 percent of the average deductible at the end of my sample.[9] Furthermore, it is estimated that about 1 in 3 adults in the United States, approximately 70 million individuals, were affected by hypertension in 2011.[10] About 80 percent of patients in my sample are treated in non-integrated practices. For sake of argument, suppose this ratio holds for the entire adult US population and ignore behavioral reponses, capacity constraints, and other factors that certainly affect counterfactual outcomes. My estimates indicate, in this naive counterfactual, that shifting all those adults with hypertension into integrated practices would increase medical care spending by about 30 billion dollars per year. In comparison, direct medical spending for hypertension in 2010 totaled 42.9 billion dollars.

Whether or not I have recovered the true magnitude of the effect of integration on costs, I argue I have provided abundant evidence that integration does not reduce medical care spending. This is especially damning for the most optimistic integration advocates, because hypertension – as a chronic condition – is one of those conditions for which care coordination is said to especially matter. That said, it is still possible these spending increases are socially

[9] See Claxton et al. (2015) for deductible estimates.

[10] This is significantly higher than the 1 in 4 rate in 1999-2000. See `http://circ.ahajournals.org/content/112/11/1651.full`.

desirable, but they spending increases I observe cannot be justified on the basis of outcomes with the one- and two-year estimates I have.

Additional discussion

Estimates from the full sample show that integration affects physician behavior along many margins. For all but one outcome – the probability of a new cardiologist prescription – the estimate for integration has the same sign across both windows. For each of those outcomes, IV yields a statistically significant estimate in one of the treatment windows.

Some of my estimates are consistent with claims made by proponents of integration. The reduction in outpatient utilization along the extensive margin suggests that integration of generalists and cardiologists reduces unnecessary testing for patients who do not need it. Higher levels of hypertension-specific human capital among cardiologists should lead to lower levels of clinical uncertainty and, therefore, lower rates of testing. However, achieving these lower rates is accompanied by greater utilization along other margins. Generalist and cardiologist RVUs both appear to increase in response to integration and it appears generalists in integrated firms are also more likely to prescribe antihypertensive medications.

It is not clear that these adjustments, on net, are socially beneficial. First, integration does not lead to lower rates of inpatient admission. This highly surprising result could reflect greater elective utilization instead of greater incidence of adverse outcomes. Either way, greater rates of inpatient admission lead to unambiguously greater medical care spending due to the high cost of inpatient care.

Advocates for greater integration among generalists and specialists should be discouraged by these results. Hypertension is a very important test case for organizational reform proposals. Because hypertension is a chronic condition with potentially severe complications, interventions that presumably increase care coordination should positively affect individuals with this condition. Unfortunately, the effects I do find indicate potentially worse outcomes and higher costs when generalists and cardiologists integrate.

CHAPTER 6 : Hypertension: Evidence from low- and high-access areas

Introduction

In the following two chapters I divide patients into two contrasting subsamples, reestimate my models for each subgroup, and look for differential responses to integration. The low-access subsample contains individuals whose generalist is located "far" from a cardiologist. The high-access subsample contains individuals whose generalist is located "close" to a cardiologist. To define "far" and "close," I computed three quantiles of the distance to the nearest specialist variable. Those in the lowest quantile are in the "close" group while those in the highest quantile are in the "far" group.[1] Differential effects should emerge because distance to the nearest specialist is literally a measure of clinical isolation for both generalists and patients. For patients marginally appropriate for referral, greater distance should reduce the probability of referral due to the inconvenience of traveling, ceteris paribus. The added incentive to refer generated by integration should increase referral probability in the low-access sample.

For ease of reading, in the tables that follow I compare the IV estimate for integration across each of my samples. Interested readers will find a full complement of tables at the end of this document.

Results

Cardiologist visit

Table 18 contains IV estimates for integration for the full, low-access, and high-access subsamples. F statistics and sample sizes for each sample are also included. Because distances are computed at the ZIP code level and many patients are located in areas dense in both population and cardiologists, the high-access subsample contains far more individuals

[1] As described in Chapter 2, I attempted to divide the sample into rural and urban subsamples or groups containing a high density versus low density of specialists. However, rural subsamples were extremely small and data loss after matching ZIPs to county with free, publicly available crosswalks was significant. Though imperfect, my final subsamples are perhaps the best of the remaining alternatives.

Table 18: IV comparison, primary specialist visit

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	0.115*	0.125	0.071	0.099	0.104	0.055
F	52.8	52.3	46.1	48.3	43.8	45.6
N	180,714	59,698	107,621	112,803	37,270	67,210

than the low-access subsample.

None of the subsample estimates are statistically significant. However, they are positive and economically significant. In low-access areas the effect of integration on the probability of a specialist visit is 12.5 percentage points at 360 days and 10.4 percentage points at 720 days. These effects are greater than the corresponding full-sample estimates. In high-access areas, the effect of integration is also positive, measuring 7.1 and 5.5 percentage points within 360 and 720 days, respectively.

Physician utilization

Table 19: IV comparison, physician utilization

Panel A: Generalists

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	1.314***	0.227	1.923***	1.784***	1.129	2.206**
F	52.6	52.3	45.7	48.3	43.9	45.3
N	180,545	59,648	107,491	112,706	37,226	67,137

Panel B: Cardiologists

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	0.709	-1.572	1.804	1.862	1.093	3.280*
F	67.6	30.6	89.8	51.6	24.9	68.2
N	28,914	9,587	17,566	24,945	8,274	15,122

Panel A of Table 19 compares effects from IV regressions for generalist total RVUs, while Panel B compares effects from the regressions for cardiologist RVUs. In both panels, the effect is stronger within high-access areas. For generalists, the high-access estimates are 46.3 and 23.7 percent larger than the corresponding full-sample estimates. The estimates are statistically significant at the 99 and 95 percent levels, respectively. For specialists, the differences between the high-access and full-sample estimates are even larger. The high-access estimates are 154.4 and 76.2 percent greater than the corresponding full-sample estimates. The 720-day estimate is statistically significant at the 90 percent level.

Outpatient utilization

Table 20: IV comparison, outpatient utilization

Panel A: Probability of any outpatient utilization

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	-0.081***	-0.004	-0.137***	-0.131***	-0.034	-0.204***
F	52.8	52.3	46.1	48.3	43.8	45.6
N	180,714	59,698	107,621	112,803	37,270	67,210

Panel B: Outpatient RVUs, conditional on at least one outpatient visit

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	3.30	-1.98	6.49**	3.62	-0.464	6.636***
F	28.5	7.3	21.9	27.5	7.1	24.5
N	5,874	1,909	3,339	5,373	1,752	3,084

In Panels A and B of Table 20 I present the effect of integration on outpatient utilization on, respectively, the extensive and intensive margins of outpatient utilization. The full-sample extensive margin effect I observe is driven by behavior in the high-access sample. In the full sample, integration causes 8.1 and 13.1 percentage point decreases in the probability of any outpatient utilization within the 360- and 720-day windows, respectively. Estimates in the low-access sample are negative like the full-sample estimates, but statistically insignificant and an order of magnitude smaller than the corresponding full-sample estimates. In contrast, integration in the high-access subsample reduces the probability of any outpatient utilization by 13.7 percentage points within 360 days and 20.4 percentage points within 720 days, significantly larger than the corresponding full sample estimates.

The intensive margin effect I observe in the full sample is also driven by behavior in the high-access sample. In the full sample, I observe (statistically insignificant) increases of

3.30 and 3.62 outpatient RVUs in the 360- and 720-day windows, respectively. Effects in the low-access sample are in fact negative. That said, the estimates are unreliable due to weak instrument problems (F statistics of 7.3 and 7.1, respective to the two windows). In the high-access sample, effects are about twice the size of the estimated full-sample effects. Integration increases outpatient RVUs in the high-access sample by 6.49 RVUs within 360 days and 6.37 RVUs within 720 days. The estimates are also statistically significant at the 95 and 99 percent levels, respectively.

Prescription drug utilization

Table 21: IV comparison, probability of a new cardiovascular drug prescription

Generalists

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	0.191*	-0.005	0.241***	0.149	-0.044	0.198*
F	52.8	52.3	46.1	48.3	43.8	45.6
N	180,714	59,698	107,621	112,803	37,270	67,210

Cardiologists

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	0.031	-0.067	0.072	-0.023	-0.105	0.020
F	8.5	5.1	5.8	17.8	7.2	12.4
N	29,654	9,826	18,070	25,569	8,473	15,528

I present estimates for the effect of integration on prescribing activity in Table 21. Panel A contains the estimates for the probability of a new cardiovascular prescription among generalists while Panel B contains the corresponding estimates for cardiologists. In the full sample, I find integration increases the probability that a generalist issues a new prescription by 19.1 percentage points in the 360-day window and 14.9 percentage points in the 720-day window. The 360-day estimate is statistically significant at the 90 percent level while the 720-day estimate is statistically insignificant. I find small, negative, but statistically insignificant effects in the low-access sample, indicating reductions of 0.5 and 4.4 percentage points within their respective windows. In contrast, integration increases the probability of a new generalist prescription by 24.1 percentage points in the 360-day window and 19.8

percentage points in the 720-day window. Both these estimates are statistically significant (99 and 95 percent level, respectively). They are significantly larger than the corresponding full-sample estimates.

It is difficult to interpret the cardiologist estimates in the low- and high-access samples due to weak instrument problems, so I forego a discussion of them.

Inpatient admission

Table 22: IV comparison, inpatient admission

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	0.003	-0.006	0.009	0.017*	-0.014	0.040***
F	52.8	52.3	46.1	48.3	43.8	45.6
N	180,714	59,698	107,621	112,803	37,270	67,210

In Table 22, I present estimates for the effect of integration on the probability of inpatient admission. In the full sample, I estimate increases of 0.3 and 1.7 percentage points in the probability of an inpatient admission. Only the 720-day estimate is statistically significant (90 percent level). These full sample effects are driven by positive effects in the high-access sample. In the high-access sample, I observe a statistically insignificant increase of 0.9 percentage points within 360 days and a statistically significant increase (99 percent level) of 4.0 percentage points within 720 days. Interestingly, I observe negative, though statistically insignificant effects, within the low-access sample. The low-access estimates indicate reductions of 0.6 and 1.4 percentage points within, respectively, the 360- and 720-day windows.

If these estimates are indeed indicative of the true sign of the effect of integration on the probability of an inpatient admission, then the results in this section suggest that the most important effect of integration for outcomes is its effect on the probability of a relevant specialist visit and clinical effects due to adjustments along other margins might actually be second-order. I will return to this point in the subsequent chapter.

Medical spending

Table 23: IV comparison, medical care spending

Panel A: Physician costs

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	63.87	-9.60	96.35*	97.0*	60.6	97.3*
F	52.8	52.3	46.1	48.3	43.8	45.6
N	180,714	59,698	107,621	112,803	37,270	67,210

Panel B: Outpatient costs

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	-43.46***	-46.61*	-46.25**	-64.02***	-49.06	-66.07**
F	52.8	52.3	46.1	48.3	43.8	45.6
N	180,714	59,698	107,621	112,803	37,270	67,210

Panel C: Pharmacy costs

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	224.0**	-213.9	416.8***	381.6	-67.7	629.7**
F	52.8	52.3	46.1	48.3	43.8	45.6
N	180,714	59,698	107,621	112,803	37,270	67,210

Panel D: All non-inpatient costs

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	244.4**	-270.1	466.9***	414.6	-56.2	660.9**
F	52.8	52.3	46.1	48.3	43.8	45.6
N	180,714	59,698	107,621	112,803	37,270	67,210

Panel E: All medical spending

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	538.9*	-178.5	1041.3***	1060.2*	-367.4	1935.4***
F	52.8	52.3	46.1	48.3	43.8	45.6
N	180,714	59,698	107,621	112,803	37,270	67,210

In Table 23 I compare IV estimates in the full, low-access, and high-access samples on the effect of integration on medical care spending. Generally speaking, estimates indicate that outcomes in the high-access sample drive the full-sample estimates. Integration in high-access areas leads to greater spending on physician services, prescription drugs, all non-inpatient services, and total costs. The exception is outpatient costs, where the differences between low-access and high-access areas are relatively muted.

Interestingly, though unsurprisingly, I detect economically significant reductions in total spending due to integration in the low-access sample. These effects are statistically insignificant. They track the statistically insignificant reductions in inpatient admission for low-access areas. For sake of argument, suppose integration truly has the above effects on spending and outcomes. These estimates illustrate the importance as well as the difficulty of choosing an appropriate intervention when trying to alter patterns of physician organization. Any improvements due to integration in the low-access subgroup are swamped by worse outcomes and increased spending in high-access areas.

Additional discussion

Consistent with intuition, the positive effect of integration on the probability of a specialist referral more strongly affects generalists in low-access areas. It is possible integration in low-access areas leads to lower rates of inpatient admission and lower levels of medical spending, but none of the relevant estimates are statistically significant. In contrast, Integration in high-access areas strongly results in higher levels of physician utilization, prescription drug utilization, and total medical spending.

The divergence in results across the two samples is potentially puzzling. However, disutility for driving might rationalize the patterns I observe. A patient might be willing to make one long trip to see a cardiologist, but unwilling to see that cardiologist for multiple visits. This preference would generate both the stronger responses I observe in high-access areas if utilization is driven by the number of cardiologist visits as well as the stronger specialist visit effect I observe in low-access areas.

Note that if it is in fact the case that integration in low-access areas leads to better outcomes and lower spending – a possibility hinted at in my results – then my results would indicate that care coordination problems for hypertension are small and policies strongly incentivizing relatively severe changes in organizational form are, in a sense, overshooting. Results from the low-access sample suggest that generalists might only need minimal assistance from cardiologists in managing hypertension. A lower-intensity intervention, such as carefully adjusting anti-kickback provisions or finding ways to help private payers develop socially beneficial yet profitable utilization management policies, might achieve higher levels of referral without inducing the far higher levels of utilization observed in the high-access sample.

CHAPTER 7 : Hypertension Evidence from non-affiliated and affiliated practices

Introduction

In this chapter I re-run my regressions on hospital affiliation-based samples. In the SK&A data, there is a variable that denotes whether an office reports a hospital affiliation. Patients treated by a generalist in an office with a hospital affiliation belong to the "affiliated" group while the rest are assigned to the "not affiliated" group. Data quality for this variable is potentially lower than the other identifiers and coordinates I have used in the SK&A data. First, the variable (`CODE5`) suffers from the same missing values problem as the inter-office linking variable (`CODE3`). Though I impute values to improve data quality of the latter, I have not imputed values for the former. Furthermore, other researchers have expressed concerns that survey respondents do not have a uniform understanding of what SK&A means by "hospital affiliation."[1] That said, if the hospital affiliation variable contains useful information, differential effects should emerge among the two subsamples because generalists without a hospital affiliation are more clinically isolated relative to those with an affiliation.

In contrast to the access-based subsamples, sample sizes for the subgroup of interest – the not-affiliated practices – are far larger than the other subgroup, as most practices do not report a hospital affiliation in the SK&A data.

Results

Cardiologist visit

In Table 24 I compare estimates of the effect of integration on the probability of a cardiologist visit in the full, not-affiliated, and affiliated samples. In the full sample I detect positive effects of 11.5 and 9.9 percentage points within the 360- and 720-day windows, respectively. The not-affiliated estimates are positive, but smaller and statistically insignificant. The 360-

[1] Source: Private conversations with other SK&A users. Interested researchers may contact the author for more information.

Table 24: IV comparison, primary specialist visit

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	0.115*	0.089	0.198**	0.099	0.039	0.298**
F	52.8	33.4	36.7	48.3	31.3	27.3
N	180,714	153,185	27,529	112,803	95,029	17,774

day estimate indicates an 8.9 percentage point increase in the probability of a cardiologist visit, while the 720-day estimate indicates only a 3.9 percentage point increase. In contrast, the estimates in the affiliated subsample are significantly higher, indicating a 19.8 percentage point increase at 360 days and a 29.8 percentage point increase at 720 days. These estimates are statistically significant at the 95 percent level.

These results are contrary to intuition, that more clinically isolated generalists will respond more strongly to integration. They may be due to some underlying clinical feature I do not observe. In the diabetes analysis, I find integration increases the probability of an endocrinologist in the full sample, low-access sample, and not-affiliated sample, making this result the exception among all my results, and preserving the confidence one might currently have in my results. I further discuss the interpretation of this result in the final section of the chapter.

Physician utilization

In Table 25 I compare estimates of the effect of integration on physician utilization across my samples. Panel A contains estimates from the generalist RVU regressions while Panel B contains estimates from the cardiologist RVU regressions.

Estimates in Panel A show that integration increases generalist utilization in both subsamples. In the shorter window, the not-affiliated estimate is greater than the affiliated estimate (1.49 vs 0.90) and both are statistically significant (99 and 90 percent level, respectively). However, in the longer window the not-affiliated estimate is smaller than, though almost equal to, the affiliated estimate (1.80 vs. 1.86). As in the 360-day window, both 720-day

Table 25: IV comparison, physician utilization

Panel A: Generalists

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	1.314***	1.490***	0.900*	1.784***	1.804***	1.855**
F	52.6	33.3	36.9	48.3	31.2	27.5
N	180,545	153,027	27,518	112,706	94,939	17,767

Panel B: Cardiologists

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	0.709	1.224	-0.878	1.862	2.796**	-0.979
F	67.6	54.7	9.9	51.6	53.9	10.2
N	28,914	25,363	3,551	24,945	21,680	3,265

estimates are statistically significant (99 and 96 percent level, respectively). It is difficult to argue there is a differential effect between the samples given these estimates.

Estimates in Panel B suggest a differential effect between the not-affiliated and affiliated subsamples. In the full sample I estimates positive but statistically insignificant effects of integration on cardiologist RVUs, measuring a 0.71 RVU increase within 360 days and a 1.86 RVU increase within 720 days. The not-affiliated IV estimates are positive and larger than corresponding full-sample estimates, indicating 1.22 and 2.80 RVU increases within their respective windows. While the 360-day estimate is not statistically significant, the 720-day estimate is statistically significant at the 95 percent level. In contrast, estimates in the affiliated subsample are negative and also larger in magnitude than the corresponding full-sample estimates. However, I must be cautious in interpreting these estimates because my F statistics are marginally problematic, with values of 9.9 and 10.2 in the 360- and 720-day windows, respectively.

Outpatient utilization

I compare IV estimates of both extensive and intensive margin effects in Table 26. Panel A contains estimates for the extensive margin regression. In the full sample I find integration

Table 26: IV comparison, outpatient utilization

Panel A: Probability of any outpatient utilization

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	-0.081***	-0.092***	-0.026	-0.131***	-0.134***	-0.102
F	52.8	33.4	36.7	48.3	31.3	27.3
N	180,714	153,185	27,529	112,803	95,029	17,774

Panel B: Outpatient RVUs, conditional on at least one outpatient visit

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	3.30	4.20*	1.900	3.62	3.588*	4.19
F	28.5	21.3	7.4	27.5	20.9	9.2
N	5,874	3,956	1,918	5,373	3,642	1,731

reduces the probability of any outpatient utilization by 8.1 percentage points in the 360-day window and by 13.1 percentage points in the 720-day window. Both estimates are statistically significant at the 99 percent level. Estimates from the not-affiliated sample are negative like the full-sample estimates and slightly larger in magnitude, indicating decreases in the probability of any outpatient utilization of 9.2 and 13.4 percentage points at 360 and 720 days, respectively. These estimates are also statistically significant at the 99 percent level. Estimates from the affiliated sample are also negative, smaller in magnitude, and statistically insignificant. It is possible lack of significance is due to the comparatively smaller amount of observations in this sample (27,529 and 17,774 patients in the affiliated samples versus 153,185 and 95,029 patients in the non-affiliated samples). In the affiliated sample, integration reduces the probability of any outpatient utilizatinon by 2.6 percentage points in the 360-day window and 10.2 percentage points in the 720-day window.

Panel B contains estimates for the intensive margin regression. In the full sample integration reduces outpatient utilization by 3.30 and 3.62 RVUs in the 360- and 720-day windows, respectively. Neither estimate is statistically significant. In the not-affiliated sample, however, the increases are statistically significant at the 90 percent level. At 360 days the increase is 4.20 RVUs, which is 27.3 percent larger than the corresponding full-sample estimate. At

720 days the increase is 3.59 RVUs, which is 0.8 percent smaller than the corresponding full-sample estimate. Due to low values for the F statistic in the first stage, it is difficult to interpret estimates from the affiliated subsample.

Prescription drug utilization

Table 27: IV comparison, probability of a new cardiovascular drug prescription

Panel A: Generalists

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	0.191*	0.208*	0.189*	0.149	0.140	0.232
F	52.8	33.4	36.7	48.3	31.3	27.3
N	180,714	153,185	27,529	112,803	95,029	17,774

Panel B: Cardiologists

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	0.031	0.060	-0.091	-0.023	-0.022	0.026
F	8.5	5.8	2.1	17.8	15.6	4.5
N	29,654	25,984	3,670	25,569	22,195	3,374

I compare estimates of the effect on prescribing behavior in Table 27. Panel A reports results from the regression for new generalist prescriptions. I do not discern any systematic differences between the estimates from the non-affiliated and affiliated subsamples. In the full sample, integration increases the probability that a generalist issues a new cardiovascular drug prescription by 19.1 percentage points within 360 days and 14.9 percentage points within 720 days. Only the 360-day estimate is statistically significant (90 percent level). In the 360-day window, both the non-affiliated and affiliated estimates are positive and statistically significant at the 90 percent level. Compared to the full-sample estimate, the non-affiliated estimate is larger, indicating an increase in the probability of a generalist prescription of 20.8 percentage points. The affiliated estimate is slightly smaller, indicating an increase of 18.9 percentage points. In the 720-day window, neither the non-affiliated nor the affiliated estimate is statistically significant. And in contrast to the prior results, the affiliated estimate, 0.232, is larger than the non-affiliated estimate, 0.140.

Panel B reports results from the regression for new cardiologist prescriptions. None of the estimates are statistically significant and most of these regressions suffer from weak instrument problems, making it inadvisable to rely on these coefficients. Furthermore, the IV estimates within each sample do not have the same sign.

Inpatient admission

Table 28: IV comparison, inpatient admission

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	0.003	0.005	-0.002	0.017*	0.023*	0.009
F	52.8	33.4	36.7	48.3	31.3	27.3
N	180,714	153,185	27,529	112,803	95,029	17,774

I compare estimates of the effect of integration on the probability of inpatient admission in Table 28 In the full sample I estimate increases in this probability of 0.3 and 1.7 percentage points within 360 and 720 days, respectively. The 360-day estimate is not statistically significant while the 720-day estimate is significant at the 90 percent level. The statistically significant result I observe in the 720-day window is driven by behavior in not-affiliated practices. The estimate for that subsample is positive, statistically significant (90 percent level), and larger than the corresponding full sample estimate, indicating a 2.3 percentage point increase in the probability of an inpatient admission. In contrast, the affiliated estimates are not statistically significant and are far smaller in magnitude than the not-affiliated estimates, indicating a reduction of 0.2 percentage points within 360 days and an increase of 0.9 percentage points within 720 days.

Medical spending

I present estimates of integration on my cost measures in Table 29.

For some medical spending, differential effects of integration are not consistent across windows. The effect of integration on physician, pharmacy, and all non-inpatient spending – see, respectively, Panels A, C, and D – is greater among non-affiliated practices within the

Table 29: IV comparison, medical care spending

Panel A: Physician costs

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	63.87	71.48	52.14	97.0*	91.7	131.6
F	52.8	33.4	36.7	48.3	31.3	27.3
N	180,714	153,185	27,529	112,803	95,029	17,774

Panel B: Outpatient costs

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	-43.46***	-47.74***	-28.65	-64.02***	-73.55***	-18.0
F	52.8	33.4	36.7	48.3	31.3	27.3
N	180,714	153,185	27,529	112,803	95,029	17,774

Panel C: Pharmacy costs

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	224.0**	299.0**	51.7	381.6	408.2	524.2
F	52.8	33.4	36.7	48.3	31.3	27.3
N	180,714	153,185	27,529	112,803	95,029	17,774

Panel D: All non-inpatient costs

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	244.4**	322.7**	75.2	414.6	426.3	637.7
F	52.8	33.4	36.7	48.3	31.3	27.3
N	180,714	153,185	27,529	112,803	95,029	17,774

Panel E: All medical spending

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	538.9*	704.7*	283.4	1060.2*	1193.6	975.1*
F	52.8	33.4	36.7	48.3	31.3	27.3
N	180,714	153,185	27,529	112,803	95,029	17,774

360-day window and but greater among affiliated practices within the 720-day window. For example, I estimate increases in physician spending of 63.9 and 97.0 dollars at 360 and 720 days, respectively, in the full sample. The not-affiliated estimate for physician spending is larger than the affiliated estimate in the shorter window (71.5 vs. 52.1) but the opposite is true in the longer window (91.7 vs. 131.6).

However, for other measures, differential effects are consistent across windows. Panel B shows that reductions in outpatient spending due to integration are concentrated within the not-affiliated subsample. In the full sample I observe statistically significant reductions in outpatient spending of 43.5 and 64.0 dollars within the 360- and 720-day windows, respectively. The estimates from the affiliated subsample are substantially smaller and statistically insignificant. In contrast, the estimates from the not-affiliated subsample are larger and statistically significant at the 99 percent level, indicating reductions of 47.7 and 73.6 dollars within the respective windows.

Panel E shows that reductions in total medical spending – for hypertension this is the sum of physician, outpatient, pharmacy, and inpatient spending – are greater within the not-affiliated subsample. In the full sample I find that integration increases total medical spending for hypertension patients by 538.9 dollars within 360 days and 1,060.2 dollars within 720 days. Both estimates are statistically significant at the 90 percent level. Both estimates from the not-affiliated subsamples are positive and larger, indicating increases of 704.7 and 1,193.6 dollars within the 360- and 720-day windows, respectively. The 360-day estimate is statistically significant at the 90 percent level while the 720-day estimate is statistically insignificant. By comparison, the estimates from the affiliated subsamples are smaller. They indicate increases of 283.4 and 975.1 dollars in the 360- and 720-day windows, respectively. Only the 720-day estimate is statistically significant (90 percent level).

Discussion

It is puzzling that the full-sample effect on the probability of a cardiologist visit is driven by behavior in the affiliated sample. Because the not-affiliated generalists are more isolated, one would expect them to respond even more strongly to the addition of a cardiologist to their practices. It is possible that generalists who have obtained a hospital affiliation are subject to norms and policies that generate even stronger referral incentives once a cardiologist is added to the firm. These incentives could be non-financial, such as peer or administrator expectations that the generalist consult specialists when they are available. Clinicians I have spoken to consider an example like this highly plausible. I cannot determine the reason for this behavior with my data and both theory and external literature do not point to a single, obvious explanation.

That said, this finding is an anomaly among my results. I reported that integration more strongly increases the probability of a cardiologist visit in the previous chapter. In the diabetes analysis I find that integration increases the probability of an endocrinologist visit in the full sample and is driven by behavior in the low-access and not-affiliated subgroups. These other findings should simultaneously increase confidence that integration increases the probability of a specialist visit and illustrate the pitfalls of generalizing from the relatively narrow evidence base on generalist-specialist integration we currently possess.

That said, I am able to replicate the full-sample reduction in the probability of any outpatient service in the not-affiliated sample. This is the most robust result in my analysis and its emergence here should provide some reassurance that the patterns that emerge in this chapter are valid, if perhaps idiosyncratic to hypertension. Among the other behavioral outcomes, integration generates differentially stronger affects in the not-affiliated sample for specialist RVUs and outpatient RVUs, but does not generate systematically different effects for generalist RVUs and the probability of a new generalist prescription in the two subsamples.

As in the full sample and access-based samples, there is no evidence that integration improves outcomes or reduces spending. If my inpatient admission regression primarily captures the effect of integration on outcomes, then they show that integration is worsening outcomes for hypertension patients. If they primarily capture the effect of integration on elective utilization, then I cannot rule out the possibility that integration is cost effective. However, the remaining evidence in this dissertation and the external evidence I am aware of are not encouraging. In my data, integration increases spending for patients in both affiliation subsamples. The external evidence I am aware is dated but suggests that it could take a decade or more for strong clinical benefits to materialize (see Greenfield et al., 1995).

CHAPTER 8 : Diabetes mellitus: Evidence from the full sample

Introduction

Diabetes mellitus (diabetes) is a condition affecting the endocrine system, causing individuals to have too much glucose in their blood.[1] Clinicians may employ a wide range of strategies in treating this condition. The choice set includes insulin therapy, blood sugar monitoring, improved diet and exercise, other oral or injected medications, pancreas transplantation, and bariatric surgery. The latter two options are far less frequently used. Complications of diabetes develop gradually, but are potentially severe. These complications include cardiovascular disease, nerve damage, kidney damage, eye damage, foot damage, skin damage, hearing impairment, and Alzheimer's disease. Individuals who have trouble achieving control over their diabetes or develop severe complications are likely candidates for referral.[2]

Full-sample results

Endocrinologist visit

Table 30 contains estimates from the regression of the endocrinologist visit indicator on my regressors. OLS estimates indicate statistically significant decreases in the probability of an endocrinologist visit of 4.9 and 4.2 percentage points at 360 and 720 days, respectively. The OLS FE estimates are also negative and statistically significant. The 360-day estimate indicates a 1.9 percentage point decrease while the 720-day estimate indicates a 4.1 percentage point decrease. In contrast, the IV estimates indicate large, positive, and statistically significant increases in the probability of an endocrinologist visit. Within the 360-day window, integration increases the probability of an endocrinologist visit by 41.1 percentage points.

[1] Clinically speaking, the term diabetes mellitus refers to several conditions, including Type I diabetes (disorder of the pancreas), Type II diabetes (insulin resistance), gestational diabetes (onset of diabetes during pregnancy). (See `http://www.diabetes.ca/about-diabetes/types-of-diabetes` for descriptions of these types of diabetes.) Individuals with gestational diabetes are likely excluded from my sample, provided physicians are coding their condition with the ICD-9 code for gestational diabetes.

[2] For example, see `http://www.diabetes.org/living-with-diabetes/treatment-and-care/whos-on-your-health-care-team/your-health-care-team.html`.

Table 30: Effect of integration on the probability of an endocrinologist visit (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.049***	-0.019**	0.411***	-0.042**	-0.041***	0.397**
	-2.9671	-2.0813	3.2688	-2.3504	-2.6227	2.2112
Severity	0.036***	0.036***	0.036***	0.046***	0.047***	0.047***
	10.3883	9.3661	10.7455	10.8632	9.3109	11.0365
Severity squared	-0.002***	-0.002***	-0.002***	-0.002***	-0.003***	-0.002***
	-5.4014	-5.7622	-5.6966	-6.2086	-5.7775	-6.4443
Male	-0.004*	-0.003	-0.004	-0.011***	-0.010***	-0.011***
	-1.6551	-0.9126	-1.4025	-3.7544	-2.6727	-3.4587
Age	-0.004***	-0.005***	-0.005***	-0.003***	-0.003**	-0.004***
	-4.4029	-4.1329	-4.7741	-2.7463	-2.1438	-2.9555
Age squared	0.000***	0.000***	0.000***	0.000	0.000	0.000
	2.6978	2.7138	2.8924	1.0715	0.8636	1.1179
Distance to nearest specialist	-0.002***	-0.000	-0.001**	-0.003***	-0.002	-0.002***
	-6.4333	-0.3221	-2.4966	-8.0346	-1.3510	-3.6696
Distance to nearest specialist squared	0.000**	0.000	0.000	0.000***	0.000	0.000**
	2.4738	1.1647	0.9977	4.0374	1.5962	2.5279
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

In the 720-window, the increase is 39.7 percentage points. These estimates are significant at the 99 and 95 percent levels, respectively.

