Better Late than Never? Physician Response to Schedule Disruptions

Hannah T. Neprash*

November 15, 2016

Abstract

Many physicians face increasing stress to see more patients in the same or less time. This leads to crowded appointment schedules and increased schedule disruptions. I examine how physicians respond to schedule disruptions, instrumenting for appointment start time with the office arrival time of the physician's previous patient. I use novel data from athenahealth, Inc., a national provider of electronic health records, medical billing, and practice management services. I find that when primary care physicians fall behind schedule, they truncate appointment duration, perform fewer in-office procedures, and record fewer diagnoses. The likelihood of a patient revisiting the primary care practice within two weeks significantly increases as a function of delayed appointment start time. Physician ordering behavior also responds to a schedule disruption. In particular, physicians who run behind schedule increase antibiotic and opioid painkiller prescribing and increase referrals of a new patient to a specialist. For patients with preexisting prescription drug regimens, physicians running behind schedule are less likely to change the existing course of treatment. These findings suggest possible unintended consequences of the increasing time pressures placed on physicians by policymakers and private payers. Implications may include higher health care spending and lower quality care.

* Correspondence: Department of Health Policy, Harvard University, Cambridge, MA 02138. Tel.: (651) 587-7250. Email: hneprash@fas.harvard.edu. Web: http://scholar.harvard.edu/hannahneprash. I am indebted to my committee: Michael Chernew, David Cutler, and Michael McWilliams for input at all stages of this project. I also thank Michael Barnett, Caitlin Carroll, David Chan, Amitabh Chandra, Katherine Donato, Bapu Jena, Tim Layton, Nicole Maestas, Tom McGuire, Ateev Mehrotra, Ellen Montz, Joe Newhouse, Daria Pelech, Alan Rozenshtein, Adam Sacarny, Zirui Song, Teddy Svoronos, Jacob Wallace, and seminar participants for helpful comments and suggestions. I also thank athenahealth, Inc., particularly the research team, for their generous assistance in obtaining the data. I gratefully acknowledge fellowship support through the Agency for Healthcare Research and Quality (AHRQ) trainee program (T32), the National Science Foundation Graduate Research Fellowship, and the AHRQ (IR36H202445-01) Dissertation Grant. All errors are my own.

1 Introduction

Primary care physicians' time has become an increasingly scarce resource, as they are pushed to see more patients, comply with more complex documentation and quality reporting requirements, monitor a greater number of medications and chronic conditions, and follow more screening and preventive service recommendations. This pressure is increasingly felt in primary care. Primary care physicians are expected to attend to a range of acute and chronic medical and psychosocial issues, provide preventive care, coordinate care with specialists, and encourage informed decision-making that respects patients' preferences [Fiscella and Epstein, 2008].

Primary care physicians generally organize their time by keeping appointment schedules that dictate whom they see, when, and for how long [Brandenburg et al., 2015, Konrad et al., 2010]. Although survey estimates suggest that average visit duration has increased over the past two decades [Shaw et al., 2014], increased visit complexity and added documentation burdens may have led to the perceived increased time pressures on physicians discussed in the popular press [Brownlee, 2012, Rabin, 2014a]. In an era of more compressed visits, keeping to schedule becomes very important.

A tight timeframe for primary care visits raises issues of quality and cost. On the quality side, the concern is that shorter visit durations will lead to more missed problems, less recommended care, and less adherence to chronic care needs. On the cost dimension, the concern is that harried physicians will schedule expensive referrals or diagnostic tests to minimize appointment duration, leading to greater total spending.

In this paper, I study how appointment schedule disruptions affect the input choices and decision-making of primary care physicians. I ask whether physicians respond to a schedule disruption by spending less time with patients or changing the inputs they provide, and what implications these changes have for health care spending and quality. To answer these questions, I use claims and electronic health record data from athenahealth, Inc., a national provider of electronic health records, medical billing, and practice management services. With this novel dataset, I am able to observe intended and realized appointment timing, in addition to detailed insurance claims and orders placed by the physician for follow-up care.