Though the IV estimator is potentially biased upward due to physician selection and the effects I observe among diabetes patients are far larger than those observed in the hypertension sample, it is highly unlikely the true effect is negative and my estimates may indeed be highly accurate. In the 720-day sample, OLS and OLS FE estimates are are almost identical. Furthermore, the effect of severity is positive and invariant to estimator in the 360-day window and hardly changes in the 720-day window.

Physician utilization

In this section I report estimates from regressions of generalist and endocrinologist RVUs on my regressors. The effect of integration on generalist RVUs is indeterminate. I find strongly suggestive evidence that the effect of integration on specialist RVUs is positive and

Table 31: Effect of integration on generalist RVUs (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.540***	-0.065	0.667	-0.710***	0.115	-0.822
	-7.5129	-0.8543	1.0484	-5.4434	0.8646	-0.5163
Severity	0.186***	0.155***	0.186***	0.308***	0.254***	0.310***
	10.0154	10.4053	9.5053	9.4560	10.6724	8.8224
Severity squared	-0.007**	-0.006***	-0.007**	-0.013***	-0.009***	-0.012***
	-2.3476	-2.7798	-2.2313	-2.7834	-2.8393	-2.6220
Male	0.105***	0.122***	0.102***	0.205***	0.219***	0.198***
	3.9329	4.9181	3.7647	5.0411	7.6129	5.0232
Age	0.004	0.009	0.005	0.055***	0.067***	0.057***
	0.3873	0.8209	0.4925	2.6661	2.8591	2.6729
Age squared	-0.000	-0.000	-0.000	-0.000**	-0.001**	-0.001**
	-0.5367	-0.5967	-0.6656	-2.2944	-2.3119	-2.3418
Distance to nearest specialist	-0.012***	-0.008	-0.009***	-0.014***	-0.027*	-0.013*
	-4.6591	-1.3299	-2.9266	-2.6092	-1.8024	-1.6510
Distance to nearest specialist squared	0.000**	0.000	0.000	0.000	0.000	0.000
	2.3306	1.0571	1.6143	0.9309	1.5639	0.5780
F	–	–	23.7	–	–	20.6
N	188,653	188,653	185,669	130,198	130,198	128,002

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

large.

Estimates for the generalist RVU regressions are presented in Table 31. OLS estimates are negative and statistically significant at the 99 percent level, indicating reductions in generalist utilization by 0.54 RVUs at 360 days and 0.71 RVUs at 720 days. OLS FE estimates are small and imprecisely estimated, indicating a decrease of 0.07 RVUs within 360 days and an increase of 0.82 RVUs within 720 days. IV estimates are also statistically insignificant, indicating an increase of 0.67 RVUs within 360 days and a decrease of 0.82 RVUs within 720 days. If there is indeed an effect of integration on generalist RVUs, it may be small.

Estimates for the endocrinologist RVU regressions are presented in Table 32. OLS estimates are negative and statistically significant at the 99 percent level, indicating reductions in generalist utilization by 0.63 RVUs at 360 days and 0.35 RVUs at 720 days. OLS FE estimates are small and imprecisely estimated, indicating a decrease of 0.35 RVUs within

Table 32: Effect of integration on endocrinologist RVUs (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.627***	-0.346	4.418	0.995***	0.173	8.900
	2.7095	-1.0035	1.0492	2.7619	0.4065	1.2265
Severity	0.770***	0.551***	0.770***	1.38***	1.17***	1.39***
	9.3668	7.1149	9.2893	13.489	9.912	13.638
Severity squared	-0.015	0.003	-0.015	-0.051***	-0.033***	-0.052***
	-1.0053	0.1860	-1.0113	-4.6811	-2.7691	-4.7690
Male	-0.050	-0.100	-0.036	-0.100	-0.232	-0.095
	-0.8209	-1.2935	-0.5706	-0.8493	-1.5407	-0.7452
Age	-0.265***	-0.166***	-0.265***	-0.362***	-0.370***	-0.382***
	-9.0487	-3.4553	-9.4635	-8.3788	-4.6334	-8.6392
Age squared	0.002***	0.001***	0.002***	0.003***	0.003***	0.004***
	7.9979	3.0549	8.3660	7.1437	4.0338	7.3637
Distance to nearest specialist	-0.025***	-0.002	-0.020**	-0.038***	-0.049	-0.025
	-4.1299	-0.0746	-2.4712	-4.7324	-0.8737	-1.6230
Distance to nearest specialist squared	0.000***	-0.000	0.000***	0.000***	0.001	0.000*
	3.4199	-0.4175	2.7525	3.3874	0.9638	1.7036
F	–	–	22.6	–	–	21.3
N	49,719	49,719	49,136	45,382	45,382	44,775

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

360 days and an increase of 0.17 RVUs within 720 days.

IV estimates are also statistically insignificant, but positive and large. They indicate increases of 4.42 and 8.90 endocrinologist RVUs within 360 and 720 days, respectively. The lack of significance may be caused by smaller sample sizes, as the sample underlying these results contains only about 25 percent of the full sample. Furthermore, if there is bias induced by physician selection, the IV estimator is biased down, implying I am biased against finding a result.

To help size these estimates, I repeat the calculation performed in Chapter 5. I find that even though diabetes is far less prevalent than hypertension, integration on a large scale would cause an even greater increase in physician work. As before, assume a routine office visit generates 1 RVU on average. Therefore, one might think of integration among generalists and endocrinologists as increasing utilization by five routine office visits within 360 days (4.42 RVUs for endocrinologists and 0.67 RVUs for generalists) and eight routine office visits within 720 days (8.90 RVUs for endocrinologists and -0.82 RVUs for generalists). An estimated 29.1 million Americans had diabetes in 2012.[3] For sake of argument, suppose 28.8 of those Americans are adults.[4] Furthermore, suppose all these adult Americans are in medical care and 17 percent of these adult Americans are treated in integrated practices. If we shifted the remaining adult Americans from non-integrated to integrated practices, it would induce an increase in physician work equivalent to 50.0 million office visits, or 5.4 percent of total office visits in 2012. This effect is actually greater than the one I compute for hypertension, even though far less people are diagnosed with diabetes.

Outpatient utilization

There is strong evidence that integration reduces outpatient utilization on the extensive margin and highly suggestive evidence that integration increases outpatient utilization on the intensive margin.

In Table 33 I present results from the any outpatient claim regression. All estimates are negative and the IV estimates are particularly large. OLS estimates are small and statistically insignificant, indicating a reduction of 0.3 percentage points at 360 days and 0.4 percentage points at 720 days. OLS FE estimates are slightly larger and the 720-day estimate is statistically significant (90 percent level). They indicate reductions of 1.0 and 2.5 percentage points in the 360- and 720-day windows, respectively. IV estimates are larger by an order of magnitude. Both are statistically significant at the 99 percent level. These estimates indicate reductions in the probability of an outpatient claim by 21.5 percentage points at 360 days and 31.3 percentage points at 720 days. These effects are quantitatively large, even larger than the effects I observed in the hypertension sample.

Estimates for the outpatient RVU regressions are presented in Table 34. OLS estimates are

[3] All external data on prevalence and diagnosis rates for the calculation that follows is due to `http://www.diabetes.org/diabetes-basics/statistics/`.

[4] This is a rough estimate. Of the 29.1 million with diabetes, 21.0 million were diagnosed. About 0.21 million Americans under the age of 20 are estimated to have diagnosed diabetes. Assuming rates of diagnosis are equal in the older and younger populations, this implies 0.29 million Americans under the age of 20 have diabetes and 28.8 million Americans over the age of 20 have diabetes.

Table 33: Effect of integration on the probability of outpatient service (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.003	-0.010	-0.215***	-0.004	-0.025*	-0.313***
	-0.3233	-0.8645	-3.3518	-0.3584	-1.8709	-3.7822
Severity	0.009***	0.007***	0.008***	0.017***	0.014***	0.016***
	5.8976	7.1256	5.3855	8.3122	9.8960	7.5407
Severity squared	-0.000**	-0.000***	-0.000*	-0.001***	-0.001***	-0.001***
	-2.0696	-2.8891	-1.8094	-3.8491	-4.8736	-3.2153
Male	-0.003***	-0.000	-0.003***	-0.004***	-0.000	-0.004***
	-3.0076	-0.2471	-3.1041	-2.7083	-0.0320	-2.7198
Age	-0.001**	-0.001	-0.001	-0.002*	-0.002*	-0.001
	-2.0266	-1.2397	-1.5968	-1.8094	-1.8840	-1.5067
Age squared	0.000	0.000	0.000	0.000	0.000	0.000
	1.6242	0.8291	1.3579	1.6203	1.5449	1.4335
Distance to nearest specialist	0.002***	0.001	0.001***	0.002***	0.000	0.001***
	6.2358	1.6167	3.5067	6.4365	0.2725	3.2136
Distance to nearest specialist squared	-0.000***	-0.000	-0.000***	-0.000***	0.000	-0.000**
	-4.2203	-1.4894	-2.6204	-4.4427	0.1172	-2.5338
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

small and negative. They indicate reductions of 0.65 and 0.55 RVUs within the 360- and 720-day windows, respectively. The 360-day estimate is significant at the 90 percent level and the 720-day estimate is significant at the 95 percent level. OLS FE estimates are also negative, indicating a decrease of 2.41 and 0.38 RVUs in the 360- and 720-day windows, respectively. Only the 360-day estimate is statistically significant (90 percent level). In contrast, the IV estimates are large and positive. The 360-day estimate indicates a 10.14 RVU increase in utilization among integrated firms. This estimate is statistically significant at the 95 percent level. The 720-day estimate indicates a smaller 3.19 RVU increase in intensive margin utilization. This estimate is not statistically significant.

As in my hypertension sample, it is possible lack of significance is a result of small sample sizes. Because I estimate the model only for individuals with an outpatient claim and outpatient service rates are small, sample sizes are merely $11,036$ at 360 days and $10,729$ at 720 days. Like the extensive margin estimates, the intensive margin estimates are also

Table 34: Effect of integration on outpatient RVUs (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.65*	-2.41*	10.14**	-0.547**	-0.379	3.189
	-1.951	-1.790	2.030	-2.1792	-0.2842	0.5328
Severity	1.03***	0.39	1.15***	0.972***	0.202	1.021***
	5.428	1.411	6.089	5.7105	0.3752	5.8210
Severity squared	-0.052***	-0.029	-0.059***	-0.024	0.026	-0.026
	-4.0023	-1.2381	-3.9734	-1.3434	0.4600	-1.3595
Male	0.731***	0.712*	0.777***	0.644***	0.826*	0.662***
	2.8910	1.6813	3.0149	2.9353	1.8636	2.9297
Age	-0.029	0.074	-0.024	0.073	0.670**	0.082
	-0.3021	0.3613	-0.2363	0.9104	1.9869	0.9693
Age squared	0.000	-0.000	0.000	-0.001	-0.006*	-0.001
	0.2917	-0.2062	0.2156	-0.9293	-1.9133	-0.9938
Distance to nearest specialist	-0.041***	-0.046	-0.006	-0.035***	0.109	-0.022
	-2.6918	-0.2494	-0.2334	-2.9464	1.3886	-1.0523
Distance to nearest specialist squared	0.000**	0.001	0.000	0.000*	-0.003	0.000
	2.0122	0.1580	0.5369	1.7867	-1.6197	1.1025
F	–	–	20.0	–	–	27.2
N	11,036	11,036	10,768	10,789	10,789	10,504

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 35: Effect of integration on the probability of inpatient admission (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.002	-0.004	-0.042***	-0.002	0.008	0.009
	-1.3618	-1.0219	-2.8192	-0.6743	1.6288	0.2652
Severity	0.011***	0.010***	0.011***	0.020***	0.016***	0.020***
	13.2381	11.8198	12.8710	14.6300	10.9812	15.1202
Severity squared	0.000	0.000	0.000	0.000	0.000	0.000
	1.2221	0.7144	1.1576	0.7549	1.0789	0.5958
Male	0.007***	0.006***	0.007***	0.010***	0.008***	0.010***
	11.0817	9.4988	11.0195	12.5368	6.2211	13.0584
Age	-0.002***	-0.002***	-0.002***	-0.002***	-0.001**	-0.002***
	-6.1414	-6.6997	-6.0587	-3.7763	-2.1813	-3.7988
Age squared	0.000***	0.000***	0.000***	0.000***	0.000**	0.000***
	6.2302	6.8683	6.1376	3.9237	2.4545	3.9180
Distance to nearest specialist	-0.000	-0.000	-0.000**	-0.000*	-0.001	-0.000
	-1.4497	-0.0577	-2.4768	-1.9534	-1.4445	-1.1888
Distance to nearest specialist squared	0.000	-0.000	0.000	0.000	0.000	0.000
	0.3715	-0.6302	0.8842	1.6410	1.4304	1.0373
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

large. Mean outpatient RVUs in the full sample are 2.28 in the 360-day window and 2.43 in the 720-day window. Even if the 360-day estimate is implausibly large, both estimates provide suggestive evidence that integration does increase outpatient utilization along the intensive margin.

Inpatient admission

In Table 35 I present estimates of the effect of integration on the probability of inpatient admission. OLS estimates indicate a decrease of 0.2 percentage points in both windows. Both estimates are statistically insignificant. OLS FE estimates are larger in magnitude, but have opposite signs. They are also statistically insignificant. In the 360-day window, the estimate indicates a reduction in the probability of an inpatient admission of 0.4 percentage points. In the 720-day window, a 0.8 percentage point increase is indicated.

In the 360-day window, the IV estimate is negative, an order of magnitude larger than the corresponding OLS FE estimate, and statistically significant at the 99 percent level. It indicates a 4.2 percentage point decrease in the probability of an inpatient admission. However, the effect fades by 720 days. In the longer window, the IV estimate is positive and statistically insignificant. It is very close to the OLS FE estimate. The 720-day IV estimate indicates a 0.9 percentage point increase in the probability of an inpatient admission.

The effect of severity is right-signed. This effect is also economically snd statistically significant (99 percent level) in both windows. The severity coefficients suggest the clinical improvement due to integration within 360 days is economically significant. In this window, the effect of severity is far smaller than the effect of integration for most values of the severity score. However, the effect of severity dominates the effect of integration in the 720-day window.

Though there are many possible interpretations of my results, I believe the most plausible one is that any clinical gains for diabetes patients due to integration are not long-lived. This is primarily a methodological issue. If one is willing to accept the 360-day result, that integration reduces admissions in the shorter window, then one should be surprised that the effect does not appear in the longer window. Save for the difference in claims window, the 720-day analysis is exactly the same. I also argue that the strength of my severity estimates, measured by the size of my t-statistics, lends great credibility to my results. Intuitively, severity should predict hospitalization rates very strongly. Not only are the estimated effects right-signed, the values of the t-statistics on the linear term – the economically significant term – are extremely large: 12.9 in the 360-day window and 15.1 in the 720-day window.[5]

Table 36: Effect of integration on physician services spending (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	16.2***	-21.8**	246.6**	31.1***	-44.6*	307.3**
	2.99	-2.09	2.54	2.87	-1.68	2.03
Severity	38.4***	35.6***	38.8***	76.6***	72.6***	77.8***
	12.88	10.43	12.85	10.58	8.97	10.58
Severity squared	-1.57***	-1.53***	-1.60***	-3.53***	-3.27***	-3.60***
	-4.850	-5.734	-4.812	-8.599	-7.119	-8.600
Male	5.49**	2.82	5.34**	5.12	-2.89	4.47
	2.319	0.945	2.339	1.372	-0.541	1.189
Age	-8.40***	-7.38***	-8.57***	-12.7***	-12.8***	-13.1***
	-5.709	-3.476	-5.848	-5.66	-4.69	-5.88
Age squared	0.059***	0.051**	0.060***	0.083***	0.087***	0.086***
	4.0993	2.5063	4.2124	3.7247	3.4581	3.8708
Distance to nearest specialist	-0.878***	-0.625	-0.356	-1.78***	-1.77	-1.06**
	-3.9712	-0.9180	-1.3978	-3.584	-1.063	-2.140
Distance to nearest specialist squared	0.008***	0.015*	0.006**	0.015**	0.032*	0.012**
	3.1702	1.9512	2.1368	2.5130	1.6499	2.0408
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Medical spending

Generalist and endocrinologist spending

In Table 36, the explanatory variable is the sum of spending attributable to generalists and endocrinologists. OLS estimates are positive and relatively small, indicating spending increases of 16.2 dollars at 360 days and 31.1 dollars at 720 days. Both estimates are significant at the 99 percent level. OLS FE estimates are larger in magnitude but negative, indicating decreases of 21.8 at 360 days and 44.6 dollars at 720 days. These estimates are also statistically significant.

IV estimates are positive, and larger by about an order of magnitude, which is consistent with higher rates of endocrinologist utilization on both the extensive and intensive margins. The 360-day estimate is statistically significant at the 95 percent level, indicating a 246.6 dollar increase in spending. The 720-day estimate is also statistically significant at the 95 percent level, indicating a 307.3 dollar increase in spending. These IV estimates seem large in absolute terms. The estimated effects are also as important as, if not more important than, the effect of severity. In the 360-day sample, the effect of integration is greater than the effect of severity for patients with severity up to 2.54. In the 720-day sample, the threshold is lower, at 2.01. Both these thresholds are above the 75th percentile for severity in my diabetes samples.

Outpatient spending

Table 36 contains the estimates from the regression in which the outcome is the sum of spending across all outpatient claims. OLS estimates are extremely small and statistically insignificant in both windows. Within the 360-day window, the integration coefficient indicates a 3.1 dollar decrease in outpatient services spending. In the 720-day window, the

[5] As in the hypertension analysis, it is possible that positive effects I observe are an artifact of non-emergency utilization. I again tried to study this issue by using admission type classifiers in the data – e.g. emergency, urgent, or emergent admission – but my estimates lacked precision.

Table 37: Effect of integration on outpatient services spending (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-3.09	-1.20	-47.04	0.75	4.14	-197.05
	-0.440	-0.218	-0.530	0.076	0.469	-1.063
Severity	12.6***	10.3***	12.7***	13.0	7.3	12.6
	5.47	2.59	5.52	1.31	0.58	1.26
Severity squared	-0.055	0.155	-0.047	1.96	2.30	2.05
	-0.1477	0.2732	-0.1241	1.073	1.019	1.109
Male	9.84***	9.97***	10.01***	10.9***	8.1**	11.2***
	3.319	3.486	3.321	2.67	2.06	2.73
Age	0.215	0.391	0.266	-1.67*	-0.34	-1.52
	0.2342	0.2839	0.2869	-1.698	-0.222	-1.569
Age squared	-0.008	-0.006	-0.008	0.011	0.001	0.010
	-0.7277	-0.4544	-0.7646	1.0854	0.0564	0.9799
Distance to nearest specialist	0.406	0.067	0.306	1.029***	-0.444	0.548
	1.3896	0.1196	0.8886	3.1464	-0.3612	1.1908
Distance to nearest specialist squared	-0.003	-0.007	-0.003	-0.010***	-0.001	-0.007
	-1.0825	-0.8783	-0.7795	-2.7419	-0.0345	-1.6142
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

integration coefficient indicates a 0.8 dollar increase in spending. OLS FE estimates are also small and statistically insignificant, indicating a 1.2 dollar decrease and a 4.1 dollar increase in outpatient spending in the 360- and 720-day windows, respectively.

IV estimates are also statistically insignificant, but they are comparatively large in magnitude and have the same sign in both windows. They indicate a 47.0 dollar decrease within 360 days and a 197.1 dollar decrease within 720 days. Similar to the hypertension estimates, these IV estimates show that spending decreases due to reductions along the extensive margin more than offset the spending increases due to greater intensive margin utilization. They also show that the effect of integration is more important than the effect of severity for all but very high scores. In the 360-day window, the effect of integration is greater than the effect of severity when severity is less than or equal to 3.76. In the 720-day window, the threshold value is 7.22.

Total spending

In Table 38 I set my outcome as the sum of physician and outpatient spending. Recall, in the hypertension analyses my total non-inpatient spending measure included physician, outpatient services, and prescription drug spending. Because I do not have a therapeutic class identifier that can be used to flag diabetes medications, I do not analyze any prescription drug measures for diabetes patients. OLS estimates are small and positive, indicating increases of 13.1 and 31.8 dollars within 360 and 720 days, respectively. The former estimate is statistically insignificant while the latter estimate is statistically significant at the 95 percent level. OLS FE estimates are negative and slightly larger in size, indicating decreases of 23.0 and 40.5 dollars within the 360- and 720-day windows, respectively. The 360-day estimate is statistically significant at the 95 percent level while the 720-day estimate is not.

In contrast to the OLS FE estimates, IV estimates are positive and significantly larger, indicating spending increases of 199.5 and 110.3 dollars within 360 and 720 days, respectively. Only the 360-day estimate is statistically significant (99 percent level). The effect

Table 38: Effect of integration on non-inpatient spending (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	13.1	-23.0**	199.5***	31.8**	-40.5	110.3
	1.52	-1.97	2.59	2.35	-1.50	0.85
Severity	51.0***	45.8***	51.4***	89.6***	79.8***	90.3***
	13.05	10.04	12.91	7.86	5.12	7.99
Severity squared	-1.62***	-1.37**	-1.64***	-1.56	-0.97	-1.56
	-2.601	-2.045	-2.605	-0.858	-0.390	-0.856
Male	15.3***	12.8***	15.3***	16.0***	5.2	15.7***
	3.84	3.13	3.83	3.16	0.82	3.18
Age	-8.18***	-6.98***	-8.31***	-14.4***	-13.2***	-14.6***
	-4.179	-2.586	-4.139	-6.29	-4.33	-6.37
Age squared	0.051**	0.045*	0.052**	0.094***	0.087***	0.096***
	2.5144	1.6812	2.5081	3.8333	3.0242	3.9262
Distance to nearest specialist	-0.472	-0.559	-0.050	-0.751	-2.216	-0.517
	-1.3190	-0.7983	-0.1306	-1.4337	-1.0896	-0.7947
Distance to nearest specialist squared	0.005	0.008	0.003	0.005	0.032	0.005
	1.2263	0.6831	0.7718	0.7573	1.1109	0.6665
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 39: Effect of integration on physician, outpatient, and inpatient spending (full sample)

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-69.2	43.1	-686.1	-68.4	54.5	852.4
	-1.62	0.37	-1.52	-0.76	0.39	1.05
Severity	240.2***	190.5***	240.6***	372.1***	299.7***	379.7***
	5.84	4.70	5.68	4.12	3.02	4.17
Severity squared	6.31	7.03	6.34	27.2	26.7	26.3
	0.946	1.134	0.938	1.49	1.37	1.42
Male	181.7***	150.0***	181.2***	318.3***	263.2***	316.2***
	8.22	5.66	8.48	10.24	5.82	10.16
Age	-34.4***	-35.1**	-33.3***	-47.8**	-51.8**	-48.5***
	-2.66	-2.53	-2.59	-2.49	-2.39	-2.60
Age squared	0.327**	0.355**	0.316**	0.423**	0.490**	0.424**
	2.3648	2.4091	2.2976	2.0775	2.2364	2.1540
Distance to nearest specialist	-3.76**	1.99	-5.12***	-7.11***	-14.22	-4.58
	-2.270	0.327	-2.796	-2.775	-1.345	-1.554
Distance to nearest specialist squared	0.026	-0.148*	0.036**	0.077**	0.057	0.060*
	1.5009	-1.8905	2.0591	2.5096	0.4432	1.8125
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

of integration is also large compared to the effect of severity within the 360-day window. The effect of severity is positive, but only greater than or equal to the effect of severity for patients with severity scores of 4.54 or greater. Within the 720-day window, the threshold value is much lower, indicating the severity effect dominates for severity scores greater than or equal to 1.25.

In Table 39 I add inpatient spending to my cost measure. OLS estimates are negative but statistically insignificant in both windows. The estimates indicate 69.2 and 68.4 dollar reductions in spending in the 360- and 720-day windows, respectively. OLS FE estimates are positive, but smaller in magnitude. They indicate a 43.1 dollar increase in spending within 360 days and a 54.5 dollar increase in spending within 720 days. IV estimates are far larger, but statistically insignificant. In the 360-day window, integration reduces spending by 686.1 dollars. However, in the 720-day window, integration increases spending by 852.4 dollars. As in the hypertension analysis, it seems the effect of integration on inpatient outcomes drives the effect of integration on total spending. So again I find that achieving better health outcomes is extremely important for the success of policies increasing generalist and specialist integration. These estimates also suggest that spending reductions for diabetes patients due to integration might be short-lived, a point I further discuss in the following section.

Additional discussion

Estimates from the full sample of diabetes patients show that generalist-endocrinologist integration has economically important effects along multiple margins. I recover strong evidence that integration increases the probability of an endocrinologist visit and that it also reduces the probability of any outpatient utilization. These effects are significantly larger than the effects I observe in the corresponding hypertension analysis. The effect of integration on the probability of an endocrinologist visit is about four times as large as the effect of integration on the probability of a cardiologist visit. The effect on the probability of any outpatient service for diabetes patients is roughly three times as large as the effect for hypertension patients.

The differences are so large as to almost seem implausible, but to the best of my knowledge there is a clinical basis for these differences. Informal conversations with practicing physicians reveals that management of diabetes patients can be far more difficult than management of hypertension patients. There is some evidence consistent with this in my data, as mean severity in the diabetes samples is greater than mean severity in the hypertension samples. Furthermore, I do observe a substantially larger effect of integration on specialist RVUs in the diabetes sample (though all estimates are statistically insignificant). These estimates are consistent with a story in which generalists with formal ties to an endocrinologist are more eager to exploit that connection for clinical reasons. Furthermore, if outpatient utilization on the extensive margin is driven by the generalist's desire to resolve clinical uncertainty, the large negative effect of integration on the probability of any outpatient service would be consistent with greater use of endocrinologists. After referral, the burden would be transferred to the provider with greater clinical expertise in diabetes.

Though my estimates for both inpatient admission and spending are opposite-signed in the 360- and 720-day windows, they cannot encourage those advocating for greater integration among generalists and specialists. Diabetes is another important test case for policies attempting to increase care coordination by increasing rates of integration. The complications of diabetes are greater in number and are in some ways more severe than those for hypertension. Therefore, if there are any clinical benefits to integration, one should have a very strong prior those benefits would emerge from my diabetes data. Unfortunately, my estimates suggest that clinical benefits from generalist-endocrinologist integration are not long-lived. While integration causes a statistically significant reduction in inpatient admissions within 360 days, the reduction dissipates by 720 days. My cost estimates suggest that this failure to achieve longer-lasting clinical improvements is the reason why integration does not reduce total medical spending in the longer window of my study.

CHAPTER 9 : Diabetes mellitus: Evidence from low- and high-access areas

Introduction

The plan of this chapter is identical to the plan of the corresponding chapter in my hypertension analysis, save for analyses of prescription drug utilization and outcomes. Recall, I define a low-access subsample that contains individuals whose generalist is located "far" from a cardiologist and a high-access subsample containing individuals whose generalist is located "close" to a cardiologist. An individual is assigned to one of these groups if the distance between the patient's generalist and the nearest endocrinologist is sufficiently high or low. In the hypertension sample, the number of cardiologists was sufficiently high to make the high-access group about two times as large as the low-access group. In the diabetes sample, the low- and high-access groups are almost equal. This is consistent with the fact that while there were about 23,000 cardiologists in the United States in 2010, there were less than 6,000 endocrinologists at the same time.[1]

As before I search for differential responses to integration between the low- and high-access groups. Because generalists in low-access areas are clinically isolated, relative to their high-access counterparts, they should respond more strongly to the addition of an endocrinologist to their practices. Many of my full-sample results are corroborated in the low-access sample. However for some outcomes the full-sample effect is driven by the high-access sample. I first present my results below. In the final section I further develop the implications of my results and discuss whether the relatively strong high-access responses are easily rationalized.

Table 40: IV comparison, endocrinologist visit

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	0.411***	0.519***	0.125	0.397**	0.719***	-0.011
F	23.7	17.3	30.2	20.6	14.5	31.0
N	185,697	60,727	64,546	128,021	41,829	44,531

Results

Endocrinologist visit

Table 40 contains IV estimates for the effect of integration on the probability of an endocrinologist visit. In the full sample I detect statistically significant increases in both windows. Within 360 days I observe a 41.1 percentage point increase and within 720 days I observe a 39.7 percentage point increase. These estimates are significant at the 99 and 95 percent level, respectively. The estimated effects in low-access areas are also statistically significant and are, in fact, even greater. They indicate 51.9 and 71.9 percentage point increases in the probability of an endocrinologist visit, respectively. Both estimates are significant at the 99 percent level. In contrast, the high-access estimates are far smaller and statistically insignificant.

Physician utilization

In Table 41 I compare IV estimates from the generalist and endocrinologist RVU regressions. Only one of these estimates is significant at any conventional level, so the results in this table are at best suggestive.

Panel A contains estimates for the generalist RVU regressions. In the full sample IV estimates are statistically insignificant, indicating an increase of 0.67 RVUs in the 360-day window and a decrease of 0.82 RVUs in the 720-day window. Integration may differentially affect generalists in low- and high-access areas. In low-access areas, I estimate statisti-

[1] See `http://www.cardiosolution.com/market-situation/` and `http://www.healio.com/endocrinology/diabetes/news/print/endocrine-today/\%7B511d7427-678b-42e0-9b7b-4e374fabc62a\%7D/us-endocrinologist-shortage-affects-access-to-care-physician-satisfaction`.

Table 41: IV comparison, physician utilization

Panel A: Generalists

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	0.667	1.249	-1.132	-0.822	2.344	-3.788*
F	23.7	17.3	30.2	20.6	14.5	30.9
N	185,669	60,721	64,533	128,002	41,827	44,522

Panel B: Endocrinologists

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	4.418	7.672	3.517	8.900	13.153	5.55
F	22.6	13.6	23.6	21.3	16.1	25.0
N	49,136	16,183	17,988	44,775	14,716	16,314

cally significant increases of 1.25 and 2.34 RVUs within 360 and 720 days, respectively. In high-access areas, I observe decreases of 1.13 and 3.79 RVUs within the 360- and 720-day windows, respectively. Only the 720-day high-access estimate is statistically significant (90 percent level). These estimates indicate that lack of significance in the full sample may be due in part to divergent behavior across low- and high-access areas.

Panel B contains estimates for the endocrinologist RVU regressions. None of these estimates are statistically significant. In the full sample IV estimates are large and positive, indicating increases of 4.42 and 8.90 RVUs for endocrinologists within 360 and 720 days, respectively. The low-access estimates are greater than the high-access estimates within both windows (360 days: 7.67 vs. 3.52, 720 days: 13.14 vs. 5.55).

Outpatient utilization

Table 42 contains estimates for the outpatient utilization regressions. Panel A compares estimates of the extensive margin effect for outpatient utilization. In the full sample I detect large, negative, and statistically significant decreases in the probability of any outpatient utilization. At 360 days the reduction is 21.5 percentage points while at 720 days the reduction is 31.3 percentage points. These estimates are statistically significant at the 99

Table 42: IV comparison, outpatient utilization

Panel A: Probability of any outpatient utilization

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	-0.215***	-0.266**	-0.164***	-0.313***	-0.355***	-0.236***
F	23.7	17.3	30.2	20.6	14.5	31.0
N	185,697	60,727	64,546	128,021	41,829	44,531

Panel B: Outpatient RVUs, conditional on at least one outpatient visit

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	10.14**	0.510	15.4***	3.189	7.83	4.91
F	20.0	8.3	20.4	27.2	10.8	25.7
N	10,768	3,478	3,644	10,504	3,384	3,555

percent level. Both sets of subsample estimates are negative and statistically significant at the 99 percent level. However, in both windows only the low-access estimate is larger than the full-sample estimate. At 360 days, the low-access estimate indicates a 26.6 percentage point reduction while the high-access estimate indicates a 16.4 percentage point reduction in the probability of any outpatient utilization. At 720 days, the low-access estimate indicates a 35.5 percentage point reduction while the high-access estimate indicates a 23.6 percentage point reduction in the same probability.

Panel B compares estimates of the intensive margin effect for outpatient utilization. In the full sample I estimate large, positive effects. The 720-day estimate is substantially smaller than the 360-day estimate. At 360 days I detect a statistically significant (95 percent level) increase of 10.14 RVUs while at 720 days I estimate a statistically insignificant increase of 3.19 RVUs. If the intensive margin effect of integration is indeed positive, it is not clear from Panel B whether the effect is concentrated within low- or high-access areas. In the 360-day window the low-access estimate, 0.510, is far smaller than the high-access estimate, 15.40. But in the 720-day window, the low-access estimate, 7.83, is larger than the high-access estimate.

Table 43: IV comparison, inpatient admission

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	-0.042***	0.002	-0.063	0.009	0.084	-0.041
F	23.7	17.3	30.2	20.6	14.5	31.0
N	185,697	60,727	64,546	128,021	41,829	44,531

Inpatient admission

In Table 43 I present estimates for the inpatient admission regressions. In the full sample I detect a negative, large, and statistically significant decrease in the probability of an inpatient admission at 360 days, a reduction of 4.2 percentage points. At 720 days, the estimated effect is positive but statistically insignificant at 0.9 percentage points. In the low-access subsamples I estimate positive but statistically insignificant effects of 0.2 percentage points within the 360-day windows and 8.4 percentage points in the 720-day window. In contrast, the high-access estimates are negative and large, though also statistically insignificant. Within 360 days I detect a 6.3 percentage point decrease and within 720 days I detect a 4.1 percentage point decrease.

Recall, integration in low-access areas increases the probability of an endocrinologist visit and the point estimates suggest endocrinologist RVUs might also increase as well. Yet I fail to detect a decrease in the probability of an inpatient admission in low-access areas. This should be a discouraging result for those advocating for greater generalist-specialist integration.

Medical spending

I present estimates of the effect of integration on my cost measures in Table 44.

Panel A contains estimates for physician costs. The statistically significant spending increases I observe in the full sample are concentrated within low-access areas. In the full sample, integration increases spending by 246.6 dollars within 360 days and 307.3 dollars within 720 days. Both estimates are statistically significant at the 95 percent level. In the

Table 44: IV comparison, medical care spending

Panel A: Physician costs

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	246.6**	271.2**	29.5	307.3**	438.4	85.2
F	23.7	17.3	30.2	20.6	14.5	31.0
N	185,697	60,727	64,546	128,021	41,829	44,531

Panel B: Outpatient costs

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	-47.04	-168.04	214.6***	-197.05	-327.4	186.8**
F	23.7	17.3	30.2	20.6	14.5	31.0
N	185,697	60,727	64,546	128,021	41,829	44,531

Panel C: All non-inpatient costs

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	199.5***	103.2	244.1**	110.3	111.0	272.0*
F	23.7	17.3	30.2	20.6	14.5	31.0
N	185,697	60,727	64,546	128,021	41,829	44,531

Panel D: All medical spending

	360-day window			720-day window		
	Full sample	Low-access	High-access	Full sample	Low-access	High-access
Integrated	-686.1	-200.4	-1467.8	852.4	2029.9	-267.0
F	23.7	17.3	30.2	20.6	14.5	31.0
N	185,697	60,727	64,546	128,021	41,829	44,531

low-access sample, the effect of integration is larger, indicating spending increases of 271.2 and 438.4 dollars within 360 and 720 days, respectively. The 360-day estimate is statistically significant at the 95 percent level while the 720-day estimate is statistically insignificant. In the high-access sample, the effect of integration is positive, small by comparison, and statistically insignificant, indicating increases of 29.5 and 85.2 dollars within the 360- and 720-day windows, respectively.

Panel B contains estimates for outpatient costs. I observe strongly suggestive evidence of a differential effect between low- and high-access areas. In the full sample I observe negative but statistically insignificant decreases in outpatient spending. Integration reduces outpatient services spending by 47.0 dollars within 360 days and 197.1 dollars within 720 days. I observe greater spending reductions in low-access areas. My low-access estimates indicate spending reductions of 168.0 dollars in the 360-day window and 327.4 dollars in the 720-day window. As in the full sample, these estimates are not statistically significant. In contrast, I observe statistically significant spending increases in high-access areas for both windows. In the 360-day window, integration increases spending by 214.6 dollars while in the 720-day window the increase is 186.8 dollars.

Panel C contains estimates for the sum of physician and outpatient costs. In the full sample I detect a statistically significant (99 percent level) spending increase of 199.5 dollars within 360 days and a statistically insignificant increase of 110.3 dollars within 720 days. For both subsamples my estimates indicate integration increases spending in both windows. Only the high-access estimates are statistically significant, indicating spending increases of 244.1 and 272.0 dollars in the 360- and 720-day windows, respectively. The 360-day estimate is significant at the 95 percent level while the 720-day estimate is significant at the 90 percent level.

Panel D contains estimates for all medical spending, exclusive of pharmacy costs. Though none of the estimates are statistically significant, they again demonstrate that the effect of integration on inpatient utilization strongly determines the effect of integration on total medical spending. Recall, in the full sample I find integration reduces the probability of an inpatient admission in the 360-day window but do not replicate the result in the 720-day window. These admission effects are reflected in the (statistically insignificant) full-sample estimates for the effect of integration on the spending measure in Panel D. The 360-day full-sample estimate is negative (-686.1 dollars) while the 720-day estimate is positive (852.4). In the low-access sample the effect of integration on the probability of an inpatient admission increases from 0.2 percentage points within 360 days to 8.4 percentage points within 720 days. The effect of integration on low-access medical spending increases from -200.4 dollars within 360 days to 2,029.9 dollars within 720 days. Similarly, in the high-access sample the effect of integration on the probability of an inpatient admission increases from -6.3 percentage points to -4.1 percentage points within the 360- and 720-day windows, respectively. The effect of integration on high-access medical spending increases from -1,467.8 dollars within 360 days to -267.0 dollars within 720 days.

Additional discussion

The strongest evidence I have indicates that the effects of integration on the probability of an endocrinologist visit as well as the probability of any outpatient utilization are differentially stronger in the low-access sample. There is at best suggestive evidence that endocrinologist RVUs increase more strongly in low-access areas. These results are consistent with my findings in the full sample.

Some of my results are more difficult to rationalize. In low-access areas, integration strongly increases the probability of an endocrinologist visit and seems to increase both generalist and endocrinologist RVUs. In high-access areas, integration has much weaker effects on the probability of an endocrinologist visit, generalists reduce RVUs, and endocrinologists increase RVUs. The travel time argument I make in Chapter 6, to rationalize the differential patterns I observe between low- and high-access hypertension patients, cannot rationalize the results in this chapter. This explanation fails specifically due to the higher levels of endocrinologist intensive margin usage in low-access areas. If patients are willing to see a relatively distant endocrinologist, but are unwilling to make very many trips to see him or her, endocrinologists RVUs should be lower in low-access areas. To the best of my knowledge, there are no obvious, clinically plausible assumptions that can explain the differential utilization patterns I observe in these low- and high-access samples but also generate the behavior I have already rationalized in the hypertension sample. It seems reasonable to defer this question to future research.

The inpatient admission and medical spending estimates provide further evidence that integration will not achieve the goals of its most strident advocates. First, in both the low- and high-access samples the (statistically insignificant) effect of integration on the probability of inpatient admission is worse in the 720-day window compared to the 360-day window. Second, though referral and utilization strongly increase in low-access areas, I do not observe decreases in the probability of an inpatient admission. Because spending effects in the aggregate are highly responsive to inpatient expenditures, integration's failure to reduce inpatient admissions in low-access areas again leads to higher total medical spending. In contrast, the effect of integration on specialist utilization is much less strong in high-access areas, yet I observe (statistically significant) decreases in the probability of an inpatient admission in both windows. I argue these results provide further evidence that policies encouraging formal integration are too blunt and too potent for the generalist-specialist coordination problems that actually confront us.

CHAPTER 10 : Diabetes mellitus: Evidence from non-affiliated and affiliated practices

Introduction

With the exception of prescription drug analyses, the plan of this chapter follows the plan of Chapter 7. I have again divided patients into two groups. The not-affiliated group contains patients whose primary care physician does not report a hospital affiliation. The affiliated group contains the remaining patients. Those in the not-affiliated group should experience a stronger response to integration than those outside it due to the greater clinical isolation of not-affiliated generalists. In contrast to the corresponding hypertension analysis, I do indeed estimate stronger effects in the more clinically isolated subgroup. I present my results in the next section and provide additional discussion in the final section.

Results

Endocrinologist visit

Table 45: IV comparison, primary specialist visit

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	0.411***	0.642***	0.092	0.397**	0.604**	0.113
F	23.7	13.9	56.3	20.6	13.2	49.8
N	185,697	160,974	24,723	128,021	110,355	17,666

I compare estimates for the effect of integration on the probability of an endocrinologist visit for the full, not-affiliated, and affiliated samples in Table 45 In the full sample I detect large, positive, and statistically significant effects in both windows. The 360-day estimate indictates a 41.1 percentage point increase and the 720-day estimate indicates a 39.7 percentage point increase. Behavior in the not-affiliated subsample drives these estimates. Among patients of not-affiliated practices, integration increases the probability of an endocrinologist visit by 64.2 percentage points within 360 days and 60.4 percentage points within 720 days. The full-sample estimates are large, but these not-affiliated estimates are even larger. In contrast, integration causes smaller, statistically significant increases in the probability of an endocrinologist visit of 9.2 percentage points within 360 days and 11.3 percentage points within 720 days.

Physician utilization

Table 46: IV comparison, physician utilization

Panel A: Generalists

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	0.667	1.200	0.442	-0.822	-1.135	0.729
F	23.7	13.9	56.3	20.6	13.2	49.7
N	185,669	160,947	24,722	128,002	110,339	17,663

Panel B: Endocrinologists

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	4.418	5.923	1.334	8.900	10.908	5.066
F	22.6	9.8	37.1	21.3	13.3	30.4
N	49,136	43,635	5,501	44,775	39,274	5,501

I compare estimates for the effect of integration on generalist and endocrinologist RVUs in Table 46. Panel A contains estimates for the generalist RVU regressions. In the full sample I estimate a 0.67 RVU increase in the 360-day window and a 0.82 RVU decrease in the the 720-day window. The change in sign appears to be driven by the response in the not-affiliated sample. Similar to the full sample, I observe an increase in RVUs in the 360-day window (1.20 RVUs) and a decrease in RVUs in the 720-day window (1.14 RVUs). In contrast, the estimates for the affiliated subsamples are positive, though smaller in magnitude, in both windows. In the 360-day window integration causes increases of 0.44 and 0.73 RVUs in the 360- and 720-day windows, respectively. Because none of these estimates are statistically significant, it is difficult to make any strong inferences from these results.

Panel B contains estimates for the specialist RVU regressions. In the full sample I observe increases of 4.42 and 8.90 RVUs within 360 days and 720 days, respectively. The not-affiliated estimates are substantially larger than the non-affiliated estimates in both windows. In the not-affiliated subsample, integration causes an increase of 5.92 RVUs within 360 days and 8.90 RVUs within 720 days. The 360-day estimate may be suspect due to weak instrument problems (F statistic of 9.8). These estimates are larger than the corresponding full-sample estimates and significantly larger than the affiliated subsample estimates. In the affiliated subsample, integration causes an increase of 1.33 RVUs in the 360-day window and an increase of 5.07 RVUs in the 720-day window. As in Panel A, the estimates in Panel B are not statistically significant, so it is inadvisable to rely heavily on these results.

Outpatient utilization

Table 47: IV comparison, outpatient utilization

Panel A: Probability of any outpatient utilization

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	-0.215***	-0.342***	-0.077	-0.313***	-0.513***	-0.084
F	23.7	13.9	56.3	20.6	13.2	49.8
N	185,697	160,974	24,723	128,021	110,355	17,666

Panel B: Outpatient RVUs, conditional on at least one outpatient visit

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	10.14**	19.47**	5.44	3.189	5.94	1.090
F	20.0	9.9	11.2	27.2	15.4	16.0
N	10,768	7,680	3,088	10,504	7,562	2,942

I compare estimates of the effect of integration on extensive margin utilization in Panel A and intensive margin utilization in Panel B. There is strong evidence that full-sample extensive margin effects are driven by behavior in the not-affiliated subsample, while there is at best suggestive evidence that intensive margin effects are also driven by the not-affiliated subsample.

With respect to extensive margin estimates in the full sample, integration reduces the probability of any outpatient utilization by 21.5 and 31.3 percentage points within 360 and 720 days, respectively. Both estimates are statistically significant at the 99 percent level. The estimates for the not-affiliated subsample are negative, larger in magnitude, and also statistically significant at the 99 percent level, indicating decreases of 34.2 and 51.3 percentage points within 360 and 720 days. In the affiliated subsample, estimates are also negative, but much smaller and statistically insignificant, indicating decreases of 7.7 and 8.4 percentage points within the 360- and 720-day windows.

In Panel B, there is evidence suggesting that intensive margin effects are stronger in the not-affiliated subsample, but it is weak. In the full sample I observe increases of 10.14 and 3.19 RVUs in response to integration within the 360- and 720-day windows, respectively. The 360-day estimate is statistically significant at the 95 percent level while the 720-day estimate is statistically insignificant. The not-affiliated estimates are significantly larger than the affiliated estimates. In the not-affiliated subsample, I observe increases of 19.47 and 5.94 RVUs within the 360- and 720-day windows. The 360-day estimate is statistically significant at the 95 percent level, but the instrument may be weak, as the F statistic for the first stage is 9.9. The 720-day estimate is not statistically significant. In the affiliated subsample, estimates are also positive, but they are far smaller as well as statistically insignificant. The estimates indicate increases of 5.44 RVUs within 360 days and 1.09 RVUs within 720 days.

Inpatient admission

Table 48: IV comparison, inpatient admission

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	-0.042***	-0.068***	-0.000	0.009	-0.002	0.036
F	23.7	13.9	56.3	20.6	13.2	49.8
N	185,697	160,974	24,723	128,021	110,355	17,666

Table 48 compares estimates for the effect of integration on the probability of an inpatient admission. In the 360-day window for the full sample, I estimate a strong negative effect that indicates a 4.2 percentage point decrease in the probability of an inpatient admission. The estimate is significant at the 99 percent level. This effect dissipates by the end of the 720-day window, as I observe a small, statistically insignificant increase of only 0.9

percentage points.

The improvement in outcomes I observe within 360 days is driven by gains in the not-affiliated subsample. In this window, integration for not-affiliated practices reduces the probability of an inpatient admission by 6.8 percentage points. The estimate is significant at the 99 percent level. However, mirroring the full-sample results, the 360-day gains for the non-affiliated subsample do not persist int othe 720-day window, the estimate indicating only a 0.2 percentage point decrease in the probability of an inpatient admission. Though estimates from the affiliated subsample are not statistically significant, they do suggest that improvement in outcomes due to integration are indeed restricted to the 360-day window. At 360 days I do not detect any effect, with a point estimate of 0.000. At 720 days I detect an increase in the probability of inpatient admission of 3.6 percentage points.

Medical spending

I compare spending estimates for physician, outpatient, and my two cost aggregates in Table 49. Panel A provides strong evidence that integration increases physician spending more strongly in the not-affiliated subsample. In the full sample I observe increases of 246.6 dollars in the 360-day window and 307.3 dollars in the 720-day window. Both estimates are statistically significant at the 95 percent level. In the not-affiliated subsample, both estimates are greater than the corresponding full sample estimates, indicating spending increases of 355.1 and 378.3 dollars at 360 and 720 days, respectively. The 360-day estimate is statistically significant at the 95 percent level while the 720-day estimate is statistically insignificant. In the affiliated subsample, both estimates are positive, smaller than the corresponding full sample estimate, and statistically significant. The 360-day estimate indicates a 96.8 dollar increase in physician services spending and is significant at the 90 percent level. The 720-day estimate indicates a 238.1 dollar increase in physician spending and is significant at the 95 percent level.

Panel B does not provide evidence that reductions in outpatient spending are concentrated

Table 49: IV comparison, medical care spending

Panel A: Physician costs

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	246.6**	355.1**	96.8*	307.3**	378.3	238.1**
F	23.7	13.9	56.3	20.6	13.2	49.8
N	185,697	160,974	24,723	128,021	110,355	17,666

Panel B: Outpatient costs

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	-47.04	-52.21	-63.3	-197.05	-279.73	-129.7
F	23.7	13.9	56.3	20.6	13.2	49.8
N	185,697	160,974	24,723	128,021	110,355	17,666

Panel C: All non-inpatient costs

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	199.5***	302.93**	33.5	110.3	98.6	108.4
F	23.7	13.9	56.3	20.6	13.2	49.8
N	185,697	160,974	24,723	128,021	110,355	17,666

Panel D: All medical spending

	360-day window			720-day window		
	Full sample	Not affiliated	Affiliated	Full sample	Not affiliated	Affiliated
Integrated	-686.1	-962.9	-135.0	852.4	743.7	1248.1
F	23.7	13.9	56.3	20.6	13.2	49.8
N	185,697	160,974	24,723	128,021	110,355	17,666

in either subsample. None of the estimates are statistically significant. In the 360-day window, the not-affiliated estimate (-52.2 dollars) is smaller in magnitude than the affiliated estimate (-63.3 dollars), but in the 720-day window, the not-affiliated estimate (-279.7 dollars) is instead larger in magnitude than the affiliated estimate (-129.7 dollars). Similarly, there is not enough evidence in Panel C for me to conclude that the not-affiliated and affiliated subsamples respond differentially to integration.

Panel D estimates the effect of integration on the sum of physician, outpatient, and inpatient spending. None of the estimates are statistically significant. But similar to the estimates for the low- and high-access subsamples, the spending decreases I observe in the 360-day window dissipate by 720 days. In the full sample, integration reduces spending by 686.1 dollars within 360 days and increases spending within 852.4 dollars within 720 days. The not-affiliated estimates follow this pattern, indicating a spending decrease of 962.9 dollars within 360 days and a spending increase of 743.7 dollars within 720 days. The same holds for the affiliated estimates, with a 135.0 dollar decrease within the 360-day window and a 1,248.1 dollar increase within the 720-day window.

Additional discussion

As in the full sample and access-based samples, I confirm that integration more strongly increases the probability of an endocrinologist visit and more strongly decreases the probability of an outpatient claim among not-affiliated generalists. Though I do not detect differential effects for generalist RVUs, there is suggestive evidence that integration increases endocrinologist RVUs more strongly among not-affiliated practices.

The inpatient admission and medical spending results still argue against integration-centric policies. On one hand, integration reduces the probability of inpatient admission for not-affiliated practices in both windows. This contrasts with the low-access results. In that sample, utilization and admission move in the same direction. In the not-affiliated sample, greater utilization among not-affiliated practices results in lower rates of inpatient admission. On the other hand, the 720-day estimate for the not-affiliated sample (-0.002) is more than an order of magnitude smaller than the 360-day estimate (-0.068). Outcomes also worsen in the affiliated sample, with no effect at 360 days and an increase of 3.6 percentage points at 720 days. Because spending effects follow inpatient admission effects, savings from integration in the 360-day regressions fail to appear in the 720-day regressions.

CHAPTER 11 : Conclusion

This dissertation sheds new light on a historically neglected facet of health care organizations: integration among generalist and specialist physician. I estimate the effect of this integration on the provision of health care, health care outcomes, and health care spending.

Implications for scholarship

Understanding these effects is an important scholarly matter because physicians are central actors in the US health care system. In this section I evaluate existing theories for generalist-specialist integration with the evidence I have generated.

The popular view in health services research that integration is unambiguously good, particularly for chronic conditions, is strongly rejected by my results. First, integration fails to generate long-lasting reductions in inpatient admissions. Though the probability of an inpatient admission for diabetes patients decreases with integration in the 360-day window, the decrease does not persist into the 720-day window. I do not detect an improvement for hypertension patients in the 360-day window and outcomes appear to worsen in the 720-day window. Second, because spending is so highly driven by inpatient admissions, the failure to improve outcomes results in a failure to reduce spending. The failure to detect improvements should be especially striking for concerned observers. Diabetes and hypertension are two conditions most likely to benefit from integration if coordination is only a clinical issue. If there are improvements to be found, I would find them in my data given my sample sizes.

Rosen (1983) shows how economies of scope arising from skill complementarities can affect generalist-specialist integration patterns. I argue this theory is also an awkward fit for the data. Suppose outpatient care for diabetes and hypertension consists of a low-intensity task and a high-intensity task, of which generalists mainly do the former, specialists the latter. If the cost of acquiring skill in the high-intensity task decreases in skill in the low-intensity task, then generalists are less likely to integrate with specialists, as the net benefit to integration decreases with the skill complementarity. If Rosen's theory explained generalist-specialist integration, generalist output would be greater in non-integrated practices than integrated practices. Generalists who opted for smaller firm boundaries would have higher levels of skill complementarity and, therefore, higher rates of completing the high-intensity task. In my setting this means generalists would generate higher RVUs in non-integrated practices. However, integration generates a statistically significant increase in generalist RVUs in the hypertension sample. (The effect of integration on generalist RVUs in the diabetes sample are statistically insignificant.)

Stigler (1951) theorizes that specialization among firms is a function of market size. This theory is likely ruled out by integration patterns both before and during my observation period. Specialization in my setting entails firms that contain only generalists, only endocrinologists, or only cardiologists. This theory is inconsistent with my descriptive analysis. While incidence and prevalence of diabetes and hypertension increased over my observation period, integration rates for hypertension increased slightly in the more dependable years of my SK&A data, 2010 through 2012, and remained effectively unchanged for diabetes in that same period. Other surveys suggest that even if this theory had some explanatory power, its impact would be much less important in comparison to integrating for negotiating power, as observed during the rapid growth of HMOs in the 1990s.

Becker and Murphy (1992) posit a link between coordination costs and integration. The data suggest that a Becker and Murphy (1992) type of rationalization for integration patterns does not fit my setting well. The greater are coordination costs, the smaller are firms. If generalist-specialist coordination is at least as costly as coordination among generalists, then higher coordination costs lead to less generalist-specialist integration. I observe slight increases in the rate of both generalist-endocrinologist and generalist-cardiologist integration over my observation period. Thus according to the theory, coordination costs must have declined over my observation period, the gains to generalist-specialist integration have increased over my observation period, or both. I would not argue that coordination costs have declined. At the time of this writing, several years after the end of my observation period, health IT competition is still unsettled and many transactions (e.g. medical record transfer and health insurer billing) are handled via facsimile. I am also unaware of literature suggesting that the gains to generalist-specialization have increased in recent years. The most important study of provider effectiveness in treating hypertension finds no mortality effect of being treated by a generalist or specialist, which suggests that gains from coordination between physicians might indeed be small.[1] If one takes my results at face value, one would conclude that integration worsens outcomes over time. At the very least there is no evidence in my results that suggests Becker and Murphy (1992) is the most important model for understanding generalist-specialist integration.

What remains is to choose between the two referral-based theories of integration. In Garicano and Santos (2004), integration affects patterns of care because it allows side payments that reduce the upstream physician's incentive to exploit the hidden information he gains from his initial encounters with the patient. Save for those who should receive palliative care, higher severity levels justify more treatment and more spending. In this model, we expect severity-based specialization among generalist and specialists, with the former treating the lower-severity individuals and the latter treating those with greater severity. In David and Helmchen (2011), integration affects patterns of care because it increases the upstream physician's incentive to release tasks – steps in what are commonly agreed-upon treatment algorithms – earlier in patient treatment. In this model, a practice that integrates might implement a relatively indiscriminate referral policy in which most hypertension patients are referred to a cardiologist. To my knowledge, it is not difficult to defend this behavior on the basis of quality of care.[2] As noted earlier, one of the most striking patterns in the data is that organizational form has a much stronger impact than severity in nearly all my regressions. I interpret this finding as strong evidence in favor of David and Helmchen

[1] In other words, physician integration might not be the relevant margin for hypertension patients. Coordination between generalists and lower-credentialed providers such as nurses or case managers, however, is an alternate source of gains

[2] For example, many preventive medicine interventions are cost effective, but not cost saving. Relatively high utilization in chronic disease management might fall into this classification.

(2011) as opposed to Garicano and Santos (2004), though elements of the latter might still apply as a secondary concern.

Implications for policy

This research is also relevant for policy due to the long-standing interest in reducing provider fragmentation in the health services research and policy communities.

I provide extremely strong evidence that physician behavior responds to organizational form. When generalist and specialists integrate, they are more likely to make referrals and they are less likely to utilize outpatient services. Work intensity also seems to respond to integration. There is suggestive evidence that patients diagnosed by integrated generalists receive greater amounts of specialist care. There is also suggestive evidence that integration induces greater generalist effort for hypertension patients, but I do not find this is the case for diabetes patients. I also find strong evidence that integration induces greater use of prescription drugs in the treatment of hypertension.

However, clinical and financial gains from these behavioral changes fail to emerge. On one hand, this is surprising. Hypertension and diabetes are two of the most important test cases for policies increasing integration. Both conditions are potentially long-lasting and can lead to serious complications. Treatment for these conditions often involves more than one physician, particularly a generalist and the most relevant specialist. Therefore if there are clinical benefits to integration, one would expect this study to document them. On the other hand, most physicians acquire scant on-the-job training in non-medical areas such as management, accounting, and operations. Groups of very highly skilled and independent individuals like physicians may still lack the complementary abilities needed to achieve their full potential. In my literature review I reported findings from the CTS that physicians can be extremely difficult to manage and the supply of necessary leadership may be low. Though the underlying data is now old, dating to the late 1990s, it is not clear that managing physician organizations now is appreciably less complicated than it was then. This is a potentially important problem given the impact better management can have on hard clinical outcomes like mortality (see Bloom et al., 2012, and the cites therein).

Because clinical and financial gains do not emerge from my data, my research strongly suggests policymakers should decelerate initiatives that integrate generalist and specialist physicians. In fact, I further recommend that health care policymakers reimplement the analysis in this dissertation for a much wider range conditions to avoid extrapolating across specialties. Though cardiology and endocrinology are important specialties, they differ from other specialties such as gastroenterology and psychiatry along several important dimensions, including reliance on procedures and comparative advantage of the generalist and specialist. The analysis in my dissertation is proof of concept and highly scalable at minimal cost. The data sources I relied on for this project are easily accessible. The missing data problems I needed to solve before computing any estimates seem much less severe in newer releases of the data. Based on first-stage estimates, the differential distances instrument I employed for my analysis appears to work very well, and the weak instruments problem caused by the inclusion of fixed effects may well disappear if physician demographic data are utilized instead. Given the high costs of ineffectual policy and the relatively low cost of scaling up this analysis, it seems impossible that generating a comprehensive set of estimates would fail any reasonable cost-benefit calculation.

Though most of the implications of this research are dour, I do assume physicians will be paid on a fee-for-service basis in my back-of-the-envelope calculations. This is a reasonable assumption for the near future. Yet this also leaves room for the possibility that integration induces beneficial outcomes under different contractual arrangements. My estimates are not informative on this point. However, increasing cooperation between market participants and scholars in health economics and allied fields provides some hope that the features of an effective contract might be identified.[3]

[3] See `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3526936/` for an example of this research.