In the first step of my analysis, I develop a model of physician disutility arising from appointment schedule disruptions. The primary insight of this model is that the shadow price of time is increasing in the difference between the observed and scheduled appointment start time. In response to a schedule disruption, the physician cuts back on total time spent with the patient, other time-costly inputs, and the number of complaints or conditions (e.g., a persistent cough or hypertension) addressed. Less time spent with the patient decreases the probability of accurately observing the patient's true health state, which in turn affects ordering behavior for follow-up care. The model predicts that a schedule disruption may increase orders for follow-up care regarding new or acute conditions, while potentially decreasing orders for follow-up care that modify an existing treatment course for an established condition.

The remainder of my analysis applies this conceptual model to more than one million primary care physician appointments in office-based settings. Consistent with concerns about being rushed, appointments start almost half an hour late on average, however much of this may be endogenous. My empirical strategy employs an instrumental variables framework to circumvent the endogeneity of appointment start time. Specifically, I use the physician's previous patient's office arrival time as an instrument for their current appointment start time. I find that the office arrival time of the previous patient is highly predictive of the start time of the current appointment. A previous patient who arrives 15 minutes late to her appointment delays the physician by 2 additional minutes.

I find that physicians respond to schedule disruptions by significantly shortening appointment duration; the 2 minute delay in appointment start time caused by a 15 minute late previous patient results in a roughly half minute shorter appointment. During these shortened appointments, physicians bill significantly fewer procedures and record fewer diagnoses, saving appointment documentation for after-hours. A delayed appointment start increases the likelihood that the patient will revisit the same physician within two weeks - possibly due to worsening symptoms or at the urging of the physician, who did not have time to adequately address care needs during the initial appointment.

A schedule disruption also affects physician decision-making regarding follow-up care. For new patients, I find that a 2 minute delay in appointment start increases the likelihood a physician will refer the patient to a specialist by one percentage point, or 4% relative to a base rate referral probability of 13%. I focus on two frequently scrutinized decisions: antibiotic and opioid painkiller prescribing. For both, I find that the likelihood of receiving a prescription increases with a delayed appointment start. A 2 minute delay in appointment start time results in a 2 percentage point increase in antibiotic prescribing, or 3% relative to the base rate of 58% among patients with upper respiratory infections. For patients with a new diagnosis of spinal disorder, arthropathy, or rheumatism, the same 2 minute delayed appointment start increases the likelihood of an opioid painkiller prescription by 0.2 percentage points, or 3% relative to the base rate of 7%. I also find some evidence that physicians are less likely to alter an existing course of treatment as they run increasingly behind schedule; a 2 minute delay in appointment start time resulting in a 0.1 percentage point decrease in the likelihood of a modification to an existing prescription, or 1% relative to the base rate of 9%. Overall, the results suggest that quality of care may suffer when physicians are unexpectedly behind schedule.

This paper contributes to several strands of literature. First, this paper relates to the literature on productivity within health care. In particular, it introduces visit duration to the study of primary care physician productivity. The literature concerning non-health worker productivity frequently uses throughput time, but this metric is rarely used to examine physician productivity [Mas and Moretti, 2009, Shunko et al., 2015, Coviello et al., 2010]. Recent studies have begun to measure physician productivity using throughput time, but exclusively in the emergency department setting [Chan, 2015b, Silver, 2016, Song et al., 2016]. Most existing research generally quantifies physician productivity using metrics calculated with claims or survey data. These include visit or patient count, payment, and service intensity [Medicare Payment Advisory Commission, 2016, Zismer et al., 2015]. These metrics are unlikely to capture all aspects of physician productivity. I build on traditional productivity measurement in health care by using a rich dataset containing claims and time-stamped electronic health record information for primary care visits.

Second, this paper contributes to the literature on physician labor supply. In health care, appointment schedules structure physician labor supply. Instead of a real-time optimization between labor and leisure, physicians prespecify - frequently a month or more in advance -

whom they see, when, and for how long. Limited existing research suggests that appointment schedules matter a great deal, with evidence pointing to a behavioral norm of equalizing time across patients and "slacking off" as a physician nears the end of a work day [Tai-Seale and McGuire, 2011, Chan, 2015a]. I build on this research by examining intended and realized labor supply decisions in the context of primary care physicians' appointment schedules.