APPENDIX A : Full-sample estimates

Table 50: Primary specialist visit, full sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.013	0.000	0.115*	-0.012	0.017	0.099
	-1.5578	0.0166	1.8587	-1.0207	1.0209	1.1674
Severity	0.021***	0.021***	0.020***	0.032***	0.034***	0.032***
	7.4166	6.4125	7.4929	9.2295	9.0450	9.3042
Severity squared	-0.001*	-0.001**	-0.001*	-0.001***	-0.002***	-0.001***
	-1.8903	-2.4353	-1.6450	-3.0610	-4.5398	-2.8362
Male	-0.001	0.003	-0.001	0.001	0.006*	0.001
	-0.5669	1.0745	-0.4055	0.4932	1.6732	0.5823
Age	-0.001	-0.001	-0.001	0.001	0.001	0.001
	-1.1130	-0.9580	-1.0891	0.8630	0.7970	0.9325
Age squared	0.000	0.000	0.000	-0.000	-0.000	-0.000
	1.5159	1.5007	1.4427	-0.0862	-0.0457	-0.1857
Distance to nearest specialist	-0.000	0.000	0.000	-0.001	0.001	-0.000
	-1.2506	0.1392	0.2255	-1.4083	0.7829	-0.2306
Distance to nearest specialist squared	-0.000	0.000	-0.000	0.000	-0.000*	-0.000
	-0.3386	0.6409	-1.2106	0.0109	-1.9288	-0.5674
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.049***	-0.019**	0.411***	-0.042**	-0.041***	0.397**
	-2.9671	-2.0813	3.2688	-2.3504	-2.6227	2.2112
Severity	0.036***	0.036***	0.036***	0.046***	0.047***	0.047***
	10.3883	9.3661	10.7455	10.8632	9.3109	11.0365
Severity squared	-0.002***	-0.002***	-0.002***	-0.002***	-0.003***	-0.002***
	-5.4014	-5.7622	-5.6966	-6.2086	-5.7775	-6.4443
Male	-0.004*	-0.003	-0.004	-0.011***	-0.010***	-0.011***
	-1.6551	-0.9126	-1.4025	-3.7544	-2.6727	-3.4587
Age	-0.004***	-0.005***	-0.005***	-0.003***	-0.003**	-0.004***
	-4.4029	-4.1329	-4.7741	-2.7463	-2.1438	-2.9555
Age squared	0.000***	0.000***	0.000***	0.000	0.000	0.000
	2.6978	2.7138	2.8924	1.0715	0.8636	1.1179
Distance to nearest specialist	-0.002***	-0.000	-0.001**	-0.003***	-0.002	-0.002***
	-6.4333	-0.3221	-2.4966	-8.0346	-1.3510	-3.6696
Distance to nearest specialist squared	0.000**	0.000	0.000	0.000***	0.000	0.000**
	2.4738	1.1647	0.9977	4.0374	1.5962	2.5279
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 51: Primary specialist RVUs, full sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.115	-0.768**	0.709	0.181	-0.715	1.862
	0.8114	-2.1075	0.5028	1.4184	-1.3990	1.3508
Severity	0.446***	0.461***	0.430***	0.521***	0.435***	0.495***
	6.3418	4.4826	6.0344	6.2632	2.8906	5.8232
Severity squared	-0.027***	-0.035***	-0.026***	-0.026***	-0.018	-0.023***
	-4.4724	-3.4253	-3.9487	-3.4478	-0.9929	-3.0215
Male	0.217***	0.264***	0.221***	0.304***	0.377***	0.316***
	3.6843	2.6841	3.7571	5.7829	4.8433	5.8003
Age	-0.092***	-0.031	-0.088***	-0.079***	-0.059	-0.076***
	-4.1687	-0.8577	-4.2878	-2.8973	-1.4092	-2.8630
Age squared	0.001***	0.000	0.001***	0.001***	0.001*	0.001***
	4.4263	0.9720	4.6497	3.2204	1.6830	3.1856
Distance to nearest specialist	-0.015	-0.058	-0.013	-0.020**	-0.006	-0.014*
	-1.5499	-1.4009	-1.2538	-2.1081	-0.1676	-1.9000
Distance to nearest specialist squared	0.000**	0.002	0.000**	0.000**	0.000	0.000*
	2.5302	1.2188	2.2715	2.1914	0.3147	1.7199
F	–	–	67.6	–	–	51.6
N	29,297	29,297	28,914	25,291	25,291	24,945

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.627***	-0.346	4.418	0.995***	0.173	8.900
	2.7095	-1.0035	1.0492	2.7619	0.4065	1.2265
Severity	0.770***	0.551***	0.770***	1.38***	1.17***	1.39***
	9.3668	7.1149	9.2893	13.489	9.912	13.638
Severity squared	-0.015	0.003	-0.015	-0.051***	-0.033***	-0.052***
	-1.0053	0.1860	-1.0113	-4.6811	-2.7691	-4.7690
Male	-0.050	-0.100	-0.036	-0.100	-0.232	-0.095
	-0.8209	-1.2935	-0.5706	-0.8493	-1.5407	-0.7452
Age	-0.265***	-0.166***	-0.265***	-0.362***	-0.370***	-0.382***
	-9.0487	-3.4553	-9.4635	-8.3788	-4.6334	-8.6392
Age squared	0.002***	0.001***	0.002***	0.003***	0.003***	0.004***
	7.9979	3.0549	8.3660	7.1437	4.0338	7.3637
Distance to nearest specialist	-0.025***	-0.002	-0.020**	-0.038***	-0.049	-0.025
	-4.1299	-0.0746	-2.4712	-4.7324	-0.8737	-1.6230
Distance to nearest specialist squared	0.000***	-0.000	0.000***	0.000***	0.001	0.000*
	3.4199	-0.4175	2.7525	3.3874	0.9638	1.7036
F	–	–	22.6	–	–	21.3
N	49,719	49,719	49,136	45,382	45,382	44,775

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 52: Generalist RVUs, full sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.516***	0.044	1.314***	-0.620***	0.225*	1.784***
	-7.8311	0.6890	3.3685	-6.4062	1.7899	3.2531
Severity	0.059***	0.015	0.054***	0.083***	0.087***	0.080***
	3.8077	0.9491	2.9755	3.6206	2.6296	3.0253
Severity squared	-0.002	0.001	-0.001	-0.005	-0.007	-0.004
	-0.8323	0.3507	-0.2939	-1.6084	-1.2850	-1.2847
Male	-0.015	-0.001	-0.015	0.037	0.074**	0.036
	-0.9242	-0.0595	-0.9312	1.1524	2.2553	1.0701
Age	-0.010*	-0.003	-0.010*	0.018	0.023*	0.016
	-1.7958	-0.3781	-1.8547	1.6083	1.9037	1.5259
Age squared	0.000	0.000	0.000	-0.000	-0.000	-0.000
	1.3709	0.6776	1.3534	-0.8954	-0.8623	-0.8046
Distance to nearest specialist	-0.026***	-0.006	-0.019***	-0.030***	-0.012	-0.020***
	-9.9631	-0.8612	-5.2159	-7.4358	-0.7582	-3.9925
Distance to nearest specialist squared	0.000***	0.000	0.000***	0.000***	-0.000	0.000
	5.1907	0.8308	2.7761	3.2357	-0.5231	1.4219
F	–	–	52.6	–	–	48.3
N	183,467	183,467	180,545	114,544	114,544	112,706

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.540***	-0.065	0.667	-0.710***	0.115	-0.822
	-7.5129	-0.8543	1.0484	-5.4434	0.8646	-0.5163
Severity	0.186***	0.155***	0.186***	0.308***	0.254***	0.310***
	10.0154	10.4053	9.5053	9.4560	10.6724	8.8224
Severity squared	-0.007**	-0.006***	-0.007**	-0.013***	-0.009***	-0.012***
	-2.3476	-2.7798	-2.2313	-2.7834	-2.8393	-2.6220
Male	0.105***	0.122***	0.102***	0.205***	0.219***	0.198***
	3.9329	4.9181	3.7647	5.0411	7.6129	5.0232
Age	0.004	0.009	0.005	0.055***	0.067***	0.057***
	0.3873	0.8209	0.4925	2.6661	2.8591	2.6729
Age squared	-0.000	-0.000	-0.000	-0.000**	-0.001**	-0.001**
	-0.5367	-0.5967	-0.6656	-2.2944	-2.3119	-2.3418
Distance to nearest specialist	-0.012***	-0.008	-0.009***	-0.014***	-0.027*	-0.013*
	-4.6591	-1.3299	-2.9266	-2.6092	-1.8024	-1.6510
Distance to nearest specialist squared	0.000**	0.000	0.000	0.000	0.000	0.000
	2.3306	1.0571	1.6143	0.9309	1.5639	0.5780
F	–	–	23.7	–	–	20.6
N	188,653	188,653	185,669	130,198	130,198	128,002

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 53: Any outpatient visit, full sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.003	-0.007	-0.081***	-0.002	-0.008	-0.131***
	-0.6830	-0.8194	-3.8415	-0.3023	-0.9343	-4.1249
Severity	0.002*	0.002*	0.002*	0.005***	0.006***	0.005***
	1.8294	1.9593	1.8970	4.4204	4.8593	4.6121
Severity squared	0.000	0.000	0.000	-0.000	-0.000*	-0.000
	1.0219	0.0480	0.9182	-0.0344	-1.8132	-0.2186
Male	-0.002**	-0.001	-0.002***	-0.002	-0.001	-0.003*
	-2.4416	-1.2185	-2.7345	-1.6235	-0.7781	-1.7595
Age	-0.001**	-0.000	-0.001*	-0.001	-0.000	-0.001
	-2.0116	-0.9616	-1.9109	-1.4796	-0.2272	-1.4555
Age squared	0.000*	0.000	0.000	0.000	0.000	0.000
	1.7180	0.7068	1.6116	1.3623	0.1633	1.3638
Distance to nearest specialist	0.001***	0.000	0.000	0.001***	-0.000	0.001
	3.5183	0.6550	1.6261	3.3481	-0.0839	1.4355
Distance to nearest specialist squared	-0.000	0.000	0.000	-0.000	0.000	0.000
	-0.3998	1.2718	0.4728	-0.5513	1.2047	0.3302
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.003	-0.010	-0.215***	-0.004	-0.025*	-0.313***
	-0.3233	-0.8645	-3.3518	-0.3584	-1.8709	-3.7822
Severity	0.009***	0.007***	0.008***	0.017***	0.014***	0.016***
	5.8976	7.1256	5.3855	8.3122	9.8960	7.5407
Severity squared	-0.000**	-0.000***	-0.000*	-0.001***	-0.001***	-0.001***
	-2.0696	-2.8891	-1.8094	-3.8491	-4.8736	-3.2153
Male	-0.003***	-0.000	-0.003***	-0.004***	-0.000	-0.004***
	-3.0076	-0.2471	-3.1041	-2.7083	-0.0320	-2.7198
Age	-0.001**	-0.001	-0.001	-0.002*	-0.002*	-0.001
	-2.0266	-1.2397	-1.5968	-1.8094	-1.8840	-1.5067
Age squared	0.000	0.000	0.000	0.000	0.000	0.000
	1.6242	0.8291	1.3579	1.6203	1.5449	1.4335
Distance to nearest specialist	0.002***	0.001	0.001***	0.002***	0.000	0.001***
	6.2358	1.6167	3.5067	6.4365	0.2725	3.2136
Distance to nearest specialist squared	-0.000***	-0.000	-0.000***	-0.000***	0.000	-0.000**
	-4.2203	-1.4894	-2.6204	-4.4427	0.1172	-2.5338
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 54: Outpatient RVUs, full sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.47**	-5.88	3.30	-0.46**	1.29	3.62
	-2.350	-0.853	1.338	-2.551	0.941	1.637
Severity	0.389***	0.195	0.260**	0.351**	0.193	0.236*
	3.0762	0.6553	1.9935	2.3561	0.5030	1.7876
Severity squared	0.000	0.040	0.022	0.006	0.031	0.027
	0.0208	0.7505	1.1152	0.2562	0.4491	1.3921
Male	-0.150	-0.068	-0.162	-0.050	0.025	-0.048
	-1.1087	-0.2575	-1.1707	-0.3435	0.0962	-0.3225
Age	-0.040	-0.176	-0.023	-0.069	-0.103	-0.072
	-0.6453	-1.5062	-0.3255	-1.0135	-0.5926	-1.0021
Age squared	0.000	0.002	-0.000	0.000	0.001	0.000
	0.2980	1.4783	-0.0188	0.6473	0.5641	0.6121
Distance to nearest specialist	-0.048***	0.090	-0.036**	-0.047***	0.274	-0.033*
	-3.0347	0.2420	-2.2132	-3.0628	0.9829	-1.9028
Distance to nearest specialist squared	0.000	0.000	0.000	0.000	-0.004	0.000
	1.4733	0.0356	0.8518	1.3824	-1.0220	0.5653
F	–	–	28.5	–	–	27.5
N	6,048	6,048	5,874	5,539	5,539	5,373

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.65*	-2.41*	10.14**	-0.547**	-0.379	3.189
	-1.951	-1.790	2.030	-2.1792	-0.2842	0.5328
Severity	1.03***	0.39	1.15***	0.972***	0.202	1.021***
	5.428	1.411	6.089	5.7105	0.3752	5.8210
Severity squared	-0.052***	-0.029	-0.059***	-0.024	0.026	-0.026
	-4.0023	-1.2381	-3.9734	-1.3434	0.4600	-1.3595
Male	0.731***	0.712*	0.777***	0.644***	0.826*	0.662***
	2.8910	1.6813	3.0149	2.9353	1.8636	2.9297
Age	-0.029	0.074	-0.024	0.073	0.670**	0.082
	-0.3021	0.3613	-0.2363	0.9104	1.9869	0.9693
Age squared	0.000	-0.000	0.000	-0.001	-0.006*	-0.001
	0.2917	-0.2062	0.2156	-0.9293	-1.9133	-0.9938
Distance to nearest specialist	-0.041***	-0.046	-0.006	-0.035***	0.109	-0.022
	-2.6918	-0.2494	-0.2334	-2.9464	1.3886	-1.0523
Distance to nearest specialist squared	0.000**	0.001	0.000	0.000*	-0.003	0.000
	2.0122	0.1580	0.5369	1.7867	-1.6197	1.1025
F	–	–	20.0	–	–	27.2
N	11,036	11,036	10,768	10,789	10,789	10,504

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 55: Any new cardiovascular prescription, full sample

Panel A: Generalists

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.003	0.003	0.191*	0.003	0.010	0.149
	0.3487	0.3037	1.9476	0.3835	0.5957	1.2263
Severity	0.250***	0.264***	0.250***	0.254***	0.277***	0.254***
	20.7245	20.1813	20.5839	29.1518	29.8085	29.0214
Severity squared	-0.022***	-0.024***	-0.022***	-0.022***	-0.025***	-0.021***
	-10.8831	-10.8842	-10.7604	-15.3139	-16.9706	-15.0855
Male	0.036***	0.035***	0.036***	0.036***	0.037***	0.036***
	8.3540	6.4379	8.0643	6.8944	5.3762	6.6959
Age	0.008***	0.006***	0.008***	0.007***	0.006***	0.006***
	6.4583	5.8946	6.6942	4.6668	5.3543	4.5236
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	-8.1335	-7.6811	-8.4191	-6.6150	-7.5880	-6.4511
Distance to nearest specialist	0.000	0.003**	0.001*	-0.000	0.001	0.001
	0.1922	2.3444	1.7707	-0.2026	0.5264	0.7894
Distance to nearest specialist squared	-0.000	-0.000***	-0.000**	-0.000	0.000	-0.000*
	-1.3569	-2.6949	-2.3528	-0.8971	0.0816	-1.8610
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Cardiologists

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.008	0.013	0.031	0.007	-0.005	-0.023
	1.1363	0.7345	0.4621	1.2027	-0.2385	-0.2906
Severity	0.045***	0.039***	0.044***	0.061***	0.062***	0.060***
	8.0071	4.9654	7.9371	11.5580	14.5721	11.8846
Severity squared	-0.003***	-0.003**	-0.003***	-0.004***	-0.004***	-0.004***
	-3.4169	-2.1672	-3.3634	-5.8719	-7.3658	-5.9398
Male	0.016***	0.014***	0.016***	0.020***	0.017***	0.020***
	5.1870	3.7692	5.2936	5.8089	3.1559	6.0583
Age	0.001	-0.001	0.002*	0.000	-0.001	0.000
	1.4950	-0.2911	1.7152	0.1388	-0.3260	0.0756
Age squared	-0.000	0.000	-0.000*	0.000	0.000	0.000
	-1.4330	0.3332	-1.6750	0.0573	0.5940	0.1301
Distance to nearest specialist	-0.000	-0.000	-0.000	-0.001	-0.003	-0.001
	-0.1221	-0.1486	-0.1561	-0.8822	-1.3300	-0.8988
Distance to nearest specialist squared	0.000	0.000	0.000	0.000	0.000	0.000
	0.4393	0.7447	0.5369	1.0042	1.1406	0.7325
F	–	–	8.5	–	–	17.8
N	30,046	30,046	29,654	25,924	25,924	25,569

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 56: Inpatient admission, full sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.001*	0.001	0.003	-0.002*	0.002	0.017*
	-1.8231	0.3018	0.2956	-1.7608	0.5926	1.6749
Severity	0.005***	0.004***	0.005***	0.006***	0.006***	0.006***
	5.7489	6.7284	5.5657	4.4685	4.2897	4.3987
Severity squared	0.000	-0.000	0.000	0.001**	0.000	0.001**
	0.1595	-0.2038	0.2452	2.0910	1.0081	2.2255
Male	0.002***	0.003***	0.002***	0.004***	0.004***	0.004***
	4.4119	4.8486	4.3467	5.1932	4.5535	5.1437
Age	-0.000	-0.000**	-0.000	-0.000***	-0.001***	-0.000***
	-1.3047	-2.4722	-1.2163	-2.6117	-3.5466	-2.6110
Age squared	0.000**	0.000***	0.000**	0.000***	0.000***	0.000***
	2.2893	3.3922	2.1866	3.9681	5.1348	3.9297
Distance to nearest specialist	-0.000**	0.000	-0.000	-0.000	-0.000	-0.000
	-2.0515	0.0982	-1.3565	-1.5380	-1.3444	-0.3278
Distance to nearest specialist squared	0.000	-0.000	0.000	0.000	0.000	-0.000
	1.0040	-0.7686	0.8083	0.5961	1.1014	-0.0083
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus patients, full sample

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.002	-0.004	-0.042***	-0.002	0.008	0.009
	-1.3618	-1.0219	-2.8192	-0.6743	1.6288	0.2652
Severity	0.011***	0.010***	0.011***	0.020***	0.016***	0.020***
	13.2381	11.8198	12.8710	14.6300	10.9812	15.1202
Severity squared	0.000	0.000	0.000	0.000	0.000	0.000
	1.2221	0.7144	1.1576	0.7549	1.0789	0.5958
Male	0.007***	0.006***	0.007***	0.010***	0.008***	0.010***
	11.0817	9.4988	11.0195	12.5368	6.2211	13.0584
Age	-0.002***	-0.002***	-0.002***	-0.002***	-0.001**	-0.002***
	-6.1414	-6.6997	-6.0587	-3.7763	-2.1813	-3.7988
Age squared	0.000***	0.000***	0.000***	0.000***	0.000**	0.000***
	6.2302	6.8683	6.1376	3.9237	2.4545	3.9180
Distance to nearest specialist	-0.000	-0.000	-0.000**	-0.000*	-0.001	-0.000
	-1.4497	-0.0577	-2.4768	-1.9534	-1.4445	-1.1888
Distance to nearest specialist squared	0.000	-0.000	0.000	0.000	0.000	0.000
	0.3715	-0.6302	0.8842	1.6410	1.4304	1.0373
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 57: Physician cost, full sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	2.81	6.25	63.87	6.6	29.1***	97.0*
	0.703	1.449	1.335	1.31	3.30	1.77
Severity	5.98***	4.94***	5.77***	9.48***	9.84***	9.34***
	5.555	4.331	5.488	6.587	5.739	6.429
Severity squared	-0.248**	-0.299*	-0.206**	-0.402**	-0.420*	-0.365*
	-2.3494	-1.8981	-1.9732	-2.0880	-1.7479	-1.8213
Male	1.89*	2.95**	1.89*	7.38***	10.66***	7.32***
	1.654	2.000	1.655	3.230	3.629	3.145
Age	0.581*	0.937**	0.593*	3.54***	3.51***	3.48***
	1.8983	2.3698	1.9281	4.505	4.319	4.523
Age squared	-0.011***	-0.012***	-0.011***	-0.038***	-0.034***	-0.037***
	-3.3082	-2.9189	-3.3028	-4.4163	-3.9291	-4.4337
Distance to nearest specialist	-1.138***	-0.957	-0.866***	-1.48***	-0.70	-1.09***
	-5.1959	-1.5463	-3.6119	-5.431	-0.634	-3.654
Distance to nearest specialist squared	0.012***	0.015	0.010**	0.015***	-0.007	0.012**
	3.6066	1.0094	2.5324	3.2576	-0.5412	2.4116
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	16.2***	-21.8**	246.6**	31.1***	-44.6*	307.3**
	2.99	-2.09	2.54	2.87	-1.68	2.03
Severity	38.4***	35.6***	38.8***	76.6***	72.6***	77.8***
	12.88	10.43	12.85	10.58	8.97	10.58
Severity squared	-1.57***	-1.53***	-1.60***	-3.53***	-3.27***	-3.60***
	-4.850	-5.734	-4.812	-8.599	-7.119	-8.600
Male	5.49**	2.82	5.34**	5.12	-2.89	4.47
	2.319	0.945	2.339	1.372	-0.541	1.189
Age	-8.40***	-7.38***	-8.57***	-12.7***	-12.8***	-13.1***
	-5.709	-3.476	-5.848	-5.66	-4.69	-5.88
Age squared	0.059***	0.051**	0.060***	0.083***	0.087***	0.086***
	4.0993	2.5063	4.2124	3.7247	3.4581	3.8708
Distance to nearest specialist	-0.878***	-0.625	-0.356	-1.78***	-1.77	-1.06**
	-3.9712	-0.9180	-1.3978	-3.584	-1.063	-2.140
Distance to nearest specialist squared	0.008***	0.015*	0.006**	0.015**	0.032*	0.012**
	3.1702	1.9512	2.1368	2.5130	1.6499	2.0408
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 58: Outpatient cost, full sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-4.37*	1.38	-43.46***	-5.52	0.52	-64.02***
	-1.838	0.351	-2.697	-1.522	0.100	-3.017
Severity	0.078	3.756**	-0.139	0.294	7.961**	-0.111
	0.0201	2.0711	-0.0367	0.0510	2.1646	-0.0198
Severity squared	1.13	0.24	1.14	1.69	0.20	1.73
	1.547	0.703	1.589	1.548	0.367	1.609
Male	-1.69	-1.56	-2.19*	-1.25	-0.96	-1.95
	-1.412	-1.376	-1.873	-0.757	-0.485	-1.246
Age	0.453	0.678	0.547	0.511	1.180	0.578
	0.9325	1.2024	1.2227	0.7276	1.6419	0.8866
Age squared	-0.007	-0.009	-0.008*	-0.008	-0.014*	-0.009
	-1.3008	-1.4288	-1.6688	-1.0454	-1.8230	-1.3037
Distance to nearest specialist	0.403**	0.941	0.194	0.289	-0.161	0.007
	2.1828	0.8963	0.9109	1.0234	-0.0948	0.0237
Distance to nearest specialist squared	0.000	0.016	0.001	0.003	0.009	0.004
	0.0667	0.4863	0.3258	0.6927	0.5296	1.0377
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-3.09	-1.20	-47.04	0.75	4.14	-197.05
	-0.440	-0.218	-0.530	0.076	0.469	-1.063
Severity	12.6***	10.3***	12.7***	13.0	7.3	12.6
	5.47	2.59	5.52	1.31	0.58	1.26
Severity squared	-0.055	0.155	-0.047	1.96	2.30	2.05
	-0.1477	0.2732	-0.1241	1.073	1.019	1.109
Male	9.84***	9.97***	10.01***	10.9***	8.1**	11.2***
	3.319	3.486	3.321	2.67	2.06	2.73
Age	0.215	0.391	0.266	-1.67*	-0.34	-1.52
	0.2342	0.2839	0.2869	-1.698	-0.222	-1.569
Age squared	-0.008	-0.006	-0.008	0.011	0.001	0.010
	-0.7277	-0.4544	-0.7646	1.0854	0.0564	0.9799
Distance to nearest specialist	0.406	0.067	0.306	1.029***	-0.444	0.548
	1.3896	0.1196	0.8886	3.1464	-0.3612	1.1908
Distance to nearest specialist squared	-0.003	-0.007	-0.003	-0.010***	-0.001	-0.007
	-1.0825	-0.8783	-0.7795	-2.7419	-0.0345	-1.6142
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 59: Pharmacy cost (hypertension), full sample

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	17.6	9.6	224.0**	35.3	75.8	381.6
	1.40	0.43	2.39	1.11	0.81	1.33
Severity	616.8***	637.0***	616.0***	1136.8***	1181.0***	1137.7***
	23.32	24.27	23.38	29.58	36.18	29.57
Severity squared	-40.9***	-43.8***	-40.7***	-74.8***	-82.2***	-74.6***
	-11.84	-11.62	-11.79	-15.25	-18.01	-15.04
Male	137.2***	136.6***	138.0***	246.7***	233.4***	246.7***
	18.49	15.83	19.50	13.66	10.69	13.80
Age	13.7***	13.4***	13.5***	27.7***	31.2***	27.6***
	4.78	4.59	4.89	5.16	5.22	4.99
Age squared	-0.138***	-0.136***	-0.136***	-0.281***	-0.329***	-0.281***
	-4.0847	-3.9423	-4.1383	-4.4553	-5.0510	-4.2601
Distance to nearest specialist	-1.67	5.81***	-0.40	-3.59	8.97**	-1.48
	-1.480	2.633	-0.341	-1.451	2.215	-0.640
Distance to nearest specialist squared	0.009	-0.095**	-0.004	0.005	-0.164**	-0.018
	0.4843	-2.3517	-0.1845	0.1870	-2.1938	-0.6537
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 60: Total cost, full sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	16.0	17.2	244.4**	36.4	105.4	414.6
	1.21	0.69	2.12	1.13	1.08	1.35
Severity	622.9***	645.7***	621.7***	1146.6***	1198.8***	1147.0***
	24.58	24.98	24.60	29.77	36.42	29.68
Severity squared	-40.0***	-43.8***	-39.7***	-73.5***	-82.5***	-73.2***
	-12.09	-11.88	-12.03	-14.45	-18.77	-14.22
Male	137.4***	138.0***	137.7***	252.8***	243.2***	252.1***
	19.14	15.94	20.04	13.98	11.28	13.90
Age	14.7***	15.0***	14.6***	31.7***	35.9***	31.6***
	4.49	4.43	4.60	5.36	5.73	5.20
Age squared	-0.156***	-0.157***	-0.155***	-0.327***	-0.377***	-0.327***
	-4.0747	-3.9383	-4.1411	-4.7825	-5.5586	-4.5938
Distance to nearest specialist	-2.41**	5.80***	-1.07	-4.77**	8.11*	-2.56
	-2.228	2.755	-0.989	-1.979	1.723	-1.194
Distance to nearest specialist squared	0.021	-0.065	0.007	0.024	-0.162**	-0.003
	1.1706	-1.4752	0.3532	0.8539	-2.0812	-0.0977
F	–	–	52.8	–	–	48.3
N	183,640	183,640	180,714	114,644	114,644	112,803

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	13.1	-23.0**	199.5***	31.8**	-40.5	110.3
	1.52	-1.97	2.59	2.35	-1.50	0.85
Severity	51.0***	45.8***	51.4***	89.6***	79.8***	90.3***
	13.05	10.04	12.91	7.86	5.12	7.99
Severity squared	-1.62***	-1.37**	-1.64***	-1.56	-0.97	-1.56
	-2.601	-2.045	-2.605	-0.858	-0.390	-0.856
Male	15.3***	12.8***	15.3***	16.0***	5.2	15.7***
	3.84	3.13	3.83	3.16	0.82	3.18
Age	-8.18***	-6.98***	-8.31***	-14.4***	-13.2***	-14.6***
	-4.179	-2.586	-4.139	-6.29	-4.33	-6.37
Age squared	0.051**	0.045*	0.052**	0.094***	0.087***	0.096***
	2.5144	1.6812	2.5081	3.8333	3.0242	3.9262
Distance to nearest specialist	-0.472	-0.559	-0.050	-0.751	-2.216	-0.517
	-1.3190	-0.7983	-0.1306	-1.4337	-1.0896	-0.7947
Distance to nearest specialist squared	0.005	0.008	0.003	0.005	0.032	0.005
	1.2263	0.6831	0.7718	0.7573	1.1109	0.6665
F	–	–	23.7	–	–	20.6
N	188,681	188,681	185,697	130,217	130,217	128,021

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

APPENDIX B : Low-access estimates

Table 61: Primary specialist visit, low-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.018	-0.009	0.125	-0.027	-0.023	0.104
	-1.3269	-0.5387	1.4187	-1.3438	-0.7105	0.8562
Severity	0.018***	0.023***	0.017***	0.028***	0.037***	0.028***
	5.4229	5.5473	4.4733	5.2170	5.8311	4.9952
Severity squared	-0.001	-0.001**	-0.000	-0.001*	-0.002***	-0.001
	-1.2702	-2.1450	-0.6540	-1.6657	-3.3311	-1.0437
Male	-0.002	-0.002	-0.002	-0.000	-0.001	-0.001
	-0.6130	-0.4334	-0.6142	-0.0999	-0.1646	-0.1775
Age	-0.002*	-0.002	-0.003*	-0.001	-0.001	-0.001
	-1.7805	-1.0354	-1.7842	-0.7964	-0.6599	-0.7817
Age squared	0.000**	0.000	0.000**	0.000	0.000	0.000
	2.1414	1.3594	2.0977	1.3568	1.1993	1.3373
Distance to nearest specialist	-0.001*	-0.000	-0.001	-0.002**	-0.005	-0.001*
	-1.8380	-0.1148	-1.1297	-2.4821	-1.0984	-1.8647
Distance to nearest specialist squared	0.000	0.000	0.000	0.000*	0.000	0.000
	0.7884	1.5443	0.2969	1.7008	0.9061	1.2557
F	–	–	52.3	–	–	43.8
N	61,203	61,203	59,698	38,210	38,210	37,270

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.049***	-0.017	0.519***	-0.045	-0.043	0.719***
	-2.9671	-0.8174	4.3834	-1.6026	-1.5696	3.5642
Severity	0.036***	0.029***	0.031***	0.036***	0.038***	0.037***
	10.3883	6.6008	7.2401	6.9992	6.2522	7.1936
Severity squared	-0.002***	-0.002***	-0.002***	-0.001***	-0.001**	-0.002***
	-5.4014	-3.2262	-3.4430	-2.8292	-2.3724	-3.0949
Male	-0.004*	-0.003	-0.004	-0.010***	-0.008	-0.012***
	-1.6551	-0.6375	-1.0923	-2.5821	-1.1029	-2.7507
Age	-0.004***	-0.003**	-0.005***	-0.006***	-0.006***	-0.007***
	-4.4029	-2.1713	-3.0496	-3.1908	-2.8513	-3.4109
Age squared	0.000***	0.000	0.000**	0.000**	0.000*	0.000**
	2.6978	1.3825	1.9747	2.0375	1.7516	2.1466
Distance to nearest specialist	-0.002***	-0.000	-0.001	-0.003***	-0.003	-0.001
	-6.4333	-0.2948	-1.0174	-7.2567	-1.2837	-1.4336
Distance to nearest specialist squared	0.000**	0.000	0.000	0.000***	0.000*	0.000
	2.4738	0.7716	0.1993	5.5276	1.8404	1.5329
F	–	–	17.3	–	–	14.5
N	188,681	62,869	60,727	43,396	43,396	41,829

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 62: Primary specialist RVUs, low-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.065	-0.057	-1.572	0.046	-0.420	1.093
	0.2905	-0.0916	-1.2886	0.1841	-0.5808	0.5685
Severity	0.390***	0.283***	0.415***	0.471***	0.420***	0.438***
	7.4091	2.7276	6.8049	5.7079	3.3519	6.0806
Severity squared	-0.025***	-0.025*	-0.030***	-0.021**	-0.025**	-0.017**
	-4.0528	-1.8603	-4.3695	-2.3039	-2.1353	-2.0162
Male	0.097	0.249**	0.088	0.219**	0.343***	0.225***
	1.4207	2.5052	1.3584	2.5299	3.3367	2.5947
Age	-0.047	-0.046	-0.051	-0.040	-0.077	-0.038
	-1.2673	-0.8182	-1.4248	-0.6531	-1.6412	-0.6390
Age squared	0.001	0.001	0.001	0.001	0.001**	0.001
	1.4483	1.0744	1.6166	0.9220	2.0331	0.9094
Distance to nearest specialist	0.007	-0.086	0.005	0.004	-0.143	0.005
	0.6077	-1.3748	0.3991	0.4374	-1.0850	0.4522
Distance to nearest specialist squared	0.000	0.003*	0.000	0.000	0.003	0.000
	0.6787	1.9279	0.6997	0.3877	1.3144	0.1127
F	–	–	30.6	–	–	24.9
N	9,765	9,765	9,587	8,430	8,430	8,274