Third and finally, this paper joins the literature on variation in health care utilization. Considerable evidence demonstrates the influence of physicians on health care consumption, both theoretically and empirically [Arrow, 1963, Cutler et al., 2013, Finkelstein et al., 2016, Molitor, 2016]. Policymakers and public health experts express concern about certain problematic utilization patterns, including the overuse of antibiotics and the ongoing opioid epidemic [Centers for Disease Control and Prevention, 2016, White House, 2015, US Department of Health and Human Services, 2016]. In this paper, I document another factor that may contribute to these concerning trends and to variation in health care expenditures in general: schedule disruptions and time pressure.

The paper proceeds as follows. Section 2 discusses the background and institutional setting. In Section 3, I present a conceptual framework linking physician disutility of appointment schedule disruptions to input use and ordering behavior. Section 4 discusses data and sample selection. Section 5 details my empirical strategy. Section 6 presents results and robustness checks. Section 7 concludes.

2 Background and Institutional Setting

In this section, I describe the institutional setting of office-based primary care, including the role of primary care physicians in the U.S. health care system, the nature of primary care appointments, and the supply of primary care physicians.

2.1 The Role of Primary Care Physicians

Every year, more than four out of five of adults in the United States visit a health care professional [National Center for Health Statistics, 2014]. The vast majority of these visits occur in an office setting, with an estimated 929 million physician office visits in 2012. These physician visits and clinical services account for 19.9% of national health care expenditures, or $597 billion [National Center for Health Statistics, 2016]. Of the nearly one billion physician office visits, more than half (54.6%) were to a primary care physician [National Center for Health Statistics, 2015].

Primary care is considered a cognitive rather than a procedural specialty, and primary care physicians are expected to perform a wide range of functions. Historically, these functions included serving as a patient's point of first contact with the health care system, maintaining the continuity of care, providing comprehensive care, and coordinating with other health care providers [Starfield, 1992]. These tasks have grown increasingly complicated, due to the growing burden of chronic conditions, a greater number of medications to monitor, and a higher volume of screening and preventive recommendations. The scope of primary care work has also widened as documentation and reporting requirements expand and quality reporting and EHR adoption increasingly affect reimbursement.

2.2 Primary Care Appointment Duration and Composition

The average primary care appointment includes 21-24 minutes of interaction with the patient [National Center for Health Statistics, 2015]. Increasingly short appointments have received much attention in the popular press, as patients complain about feeling rushed with their physician [Brownlee, 2012, Chen, 2013, Rabin, 2014a, Rabin, 2014b]. However, survey estimates suggest that primary care appointment duration has increased over the past two decades, from an average of 17.9 minutes in 1993 to 22.6 minutes in 2012 [Shaw et al., 2014, National Center for Health Statistics, 2015].

Increased documentation time is one possible explanation for the discrepancy between the perceived and reported trend in appointment duration, if surveyed physicians include this in reported visit length. Estimates of the time spent on documentation (increasingly using an electronic health record) range from a quarter to two thirds of physicians' total in-office time and rising [Sinsky et al., 2016, Clynch and Kellett, 2015, Oxentenko et al., 2010]. In addition to clinical documentation, physicians spend time complying with quality reporting requirements imposed by private and public payers. A study of two dozen health insurers found that more than 500 quality measures in use, few of which matched across insurers or with the more than three times the number of measures used by federal agencies [Higgins et al., 2013, Blumenthal et al., 2015]. Estimates suggest that primary care practices devote more time and staff resources to quality reporting than any other specialty, spending an estimated 19.1 hours per week, at an annual cost of $50,468 in wages [Casalino et al., 2016].

Increasing visit complexity is another possible explanation that would explain the differently perceived trend in visit length. The average primary care appointment covers six "topics" or conditions (e.g., chronic hypertension or a persistent cough) during the visit. Videotape analysis reveals that topics do not receive equal time. The longest topic receives roughly five minutes, while remaining topics receive an average of 1.1 minutes [Tai-Seale et al., 2007]. The number of topics covered during the average primary care visit grew more than 30% between 1997 and 2005, while visit duration grew less than 10% - resulting in a decrease of time per topic [Abbo et al., 2008].