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.592*	-0.226	7.672	0.934*	0.739	13.153
	1.6915	-0.6233	1.0621	1.7820	0.8900	1.1024
Severity	0.754***	0.572***	0.778***	1.14***	0.97***	1.16***
	9.9105	5.5227	10.0285	11.379	6.189	12.824
Severity squared	-0.028***	-0.012	-0.031***	-0.037***	-0.033**	-0.039***
	-3.2173	-1.0607	-3.4965	-2.6479	-2.3553	-3.1642
Male	0.017	-0.019	0.059	0.027	-0.170	0.037
	0.1723	-0.0967	0.4767	0.1832	-0.5543	0.2087
Age	-0.237***	-0.165***	-0.237***	-0.300***	-0.433***	-0.329***
	-5.9918	-2.8612	-6.1758	-3.5542	-2.7916	-3.4470
Age squared	0.002***	0.001**	0.002***	0.003***	0.004**	0.003***
	5.4009	2.4620	5.5018	2.9085	2.4917	2.8461
Distance to nearest specialist	-0.025***	0.064**	-0.014	-0.038***	-0.041	-0.020
	-2.7022	2.5018	-1.1738	-2.6719	-0.3077	-0.9782
Distance to nearest specialist squared	0.000**	-0.001***	0.000	0.000**	0.000	0.000
	2.1805	-3.0700	1.3301	2.0111	0.1259	0.8909
F	–	–	13.6	–	–	16.1
N	16,573	16,573	16,183	15,127	15,127	14,716

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 63: Generalist RVUs, low-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.397***	0.237**	0.227	-0.503***	0.589*	1.129
	-7.9392	2.5436	0.5673	-6.0699	1.9525	1.5361
Severity	0.059***	0.023	0.052**	0.120***	0.125**	0.111**
	2.7529	0.8913	2.3037	3.0096	2.3156	2.2616
Severity squared	-0.004	-0.001	-0.002	-0.008*	-0.012**	-0.005
	-1.1786	-0.4355	-0.7141	-1.8702	-1.9979	-0.8853
Male	-0.026	-0.013	-0.035*	0.013	0.070	0.004
	-1.3425	-0.5380	-1.8461	0.2806	1.0448	0.0934
Age	-0.017**	-0.010	-0.018**	-0.002	0.020	-0.008
	-2.1209	-0.7610	-2.2111	-0.0908	1.1731	-0.4419
Age squared	0.000	0.000	0.000	0.000	-0.000	0.000
	1.4761	0.6268	1.5425	0.2277	-0.8384	0.5866
Distance to nearest specialist	-0.025***	0.004	-0.023***	-0.027***	0.014	-0.022***
	-5.7956	0.2721	-5.5464	-4.3405	0.3539	-3.4960
Distance to nearest specialist squared	0.000***	0.000	0.000***	0.000**	-0.001	0.000*
	3.5086	0.5290	3.3098	2.3499	-0.8247	1.6471
F	–	–	52.3	–	–	43.9
N	61,153	61,153	59,648	38,166	38,166	37,226

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.389***	0.170	1.249	-0.650***	0.644***	2.344
	-4.6661	1.2086	0.7965	-3.8639	2.9420	0.9167
Severity	0.158***	0.153***	0.151***	0.280***	0.258***	0.277***
	7.7044	7.2175	6.5322	6.4967	5.0771	6.1402
Severity squared	-0.004	-0.005*	-0.004	-0.010	-0.007	-0.010
	-1.2487	-1.6956	-0.9340	-1.3143	-0.8936	-1.1834
Male	0.097***	0.159***	0.086**	0.138**	0.226***	0.119**
	2.8559	5.0095	2.4535	2.4847	3.0235	1.9800
Age	-0.013	-0.011	-0.010	0.021	0.069**	0.024
	-0.7817	-0.5920	-0.5748	0.6817	2.0640	0.7219
Age squared	0.000	0.000	0.000	-0.000	-0.001*	-0.000
	0.7217	0.7830	0.5237	-0.4069	-1.7371	-0.4958
Distance to nearest specialist	-0.010***	-0.001	-0.006	-0.011**	-0.017	-0.002
	-3.9958	-0.0746	-1.2802	-2.2156	-0.5790	-0.2061
Distance to nearest specialist squared	0.000***	0.000	0.000	0.000	0.000	-0.000
	2.6240	0.0892	0.9399	1.0906	0.8802	-0.1574
F	–	–	17.3	–	–	14.5
N	62,863	62,863	60,721	43,394	43,394	41,827

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 64: Any outpatient visit, low-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.005	0.017	-0.004	-0.006	0.015	-0.034
	-0.9936	1.3852	-0.1479	-0.6243	0.9734	-0.8124
Severity	0.001	0.002	0.001	0.004	0.005**	0.004
	0.4406	1.3848	0.2503	1.3410	2.3626	1.3637
Severity squared	0.000	0.000	0.000	0.000	-0.000	0.000
	1.0616	0.2551	1.1181	0.6943	-0.6654	0.6459
Male	-0.002	-0.000	-0.002	-0.001	-0.002	-0.001
	-1.2066	-0.1560	-1.4278	-0.4689	-1.0016	-0.5668
Age	-0.001	0.000	-0.001	-0.001	0.000	-0.001
	-1.1632	0.0019	-1.0044	-1.5330	0.2526	-1.2449
Age squared	0.000	-0.000	0.000	0.000	-0.000	0.000
	0.8584	-0.3541	0.6904	1.3500	-0.4078	1.0288
Distance to nearest specialist	0.001***	-0.001	0.001***	0.002***	-0.001	0.002***
	3.4107	-1.0491	3.0579	3.3268	-0.2558	2.8992
Distance to nearest specialist squared	-0.000	0.000**	-0.000	-0.000	-0.000	-0.000
	-1.1948	2.2791	-1.1224	-1.3383	-0.1689	-1.2436
F	–	–	52.3	–	–	43.8
N	61,203	61,203	59,698	38,210	38,210	37,270

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.001	-0.026	-0.266**	-0.003	-0.042	-0.355***
	-0.0536	-1.4985	-2.4674	-0.1903	-1.6409	-2.6169
Severity	0.011***	0.009***	0.010***	0.019***	0.016***	0.019***
	6.0508	5.7896	5.0020	6.7563	7.1256	6.4989
Severity squared	-0.001**	-0.001**	-0.000*	-0.001**	-0.001***	-0.001**
	-2.5428	-2.4668	-1.8755	-2.4401	-2.9818	-2.2197
Male	-0.001	0.000	-0.001	-0.000	0.003	0.000
	-0.5976	0.0138	-0.5418	-0.0328	0.7908	0.1529
Age	-0.002**	-0.001	-0.001	-0.001	-0.000	-0.000
	-1.9987	-0.4652	-1.6175	-0.4867	-0.2302	-0.1381
Age squared	0.000*	0.000	0.000	0.000	0.000	0.000
	1.6494	0.2355	1.4274	0.3837	0.0470	0.1069
Distance to nearest specialist	0.002***	0.003***	0.001***	0.003***	0.003**	0.002***
	5.8756	2.6187	2.5763	6.6363	2.4300	3.1875
Distance to nearest specialist squared	-0.000***	-0.000**	-0.000**	-0.000***	-0.000	-0.000**
	-4.0778	-2.0928	-2.0010	-4.6201	-1.0646	-2.5030
F	–	–	17.3	–	–	14.5
N	62,869	62,869	60,727	43,396	43,396	41,829

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 65: Outpatient RVU, low-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.15	11.78	-1.98	-0.303	0.000	-0.464
	-0.283	1.054	-0.635	-0.5609	nan	-0.1394
Severity	0.541***	0.645	0.656***	0.496**	0.960**	0.520*
	3.7746	1.5016	2.6793	2.5766	2.0059	1.6779
Severity squared	-0.020	-0.007	-0.038	-0.014	-0.090	-0.016
	-1.6184	-0.0914	-1.0794	-0.8018	-1.6168	-0.3827
Male	-0.172	0.350	-0.260	0.002	0.325	-0.118
	-0.7196	0.7148	-1.3163	0.0081	0.6566	-0.5622
Age	-0.101	-0.031	-0.106	-0.076	0.047	-0.078
	-1.2172	-0.1811	-1.2337	-0.6601	0.2013	-0.6951
Age squared	0.001	0.000	0.001	0.001	-0.000	0.001
	1.0355	0.2704	0.9935	0.5960	-0.1441	0.5916
Distance to nearest specialist	-0.065***	0.371	-0.063***	-0.083***	0.750	-0.077***
	-2.5784	0.3760	-2.6056	-3.0897	0.7540	-2.9553
Distance to nearest specialist squared	0.000	-0.004	0.000	0.001**	-0.011	0.000*
	1.5159	-0.2040	1.3039	2.2071	-0.8086	1.6901
F	–	–	7.3	–	–	7.1
N	2,016	2,016	1,909	1,846	1,846	1,752

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.331	-3.596**	0.510	-0.45	-4.88	7.83
	-0.7195	-2.3670	0.0390	-1.099	-0.750	1.262
Severity	0.968***	0.281	0.990***	0.733***	-1.017	0.817***
	4.2563	0.9446	3.6016	2.6920	-1.5619	2.6532
Severity squared	-0.073***	-0.050	-0.073**	-0.024	0.151**	-0.030
	-3.1002	-0.8376	-2.5338	-0.6903	2.1782	-0.7490
Male	0.252	0.145	0.271	0.385*	0.361	0.449*
	0.8944	0.4101	0.7142	1.7177	1.2182	1.8459
Age	-0.078	-0.163	-0.089	-0.042	0.343	-0.082
	-0.6206	-0.8190	-0.7838	-0.2804	1.6037	-0.5825
Age squared	0.001	0.002	0.001	0.000	-0.003	0.001
	0.3969	0.8611	0.5229	0.0947	-1.5419	0.3754
Distance to nearest specialist	-0.011	0.324	-0.001	-0.004	-0.776	0.026
	-0.4221	0.8449	-0.0276	-0.2014	-1.4463	0.7740
Distance to nearest specialist squared	0.000	-0.012	-0.000	-0.000	0.006	-0.000
	0.1544	-1.2304	-0.1168	-0.5143	0.6945	-1.0965
F	–	–	8.3	–	–	10.8
N	3,678	3,678	3,478	3,591	3,591	3,384

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 66: Any new cardiovascular prescription, low-access sample

Panel A: Generalists

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.011	-0.009	-0.005	0.013	0.026	-0.044
	0.8374	-0.4416	-0.0350	0.8229	0.8118	-0.2468
Severity	0.264***	0.276***	0.264***	0.262***	0.286***	0.262***
	21.3476	23.1028	22.0441	18.9042	22.5953	20.4617
Severity squared	-0.024***	-0.026***	-0.024***	-0.023***	-0.027***	-0.023***
	-12.0505	-14.5327	-12.0683	-9.8729	-12.6244	-10.3029
Male	0.037***	0.030***	0.037***	0.039***	0.041***	0.039***
	7.7441	4.5584	7.7859	6.9136	6.2517	6.8874
Age	0.008***	0.008***	0.008***	0.007***	0.008***	0.007***
	4.6134	4.7477	4.2563	3.5550	3.7787	3.2300
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	-5.9562	-6.3168	-5.6196	-4.9881	-5.8144	-4.6885
Distance to nearest specialist	-0.002*	-0.001	-0.002**	-0.002**	-0.001	-0.002**
	-1.9565	-0.3123	-2.0113	-2.0908	-0.0949	-2.3442
Distance to nearest specialist squared	0.000	-0.000	0.000	0.000	0.000	0.000
	0.8192	-0.4788	0.9324	1.1711	0.0952	1.3370
F	–	–	52.3	–	–	43.8
N	61,203	61,203	59,698	38,210	38,210	37,270

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Cardiologists

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.011	0.026	-0.067	-0.003	-0.013	-0.105
	-0.8552	1.4388	-0.5997	-0.2781	-0.4172	-0.9684
Severity	0.047***	0.045***	0.048***	0.058***	0.056***	0.059***
	6.9481	6.1577	6.9420	7.3028	5.1597	7.6412
Severity squared	-0.003***	-0.003***	-0.003***	-0.004***	-0.004***	-0.004***
	-2.8299	-3.7489	-3.0396	-3.4635	-3.4036	-3.8211
Male	0.010**	0.009**	0.010**	0.015***	0.013**	0.015***
	2.0165	1.9648	2.0774	3.7095	2.1122	3.8218
Age	0.002	0.001	0.002	0.003	0.000	0.002
	1.0819	0.3608	1.0877	1.5669	0.1194	1.3259
Age squared	-0.000	-0.000	-0.000	-0.000	0.000	-0.000
	-0.9758	-0.3935	-0.9946	-1.2743	0.1594	-1.0441
Distance to nearest specialist	0.000	0.009*	0.000	0.000	0.003	0.000
	0.4276	1.7789	0.1681	0.2365	0.3827	0.0884
Distance to nearest specialist squared	-0.000	-0.000	0.000	-0.000	0.000	-0.000
	-0.2143	-1.0043	0.0498	-0.1886	0.1003	-0.5245
F	–	–	5.1	–	–	7.2
N	10,008	10,008	9,826	8,632	8,632	8,473

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 67: Inpatient admission, low-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.002	0.004	-0.006	-0.004*	0.004	-0.014
	-1.3270	1.0747	-0.7908	-1.7349	0.5496	-0.9512
Severity	0.002	0.003*	0.002	0.004*	0.005**	0.003*
	1.1473	1.7153	1.1294	1.8779	2.1332	1.8548
Severity squared	0.001	0.000	0.001*	0.001***	0.001	0.001***
	1.6361	0.8964	1.6620	3.0363	1.3397	3.1999
Male	0.002***	0.002***	0.002***	0.004***	0.003***	0.004***
	3.1959	2.6337	3.1266	4.5630	2.7925	4.5823
Age	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	-0.4726	-1.0778	-0.4288	-0.3048	-0.0390	-0.3250
Age squared	0.000	0.000	0.000	0.000	0.000	0.000
	1.2699	1.5936	1.2180	1.0880	0.6630	1.0971
Distance to nearest specialist	-0.000	0.000	-0.000	-0.000**	-0.001*	-0.000**
	-1.4201	0.5728	-1.6205	-2.5656	-1.6872	-2.2576
Distance to nearest specialist squared	0.000	-0.000	0.000	0.000*	0.000	0.000*
	0.5002	-0.8083	0.7027	1.7055	1.2407	1.7016
F	–	–	52.3	–	–	43.8
N	61,203	61,203	59,698	38,210	38,210	37,270

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.000	-0.002	0.002	-0.002	0.007	0.084
	0.1729	-0.1934	0.0527	-0.5673	0.8486	1.2324
Severity	0.010***	0.009***	0.010***	0.020***	0.019***	0.021***
	9.0202	6.3866	8.6207	11.7315	7.4736	12.7353
Severity squared	0.000	0.000	0.000	-0.000	-0.000	-0.000
	0.8976	0.9136	0.6402	-0.0325	-0.2740	-0.4754
Male	0.004***	0.003***	0.004***	0.008***	0.006**	0.008***
	4.0945	2.7273	4.1861	5.0446	2.1283	5.5995
Age	-0.001**	-0.002**	-0.001**	-0.001	0.000	-0.001
	-2.1476	-2.5407	-1.9643	-1.3465	0.5354	-1.3552
Age squared	0.000**	0.000***	0.000**	0.000	-0.000	0.000
	2.1885	2.5952	2.0228	1.3811	-0.4939	1.3508
Distance to nearest specialist	0.000	0.001	0.000	0.000	0.000	0.000
	0.8705	1.0731	0.7319	0.6033	0.1014	1.3412
Distance to nearest specialist squared	-0.000	-0.000	-0.000	-0.000	0.000	-0.000
	-1.0355	-1.2688	-0.9819	-0.4413	0.2752	-1.2327
F	–	–	17.3	–	–	14.5
N	62,869	62,869	60,727	43,396	43,396	41,829

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 68: Physician cost, low-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	3.12	14.98***	-9.60	0.8	42.2***	60.6
	0.758	2.982	-0.194	0.15	3.09	0.98
Severity	6.40***	4.37***	6.35***	11.8***	13.8***	11.6***
	5.124	3.196	4.885	6.24	5.60	6.16
Severity squared	-0.561***	-0.405*	-0.559***	-0.742***	-1.127***	-0.643**
	-3.2750	-1.8736	-3.0483	-2.9831	-3.4236	-2.3161
Male	-1.29	0.18	-1.69	1.84	6.59*	1.33
	-0.792	0.089	-1.089	0.834	1.704	0.590
Age	0.239	0.781	0.255	2.52***	3.24***	2.24**
	0.5480	1.2967	0.5726	2.593	3.025	2.403
Age squared	-0.009**	-0.011*	-0.009**	-0.030***	-0.035***	-0.028***
	-1.9853	-1.7951	-2.0024	-3.1581	-3.1681	-3.0161
Distance to nearest specialist	-1.31***	1.08	-1.34***	-1.53***	1.51	-1.34***
	-5.142	0.913	-4.944	-4.835	0.713	-3.922
Distance to nearest specialist squared	0.014***	0.003	0.014***	0.016***	-0.031	0.014***
	3.8968	0.2018	3.9520	3.5916	-0.9523	3.0271
F	–	–	52.3	–	–	43.8
N	61,203	61,203	59,698	38,210	38,210	37,270

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	17.8**	-6.1	271.2**	12.9	-8.3	438.4
	2.06	-0.40	2.11	1.01	-0.26	1.42
Severity	34.3***	31.6***	35.1***	65.4***	65.8***	66.7***
	6.84	5.21	7.03	5.23	4.08	5.28
Severity squared	-1.72***	-1.42***	-1.83***	-3.00***	-2.82**	-3.11***
	-4.219	-3.035	-4.665	-3.318	-2.214	-3.488
Male	6.21**	4.98	4.89*	4.84	3.55	2.49
	2.476	1.264	1.907	0.887	0.411	0.434
Age	-8.03***	-7.13**	-8.21***	-15.8***	-21.4***	-16.3***
	-3.434	-2.234	-3.502	-4.33	-5.51	-4.41
Age squared	0.054**	0.051	0.056**	0.111***	0.169***	0.116***
	2.3667	1.6261	2.4231	3.0052	4.4030	3.0511
Distance to nearest specialist	-0.665**	0.391	-0.093	-1.578***	0.100	-0.569
	-2.4518	0.4175	-0.2244	-2.8911	0.0414	-0.5598
Distance to nearest specialist squared	0.008***	0.009	0.005	0.016***	0.028	0.010
	2.7459	0.9871	1.4239	2.6948	1.1099	1.2595
F	–	–	17.3	–	–	14.5
N	62,869	62,869	60,727	43,396	43,396	41,829

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 69: Outpatient cost, low-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-2.02	9.54	-46.61*	-1.41	8.12	-49.06
	-0.288	1.114	-1.677	-0.118	0.763	-1.345
Severity	-1.14	5.98	-1.66	-0.82	10.94*	-1.40
	-0.136	1.435	-0.210	-0.076	1.811	-0.138
Severity squared	1.67	0.41	1.68	2.26	0.14	2.26
	1.163	0.699	1.237	1.258	0.165	1.344
Male	-0.61	-1.02	-1.90	3.40	-1.46	1.10
	-0.240	-0.425	-0.796	0.949	-0.363	0.317
Age	-1.253	0.237	-0.973	-1.24	1.67	-0.85
	-1.6063	0.2520	-1.3845	-0.994	1.254	-0.875
Age squared	0.011	-0.003	0.008	0.013	-0.016	0.008
	1.3257	-0.3129	1.0123	0.9690	-1.2552	0.7527
Distance to nearest specialist	0.312	-1.237	0.208	0.202	-0.585	0.143
	1.2346	-0.5734	0.7421	0.5147	-0.3076	0.3590
Distance to nearest specialist squared	0.001	0.064	0.001	0.005	0.003	0.003
	0.3657	1.0220	0.2884	0.8477	0.1119	0.6794
F	–	–	52.3	–	–	43.8
N	61,203	61,203	59,698	38,210	38,210	37,270

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-8.36	-5.40	-168.04	-3.7	19.8	-327.4
	-1.157	-0.640	-1.440	-0.24	0.58	-1.54
Severity	16.6***	16.9***	16.6***	-4.81	-11.39	-5.97
	4.56	4.31	4.43	-0.153	-0.301	-0.180
Severity squared	-0.864**	-0.773**	-0.814**	5.22	6.10	5.57
	-2.4333	-2.2719	-2.1704	0.874	0.839	0.889
Male	11.6***	9.8***	11.8***	9.07	7.45	9.92
	2.70	2.62	2.71	1.386	1.109	1.483
Age	-0.822	-0.775	-0.542	-2.43	-2.22	-1.87
	-0.4129	-0.3295	-0.2663	-0.882	-0.662	-0.696
Age squared	-0.001	0.003	-0.004	0.012	0.016	0.007
	-0.0481	0.1279	-0.1626	0.4460	0.4489	0.2541
Distance to nearest specialist	0.890***	-0.273	0.495	1.69***	1.67	0.95*
	3.3269	-0.1644	1.1460	5.358	1.401	1.708
Distance to nearest specialist squared	-0.008***	0.002	-0.005	-0.016***	-0.006	-0.011**
	-3.0472	0.2522	-1.4265	-4.3502	-0.7599	-2.1402
F	–	–	17.3	–	–	14.5
N	62,869	62,869	60,727	43,396	43,396	41,829

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 70: Pharmacy cost (hypertension), low-access sample

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-5.4	-72.7	-213.9	-20.9	-52.9	-67.7
	-0.33	-1.54	-1.22	-0.55	-0.36	-0.19
Severity	617.1***	629.0***	619.4***	1130.7***	1227.7***	1133.8***
	21.59	24.46	22.83	19.83	26.46	20.25
Severity squared	-38.3***	-40.1***	-38.4***	-71.1***	-85.3***	-70.8***
	-8.59	-7.72	-9.02	-7.69	-10.61	-7.74
Male	121.8***	118.0***	123.8***	230.3***	213.4***	233.7***
	8.85	8.61	9.48	7.67	6.02	7.86
Age	14.3***	14.0***	14.5***	14.9	20.6***	15.1
	3.74	3.90	3.79	1.58	2.70	1.64
Age squared	-0.145***	-0.144***	-0.146***	-0.131	-0.206**	-0.132
	-3.5223	-3.7525	-3.5859	-1.1985	-2.4107	-1.2404
Distance to nearest specialist	-4.12***	-1.31	-4.40***	-7.74**	-18.53	-6.85**
	-2.657	-0.157	-2.762	-2.351	-1.331	-2.151
Distance to nearest specialist squared	0.036	0.009	0.039*	0.047	0.245	0.034
	1.5690	0.1174	1.6885	1.2505	1.0567	0.9441
F	–	–	52.3	–	–	43.8
N	61,203	61,203	59,698	38,210	38,210	37,270

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 71: Total cost, low-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-4.3	-48.2	-270.1	-21.5	-2.6	-56.2
	-0.26	-0.96	-1.58	-0.54	-0.02	-0.15
Severity	622.4***	639.4***	624.1***	1141.7***	1252.5***	1144.0***
	19.89	23.61	21.25	17.88	26.72	18.13
Severity squared	-37.2***	-40.0***	-37.3***	-69.6***	-86.3***	-69.2***
	-7.19	-7.42	-7.72	-6.61	-11.04	-6.63
Male	119.9***	117.1***	120.2***	235.5***	218.6***	236.2***
	8.22	8.13	8.74	7.53	6.33	7.56
Age	13.3***	15.0***	13.8***	16.2*	25.5***	16.5*
	3.24	3.94	3.36	1.77	3.33	1.82
Age squared	-0.143***	-0.158***	-0.148***	-0.148	-0.257***	-0.152
	-3.2346	-3.8976	-3.3968	-1.4010	-3.0030	-1.4622
Distance to nearest specialist	-5.12***	-1.47	-5.53***	-9.06***	-17.60	-8.05***
	-3.123	-0.162	-3.377	-2.760	-1.165	-2.633
Distance to nearest specialist squared	0.051**	0.076	0.053**	0.068*	0.217	0.051
	2.0919	0.6719	2.2269	1.8389	0.9046	1.4611
F	–	–	52.3	–	–	43.8
N	61,203	61,203	59,698	38,210	38,210	37,270

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	9.4	-11.5	103.2	9.2	11.4	111.0
	0.81	-0.61	0.74	0.42	0.24	0.37
Severity	51.0***	48.5***	51.6***	60.6	54.4	60.8
	8.19	6.53	8.20	1.63	1.16	1.56
Severity squared	-2.58***	-2.19***	-2.64***	2.22	3.28	2.46
	-4.449	-3.537	-4.471	0.339	0.407	0.360
Male	17.8***	14.8**	16.7***	13.9	11.0	12.4
	4.13	2.51	3.84	1.41	0.82	1.26
Age	-8.85***	-7.91	-8.75***	-18.2***	-23.6***	-18.2***
	-2.857	-1.610	-2.739	-4.13	-5.28	-4.10
Age squared	0.053	0.054	0.052	0.123***	0.185***	0.122***
	1.6107	1.0749	1.5321	2.6927	4.0600	2.6507
Distance to nearest specialist	0.225	0.117	0.403	0.109	1.774	0.380
	0.5617	0.0674	0.7192	0.1707	0.6527	0.3500
Distance to nearest specialist squared	-0.000	0.012	-0.001	0.000	0.022	-0.001
	-0.0386	1.0032	-0.1194	0.0148	0.8315	-0.0790
F	–	–	17.3	–	–	14.5
N	62,869	62,869	60,727	43,396	43,396	41,829

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

APPENDIX C : High-access estimates

Table 72: Primary specialist visit, high-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.011	-0.008	0.071	-0.006	0.023	0.055
	-1.2316	-0.6663	1.2636	-0.4928	1.0975	0.6615
Severity	0.023***	0.022***	0.023***	0.034***	0.034***	0.034***
	6.8197	5.5519	6.9568	9.1311	7.5319	9.2342
Severity squared	-0.001*	-0.001**	-0.001*	-0.001***	-0.001***	-0.001***
	-1.9513	-2.1651	-1.8658	-3.1728	-2.9897	-3.1470
Male	-0.000	0.005	0.000	0.003	0.011**	0.003
	-0.0512	1.4869	0.0799	0.8060	2.1106	0.9640
Age	-0.001	-0.001	-0.001	0.002	0.002	0.002
	-0.7849	-0.9290	-0.7853	1.1111	0.8023	1.2047
Age squared	0.000	0.000	0.000	-0.000	-0.000	-0.000
	1.0921	1.3444	1.0551	-0.3903	-0.1870	-0.5021
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	46.1	–	–	45.6
N	108,916	108,916	107,621	68,037	68,037	67,210

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.049***	-0.020	0.125	-0.043**	-0.039	-0.011
	-2.9671	-1.2752	0.9242	-2.4210	-1.2852	-0.0633
Severity	0.036***	0.040***	0.040***	0.052***	0.047***	0.052***
	10.3883	8.0131	9.5627	9.6489	7.2091	9.9477
Severity squared	-0.002***	-0.002***	-0.002***	-0.003***	-0.002***	-0.003***
	-5.4014	-4.8959	-5.3069	-6.3903	-4.8096	-6.7176
Male	-0.004*	-0.005	-0.006	-0.009*	-0.016**	-0.009*
	-1.6551	-1.0047	-1.4598	-1.9365	-2.5496	-1.8817
Age	-0.004***	-0.006***	-0.006***	-0.004*	-0.003	-0.004*
	-4.4029	-3.4002	-4.3289	-1.8533	-0.7309	-1.8463
Age squared	0.000***	0.000**	0.000***	0.000	0.000	0.000
	2.6978	2.4821	2.8353	0.6556	0.1306	0.7038
Distance to nearest specialist	-0.002***	0.000	nan	0.000	0.000	nan
	-6.4333	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000**	0.000	nan	0.000	0.000	nan
	2.4738	nan	nan	nan	nan	nan
F	–	–	30.2	–	–	31.0
N	188,681	65,049	64,546	44,906	44,906	44,531

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 73: Primary specialist RVU, high-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.155	-0.693	1.804	0.226	-0.889	3.280*
	1.0568	-1.4419	1.0392	1.4569	-1.1240	1.6693
Severity	0.445***	0.536***	0.428***	0.590***	0.573***	0.570***
	4.7358	3.9553	4.6030	5.9278	3.0524	5.4601
Severity squared	-0.027***	-0.042***	-0.025***	-0.037***	-0.037**	-0.036***
	-3.3271	-3.5778	-3.2515	-3.8227	-2.1114	-3.7353
Male	0.255***	0.213	0.261***	0.338***	0.424***	0.349***
	3.0992	1.4007	3.3449	3.6011	2.8472	3.9045
Age	-0.113***	-0.026	-0.105***	-0.106***	-0.064	-0.097***
	-3.0553	-0.3868	-2.9078	-3.4784	-0.8964	-3.4508
Age squared	0.001***	0.000	0.001***	0.001***	0.001	0.001***
	3.2270	0.3422	3.1239	3.7189	0.9745	3.7955
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	89.8	–	–	68.2
N	17,751	17,751	17,566	15,300	15,300	15,122

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.552**	-0.021	3.517	0.79**	1.02	5.55
	1.9923	-0.0337	0.7364	2.033	1.099	0.800
Severity	0.822***	0.517***	0.826***	1.64***	1.29***	1.65***
	7.2097	4.5482	7.0919	10.876	7.657	11.331
Severity squared	-0.015	0.001	-0.016	-0.074***	-0.033	-0.076***
	-0.7853	0.0440	-0.8231	-4.7751	-1.5696	-4.9835
Male	-0.127	-0.171	-0.114	-0.246	-0.261	-0.240
	-1.4993	-1.2981	-1.4310	-1.4593	-1.3728	-1.4433
Age	-0.311***	-0.124	-0.306***	-0.407***	-0.387***	-0.406***
	-3.9563	-1.0449	-4.0067	-5.1184	-3.1866	-5.2337
Age squared	0.003***	0.001	0.003***	0.004***	0.004***	0.004***
	3.6222	0.8802	3.6292	4.6460	2.8002	4.7152
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	23.6	–	–	25.0
N	18,097	18,097	17,988	16,428	16,428	16,314