Depending on the topics that surface during a primary care visit, the physician may recommend follow-up care. More than two thirds of office visits result in instructions to return at a specified time, while 8% include a referral to another physician, and 0.5% include a referral to the emergency room or admission to the hospital [National Center for Health Statistics, 2015]. Other types of recommended follow-up care include prescription medications, lab and imaging tests, non-medication treatment (e.g., physical therapy, wound care), and health education (e.g., nutrition, stress management, tobacco cessation), but statistics on the frequency of these orders are not reported.[1]

[1] The National Ambulatory Medicare Care Survey (NAMCS) tracks these categories, but does not separate services provided during the appointment from services ordered as follow-up.

2.3 Supply of Primary Care Physicians

There are roughly 305,000 primary care physicians practicing in the United States today, representing one third of all active physicians [Kaiser Family Foundation, 2016].[2] This number has been relatively steady in recent years, with fewer than one in ten domestic medical school graduates entering a primary care residency program [Bodenheimer and Pham, 2010, National Resident Matching Program, 2016]. Policymakers and researchers express increasing concern about the supply of primary care physicians within the next decade, with estimates ranging from a shortage of 6,400 to 52,000 PCPs [Health Resources and Services Administration, 2013, IHS, 2015, Petterson et al., 2012, Bodenheimer and Bauer, 2016].

Two main factors motivate concern about primary care physician supply. First, primary care is relatively unattractive relative to other specialties. Primary care physicians have relatively low payment rates and incomes, with an average primary care salary of $195,000, or 69% of the $284,000 that specialists receive [Peckham, 2015]. Primary care physicians also have lower job satisfaction than their specialist peers, with more than half reporting dissatisfaction with their work-life balance [Shanafelt et al., 2012]. This is reflected in a higher and increasing rate of burnout, which poses the risk that the annual number of retiring primary care physicians will soon exceed the number of new entrants [Shanafelt et al., 2015, Bodenheimer and Bauer, 2016].

Second, researchers project demand for primary care services to rise as the U.S. population grows and ages. Most recently, the Affordable Care Act's expansion of health insurance coverage to more than 16 million individuals increased the pool of those seeking primary care. Nearly half of primary care physicians reported an influx of new patients due to the Affordable Care Act, compared to only 30% of specialists [Peckham, 2016]. There is considerable uncertainty about how physicians may respond to an increase in demand for their services. It is possible that the time between appointment scheduling and appointment occurrence (currently an average of 19.5 days to see a family practice physician) may increase [Merritt Hawkins, 2014]. Appointment schedules may also change, with existing research finding that health care providers responded to public insurance expansions by increasing their program participation (i.e., seeing more Medicaid patients), increasing the number of weekly appointments, modifying the number of hours spent

[2] This number includes internal medicine, family practice, and general practice specialties.

with patients, and shortening appointments [Garthwaite, 2012, Buchmueller et al., 2014].

3 Conceptual Framework

Office-based primary care is ideal for studying the importance of work schedules and disruptions to those schedules. Physicians control their schedules in a broad sense - in terms of defining a daily appointment template, deciding whether to accept new patients, and allocating their time between patient interaction and other tasks. However, on any given day, a physician has committed to see a certain list of patients. Bumping someone from that list is costly, both to the physician and the patient. In this section, I propose a simple model to consider how appointment schedule disruptions may affect decision-making among primary care physicians.

I model appointments in two stages. First, the primary care physician allocates time to an appointment, where total time is composed of time spent on one or more conditions (e.g., persistent cough, hypertension). After time spent on each condition, the physician observes the patient's health state and decides on necessary follow-up care for that patient, including lab tests, imaging, medication, or referral to a specialist. Because an office-based physician faces a daily time constraint, the shadow price of time increases when a schedule disruption occurs, with implications for both appointment time and follow-up care.

3.1 Efficient Time Allocation

Consider first what efficient time allocation would look like for a rational physician. Over the course of a day, this physician maximizes the sum of the expected net benefits of appointment time to patients, subject only to a 24-hours-in-the-day constraint or alternatively, the value of leisure or another activity. An individual appointment ends when the marginal cost of an additional minute exceeds the marginal benefit to that patient. For this physician, a schedule disruption (i.e., a later-than-anticipated appointment start) would have no effect on the time spent with subsequent patients that day.