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 74: Generalist RVUs, high-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.552***	-0.027	1.923***	-0.639***	0.279**	2.206**
	-6.3677	-0.2869	3.0448	-4.7854	2.3189	2.2437
Severity	0.064***	0.013	0.062***	0.070***	0.075	0.068**
	3.2970	0.6633	3.0037	2.7110	1.6256	2.5037
Severity squared	-0.001	0.002	-0.001	-0.004	-0.006	-0.004
	-0.5569	0.6092	-0.3284	-1.0681	-0.7031	-1.0397
Male	-0.014	-0.003	-0.012	0.044	0.069*	0.047
	-0.5999	-0.1511	-0.4932	0.9747	1.8456	0.9612
Age	-0.007	-0.001	-0.007	0.030	0.034*	0.031*
	-1.1420	-0.1009	-0.9736	1.6067	1.6490	1.7179
Age squared	0.000	0.000	0.000	-0.000	-0.000	-0.000
	0.9351	0.6268	0.6641	-1.0425	-0.8951	-1.1755
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	45.7	–	–	45.3
N	108,782	108,782	107,491	67,961	67,961	67,137

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.680***	-0.264*	-1.132	-0.923***	-0.152	-3.788*
	-6.6659	-1.7744	-0.9146	-4.2557	-0.5565	-1.6659
Severity	0.181***	0.136***	0.181***	0.247***	0.194***	0.239***
	7.0131	6.7667	7.1280	5.2225	4.8171	4.9931
Severity squared	-0.005	-0.004	-0.005	-0.007	-0.005	-0.006
	-1.3458	-1.1767	-1.3324	-0.9420	-0.8527	-0.8107
Male	0.111***	0.075*	0.110***	0.255***	0.230***	0.249***
	3.0664	1.7136	3.0450	3.7604	4.8766	3.6178
Age	0.004	0.024	0.004	0.063***	0.048*	0.064***
	0.2496	1.4327	0.2353	2.9069	1.9266	2.7466
Age squared	-0.000	-0.000	-0.000	-0.001**	-0.000	-0.001**
	-0.3800	-1.3514	-0.3591	-2.4558	-1.2747	-2.3443
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	30.2	–	–	30.9
N	65,036	65,036	64,533	44,897	44,897	44,522

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 75: Any outpatient visit, high-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.003	-0.021*	-0.137***	-0.001	-0.022*	-0.204***
	-0.7102	-1.8145	-4.3436	-0.1267	-1.9133	-5.0611
Severity	0.002	0.002	0.002	0.006***	0.007***	0.006***
	1.5543	1.2048	1.4851	4.6380	3.9942	4.7110
Severity squared	0.000	0.000	0.000	-0.000	-0.001**	-0.000
	0.7480	0.0423	0.6885	-0.6919	-1.9859	-0.7443
Male	-0.002*	-0.001	-0.002*	-0.002	0.000	-0.003
	-1.6753	-0.6265	-1.8896	-1.2553	0.2216	-1.4228
Age	-0.001*	-0.000	-0.001	-0.001	-0.000	-0.001
	-1.7775	-0.7147	-1.5947	-1.0767	-0.2456	-1.0966
Age squared	0.000*	0.000	0.000	0.000	0.000	0.000
	1.6563	0.7340	1.5590	0.9771	0.2693	1.0950
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	46.1	–	–	45.6
N	108,916	108,916	107,621	68,037	68,037	67,210

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.008	-0.016	-0.164***	-0.012	-0.033	-0.236***
	-0.8810	-0.7899	-2.7073	-1.2483	-1.4616	-3.5829
Severity	0.008***	0.006***	0.008***	0.016***	0.012***	0.014***
	4.1097	4.2570	4.0595	5.6213	4.9214	5.2288
Severity squared	-0.000	-0.000	-0.000	-0.001***	-0.001**	-0.001***
	-1.2059	-1.3510	-1.1316	-3.1451	-2.5142	-2.7706
Male	-0.005**	0.002	-0.005**	-0.007***	-0.001	-0.007***
	-2.3054	0.7340	-2.3443	-3.1823	-0.2983	-3.1560
Age	-0.001	-0.000	-0.000	-0.003**	-0.002*	-0.002**
	-0.7799	-0.5166	-0.5138	-2.5278	-1.8508	-2.2489
Age squared	0.000	0.000	0.000	0.000**	0.000	0.000**
	0.4526	0.0814	0.2472	2.3016	1.5263	2.0867
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	30.2	–	–	31.0
N	65,049	65,049	64,546	44,906	44,906	44,531

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 76: Outpatient RVUs high-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.69***	191.30	6.49**	-0.744***	-0.795	6.636***
	-2.712	0.735	1.996	-2.7903	-0.3879	2.5844
Severity	0.205	0.203	0.205	0.159	-0.100	0.186
	1.0707	0.5886	0.9366	1.0062	-0.1927	1.0483
Severity squared	0.036	0.029	0.038	0.037	0.073	0.039
	1.0195	0.6149	1.0598	1.1973	0.8267	1.2128
Male	-0.120	-0.346	-0.125	-0.126	-0.250	-0.052
	-0.6469	-1.2015	-0.5492	-0.6742	-0.9389	-0.2339
Age	-0.003	-0.389**	0.026	-0.095	-0.371*	-0.104
	-0.0336	-2.1303	0.2531	-0.9008	-1.6890	-0.9219
Age squared	-0.000	0.004**	-0.001	0.001	0.003*	0.001
	-0.2239	2.0404	-0.5223	0.5880	1.6674	0.5961
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	21.9	–	–	24.5
N	3,396	3,396	3,339	3,146	3,146	3,084

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-1.0	-11.6***	15.4***	-0.26	2.49	4.91
	-1.51	-4.92	2.83	-0.479	0.665	1.142
Severity	0.723***	-0.032	0.762**	0.908***	0.191	0.932***
	2.6212	-0.0588	2.4322	4.7402	0.2665	4.8319
Severity squared	-0.025	0.022	-0.024	-0.003	-0.019	-0.003
	-1.2863	0.5715	-1.1617	-0.1345	-0.2964	-0.1035
Male	0.603	0.741	0.641	0.871**	1.255*	0.883**
	1.2803	0.9721	1.4036	2.3378	1.6704	2.3386
Age	0.104	0.530	0.080	0.059	2.036	0.089
	0.5621	0.7794	0.3670	0.3193	1.6084	0.4813
Age squared	-0.001	-0.004	-0.001	-0.001	-0.019	-0.001
	-0.5363	-0.6705	-0.3751	-0.3163	-1.5581	-0.4869
Distance to nearest specialist	-1.03	0.00	-1.28	-4.20	-11.90	-4.43
	-0.563	nan	-0.497	-1.233	-0.886	-1.379
Distance to nearest specialist squared	1.59	0.00	1.75	3.98	11.18	4.27
	1.000	nan	0.913	1.137	0.790	1.302
F	–	–	20.4	–	–	25.7
N	3,684	3,684	3,644	3,602	3,602	3,555

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 77: Any new cardiovascular prescription, high-access sample

Panel A: Generalists

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.002	0.017	0.241***	0.001	0.015	0.198*
	-0.3199	1.2145	2.8703	0.1724	0.6834	1.9164
Severity	0.243***	0.261***	0.243***	0.250***	0.280***	0.250***
	15.5182	13.8345	15.4614	22.3660	20.7127	22.4230
Severity squared	-0.020***	-0.023***	-0.020***	-0.021***	-0.025***	-0.021***
	-7.8248	-7.2193	-7.8071	-10.9139	-11.4048	-10.9847
Male	0.036***	0.038***	0.036***	0.035***	0.034***	0.035***
	6.6886	5.8729	6.4371	5.3082	3.7981	5.0733
Age	0.007***	0.005***	0.007***	0.005***	0.004**	0.006***
	5.3904	3.4231	5.3596	3.1165	2.5371	3.1233
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	-7.4141	-5.4758	-7.4242	-5.1340	-4.6406	-5.0694
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	46.1	–	–	45.6
N	108,916	108,916	107,621	68,037	68,037	67,210

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Cardiologists

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.016**	0.003	0.072	0.010	-0.008	0.020
	2.1224	0.1249	0.7627	1.5452	-0.2421	0.1988
Severity	0.043***	0.036***	0.043***	0.065***	0.069***	0.065***
	6.1918	3.5153	6.1245	15.8692	11.3378	15.9398
Severity squared	-0.003**	-0.002	-0.003**	-0.005***	-0.005***	-0.005***
	-2.5669	-1.5577	-2.5689	-8.0469	-5.8718	-8.0276
Male	0.019***	0.016***	0.019***	0.023***	0.019**	0.023***
	4.0985	2.9298	4.1885	5.3842	2.5273	5.4956
Age	0.001	-0.002	0.001	-0.002	-0.002	-0.002
	0.5794	-0.7578	0.7763	-1.1301	-0.7527	-1.1302
Age squared	-0.000	0.000	-0.000	0.000	0.000	0.000
	-0.5709	0.8914	-0.7805	1.1933	0.9532	1.2054
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	5.8	–	–	12.4
N	18,262	18,262	18,070	15,711	15,711	15,528

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 78: Inpatient admission, high-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.002*	-0.003	0.009	-0.002	-0.002	0.040***
	-1.8157	-1.2485	0.7348	-1.5053	-0.4909	2.9412
Severity	0.007***	0.005***	0.007***	0.008***	0.006***	0.007***
	9.3194	7.3998	9.1409	4.1410	3.2540	4.1232
Severity squared	-0.000***	-0.000**	-0.000***	0.000	0.000	0.000
	-2.9025	-2.0521	-2.7514	0.6109	0.3267	0.6837
Male	0.002***	0.003***	0.002***	0.003***	0.003***	0.003***
	3.5257	4.2115	3.4413	3.3643	2.9923	3.4007
Age	-0.000	-0.000**	-0.000	-0.001**	-0.001***	-0.001*
	-1.3556	-2.2543	-1.1997	-2.0369	-3.0161	-1.9493
Age squared	0.000*	0.000***	0.000*	0.000***	0.000***	0.000***
	1.9485	2.7401	1.7820	2.7942	3.5912	2.6735
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	46.1	–	–	45.6
N	108,916	108,916	107,621	68,037	68,037	67,210

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.001	-0.006	-0.063	-0.002	0.009	-0.041
	-0.8868	-0.9287	-1.0868	-0.9038	1.0901	-0.7850
Severity	0.012***	0.011***	0.012***	0.021***	0.017***	0.021***
	8.1072	7.7672	7.9057	9.9987	7.1403	9.7576
Severity squared	0.000	-0.000	0.000	0.000	0.000	0.000
	0.6568	-0.1230	0.6997	0.2784	0.4513	0.3375
Male	0.007***	0.007***	0.008***	0.010***	0.007***	0.011***
	6.0469	4.7352	6.0131	5.2455	3.5533	5.3550
Age	-0.002***	-0.002***	-0.002***	-0.002***	-0.003***	-0.002***
	-4.0766	-3.7602	-4.1621	-2.9338	-3.3981	-3.0346
Age squared	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	4.1828	4.0405	4.2904	3.0407	3.6906	3.1350
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	30.2	–	–	31.0
N	65,049	65,049	64,546	44,906	44,906	44,531

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 79: Physician cost, high-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	2.42	1.70	96.35*	7.9	31.7**	97.3*
	0.492	0.290	1.863	1.24	2.42	1.65
Severity	5.90***	4.94***	5.75***	9.01***	8.36***	8.77***
	4.478	3.277	4.788	5.285	4.519	5.335
Severity squared	-0.159	-0.244	-0.125	-0.350	-0.198	-0.322
	-1.1255	-1.2529	-0.9283	-1.5453	-0.7030	-1.4629
Male	3.05**	3.80**	3.10**	9.44***	12.34***	9.52***
	2.189	2.404	2.185	3.127	3.969	3.051
Age	0.682	1.018*	0.743	3.87***	3.98***	3.97***
	1.4103	1.7257	1.4947	3.464	3.035	3.719
Age squared	-0.012**	-0.013**	-0.013**	-0.039***	-0.037***	-0.040***
	-2.2319	-2.0102	-2.2935	-3.2767	-2.7412	-3.5168
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	46.1	–	–	45.6
N	108,916	108,916	107,621	68,037	68,037	67,210

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	9.7	-35.1***	29.5	27.1	3.0	85.2
	1.01	-3.33	0.29	1.58	0.08	0.62
Severity	42.7***	38.8***	42.6***	84.3***	78.5***	85.2***
	10.20	7.61	10.33	7.91	8.18	8.21
Severity squared	-1.69***	-1.81***	-1.68***	-4.04***	-3.63***	-4.09***
	-3.096	-3.742	-3.041	-4.407	-6.122	-4.560
Male	4.00	-0.59	4.08	8.13	-11.27	7.93
	1.136	-0.153	1.178	1.236	-1.400	1.179
Age	-11.6***	-8.8***	-11.6***	-13.2***	-11.3***	-13.3***
	-6.09	-4.12	-5.97	-4.63	-3.15	-4.82
Age squared	0.090***	0.063***	0.089***	0.086***	0.068*	0.087***
	4.8285	2.9287	4.7483	2.9989	1.8858	3.1339
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	30.2	–	–	31.0
N	65,049	65,049	64,546	44,906	44,906	44,531

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 80: Outpatient cost, high-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-5.33***	-4.58	-46.25**	-7.02**	-4.64	-66.07**
	-2.608	-0.716	-2.521	-2.197	-0.457	-2.461
Severity	-0.173	2.145	-0.234	0.731	6.852*	0.478
	-0.0374	1.2346	-0.0509	0.1096	1.7069	0.0731
Severity squared	0.961	0.190	0.971	1.34	-0.05	1.37
	1.0158	0.5293	1.0292	0.981	-0.064	1.016
Male	-1.39	-0.44	-1.50	-2.33	0.41	-2.26
	-0.819	-0.277	-0.897	-1.039	0.157	-1.068
Age	1.17**	0.99	1.18**	1.22	0.42	1.13
	2.014	1.281	2.103	1.614	0.394	1.489
Age squared	-0.014**	-0.012	-0.015**	-0.016**	-0.007	-0.015**
	-2.3338	-1.4640	-2.4299	-2.1104	-0.6226	-1.9697
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	46.1	–	–	45.6
N	108,916	108,916	107,621	68,037	68,037	67,210

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-8.4	14.7	214.6***	2.1	-13.0	186.8**
	-0.87	1.10	2.73	0.17	-0.97	2.37
Severity	9.03**	2.91	9.54**	22.8***	11.9	23.8***
	2.515	0.373	2.538	4.67	1.05	4.99
Severity squared	0.539	1.087	0.485	0.156	0.794	0.066
	0.7944	0.8598	0.6997	0.1420	0.5205	0.0615
Male	7.28	8.92	7.99	11.3	8.3	11.6
	1.040	1.259	1.124	1.47	1.30	1.54
Age	0.750	-0.092	0.442	-4.46	0.89	-4.76*
	0.3493	-0.0404	0.1946	-1.569	0.167	-1.667
Age squared	-0.014	-0.003	-0.012	0.040	-0.012	0.043
	-0.5967	-0.1327	-0.4763	1.3019	-0.2400	1.3735
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	30.2	–	–	31.0
N	65,049	65,049	64,546	44,906	44,906	44,531

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 81: Pharmacy cost (hypertension), high-access sample

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	23.7	80.6**	416.8***	51.8	76.2	629.7**
	1.41	2.34	3.89	1.44	0.75	2.31
Severity	609.6***	637.2***	608.4***	1135.0***	1179.3***	1136.8***
	18.50	17.25	18.62	26.42	27.82	26.81
Severity squared	-41.9***	-45.8***	-41.6***	-75.5***	-84.8***	-75.6***
	-8.99	-8.30	-9.00	-13.71	-14.30	-13.89
Male	139.7***	140.1***	139.8***	256.5***	244.3***	255.2***
	15.98	12.42	16.12	14.13	10.27	13.56
Age	13.0***	11.7***	12.8***	32.9***	35.3***	33.0***
	4.05	2.98	3.98	5.12	4.41	5.06
Age squared	-0.132***	-0.121***	-0.130***	-0.345***	-0.382***	-0.348***
	-3.4135	-2.6363	-3.3543	-4.8417	-4.5747	-4.7182
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	46.1	–	–	45.6
N	108,916	108,916	107,621	68,037	68,037	67,210

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 82: Total cost, high-access sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	20.8	77.7**	466.9***	52.7	103.3	660.9**
	1.20	2.27	3.30	1.43	1.00	2.18
Severity	615.3***	644.3***	613.9***	1144.8***	1194.5***	1146.0***
	19.87	17.94	19.99	27.89	28.65	28.28
Severity squared	-41.1***	-45.8***	-40.8***	-74.5***	-85.0***	-74.5***
	-9.45	-8.57	-9.45	-13.92	-14.71	-14.07
Male	141.4***	143.5***	141.4***	263.6***	257.1***	262.4***
	17.05	13.31	17.01	13.61	10.50	12.94
Age	14.9***	13.7***	14.8***	38.0***	39.7***	38.1***
	4.02	3.00	3.95	5.36	4.74	5.31
Age squared	-0.158***	-0.145***	-0.157***	-0.401***	-0.426***	-0.404***
	-3.5729	-2.7366	-3.5029	-5.1394	-4.8909	-5.0241
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	46.1	–	–	45.6
N	108,916	108,916	107,621	68,037	68,037	67,210

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	1.3	-20.4	244.1**	29.2	-9.9	272.0*
	0.10	-1.06	2.16	1.53	-0.25	1.86
Severity	51.7***	41.7***	52.1***	107.1***	90.5***	108.9***
	7.74	5.71	7.66	10.11	7.48	10.25
Severity squared	-1.15	-0.72	-1.20	-3.88***	-2.83**	-4.02***
	-1.015	-0.586	-1.028	-3.887	-2.018	-3.876
Male	11.3	8.3	12.1	19.4**	-3.0	19.6**
	1.39	0.92	1.49	2.23	-0.29	2.26
Age	-10.9***	-8.9***	-11.1***	-17.7***	-10.4	-18.0***
	-3.96	-3.07	-4.02	-3.99	-1.52	-4.22
Age squared	0.076**	0.060**	0.077***	0.127***	0.055	0.130***
	2.5626	2.0796	2.5986	2.7178	0.8273	2.8827
Distance to nearest specialist	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
Distance to nearest specialist squared	0.000	0.000	nan	0.000	0.000	nan
	nan	nan	nan	nan	nan	nan
F	–	–	30.2	–	–	31.0
N	65,049	65,049	64,546	44,906	44,906	44,531

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

APPENDIX D : Hospital-affiliated estimates

Table 83: Primary specialist visit, hospital-affiliated

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.010	0.001	0.198**	-0.001	0.017	0.298**
	-0.8319	0.0907	2.3522	-0.0459	0.6185	2.2990
Severity	0.028***	0.027***	0.028***	0.028***	0.030***	0.030***
	6.3422	5.0806	6.1063	6.8209	5.6332	7.0180
Severity squared	-0.002***	-0.002**	-0.002***	-0.001	-0.001**	-0.001
	-2.7740	-2.4908	-2.6340	-1.0704	-1.9884	-1.0637
Male	-0.000	-0.000	0.001	0.009	0.007	0.008
	-0.0578	-0.0083	0.1352	1.2960	0.8614	1.1834
Age	-0.002	-0.002	-0.001	-0.004	-0.005	-0.003
	-1.1047	-0.7975	-0.8062	-1.3594	-1.2153	-1.1054
Age squared	0.000	0.000	0.000	0.000	0.000	0.000
	1.1441	0.9799	0.7559	1.5948	1.5065	1.2626
Distance to nearest specialist	-0.001	-0.001	0.001	-0.001	0.008**	0.001
	-1.4011	-0.7853	1.0720	-1.3318	2.0388	0.9794
Distance to nearest specialist squared	-0.000	0.000*	-0.000	0.000	-0.000	-0.000
	-0.2253	1.8486	-1.6256	0.2656	-1.6299	-1.0464
F	–	–	36.7	–	–	27.3
N	28,118	28,118	27,529	18,160	18,160	17,774

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.049***	-0.014	0.092	-0.008	-0.054*	0.113
	-2.9671	-0.7178	1.0411	-0.3392	-1.7219	0.9342
Severity	0.036***	0.036***	0.035***	0.048***	0.052***	0.051***
	10.3883	5.7108	6.6233	7.9287	6.2403	8.5332
Severity squared	-0.002***	-0.002***	-0.002***	-0.002***	-0.003***	-0.003***
	-5.4014	-2.7263	-2.7870	-3.1614	-3.0334	-3.7459
Male	-0.004*	-0.014***	-0.010*	-0.009	-0.006	-0.009
	-1.6551	-2.5815	-1.9397	-1.4003	-0.6895	-1.3144
Age	-0.004***	-0.006**	-0.007***	-0.006**	-0.004	-0.007**
	-4.4029	-2.1837	-3.5670	-2.4596	-1.3722	-2.4970
Age squared	0.000***	0.000*	0.000***	0.000*	0.000	0.000*
	2.6978	1.8151	3.3224	1.7014	0.6130	1.7543
Distance to nearest specialist	-0.002***	0.000	-0.002***	-0.004***	0.002	-0.003***
	-6.4333	0.0978	-3.3520	-5.2443	0.5049	-3.1289
Distance to nearest specialist squared	0.000**	-0.000	0.000***	0.000***	-0.000	0.000**
	2.4738	-0.4754	2.5903	3.2007	-0.8998	2.5194
F	–	–	56.3	–	–	49.8
N	188,681	25,268	24,723	18,070	18,070	17,666

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 84: Primary specialist RVUs, hospital-affiliated

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.478**	-0.862	-0.878	-0.652***	-0.411	-0.979
	-2.4741	-1.1578	-0.5525	-4.7508	-0.6804	-0.3851
Severity	0.386***	0.468**	0.379***	0.219**	-0.042	0.212**
	3.5084	2.0771	3.2839	2.0561	-0.2198	2.0825
Severity squared	-0.036***	-0.043*	-0.036***	-0.001	0.038	-0.001
	-3.1991	-1.7425	-3.0793	-0.1339	1.3774	-0.0784
Male	0.265**	0.470**	0.262*	0.390***	0.341	0.402***
	1.9678	1.9862	1.9474	2.5958	1.4681	2.6654
Age	-0.028	0.004	-0.030	-0.059	0.014	-0.060
	-0.5843	0.0523	-0.6370	-0.9867	0.1319	-1.0129
Age squared	0.001	0.000	0.001	0.001	0.000	0.001
	1.0318	0.3368	1.1068	1.4599	0.2398	1.5092
Distance to nearest specialist	0.003	0.021	-0.000	0.019	0.134	0.014
	0.2257	0.1707	-0.0230	1.3781	0.9511	0.8711
Distance to nearest specialist squared	0.000	-0.003	0.000	-0.000	-0.005	-0.000
	0.5800	-1.2392	0.8339	-0.6122	-1.3095	-0.6418
F	–	–	9.9	–	–	10.2
N	3,601	3,601	3,551	3,320	3,320	3,265

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.255	-0.577	1.334	0.809	0.722	5.066
	0.8126	-1.0945	0.4046	1.3304	1.5296	1.2035
Severity	0.768***	0.401	0.793***	1.68***	0.88***	1.74***
	3.2934	1.2798	3.9062	7.361	2.878	7.770
Severity squared	-0.012	0.057	-0.015	-0.086***	0.012	-0.091***
	-0.3686	1.0765	-0.5234	-3.4086	0.2961	-3.4025
Male	-0.021	-0.059	-0.027	0.140	0.207	0.158
	-0.1096	-0.2110	-0.1423	0.5589	0.4499	0.6286
Age	-0.412***	-0.335**	-0.404***	-0.336**	-0.664***	-0.325**
	-3.2841	-2.3414	-3.1094	-2.3148	-2.8411	-2.1581
Age squared	0.004***	0.003**	0.004***	0.003*	0.006**	0.003*
	2.9760	2.2849	2.8171	1.9232	2.5003	1.7652
Distance to nearest specialist	-0.027	0.038	-0.019	-0.051***	-0.158	-0.028
	-1.5708	0.3446	-0.7760	-2.6072	-0.6088	-0.9280
Distance to nearest specialist squared	0.000	-0.001	0.000	0.000	0.001	0.000
	0.6627	-1.0043	0.3367	1.3562	0.3605	0.5862
F	–	–	37.1	–	–	30.4
N	5,583	5,583	5,501	5,585	5,585	5,501

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 85: Generalist RVUs, hospital-affiliated

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.189***	0.023	0.900*	-0.153*	0.214	1.855**
	-2.7038	0.2926	1.7769	-1.7133	1.2478	2.1145
Severity	0.006	0.064	0.005	0.019	0.206**	0.022
	0.1808	1.5509	0.1513	0.4473	2.0439	0.5131
Severity squared	0.004	-0.008	0.004	0.000	-0.027*	-0.001
	1.1277	-1.3049	1.1885	0.0219	-1.6697	-0.1436
Male	-0.103***	-0.094***	-0.107***	-0.156***	-0.022	-0.148***
	-3.3294	-2.7303	-3.3860	-3.2082	-0.3146	-2.9462
Age	-0.028*	-0.057***	-0.022	0.029*	-0.033	0.033*
	-1.8703	-3.3339	-1.5080	1.6600	-1.6385	1.8162
Age squared	0.000	0.001***	0.000	-0.000	0.000*	-0.000
	1.6428	3.0680	1.2346	-1.2366	1.9187	-1.4243
Distance to nearest specialist	-0.013***	0.003	-0.005	-0.022**	0.051	-0.011
	-3.0233	0.1696	-0.7406	-2.3882	1.0291	-0.8845
Distance to nearest specialist squared	0.000*	-0.000	0.000	0.000	-0.001	0.000
	1.8164	-0.2839	0.1659	1.6097	-1.4792	0.4695
F	–	–	36.9	–	–	27.5
N	28,106	28,106	27,518	18,152	18,152	17,767

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.044	-0.162	0.442	-0.002	-0.012	0.729
	0.4241	-0.8141	0.9924	-0.0113	-0.0551	1.0082
Severity	0.213***	0.227***	0.220***	0.267***	0.274***	0.277***
	7.0326	4.6081	7.2265	4.4222	3.6537	4.4096
Severity squared	-0.018***	-0.019***	-0.019***	-0.023***	-0.021**	-0.023***
	-5.1296	-3.5267	-5.3432	-3.0496	-2.2264	-2.9988
Male	0.119***	0.138***	0.118***	0.278***	0.254***	0.269***
	2.6490	3.0045	2.6171	4.0138	2.6826	3.7641
Age	0.001	-0.013	0.002	0.065**	0.018	0.065**
	0.0687	-0.5926	0.0966	2.5235	0.5401	2.4861
Age squared	-0.000	0.000	-0.000	-0.001**	-0.000	-0.001**
	-0.0376	0.6904	-0.0925	-2.1201	-0.1681	-2.1094
Distance to nearest specialist	-0.009*	0.002	-0.006	-0.012	0.041	-0.006
	-1.8274	0.1479	-1.2002	-1.2724	1.0596	-0.5471
Distance to nearest specialist squared	0.000	-0.000	0.000	0.000	-0.000	0.000
	1.3079	-1.0353	0.9381	0.3449	-0.6952	0.0223
F	–	–	56.3	–	–	49.7
N	25,267	25,267	24,722	18,067	18,067	17,663

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 86: Any outpatient visit, hospital-affiliated

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.003	-0.010	-0.026	0.010	-0.004	-0.102
	0.3375	-0.5049	-0.6412	0.7777	-0.1946	-1.3840
Severity	0.001	0.003	0.000	0.009***	0.007	0.007**
	0.4091	0.6536	0.1904	2.6994	1.5029	2.3704
Severity squared	0.000	-0.000	0.000	-0.000	-0.001	-0.000
	0.7192	-0.5978	0.7606	-0.8170	-1.0998	-0.6468
Male	-0.005	-0.004	-0.006*	-0.006	-0.007	-0.007
	-1.4520	-1.1320	-1.7129	-1.1860	-1.2688	-1.2941
Age	-0.002	-0.002	-0.002	-0.003**	-0.003	-0.004**
	-1.2911	-1.3547	-1.2913	-2.0189	-1.6280	-2.1414
Age squared	0.000	0.000	0.000	0.000*	0.000	0.000**
	1.0442	1.2150	1.0127	1.8331	1.6056	1.9770
Distance to nearest specialist	0.001**	-0.002	0.001	0.002**	0.004	0.001
	2.0971	-1.2557	1.1482	2.2293	1.3053	1.0404
Distance to nearest specialist squared	-0.000	0.000	-0.000	-0.000	-0.000**	-0.000
	-0.6922	1.1606	-0.0688	-0.9405	-1.9638	-0.1180
F	–	–	36.7	–	–	27.3
N	28,118	28,118	27,529	18,160	18,160	17,774

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.022	-0.059***	-0.077	-0.023	-0.073**	-0.084
	-1.2766	-2.6879	-0.9815	-1.2168	-2.4240	-0.8820
Severity	0.010***	0.010***	0.009***	0.021***	0.018***	0.020***
	3.5078	3.7531	3.2398	4.5994	4.5709	4.3291
Severity squared	-0.001	-0.001**	-0.001	-0.001**	-0.002***	-0.001**
	-1.4259	-2.0114	-1.2847	-2.0290	-2.6738	-2.0468
Male	-0.001	0.001	-0.000	-0.002	0.006	-0.002
	-0.1674	0.1722	-0.1012	-0.3574	0.9521	-0.3066
Age	-0.002	-0.003*	-0.001	-0.004	-0.003	-0.003
	-1.0250	-1.7310	-0.8732	-1.5510	-1.1604	-1.1097
Age squared	0.000	0.000*	0.000	0.000	0.000	0.000
	0.9845	1.7134	0.8691	1.5310	1.1438	1.1735
Distance to nearest specialist	0.003***	0.004	0.002***	0.004***	-0.000	0.003***
	4.8889	1.4737	3.4859	4.6051	-0.1269	3.3909
Distance to nearest specialist squared	-0.000***	-0.000*	-0.000***	-0.000***	0.000	-0.000***
	-3.8279	-1.8298	-2.7724	-3.5529	0.5178	-2.7201
F	–	–	56.3	–	–	49.8
N	25,268	25,268	24,723	18,070	18,070	17,666

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 87: Outpatient RVUs, hospital-affiliated

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.196	0.000	1.900	0.52	18.34**	4.19
	0.3944	nan	0.3831	1.088	2.187	0.832
Severity	0.338**	0.315	0.313*	0.254**	0.540*	0.301*
	2.0045	1.1271	1.7132	2.0846	1.7871	1.7342
Severity squared	-0.021	0.004	-0.018	-0.008	-0.032	-0.013
	-0.8130	0.1169	-0.6904	-0.4922	-1.2606	-0.7950
Male	-0.531***	-0.389	-0.500***	-0.504**	-0.238	-0.427*
	-2.7230	-1.2386	-2.6958	-2.3116	-0.6640	-1.9259
Age	-0.050	-0.355*	-0.027	-0.139	-0.269	-0.100
	-0.4207	-1.6619	-0.1526	-1.0780	-0.9137	-0.5946
Age squared	0.000	0.003*	0.000	0.001	0.002	0.001
	0.3275	1.6551	0.0784	0.9927	0.8958	0.4794
Distance to nearest specialist	-0.087***	-1640.424***	-0.085**	-0.102***	3.833***	-0.095**
	-3.3295	-3.1211	-2.5390	-3.2318	5.4133	-2.5224
Distance to nearest specialist squared	0.001***	59.956***	0.001***	0.001***	-0.081***	0.001***
	2.8359	3.1227	2.7182	2.9479	-5.2605	2.6438
F	–	–	7.4	–	–	9.2
N	1,978	1,978	1,918	1,789	1,789	1,731

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.48	-6.59	5.44	-0.398	0.249	1.090
	-0.863	-1.506	1.073	-0.8868	0.1296	0.2360
Severity	1.02***	1.10	1.18***	1.17***	2.21**	1.25***
	3.237	1.284	3.199	3.676	2.041	3.357
Severity squared	-0.069**	-0.096	-0.085**	-0.075***	-0.146	-0.085**
	-2.3540	-1.1980	-2.4377	-2.6507	-1.5570	-2.5016
Male	1.01*	1.05	1.11**	0.844**	0.964	0.861**
	1.759	1.019	1.968	2.1493	0.8473	2.1244
Age	-0.220	0.087	-0.234	-0.033	1.098	-0.034
	-1.1736	0.6723	-1.2515	-0.2095	1.3034	-0.2077
Age squared	0.003	-0.001	0.003	0.000	-0.011	0.000
	1.2131	-0.6881	1.2997	0.1814	-1.3499	0.1737
Distance to nearest specialist	-0.050	0.215	-0.008	-0.029	0.565***	-0.019
	-1.1577	0.6354	-0.1269	-0.8727	3.4623	-0.4340
Distance to nearest specialist squared	0.001	-0.008	0.000	0.000	-0.009***	0.000
	1.2261	-0.7661	0.3978	0.8067	-3.0126	0.5577
F	–	–	11.2	–	–	16.0
N	3,188	3,188	3,088	3,040	3,040	2,942

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 88: Any new cardiovascular prescription, hospital-affiliated

Panel A: Generalists

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.022**	-0.014	0.189*	0.028**	0.004	0.232
	2.1545	-1.0382	1.7382	1.9872	0.1348	1.5140
Severity	0.240***	0.247***	0.241***	0.213***	0.230***	0.213***
	11.2726	12.8298	11.3348	12.8164	10.6441	12.5663
Severity squared	-0.020***	-0.022***	-0.020***	-0.015***	-0.017***	-0.015***
	-5.3725	-6.2207	-5.4207	-5.5614	-4.8363	-5.5518
Male	0.035***	0.034***	0.036***	0.030***	0.038***	0.031***
	6.9880	5.3212	7.0698	4.8450	4.4222	4.6748
Age	0.007***	0.005**	0.008***	0.008**	0.005	0.009***
	2.8140	2.1719	3.1718	2.3971	1.2611	2.6911
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000**	-0.000***
	-3.9881	-3.6809	-4.3846	-3.3808	-2.2298	-3.6924
Distance to nearest specialist	0.001	-0.002	0.002*	0.000	0.001	0.001
	0.7210	-0.8325	1.7095	0.3589	0.1428	1.2298
Distance to nearest specialist squared	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	-0.5352	-0.0518	-1.3086	-0.3721	-0.0472	-1.2214
F	–	–	36.7	–	–	27.3
N	28,118	28,118	27,529	18,160	18,160	17,774

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Cardiologists

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.016*	0.019	-0.091	-0.010	-0.124***	0.026
	-1.6726	0.9437	-0.9931	-0.8975	-3.5071	0.2283
Severity	0.054***	0.031**	0.054***	0.060***	0.042***	0.061***
	6.1815	2.4837	6.1251	7.9189	3.3292	7.6319
Severity squared	-0.004***	-0.000	-0.004***	-0.003***	-0.000	-0.003***
	-2.9368	-0.1164	-3.0045	-5.4021	-0.0297	-5.2038
Male	0.024***	0.023**	0.025***	0.032***	0.035**	0.032***
	2.7781	1.9743	3.0062	3.8139	2.4148	4.0835
Age	-0.003	-0.001	-0.003	-0.004	-0.001	-0.005*
	-0.9667	-0.1198	-0.8243	-1.4888	-0.0891	-1.6454
Age squared	0.000	0.000	0.000	0.000*	0.000	0.000*
	1.0362	0.2561	0.8958	1.7452	0.3273	1.8869
Distance to nearest specialist	0.002*	-0.006	0.002	0.001	0.006	0.001
	1.8924	-0.5897	1.4343	0.7237	0.4470	0.4673
Distance to nearest specialist squared	-0.000	0.000	-0.000	-0.000	-0.000	-0.000
	-1.4174	0.5690	-1.3648	-0.7110	-0.4753	-0.6791
F	–	–	2.1	–	–	4.5
N	3,722	3,722	3,670	3,431	3,431	3,374

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 89: Inpatient admission, hospital-affiliated

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.001	-0.002	-0.002	-0.003	-0.008	0.009
	-0.6897	-0.6943	-0.2042	-1.2760	-1.5829	0.5044
Severity	0.007***	0.007***	0.007***	0.002	0.007	0.002
	5.8039	4.6982	5.6077	0.7323	1.6442	0.7132
Severity squared	-0.000***	-0.000***	-0.000***	0.001**	0.000	0.001**
	-3.2887	-2.8486	-3.3296	2.2677	0.2184	2.3612
Male	0.002*	0.003	0.002*	0.004***	0.004*	0.005***
	1.6923	1.6279	1.7744	2.9467	1.8002	3.1582
Age	-0.000	-0.000	-0.000	-0.000	-0.001	-0.000
	-0.7809	-1.0772	-0.8462	-0.8561	-0.8742	-0.8740
Age squared	0.000	0.000	0.000	0.000	0.000	0.000
	1.1197	1.3072	1.1741	1.4889	1.0745	1.4356
Distance to nearest specialist	-0.000	-0.000	-0.000	-0.000	-0.001	0.000
	-1.5117	-0.9027	-0.9591	-0.5100	-0.8461	0.0393
Distance to nearest specialist squared	0.000	0.000	0.000	0.000	0.000	0.000
	0.4671	0.8038	0.2528	0.3189	0.7097	0.0498
F	–	–	36.7	–	–	27.3
N	28,118	28,118	27,529	18,160	18,160	17,774

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.003	-0.001	-0.000	0.005	0.009	0.036
	1.3294	-0.1724	-0.0112	1.1382	1.2671	0.9873
Severity	0.012***	0.010***	0.012***	0.016***	0.015***	0.018***
	5.0007	3.4921	4.8655	4.0651	2.8533	4.5303
Severity squared	0.000	-0.000	0.000	0.001	0.000	0.000
	0.0296	-0.1370	0.0754	0.7883	0.3591	0.5563
Male	0.007***	0.008***	0.007***	0.013***	0.010***	0.014***
	4.8582	4.6566	4.9163	5.5155	3.3888	5.8559
Age	-0.001	-0.001	-0.001	0.002*	0.002	0.002
	-1.6073	-1.4933	-1.6396	1.6475	0.9825	1.6430
Age squared	0.000	0.000	0.000	-0.000	-0.000	-0.000
	1.5260	1.4320	1.5684	-1.4940	-0.7481	-1.5050
Distance to nearest specialist	0.000	0.001	0.000	0.000	0.000	0.000
	0.2760	1.5907	0.2303	0.1829	0.1289	0.7094
Distance to nearest specialist squared	-0.000	-0.000**	-0.000	-0.000	0.000	-0.000
	-1.1859	-1.9985	-1.2372	-0.2789	0.1358	-0.5073
F	–	–	56.3	–	–	49.8
N	25,268	25,268	24,723	18,070	18,070	17,666

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 90: Physician cost, hospital-affiliated

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-8.52	7.18	52.14	-9.1	36.4***	131.6
	-1.393	1.242	0.942	-1.09	3.41	1.57
Severity	0.472	2.341	0.563	1.54	5.91*	2.39
	0.2137	0.9275	0.2623	0.604	1.845	0.970
Severity squared	0.112	-0.270	0.129	0.039	-0.453	0.005
	0.4656	-0.9194	0.5485	0.1525	-1.2498	0.0176
Male	-2.38	-0.29	-2.46	-1.08	8.83**	-0.75
	-1.043	-0.118	-1.050	-0.354	2.047	-0.229
Age	0.254	-1.509	0.566	3.92**	1.15	4.33**
	0.2677	-1.3067	0.5820	2.394	0.634	2.536
Age squared	-0.010	0.010	-0.013	-0.045**	-0.014	-0.050***
	-0.9881	0.8305	-1.2939	-2.4234	-0.6698	-2.5770
Distance to nearest specialist	-0.659*	0.014	-0.186	-1.00*	3.33	-0.18
	-1.8129	0.0126	-0.3198	-1.877	1.291	-0.224
Distance to nearest specialist squared	0.008	-0.012	0.004	0.012*	-0.063	0.004
	1.4974	-1.4187	0.4856	1.6855	-1.5024	0.3938
F	–	–	36.7	–	–	27.3
N	28,118	28,118	27,529	18,160	18,160	17,774

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	23.3***	-52.5**	96.8*	46.7***	-85.3*	238.1**
	2.63	-2.36	1.81	3.06	-1.76	2.15
Severity	32.9***	31.3***	34.0***	73.3***	71.5***	76.9***
	5.79	6.02	6.05	8.91	6.78	8.72
Severity squared	-1.97***	-1.96***	-2.06***	-4.48***	-4.57***	-4.72***
	-3.191	-2.595	-3.437	-5.589	-3.997	-5.739
Male	12.5**	9.1*	12.2**	22.8***	24.2***	23.1***
	2.35	1.67	2.45	2.76	2.63	2.72
Age	-10.5***	-11.1**	-10.8***	-9.14*	-17.91**	-9.26*
	-2.70	-2.21	-2.71	-1.772	-2.351	-1.744
Age squared	0.078**	0.088*	0.081**	0.040	0.132*	0.040
	2.0539	1.7632	2.0736	0.7407	1.6917	0.7192
Distance to nearest specialist	-1.36**	1.74	-0.87	-2.71***	2.60	-1.48
	-2.461	0.924	-1.485	-2.992	0.524	-1.408
Distance to nearest specialist squared	0.010*	-0.032	0.007	0.017	-0.025	0.010
	1.8622	-1.5463	1.1615	1.5082	-0.4727	0.7506
F	–	–	56.3	–	–	49.8
N	25,268	25,268	24,723	18,070	18,070	17,666

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 91: Outpatient cost, hospital-affiliated

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-4.14	0.74	-28.65	1.8	15.2	-18.0
	-1.069	0.079	-0.632	0.31	0.97	-0.24
Severity	7.48**	3.12	7.27**	10.9***	10.8*	10.5***
	2.482	0.706	2.437	3.04	1.76	2.92
Severity squared	-0.016	0.711	-0.001	-0.390	0.091	-0.349
	-0.0252	0.7099	-0.0010	-1.1897	0.1036	-1.0887
Male	-5.52***	-7.16**	-5.76***	-8.59***	-17.58***	-8.30***
	-2.648	-2.082	-2.687	-2.898	-4.708	-2.853
Age	2.18**	1.53	2.15**	1.10	1.51	0.96
	2.247	1.283	2.005	0.792	0.795	0.640
Age squared	-0.029***	-0.021	-0.029**	-0.021	-0.024	-0.020
	-2.7332	-1.5842	-2.4225	-1.3443	-1.0806	-1.1572
Distance to nearest specialist	0.168	0.289	-0.101	-0.170	0.291	-0.347
	0.5478	0.2300	-0.2120	-0.3896	0.1441	-0.6301
Distance to nearest specialist squared	0.002	-0.011	0.005	0.006	-0.024	0.008
	0.4634	-0.7080	0.9511	0.9030	-0.7047	1.1492
F	–	–	36.7	–	–	27.3
N	28,118	28,118	27,529	18,160	18,160	17,774

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-7.9	-16.2	-63.3	-15.1	17.1	-129.7
	-0.66	-0.67	-0.65	-0.89	0.78	-0.83
Severity	21.5***	21.9***	21.1***	44.3***	35.5***	43.4***
	3.09	3.35	2.75	3.78	4.08	3.36
Severity squared	-1.14**	-2.04***	-1.09*	-2.59**	-3.27***	-2.53**
	-2.142	-2.646	-1.798	-2.439	-3.285	-2.211
Male	16.4**	16.2*	17.4**	26.0**	24.3**	26.6**
	2.20	1.84	2.29	2.12	2.09	2.12
Age	-2.15	-2.59	-2.15	-2.56	2.96	-2.11
	-0.609	-0.963	-0.632	-0.484	0.640	-0.408
Age squared	0.018	0.025	0.019	0.011	-0.040	0.007
	0.4709	0.8191	0.5015	0.1900	-0.8403	0.1249
Distance to nearest specialist	1.16*	3.33	0.93**	2.39**	0.84	1.88*
	1.915	1.383	2.045	2.349	0.753	1.807
Distance to nearest specialist squared	-0.010	-0.062	-0.008	-0.021*	-0.008	-0.018
	-1.4217	-1.5568	-1.2949	-1.8249	-0.6372	-1.4619
F	–	–	56.3	–	–	49.8
N	25,268	25,268	24,723	18,070	18,070	17,666

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 92: Pharmacy cost (hypertension), hospital-affiliated

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	36.2	-50.8	51.7	111.6	232.7	524.2
	1.63	-0.85	0.25	1.36	1.02	0.96
Severity	615.3***	633.0***	615.1***	1143.0***	1168.1***	1140.2***
	17.42	15.60	17.54	15.38	13.18	15.09
Severity squared	-37.2***	-40.4***	-37.1***	-62.8***	-65.8***	-62.6***
	-7.24	-6.38	-7.23	-9.21	-6.46	-9.04
Male	110.7***	122.2***	110.9***	217.2***	241.5***	220.0***
	6.49	7.21	6.46	5.38	6.18	5.37
Age	17.6***	13.1***	18.4***	27.5**	10.6	30.4***
	4.04	2.69	4.54	2.41	0.63	2.61
Age squared	-0.178***	-0.131**	-0.188***	-0.260**	-0.091	-0.296**
	-3.6421	-2.5034	-4.1317	-2.0243	-0.5001	-2.2542
Distance to nearest specialist	0.79	10.21*	1.07	-2.07	11.07	-0.64
	0.437	1.713	0.513	-0.584	0.622	-0.137
Distance to nearest specialist squared	-0.012	-0.094	-0.021	0.012	0.111	-0.015
	-0.4814	-1.6095	-0.7703	0.2781	0.2820	-0.2849
F	–	–	36.7	–	–	27.3
N	28,118	28,118	27,529	18,160	18,160	17,774

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 93: Total cost, hospital-affiliated

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	23.6	-42.9	75.2	104.3	284.4	637.7
	1.05	-0.75	0.38	1.30	1.26	1.24
Severity	623.3***	638.5***	623.0***	1155.5***	1184.9***	1153.1***
	17.79	16.40	17.97	15.41	13.53	15.13
Severity squared	-37.1***	-39.9***	-37.0***	-63.2***	-66.2***	-63.0***
	-7.53	-6.86	-7.54	-9.25	-6.82	-9.07
Male	102.8***	114.7***	102.7***	207.6***	232.8***	211.0***
	5.79	6.40	5.76	5.01	6.29	5.06
Age	20.0***	13.1**	21.2***	32.5***	13.3	35.7***
	3.78	2.38	4.23	2.75	0.75	2.94
Age squared	-0.217***	-0.142**	-0.230***	-0.326**	-0.128	-0.366***
	-3.6606	-2.3876	-4.0939	-2.4308	-0.6734	-2.6427
Distance to nearest specialist	0.297	10.510*	0.785	-3.24	14.68	-1.17
	0.1571	1.8025	0.3522	-0.939	0.823	-0.261
Distance to nearest specialist squared	-0.002	-0.118**	-0.012	0.029	0.023	-0.004
	-0.0651	-1.9771	-0.4015	0.6988	0.0599	-0.0703
F	–	–	36.7	–	–	27.3
N	28,118	28,118	27,529	18,160	18,160	17,774

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	15.4	-68.8*	33.5	31.6	-68.2	108.4
	1.22	-1.90	0.30	1.47	-1.52	0.56
Severity	54.4***	53.2***	55.1***	117.6***	107.0***	120.2***
	5.34	6.62	5.14	7.55	8.43	7.10
Severity squared	-3.11***	-4.00***	-3.15***	-7.07***	-7.84***	-7.25***
	-4.087	-3.411	-3.964	-5.112	-5.878	-4.923
Male	28.8***	25.3**	29.5***	48.8***	48.4***	49.8***
	2.90	2.18	2.95	3.28	3.67	3.30
Age	-12.6**	-13.7***	-13.0***	-11.7	-14.9*	-11.4
	-2.52	-2.70	-2.61	-1.40	-1.67	-1.34
Age squared	0.096*	0.112**	0.100*	0.051	0.092	0.047
	1.8605	2.1843	1.9432	0.5663	0.9807	0.5167
Distance to nearest specialist	-0.198	5.069*	0.057	-0.318	3.446	0.396
	-0.2680	1.6985	0.0872	-0.2399	0.6756	0.3020
Distance to nearest specialist squared	0.000	-0.094**	-0.001	-0.004	-0.033	-0.008
	0.0484	-2.1742	-0.1214	-0.2579	-0.6055	-0.4741
F	–	–	56.3	–	–	49.8
N	25,268	25,268	24,723	18,070	18,070	17,666

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

APPENDIX E : Not hospital-affiliated sample estimates

Table 94: Primary specialist visit, not-affiliated sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.010	-0.000	0.089	-0.009	0.019	0.039
	-1.1160	-0.0217	1.2158	-0.6883	1.0656	0.3984
Severity	0.019***	0.020***	0.019***	0.033***	0.035***	0.033***
	6.4647	5.7783	6.5443	7.9754	7.7451	8.1223
Severity squared	-0.000	-0.001**	-0.000	-0.001***	-0.002***	-0.001***
	-1.2459	-2.0575	-1.0287	-2.8504	-4.0317	-2.7719
Male	-0.001	0.003	-0.001	-0.000	0.006	0.000
	-0.5968	1.2268	-0.5261	-0.0804	1.5350	0.0309
Age	-0.001	-0.001	-0.001	0.002	0.002	0.002
	-0.9199	-0.8260	-0.9369	1.5392	1.4852	1.5961
Age squared	0.000	0.000	0.000	-0.000	-0.000	-0.000
	1.3552	1.3907	1.3430	-0.7069	-0.7078	-0.7836
Distance to nearest specialist	-0.000	0.001	0.000	-0.000	0.000	-0.000
	-0.5993	0.9144	0.2853	-0.7477	0.2705	-0.2494
Distance to nearest specialist squared	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	-0.4627	-0.6647	-0.8285	-0.3103	-1.5220	-0.4028
F	–	–	33.4	–	–	31.3
N	155,522	155,522	153,185	96,484	96,484	95,029

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.049***	-0.021**	0.642***	-0.046***	-0.036*	0.604**
	-2.9671	-2.1483	2.8675	-2.6052	-1.8674	2.3716
Severity	0.036***	0.037***	0.037***	0.046***	0.046***	0.046***
	10.3883	9.2993	10.0613	10.1884	8.9895	9.6825
Severity squared	-0.002***	-0.002***	-0.002***	-0.002***	-0.003***	-0.002***
	-5.4014	-5.7493	-5.2510	-5.7095	-5.6368	-5.6185
Male	-0.004*	-0.001	-0.003	-0.011***	-0.010***	-0.012***
	-1.6551	-0.3843	-1.0806	-3.4852	-2.6332	-3.2600
Age	-0.004***	-0.004***	-0.005***	-0.003**	-0.003*	-0.003**
	-4.4029	-3.6678	-4.6528	-2.2631	-1.8815	-2.4940
Age squared	0.000***	0.000**	0.000***	0.000	0.000	0.000
	2.6978	2.2151	2.5977	0.6403	0.6865	0.8264
Distance to nearest specialist	-0.002***	-0.000	-0.000	-0.003***	-0.002	-0.001**
	-6.4333	-0.3476	-0.7453	-7.0512	-1.2947	-2.0441
Distance to nearest specialist squared	0.000**	0.000	0.000	0.000***	0.000	0.000
	2.4738	1.1467	0.1012	3.3636	1.5492	1.5259
F	–	–	13.9	–	–	13.2
N	188,681	163,413	160,974	112,147	112,147	110,355

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 95: Primary specialist RVUs, not-affiliated sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.262	-0.922***	1.224	0.354**	-0.764	2.796**
	1.4575	-2.8038	0.7717	2.1743	-1.1092	2.0662
Severity	0.456***	0.464***	0.435***	0.578***	0.498***	0.532***
	5.5922	3.6090	5.2697	5.9783	3.2637	5.3481
Severity squared	-0.027***	-0.034***	-0.025***	-0.032***	-0.025	-0.027***
	-3.9017	-2.8155	-3.3273	-3.6701	-1.4040	-2.9594
Male	0.212***	0.239**	0.219***	0.297***	0.379***	0.307***
	3.0325	2.0933	3.1830	5.1010	4.3983	5.2365
Age	-0.098***	-0.033	-0.093***	-0.082**	-0.069	-0.079**
	-4.0003	-0.8158	-4.0152	-2.5481	-1.4055	-2.5079
Age squared	0.001***	0.000	0.001***	0.001***	0.001	0.001***
	4.2306	0.8548	4.3022	2.7992	1.5980	2.7419
Distance to nearest specialist	-0.019**	-0.063*	-0.015	-0.027***	-0.010	-0.017**
	-1.9634	-1.6838	-1.2622	-2.7554	-0.2866	-2.1477
Distance to nearest specialist squared	0.000***	0.002*	0.000**	0.000***	0.000	0.000*
	2.5786	1.8195	2.0526	2.8097	0.4226	1.9276
F	–	–	54.7	–	–	53.9
N	25,696	25,696	25,363	21,971	21,971	21,680

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.667***	-0.329	5.923	0.969***	0.004	10.908
	2.9642	-0.7157	1.0477	2.7093	0.0066	1.0926
Severity	0.763***	0.547***	0.756***	1.34***	1.20***	1.33***
	8.6145	6.2840	8.5505	12.728	10.051	12.454
Severity squared	-0.014	0.001	-0.014	-0.047***	-0.038***	-0.047***
	-0.9080	0.0503	-0.8656	-3.9429	-2.9353	-3.9794
Male	-0.048	-0.112	-0.030	-0.132	-0.289*	-0.133
	-0.7540	-1.4874	-0.4572	-1.0576	-1.7779	-1.0060
Age	-0.247***	-0.149***	-0.251***	-0.365***	-0.344***	-0.395***
	-8.9262	-2.6822	-8.3445	-7.8237	-4.0218	-6.6702
Age squared	0.002***	0.001**	0.002***	0.003***	0.003***	0.004***
	7.9832	2.3066	7.5291	6.7795	3.5194	5.8468
Distance to nearest specialist	-0.025***	-0.014	-0.020***	-0.037***	-0.068	-0.025*
	-4.3754	-0.4015	-2.7034	-5.0457	-1.0632	-1.6691
Distance to nearest specialist squared	0.000***	0.000	0.000***	0.000***	0.001	0.000*
	3.5068	0.1607	2.9433	3.3750	1.2532	1.7724
F	–	–	9.8	–	–	13.3
N	44,136	44,136	43,635	39,797	39,797	39,274

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 96: Generalist RVUs, not-affiliated sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.542***	0.081	1.490***	-0.675***	0.181	1.804***
	-7.0501	1.0459	2.8300	-6.1656	1.0029	2.7891
Severity	0.070***	0.005	0.064***	0.100***	0.055*	0.094***
	3.4668	0.2449	2.6594	3.2751	1.7737	2.6259
Severity squared	-0.003	0.002	-0.001	-0.006	-0.002	-0.005
	-1.1617	0.9589	-0.5080	-1.6173	-0.4209	-1.0979
Male	0.001	0.015	0.001	0.074**	0.091***	0.070**
	0.0881	0.9744	0.0700	2.2296	2.9771	1.9992
Age	-0.007	0.006	-0.009	0.017	0.032**	0.012
	-1.1758	0.7180	-1.5340	1.2986	2.3498	1.0077
Age squared	0.000	-0.000	0.000	-0.000	-0.000	-0.000
	0.8286	-0.3501	1.1813	-0.6343	-1.3347	-0.3094
Distance to nearest specialist	-0.027***	-0.006	-0.019***	-0.029***	-0.014	-0.018***
	-8.6998	-0.6487	-4.9549	-6.8115	-0.8210	-3.8889
Distance to nearest specialist squared	0.000***	0.000	0.000***	0.000***	-0.000	0.000
	4.7794	0.7144	3.0295	2.6674	-0.1065	1.0839
F	–	–	33.3	–	–	31.2
N	155,361	155,361	153,027	96,392	96,392	94,939

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.642***	-0.060	1.200	-0.808***	0.177	-1.135
	-8.6154	-0.7323	1.3389	-6.0571	0.9191	-0.4742
Severity	0.188***	0.145***	0.185***	0.321***	0.252***	0.322***
	9.3093	8.9128	8.9582	9.1590	9.8431	8.9418
Severity squared	-0.006*	-0.005*	-0.006*	-0.011**	-0.007**	-0.011**
	-1.8047	-1.7393	-1.6785	-2.4732	-2.0671	-2.3723
Male	0.102***	0.118***	0.098***	0.192***	0.207***	0.186***
	3.7342	4.4313	3.5596	4.5738	7.1148	4.5746
Age	0.005	0.013	0.005	0.054**	0.075***	0.056**
	0.3933	1.0200	0.4111	2.3070	2.8134	2.2710
Age squared	-0.000	-0.000	-0.000	-0.000**	-0.001**	-0.000**
	-0.5581	-0.8304	-0.5986	-1.9668	-2.3236	-1.9757
Distance to nearest specialist	-0.012***	-0.012*	-0.008**	-0.013***	-0.035**	-0.013
	-4.6126	-1.8528	-2.3868	-2.6545	-2.1941	-1.5018
Distance to nearest specialist squared	0.000**	0.000*	0.000	0.000	0.000*	0.000
	2.1298	1.8145	1.1930	0.9281	1.9053	0.5236
F	–	–	13.9	–	–	13.2
N	163,386	163,386	160,947	112,131	112,131	110,339

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 97: Any outpatient visit, not-affiliated sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.006**	-0.005	-0.092***	-0.007**	-0.009	-0.134***
	-2.4571	-0.8582	-4.2131	-2.1526	-0.9870	-4.5665
Severity	0.002*	0.002**	0.002**	0.004***	0.005***	0.004***
	1.9049	2.2645	2.1360	3.6707	4.5722	4.2634
Severity squared	0.000	0.000	0.000	0.000	-0.000	0.000
	0.9524	0.5677	0.8001	0.3682	-1.1412	0.0367
Male	-0.001**	-0.000	-0.002**	-0.001	0.000	-0.001
	-2.0572	-0.4909	-2.3066	-1.0729	0.1443	-1.2130
Age	-0.000	0.000	-0.000	-0.000	0.000	-0.000
	-1.3203	0.0158	-1.0023	-0.4534	0.9348	-0.1896
Age squared	0.000	-0.000	0.000	0.000	-0.000	0.000
	1.1162	-0.2311	0.8242	0.3684	-1.0468	0.1314
Distance to nearest specialist	0.001**	0.001	0.000	0.001**	-0.000	0.000
	2.5372	1.5893	0.6030	2.1336	-0.3486	0.5234
Distance to nearest specialist squared	0.000	-0.000	0.000	0.000	0.000**	0.000
	0.4427	-0.0110	1.1232	0.2715	2.0179	0.8010
F	–	–	33.4	–	–	31.3
N	155,522	155,522	153,185	96,484	96,484	95,029

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.002	-0.002	-0.342***	-0.003	-0.012	-0.513***
	-0.3277	-0.2061	-3.2371	-0.3920	-0.9980	-3.8307
Severity	0.008***	0.007***	0.008***	0.015***	0.013***	0.015***
	5.9141	5.9133	4.8117	8.0258	8.8063	6.6137
Severity squared	-0.000*	-0.000**	-0.000	-0.001***	-0.001***	-0.001***
	-1.8356	-2.1239	-1.5417	-3.4758	-4.1155	-2.6755
Male	-0.003***	-0.000	-0.003***	-0.004***	-0.001	-0.004**
	-3.0340	-0.2946	-2.8703	-2.9027	-0.6848	-2.5164
Age	-0.001*	-0.000	-0.001	-0.001	-0.001*	-0.001
	-1.7539	-0.5363	-0.9801	-1.5228	-1.7917	-0.9801
Age squared	0.000	0.000	0.000	0.000	0.000	0.000
	1.3486	0.0559	0.7597	1.2760	1.2946	0.8640
Distance to nearest specialist	0.001***	0.001	0.000	0.002***	0.001	0.001
	5.3210	1.0696	1.2738	5.2098	0.7142	1.1906
Distance to nearest specialist squared	-0.000***	-0.000	-0.000	-0.000***	-0.000	-0.000
	-3.3057	-1.0967	-1.1246	-2.9948	-0.3956	-1.0444
F	–	–	13.9	–	–	13.2
N	163,413	163,413	160,974	112,147	112,147	110,355

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 98: Outpatient RVUs, not-affiliated sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.53**	8.53	4.20*	-0.816***	0.335	3.588*
	-2.297	1.536	1.776	-2.9468	0.1921	1.7667
Severity	0.480**	-0.287	0.300	0.442**	-0.628	0.290
	2.4505	-0.4277	1.5642	1.9938	-0.6694	1.3982
Severity squared	-0.002	0.144	0.029	0.005	0.188	0.033
	-0.0537	1.0079	1.1297	0.1612	0.9707	1.1547
Male	0.017	0.022	0.018	0.136	0.130	0.134
	0.0935	0.0635	0.0936	0.6559	0.4397	0.6703
Age	-0.026	-0.069	-0.033	-0.019	-0.013	-0.049
	-0.3770	-0.5833	-0.4775	-0.2704	-0.0629	-0.6757
Age squared	-0.000	0.001	-0.000	-0.000	0.000	0.000
	-0.0254	0.6266	-0.0120	-0.1193	0.0983	0.2095
Distance to nearest specialist	-0.024	-0.147	-0.006	-0.024	0.369	-0.004
	-1.0990	-0.8409	-0.2230	-1.0162	1.2734	-0.1221
Distance to nearest specialist squared	-0.000	0.003	-0.000	-0.000	-0.006	-0.000
	-0.4550	0.5265	-1.2245	-0.3730	-1.3158	-1.1207
F	–	–	21.3	–	–	20.9
N	4,070	4,070	3,956	3,750	3,750	3,642