In practice, physicians do not allocate time to patients in this way, but rather seem to apply a behavioral rule about "target" visit duration [Tai-Seale and McGuire, 2011]. In the extreme, this might (and does, in my data) look like a single appointment template of 15 minutes, regardless of what the physician knows *ex ante* about how much time specific patients might need, based on their health states. That 15-minute target may be flexible, but subject to certain behavioral norms. If the physician spends much less time with the patient, she risks causing offense. Spending much more time risks irritating subsequent patients who must wait or angering staff who expect to go home at 5pm.

3.2 Model Setup

Physicians maintain appointment schedules, such that each appointment has an intended start (ω_s) and end (ω_e) time. Realized appointment start (τ_s) and end (τ_e) time define total appointment time (T). Total appointment time is allocated to one or more conditions $c = (c_1, ...c_K)$ where K is the total number of conditions discussed and time per condition must sum to total appointment time, such that $\sum_{k=1}^{K} c_k = T$. I specify an appointment-level physician utility function, including a term expressing the disutility of a schedule disruption:

$$U = f(T) - C(T) - D(T, a; \tau_s - \omega_s, N - n) \tag{1}$$

where $f(T)$ represents the monetary and non-monetary rewards to the physician for providing care. $C(T)$ is the physician's private cost of effort. $D(T, a; \tau_s - \omega_s, N - n)$ is a customer service component, which is a function of the time allocated to the appointment (T), physician characteristics (a), appointment start time relative to scheduled start ($\tau_s - \omega_s$), and how many patients remain to be seen that day (daily patient total [N] minus the number of patients seen so far [n]).

Intuitively, the disutility of a schedule disruption arises from the physician's desire to provide good "customer service", in addition to good medical care, which is built into $f(T)$. Since most primary care physicians practice in groups with other physicians, mid-level practitioners, and administrative staff - and because patient-physician relationships are frequently maintained over the course of years, physician utility suffers when office staff or patients are inconvenienced by a schedule disruption. Note that this does not imply that a rational physician with this utility function necessarily schedules only 15-minute appointments. Rather, conditional on a pre-existing appointment schedule, the physician does not simply maximize the sum of the expected net benefits of appointment time to patients during a day.

3.3 Schedule Disruptions and the Shadow Price of Time

A schedule disruption effectively increases the shadow price of a physician's time during the subsequent appointment(s). To illustrate this, consider an appointment with two conditions and a disutility of schedule disruption component that simply scales the private cost of spending more time on that appointment. The physician's utility function is now $U = f(T) - C(T) - (\tau_s - \omega_s)T$, subject to the constraint $c_1 + c_2 = T$. The physician first decides how much time to spend on an appointment and then allocates time to maximize $f(c_1, c_2)$. Solving this two-stage problem by backwards induction, the second-stage indirect utility function is:

$$V(T) = \max f(c_1, c_2) \text{ s.t. } c_1 + c_2 <= T$$

The first-order conditions of this problem are $f_1 = f_2 = \mu$, where μ is the Lagrange multiplier on the time constraint - or the shadow price of time. Applying the envelope theorem, $V'(T) = \mu \equiv$ shadow price of and the physician's first-stage problem is:

$$\max V(T) - C(T) - (\tau_s - \omega_s)T$$

Now the first-order condition is $V' = \mu = C' + (\tau_s - \omega_s)$ and this demonstrates that the difference between scheduled and observed appointment start time, $\tau_s - \omega_s$, directly influences the shadow price of time, μ. When the physician experiences a schedule disruption, she decreases her time spent per appointment ($\frac{\partial T^*}{\partial(\tau_s - \omega_s)} < 0$).

Proposition 1. *Denote decisions that maximize physician's expected utility in Equation 1 with a * superscript. As $\tau_s - \omega_s \to \infty$, T^* decreases, $c^* = (c_1, ..., c_K)$ decreases and K^* weakly decreases.*