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.98**	-1.59	19.47**	-0.75**	1.96	5.94
	-2.569	-1.498	2.044	-2.324	1.097	0.536
Severity	1.06***	0.10	1.19***	0.899***	-0.604	0.937***
	4.753	0.548	4.991	4.0812	-1.0710	4.2584
Severity squared	-0.051***	-0.003	-0.054***	-0.008	0.096	-0.007
	-3.5911	-0.1936	-2.8723	-0.3427	1.5704	-0.3213
Male	0.668**	0.547	0.678**	0.591**	0.678	0.592**
	2.2735	1.2934	2.4526	2.0070	1.5999	2.0526
Age	0.061	0.053	0.059	0.109	0.459	0.119
	0.6146	0.1853	0.5336	1.2423	1.5482	1.2404
Age squared	-0.001	-0.000	-0.001	-0.001	-0.004	-0.001
	-0.6909	-0.0047	-0.6606	-1.2524	-1.3926	-1.2403
Distance to nearest specialist	-0.036*	-0.057	0.003	-0.034**	0.044	-0.019
	-1.7386	-0.2539	0.1451	-2.1464	0.3498	-0.8427
Distance to nearest specialist squared	0.000	0.001	0.000	0.000	-0.002	0.000
	0.9422	0.0977	0.1426	1.1805	-0.5323	0.6987
F	–	–	9.9	–	–	15.4
N	7,848	7,848	7,680	7,749	7,749	7,562

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 99: Any new cardiovascular prescription, not-affiliated sample

Panel A: Generalists

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.002	0.010	0.208*	-0.002	0.012	0.140
	-0.2756	0.7877	1.8414	-0.2879	0.6624	1.0481
Severity	0.252***	0.268***	0.252***	0.267***	0.289***	0.267***
	23.2964	21.3594	22.8108	29.9365	31.1344	29.4768
Severity squared	-0.022***	-0.024***	-0.022***	-0.024***	-0.027***	-0.024***
	-12.7921	-11.8648	-12.4141	-16.7084	-17.5233	-16.0604
Male	0.036***	0.035***	0.036***	0.037***	0.038***	0.037***
	7.9011	6.3004	7.5730	6.5018	4.9824	6.2941
Age	0.008***	0.007***	0.008***	0.006***	0.006***	0.006***
	6.8637	5.5570	6.7279	4.0037	4.1421	3.7876
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	-8.5976	-7.1467	-8.5555	-6.0877	-6.1104	-5.8764
Distance to nearest specialist	0.000	0.004***	0.001**	-0.000	0.001	0.001
	0.2972	2.7832	2.1053	-0.0364	0.4691	1.0172
Distance to nearest specialist squared	-0.000*	-0.000***	-0.000***	-0.000	0.000	-0.000**
	-1.8264	-2.8762	-2.6710	-1.3308	0.0960	-2.1317
F	–	–	33.4	–	–	31.3
N	155,522	155,522	153,185	96,484	96,484	95,029

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Cardiologists

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	0.012	0.007	0.060	0.011*	0.036	-0.022
	1.4841	0.3483	0.8834	1.8368	1.2788	-0.2754
Severity	0.043***	0.039***	0.042***	0.062***	0.064***	0.062***
	7.4413	4.6298	7.2651	9.1239	13.6119	9.2593
Severity squared	-0.003***	-0.003**	-0.002***	-0.004***	-0.005***	-0.004***
	-3.1151	-2.0348	-2.9886	-4.6680	-8.3963	-4.7067
Male	0.014***	0.013***	0.014***	0.019***	0.016**	0.019***
	4.8552	3.3828	5.0090	4.2299	2.5563	4.3233
Age	0.002**	-0.001	0.002**	0.001	-0.001	0.001
	2.1300	-0.3725	2.3360	0.5911	-0.4502	0.6015
Age squared	-0.000**	0.000	-0.000**	-0.000	0.000	-0.000
	-2.1097	0.3917	-2.3434	-0.4809	0.6215	-0.4904
Distance to nearest specialist	-0.001	0.001	-0.001	-0.001	-0.004*	-0.001
	-1.4410	0.2511	-1.1766	-1.4180	-1.6488	-1.2583
Distance to nearest specialist squared	0.000	0.000	0.000	0.000	0.000	0.000
	1.2355	0.4877	1.1815	1.3970	1.4054	0.9352
F	–	–	5.8	–	–	15.6
N	26,324	26,324	25,984	22,493	22,493	22,195

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 100: Inpatient admission, not-affiliated sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.002**	-0.000	0.005	-0.002	0.005*	0.023*
	-2.0028	-0.0124	0.4762	-1.5803	1.8367	1.7279
Severity	0.005***	0.004***	0.005***	0.007***	0.005***	0.007***
	5.1241	5.3554	4.8947	6.2219	3.5641	5.9955
Severity squared	0.000	0.000	0.000	0.000	0.000	0.000
	0.6017	0.4224	0.7064	1.4391	1.0198	1.5768
Male	0.003***	0.003***	0.002***	0.004***	0.003***	0.004***
	4.2107	4.1557	4.1425	4.3014	3.5314	4.2058
Age	-0.000	-0.000**	-0.000	-0.001**	-0.001***	-0.001**
	-1.1698	-1.9653	-1.0723	-2.5060	-3.2193	-2.4840
Age squared	0.000**	0.000***	0.000**	0.000***	0.000***	0.000***
	2.2149	2.8544	2.0828	3.6173	4.3209	3.5233
Distance to nearest specialist	-0.000	0.000	-0.000	-0.000	-0.001	-0.000
	-1.5128	0.2888	-0.9800	-1.4645	-1.3459	-0.1776
Distance to nearest specialist squared	0.000	-0.000	0.000	0.000	0.000	-0.000
	0.9972	-0.8139	0.8562	0.4140	1.1672	-0.2773
F	–	–	33.4	–	–	31.3
N	155,522	155,522	153,185	96,484	96,484	95,029

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-0.003**	-0.005	-0.068***	-0.004	0.006	-0.002
	-2.1844	-0.9958	-2.6646	-1.5708	0.8281	-0.0471
Severity	0.011***	0.010***	0.011***	0.020***	0.016***	0.020***
	13.6843	12.1997	13.1364	14.3077	10.1496	14.6206
Severity squared	0.000	0.000	0.000	0.000	0.000	0.000
	1.3898	0.9921	1.2810	0.5425	1.0152	0.4502
Male	0.007***	0.006***	0.007***	0.010***	0.008***	0.010***
	9.4714	8.5359	9.2784	12.8126	6.0598	13.1024
Age	-0.002***	-0.002***	-0.002***	-0.002***	-0.001***	-0.002***
	-6.7716	-6.2770	-6.4056	-4.6038	-2.7803	-4.6331
Age squared	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	6.8089	6.3365	6.4258	4.7134	3.0226	4.7382
Distance to nearest specialist	-0.000*	-0.000	-0.000***	-0.000**	-0.001*	-0.000
	-1.8536	-0.3081	-3.1705	-2.1502	-1.7966	-1.5293
Distance to nearest specialist squared	0.000	-0.000	0.000*	0.000**	0.000*	0.000
	0.8875	-0.2166	1.6597	2.1709	1.8238	1.5249
F	–	–	13.9	–	–	13.2
N	163,413	163,413	160,974	112,147	112,147	110,355

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 101: Physician cost, not-affiliated sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	6.96*	6.22	71.48	12.4**	24.2*	91.7
	1.833	1.316	1.332	2.43	1.94	1.60
Severity	7.07***	5.36***	6.82***	11.2***	10.3***	10.9***
	5.860	4.209	5.755	6.59	5.15	6.22
Severity squared	-0.323**	-0.309*	-0.273**	-0.510**	-0.364	-0.450*
	-2.5612	-1.7674	-2.1014	-2.1225	-1.3466	-1.7698
Male	2.66**	3.57**	2.65**	8.98***	11.04***	8.78***
	2.309	2.453	2.312	3.750	3.685	3.617
Age	0.597**	1.296***	0.537**	3.41***	3.84***	3.27***
	2.2843	3.4683	2.0720	4.088	4.576	4.007
Age squared	-0.011***	-0.015***	-0.010***	-0.036***	-0.037***	-0.034***
	-3.8607	-4.0843	-3.6898	-4.0681	-4.2664	-3.9872
Distance to nearest specialist	-1.23***	-1.18*	-0.98***	-1.53***	-0.95	-1.22***
	-4.916	-1.651	-3.661	-4.850	-0.759	-3.663
Distance to nearest specialist squared	0.013***	0.024	0.011***	0.017***	-0.003	0.014**
	3.9307	1.3675	2.9888	2.7803	-0.2153	2.4149
F	–	–	33.4	–	–	31.3
N	155,522	155,522	153,185	96,484	96,484	95,029

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	10.9	-13.9	355.1**	22.0	-32.7	378.3
	1.48	-1.30	2.34	1.60	-0.99	1.55
Severity	39.6***	36.2***	39.7***	77.5***	72.3***	78.3***
	12.01	9.91	11.90	9.99	8.45	9.80
Severity squared	-1.54***	-1.46***	-1.54***	-3.41***	-3.01***	-3.46***
	-4.081	-4.605	-4.022	-7.686	-5.220	-7.502
Male	4.31	1.49	4.07	2.11	-7.72	1.29
	1.580	0.463	1.512	0.465	-1.233	0.284
Age	-8.08***	-6.79***	-8.41***	-13.3***	-12.0***	-13.8***
	-4.973	-3.022	-5.251	-5.36	-3.50	-5.74
Age squared	0.056***	0.045**	0.059***	0.089***	0.079**	0.094***
	3.4966	2.0886	3.7461	3.7160	2.4986	4.0239
Distance to nearest specialist	-0.814***	-1.140	-0.141	-1.63***	-2.61	-0.90*
	-3.8551	-1.6130	-0.4154	-3.135	-1.595	-1.705
Distance to nearest specialist squared	0.008***	0.023***	0.005	0.015**	0.044**	0.012**
	2.9966	2.7033	1.5831	2.4754	2.2799	2.3299
F	–	–	13.9	–	–	13.2
N	163,413	163,413	160,974	112,147	112,147	110,355

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 102: Outpatient cost, not-affiliated sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-5.57**	1.44	-47.74***	-8.91**	-2.19	-73.55***
	-1.989	0.369	-2.776	-1.966	-0.378	-3.006
Severity	-1.36	3.73**	-1.57	-3.57	7.36*	-3.89
	-0.301	2.116	-0.356	-0.472	1.805	-0.530
Severity squared	1.34	0.17	1.35	2.41*	0.24	2.43*
	1.581	0.587	1.620	1.706	0.393	1.755
Male	-0.969	-0.696	-1.501	0.205	1.916	-0.705
	-0.7136	-0.5923	-1.1608	0.1106	0.8382	-0.4043
Age	0.174	0.659	0.312	0.483	1.145*	0.641
	0.3492	1.1877	0.7034	0.6967	1.7203	1.0457
Age squared	-0.003	-0.008	-0.005	-0.006	-0.012*	-0.008
	-0.5790	-1.2824	-1.0203	-0.8134	-1.8306	-1.2937
Distance to nearest specialist	0.363*	1.090	0.173	0.281	0.258	-0.005
	1.9165	0.7562	0.9367	0.8604	0.1244	-0.0172
Distance to nearest specialist squared	0.001	0.025	0.000	0.004	0.007	0.003
	0.1408	0.5590	0.0416	0.5263	0.3477	0.6662
F	–	–	33.4	–	–	31.3
N	155,522	155,522	153,185	96,484	96,484	95,029

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	-4.15	3.88	-52.21	0.78	-7.25	-279.73
	-0.580	0.493	-0.404	0.071	-0.792	-0.996
Severity	11.3***	9.1**	11.4***	8.47	3.68	8.35
	3.99	2.22	4.02	0.766	0.256	0.742
Severity squared	0.092	0.407	0.096	2.60	3.05	2.67
	0.2139	0.6745	0.2238	1.265	1.173	1.285
Male	9.06***	8.99***	9.11***	8.80**	5.87	9.16**
	2.773	2.711	2.735	1.964	1.477	1.962
Age	0.570	0.784	0.638	-1.59	-0.55	-1.37
	0.5663	0.4823	0.6252	-1.551	-0.334	-1.346
Age squared	-0.011	-0.010	-0.012	0.012	0.004	0.010
	-1.0337	-0.6359	-1.0792	1.1150	0.2509	0.9241
Distance to nearest specialist	0.214	-0.319	0.120	0.697**	-0.573	0.123
	0.6539	-0.3710	0.2808	2.3671	-0.4095	0.1842
Distance to nearest specialist squared	-0.002	-0.001	-0.001	-0.007**	-0.002	-0.004
	-0.5626	-0.0437	-0.3405	-2.1456	-0.0781	-0.6442
F	–	–	13.9	–	–	13.2
N	163,413	163,413	160,974	112,147	112,147	110,355

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 103: Pharmacy cost (hypertension), not-affiliated sample

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	17.3	37.8	299.0**	20.0	49.9	408.2
	1.41	1.27	2.34	0.79	0.56	1.29
Severity	617.1***	638.0***	616.1***	1150.2***	1193.1***	1151.6***
	23.63	24.40	23.48	28.31	40.37	27.93
Severity squared	-41.5***	-44.3***	-41.2***	-79.7***	-87.0***	-79.3***
	-12.37	-11.90	-12.21	-13.68	-19.79	-13.22
Male	141.9***	140.8***	142.8***	252.7***	236.2***	252.1***
	17.13	14.62	18.13	13.19	9.76	13.35
Age	12.9***	13.6***	12.3***	27.1***	34.8***	26.4***
	4.40	4.27	4.29	4.47	5.59	4.31
Age squared	-0.130***	-0.138***	-0.123***	-0.280***	-0.372***	-0.271***
	-3.7549	-3.6592	-3.6114	-4.0589	-5.4263	-3.8678
Distance to nearest specialist	-2.18*	5.55**	-0.62	-3.63	7.13	-1.04
	-1.807	2.163	-0.478	-1.324	1.284	-0.439
Distance to nearest specialist squared	0.015	-0.104*	0.002	0.003	-0.173*	-0.020
	0.6680	-1.9577	0.1006	0.0825	-1.7066	-0.6132
F	–	–	33.4	–	–	31.3
N	155,522	155,522	153,185	96,484	96,484	95,029

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Table 104: Total cost, not-affiliated sample

Panel A: Hypertension

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	18.7	45.4	322.7**	23.5	71.9	426.3
	1.33	1.38	2.07	0.85	0.78	1.24
Severity	622.9***	647.1***	621.3***	1157.8***	1210.8***	1158.6***
	24.99	25.15	24.72	28.08	40.97	27.59
Severity squared	-40.5***	-44.5***	-40.1***	-77.8***	-87.1***	-77.4***
	-12.44	-12.12	-12.21	-12.53	-20.28	-12.10
Male	143.6***	143.7***	144.0***	261.8***	249.2***	260.2***
	18.94	15.23	19.96	13.97	10.42	13.88
Age	13.6***	15.5***	13.1***	31.0***	39.8***	30.3***
	4.20	4.39	4.14	4.78	6.36	4.60
Age squared	-0.144***	-0.161***	-0.138***	-0.322***	-0.421***	-0.314***
	-3.7814	-3.8437	-3.6889	-4.4192	-6.1428	-4.2270
Distance to nearest specialist	-3.04***	5.47**	-1.42	-4.88*	6.44	-2.27
	-2.635	2.128	-1.173	-1.800	0.997	-1.012
Distance to nearest specialist squared	0.029	-0.055	0.014	0.024	-0.169	-0.003
	1.3821	-0.8890	0.6110	0.6963	-1.6212	-0.0846
F	–	–	33.4	–	–	31.3
N	155,522	155,522	153,185	96,484	96,484	95,029

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

Panel B: Diabetes mellitus

	360-day window			720-day window		
	OLS	OLS FE	IV	OLS	OLS FE	IV
Integrated	6.75	-9.98	302.93**	22.8	-40.0	98.6
	0.681	-0.693	2.510	1.38	-1.08	0.44
Severity	50.9***	45.2***	51.1***	86.0***	76.0***	86.6***
	11.14	9.08	11.12	6.97	4.28	7.06
Severity squared	-1.44**	-1.06	-1.44**	-0.807	0.044	-0.780
	-2.018	-1.445	-2.006	-0.3913	0.0152	-0.3790
Male	13.4***	10.5**	13.2***	10.9*	-1.9	10.4*
	2.93	2.38	2.93	1.72	-0.28	1.71
Age	-7.51***	-6.00*	-7.77***	-14.9***	-12.6***	-15.2***
	-3.421	-1.940	-3.430	-5.74	-3.30	-5.74
Age squared	0.044**	0.035	0.047**	0.101***	0.083**	0.104***
	1.9657	1.1358	2.0098	3.8544	2.3080	3.9164
Distance to nearest specialist	-0.600	-1.459	-0.021	-0.935	-3.181	-0.778
	-1.5211	-1.5132	-0.0538	-1.5948	-1.4266	-0.9849
Distance to nearest specialist squared	0.006	0.022	0.003	0.008	0.043	0.008
	1.4237	1.4766	0.8359	1.1219	1.3388	1.0320
F	–	–	13.9	–	–	13.2
N	163,413	163,413	160,974	112,147	112,147	110,355

Notes: Standard errors are clustered at the state level. Significance stars have their usual meaning (90, 95, and 99 percent confidence). The F statistic reported for the IV regression is the Kleibergen-Paap Wald rk statistic. Low values for this statistic indicate indicate the instrument is weak. I note in the text when values are dangerously low. State and year fixed effects are included in all regressions. Firm and area regressors are also included in all regressions, but due to lack of significance across nearly all specifications I do not report their values.

APPENDIX F : SK&A imputation

Introduction

Before the SK&A data can be used in my analysis, physician and firm identifiers must be imputed. I describe my procedure at a high level.

Physician identifier imputation

No single variable identifies physicians across all years in the SK&A Physician List. The List is compiled annually for marketing purposes and longitudinal identifiers do not have high value to the data set's primary users. Unfortunately, longitudinal identifiers are important for my research.

My strategy was to fill the NPI field for all physicians in my data. In later years of the data set, NPI is filled at almost a 90 percent rate. In earlier years, missing values are significant, on the order of 40 percent. However, alternative government identifiers, particularly UPIN, state license identifiers, and DEA registration numbers are available for most physicians in the data.

Therefore, to impute NPIs I begin by determining which identifiers are observed coincidentally in any year of the data and link them together. For example, if physician A is observed with NPI B in 2011 and 2012, UPIN C in 2010 and 2011, and DEA registration number D in 2009 and 2010, I draw an undirected graph with nodes physician A, NPI B, UPIN C, and DEA registration number D, and edges between A and B, B and C, and finally C and D. This clique indicates that NPI B may be associated with any observation with UPIN C or DEA registration number D. In addition, I exploit physician identifier crosswalks provided by the NBER that link UPIN and state licenses to NPIs. (These crosswalks in turn were generated from NPPES files.)

This procedure can be implemented with all rows of the data. However, a large number of observations, about 10 percent of rows of the stacked Physician List extracts, are linked to more than one NPI. This is due in large part to the use of UPINs by multiple providers. For observations linked with more than one NPI, I search for a name and state match between the SK&A data and the corresponding NPPES records. If I can find a unique exact match, I assign an NPI. If I cannot, I assign an NPI if I can match on state and the fuzzy string match score for name exceeds X. The threshold is arbitrary, set by examining the distribution of fuzzy match scores, trying different values, and examining the results for a random sample of observations.

Firm identifier imputation

The variable CODE3 identifies firms in the SK&A data. CODE3 indicates which locations in the data belong to the same firm. Offices with the same non-empty value belong to the same firm. Offices without a value might truly be stand-alone locations. The value may also be empty because the SK&A caller was not able to elicit the relevant information. Because firm composition is essential to my research, this missing value problem can cause significant bias.

The Physician List is highly valued by marketers because the contact information is verified by phone by SK&A. My imputation strategy relies on the high quality of this contact information. I also observe entity NPIs and contact information in the NPPES. Because the SK&A data is of high quality, any match on address or phone between entities in the SK&A and NPPES data files is potentially credible.

Therefore, I build another large undirected graph in which SK&A and NPPES entity NPIs, addresses, and phone numbers are nodes and links detected between them generate edges. Offices in the same clique likely belong to the same firm.

Similar to the NPI imputation, there are a significant amount of false positives. By manual inspection I discovered that unassociated groups with hospital-based locations or phone numbers were often linked incorrectly. Therefore in my imputation I only drew an edge between offices in the NPPES data if they could also be linked on authorized individual, a relatively severe requirement. When I imposed this condition in my estimation, the false positives I observed in my random sample disappeared.

BIBLIOGRAPHY

C. C. Afendulis and D. P. Kessler. Tradeoffs from integrating diagnosis and treatment in markets for health care. *The American Economic Review*, 97(3):1013–1020, 2007.

C. C. Afendulis and D. P. Kessler. Vertical integration and optimal reimbursement policy. *International journal of health care finance and economics*, 11(3):165–179, 2011.

American Medical Association. Evaluating and negotiating emerging payment options. Accessed on June 25, 2013 at `http://www.ama-assn.org/resources/doc/psa/payment-options.pdf\#page=134`, 2012.

A.-M. J. Audet and S. Sinha. To the point: The care coordination imperative: Responding to the needs of people with chronic diseases. Technical report, Commonwealth Fund, 2012.

J. R. Barro, R. S. Huckman, and D. P. Kessler. The effects of cardiac specialty hospitals on the cost and quality of medical care. *Journal of Health Economics*, 25(4):702–721, 2006.

G. S. Becker and K. M. Murphy. The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics*, 107(4):1137–1160, 1992.

R. D. Blair and J. B. Herndon. Physician cooperative bargaining ventures: An economic analysis. *Antitrust Law Journal*, 71(3):989–1016, 2004.

N. Bloom, R. Sadun, and J. Van Reenen. Does management really work? *Harvard Business Review*, 90(11):76–82, 2012.

G. Bordonaro. Doctors press legislature to allow co-op bargaining. *HartfordBusiness.com*, March 2013. Accessed June 25, 2013 at `http://www.hartfordbusiness.com/article/20130318/PRINTEDITION/303149951/doctors-press-legislature-to-allow-co-op-bargaining`.

L. R. Burns and R. W. Muller. Hospital-physician collaboration: Landscape of economic integration and impact on clinical integration. *Milbank Quarterly*, 86(3):375–434, 2008.

L. R. Burns and M. V. Pauly. Integrated delivery networks: a detour on the road to integrated health care? *Health affairs*, 21(4):128–143, 2002.

L. P. Casalino, K. J. Devers, T. K. Lake, M. Reed, and J. J. Stoddard. Benefits of and barriers to large medical group practice in the united states. *Archives of Internal Medicine*, 163(16):1958–1964, 2003.

A. Chandra and D. O. Staiger. Productivity spillovers in healthcare: evidence from the treatment of heart attacks. *The journal of political economy*, 115:103, 2007.

N. K. Choudhry, R. H. Fletcher, and S. B. Soumerai. Systematic review: The relationship between clinical experience and quality of health care. *Annals of Internal Medicine*, 142(4):260–273, 2005. doi: 10.7326/0003-4819-142-4-200502150-00008. URL +http://dx.doi.org/10.7326/0003-4819-142-4-200502150-00008.

F. Ciliberto and D. Dranove. The effect of physician–hospital affiliations on hospital prices in california. *Journal of Health Economics*, 25(1):29–38, 2006.

G. Claxton, M. Rae, M. Long, N. Panchal, A. Damico, K. Kenward, and H. Whitmore. Employer health benefits: 2015 annual survey. Technical report, The Henry J. Kaiser Family Foundation and Health Research & Educational Trust, 2015.

A. E. Cuellar and P. J. Gertler. Strategic integration of hospitals and physicians. *Journal of health economics*, 25(1):1, 2006.

D. M. Cutler. The lifetime costs and benefits of medical technology. *Journal of Health Economics*, 26(6):1081–1100, December 2007.

L. S. Dafny. Are health insurance markets competitive? *The American Economic Review*, 100(4):1399–1431, 2010.

G. David and L. A. Helmchen. The role of task adhesion in limiting specialization along the medical care continuum. *Labour*, 25(1):24–44, 2011.

G. David and M. D. Neuman. Physician division of labor and patient selection for outpatient procedures. *Journal of Health Economics*, 30(2):381–391, 2011.

G. David, E. Rawley, and D. Polsky. Temporal specificity and task alignment: Evidence from patient care. 2011.

L. C. DeFelice and W. D. Bradford. Relative inefficiencies in production between solo and group practice physicians. *Health Economics*, 6(5):455–465, 1997.

M. T. Donohoe. Comparing generalist and specialty care: discrepancies, deficiencies, and excesses. *Archives of Internal Medicine*, 158(15):1596, 1998.

A. J. Epstein, J. D. Ketcham, and S. Nicholson. Specialization and matching in professional services firms. *The RAND Journal of Economics*, 41(4):811–834, 2010.

R. G. Evans. Supplier-induced demand: Some empirical evidence and implications. In M. Perlman, editor, *The economics of health and medical care*, pages 162–173. London, Macmillan, 1974.

C. B. Forrest, A. Majeed, J. P. Weiner, K. Carroll, and A. B. Bindman. Comparison of specialty referral rates in the united kingdom and the united states: retrospective cohort analysis. *BMJ: British Medical Journal*, 325(7360):370, 2002.

C. B. Forrest, P. A. Nutting, S. von Schrader, C. Rohde, and B. Starfield. Primary care physician specialty referral decision making: patient, physician, and health care system determinants. *Medical decision making*, 26(1):76–85, 2006.

E. Gal-Or. The profitability of vertical mergers between hospitals and physician practices. *Journal of Health Economics*, 18(5):623–654, 1999.

L. Garicano and T. N. Hubbard. Managerial leverage is limited by the extent of the market: Hierarchies, specialization, and the utilization of lawyers human capital. *Journal of Law and Economics*, 50(1):1–43, 2007.

L. Garicano and T. N. Hubbard. Specialization, firms, and markets: The division of labor within and between law firms. *Journal of law, economics, and organization*, 25(2):339–371, 2009.

L. Garicano and T. Santos. Referrals. *American Economic Review*, 94(3):499–525, 2004.

M. Gaynor and P. Gertler. Moral hazard and risk spreading in partnerships. *The RAND Journal of Economics*, pages 591–613, 1995.

M. Gaynor and M. V. Pauly. Compensation and productive efficiency in partnerships: Evidence from medical groups practice. *Journal of Political Economy*, pages 544–573, 1990.

S. Greenfield, W. Rogers, M. Mangotich, M. F. Carney, and A. R. Tarlov. Outcomes of patients with hypertension and noninsulin-dependent diabetes mellitus treated by different systems and specialties: Results from the medical outcomes study. *Jama*, 274(18): 1436–1444, 1995.

L. R. Harrold, T. S. Field, and J. H. Gurwitz. Knowledge, patterns of care, and outcomes of care for generalists and specialists. *Journal of General Internal Medicine*, 14(8):499–511, 1999.

B. J. Hillman, C. A. Joseph, M. R. Mabry, J. H. Sunshine, S. D. Kennedy, and M. Noether. Frequency and costs of diagnostic imaging in office practicea comparison of self-referring and radiologist-referring physicians. *New England Journal of Medicine*, 323(23):1604–1608, 1990.

HMSA Provider Resource Center. Diagnosis coding. Web page: `https://hmsa.com/portal/PROVIDER/zav_pel.aa.DIA.300.htm`, HMSA Provider Resource Center, 2014.

K. Ho. Insurer-provider networks in the medical care market. *The American Economic Review*, pages 393–430, 2009.

B. Holmstrom. Moral hazard in teams. *The Bell Journal of Economics*, pages 324–340, 1982.

R. S. Huckman. Hospital integration and vertical consolidation: An analysis of acquisitions in new york state. *Journal of Health Economics*, 25(1):58–80, 2006.

E. M. Johnson. Ability, learning, and the career path of cardiac specialists. Mimeo. Available at `http://www.mit.edu/~erinmj/files/Draft_042011.pdf`, 2011.

D. P. Kessler and M. B. McClellan. Is hospital competition socially wasteful? *The Quarterly Journal of Economics*, 115(2):577–615, 2000.

J. T. Kolstad and M. E. Chernew. Quality and consumer decision making in the market for health insurance and health care services. *Medical Care Research and Review*, 66(1 suppl):28S–52S, 2009.

F. Lafontaine and M. Slade. Vertical integration and firm boundaries: The evidence. *Journal of Economic Literature*, pages 629–685, 2007.

A. Liebhaber and J. M. Grossman. Physicians moving to mid-sized, single-specialty practices. Center for Studying Health System Change Tracking Report No. 18, available at `http://www.hschange.com/CONTENT/941/`, 2007.

K. Madison. Hospital–physician affiliations and patient treatments, expenditures, and outcomes. *Health services research*, 39(2):257–278, 2004.

M. McClellan, B. J. McNeil, and J. P. Newhouse. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality?: analysis using instrumental variables. *Jama*, 272(11):859–866, 1994.

K. M. McDonald, E. Schultz, L. Albin, N. Pineda, J. Lonhart, V. Sundaram, C. Smith-Spangler, J. Brustrom, E. Malcolm, L. Rohn, and S. Davies. Care coordination measures atlas, 2014. Technical report, Agency for Healthcare Research and Quality, 2014.

T. G. McGuire. Physician agency. In A. J. Culyer and J. P. Newhouse, editors, *Handbook of Health Economics, Volume 1*, pages 461–536. 2000.

J. M. Mitchell and T. R. Sass. Physician ownership of ancillary services: indirect demand inducement or quality assurance? *Journal of Health Economics*, 14(3):263–289, 1995.

S. Nakamura, C. Capps, and D. Dranove. Patient admission patterns and acquisitions of "feeder" hospitals. *Journal of Economics & Management Strategy*, 16(4):995–1030, 2007.

A. Navathe and G. David. The formation of peer reputation among physicians and its effect on technology adoption. *Journal of Human Capital*, 3(4):289–322, 2009.

J. P. Newhouse. The economics of group practice. *Journal of Human Resources*, pages 37–56, 1973.

J. B. Rebitzer and L. J. Taylor. When knowledge is an asset: Explaining the organizational structure of large law firms. *Journal of Labor Economics*, 25(2):201–229, 2007.

J. B. Rebitzer and M. E. Votruba. Organizational economics and physician practices. Technical report, National Bureau of Economic Research, 2011.

U. E. Reinhardt, M. V. Pauly, and P. J. Held. Financial incentives and the economic performance of group practitioners. Analysis of Economic Performance in Medical Group Practices, Project Report 79-05, 1979.

S. Rosen. Specialization and human capital. *Journal of Labor Economics*, pages 43–49, 1983.

R. Rosenman and D. Friesner. Scope and scale inefficiencies in physician practices. *Health Economics*, 13(11):1091–1116, 2004.

P. Starr. *The social transformation of American medicine.* Basic Books, 1982.

G. J. Stigler. The division of labor is limited by the extent of the market. *The Journal of Political Economy*, 59(3):185–193, 1951.

The Johns Hopkins ACG System. Describing morbidity burden. Web page: `http://acg.jhsph.org/index.php?option=com\_content\&view=article\&id=55\&Itemid=149`, last accessed august 1, 2016, The Johns Hopkins ACG System, 2016.

R. Town and G. Vistnes. Hospital competition in hmo networks. *Journal of health economics*, 20(5):733–753, 2001.

United States Department of Agriculture, Economic Research Service. Documentation: 2010 rural-urban commuting area (ruca) codes. Web page: `http://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/documentation.aspx`, United States Department of Agriculture, Economic Research Service, 2016.