3.4 Deciding on Follow-Up Care

In this simplified model, patients vary only in their health state, $\beta \in \{0, 1\}$, for any given condition. Consider an appointment where one condition is discussed. The physician observes a patient's true health state with probability $p \in (0, 1)$. Patient care increases the probability p of observing β and making appropriate decisions regarding the patient's course of treatment. p is increasing and concave with respect to T. After time spent on the condition, the physician decides whether to order follow-up care ($\Gamma = 1$) or not ($\Gamma = 0$). Patients with $\beta = 0$ should receive no further care, while patients with $\beta = 1$ should receive follow-up care. Physicians are risk averse, such that failing to order follow up care for a sick patient is particularly harmful: $f(\beta = 1, \Gamma = 1) - f(1, 0) > f(0, 0) - f(0, 1)$. If β remains unobserved, the physician will order follow-up care if and only if $p > p^* < \frac{1}{2}$:

$$E[f|\Delta = 0, p = p^*] = E[f|\Delta = 1, p = p^*]$$

$$p^* f(1, 0) + (1 - p^*) f(0, 0) = p^* f(1, 1) + (1 - p^*) f(0, 1)$$

$$\frac{1 - p^*}{p^*} = \frac{f(1, 1) - f(1, 0)}{f(0, 0) - f(0, 1)} > 1$$

If the physician spends less time on a condition, she is more likely to over-order follow-up care for the patient.

Proposition 2 *As $\tau_s - \omega_s \to \infty$ and $T^* \to 0$, $E[\Gamma]$ weakly increases.*

Consider now a chronic condition. This is a condition where the treating physician (or another physician in the practice) has already observed β during a previous encounter. Decisions regarding time allocation and appropriate follow-up care for a chronic condition may differ from a new or acute condition. If conditions are addressed in descending order of acuity (or marginal benefit to the patient), a chronic condition may be less likely to surface during a shorter appointment. Alternatively, economizing on time by simply continuing the patient's ongoing course of treatment for this condition may be more acceptable from a customer service viewpoint than failing to address a new or acute condition. For these reasons, I will explore empirically the difference in ordering behavior by type of condition.

4 Data and Sample Selection

Physicians may also alter decisions for follow-up care in response to a delayed appointment start. In the remainder of this paper, I will provide empirical evidence regarding the relative size and direction of physician response to schedule disruptions. This section describes the data used to answer these questions, presents descriptive statistics, and delves deeper into observed appointment duration.

4.1 Data

I rely primarily on data from athenahealth, Inc., ("athenahealth"), a company that sells cloud-based medical billing, practice management, and electronic health record (EHR) services to health care providers nationwide. Clients span provider types and specialties, with a high concentration of office-based primary care providers.

These novel data contain claims information for all athenahealth providers during 2013-2014, including date of service, patient age, sex, marital status, insurance type, diagnosis and procedure codes, provider place of service, provider type and specialty, allowable charges, and patient cost-sharing. A subset of athenahealth clients also purchase practice management and EHR services. For this group of providers, I use data derived from the athenahealth EHR, including appointment date, time stamps, date of scheduling, scheduled start time, intended duration, and orders placed by the physician for prescriptions, consults, imaging, and lab tests. This combination of claims and EHR data is unique in four important ways:

Physician and Group Identifiers: These allow for observation of practice organization, from the smallest department level (i.e., office location) to the highest health system affiliation. I can then observe daily staffing patterns, presence of non-physician providers, and availability of affiliated, non-office settings.

Time Stamps: My analysis relies heavily on the reporting of scheduled and observed appointment start and end times. I use these data elements to a) calculate appointment duration,

b) construct physician schedules, and c) and observe deviations from these schedules.

All-Payer Prices: The data detail reimbursement for all payers, including those negotiated between the physician and commercial insurers. This allows me to construct a measure of spending per appointment which accurately reflects the physician's compensation.

Orders for Future Care: In addition to the procedures conducted during the appointment, I also observe physician orders for a patient's follow-up care. These include orders for lab and imaging tests, medication prescriptions, and referrals to specialists.

An important limitation of the data is the inability to track patients. If a patient visited a non-athenahealth provider, that utilization is not captured in the data. This generally means that I cannot observe whether a patient followed through on a referral to a specialist, filled a prescription, or received the physician-ordered imaging or lab test.

4.2 Sample Selection and Descriptive Statistics

The main sample comprises adult appointments with office-based primary care physicians (internal medicine, family practice, general practice) during 2013-2014. I restrict the sample to weekday appointments scheduled for 10,15, 20, or 30 minutes time blocks, with no breaks between appointments and non-anomalous, non-overlapping time stamps documenting the start and end of the exam, as I need to accurately measure a) when an appointment began, relative to the scheduled start time, and b) how much time the physician spent with the patient. Physicians with very few appointments or a high proportion of same-day appointments are dropped, as I rely on within physician variation and am interested in the behavior of the vast majority of physicians who organize their labor supply using appointment schedules. My final sample includes 1,194,617 primary care appointments for 917,797 patients, seeing 3,021 primary care physicians, working at 2,017 departments/office locations, and belonging to 792 practices or health systems. For a complete description of sample selection criteria, see Appendix B.

4.2.1 Descriptive Statistics

My data provider's clients comprise a convenience sample of all primary care appointments in the United States. Table 1 provides descriptive statistics of all appointments, as well as those I use as my analytic sample. Both datasets have a similar patient gender and age composition, with each subsequent age category representing a greater share of appointments. Insurance type distribution is also similar across datasets, with the bulk of appointments (58.3% in the full dataset and 60.2% in the analytic sample) covered by commercial insurance, followed by Medicare, Medicaid, self-pay, and workers' compensation. The analytic sample has a higher burden of chronic disease than the full dataset. Geographically, both datasets overrepresent the South and underrepresent the West. The analytic sample has a smaller share of appointments for new patients (13.7%) than the full dataset (19.4%), which likely reflects the exclusion of specialist physicians, with whom patients may not need an established relationship. Finally, the seasonal distribution of appointments is similar between datasets, with slightly more appointments occurring in the autumn than the other seasons, and 2014 shows an increase in appointment count - likely reflecting growth in the athenahealth client base. Despite being a convenience sample, patient characteristics of appointments largely mirror national survey-based estimates, as shown in Appendix Table 2.

Table 2 compares characteristics of physicians and physician practices between my analytic sample, the full dataset, and all PCPs in the full dataset. Physicians in my sample have higher average volume (weekly appointment count and panel size) than all physicians or primary care physicians in the full dataset. They also work more days per week, with an average of 4.6 days in the analytic sample, compared to 3.3 days per week for all physicians and 3.5 days per week for all primary care physicians. Practices in my analytic sample look similar to PCP practices in the full sample, with an average of 21 physicians, 4.5 nurse practitioners, and 2.2 physician assistants billing at the practice.

4.3 Appointment Duration

The data provider's EHR organizes each appointment into five stages, as depicted in Figure 1: checkin, intake, exam, checkout, and signoff - followed by any post-visit documentation that the provider may do. Checkin generally happens with front-office staff, depending on staffing arrangements, and includes insurance status confirmation and other administrative details. During intake, a non-physician provider (or in some instances, the physician) measures the patient's vitals. The exam stage encompasses all interaction between patient and physician. checkout and signoff involve all post-appointment administrative functions and are generally not conducted by the physician.

I measure observed appointment duration as an important input in patient care. This is defined by the EHR as the time elapsed between the minimum and maximum keystroke entry time during the exam stage. I define excess appointment duration as the difference between observed duration and scheduled duration[3] and trim the top and bottom 5% of appointments, to reduce the influence of outliers on my findings. I frequently observe overlapping appointments, suggesting that the physician had the EHR open in two different exam rooms. While double-booked appointments are a common occurrence in primary care practices, it is impossible to accurately observe how much time the physician spent with each patient when appointments overlap, so these are excluded from my sample.

Appointments in my sample are scheduled for 10, 15, 20, or 30 minutes ($>$ 95% of appointments in the full dataset are scheduled for one of these times). Figure 2 plots the distribution of observed duration for each scheduled duration category, excluding overlapping appointments. Average observed appointment duration is monotonically increasing in scheduled duration, but consistently shorter than the scheduled appointment duration. Including all appointments (overlapping and non-overlapping) shows a similar pattern, with observed durations exceeding the scheduled duration (Appendix Figure 1), as one would expect.

Most appointments in my sample (95%) start at or after their scheduled start time. Figure 3 plots the distribution of observed appointment start time, relative to the scheduled start time. On average, appointments start 26.6 minutes after the scheduled start time. Lateness appears to compound over the course of a day. The first scheduled appointment starts an average of 10.2 minutes late (Panel B of Figure 3), but this almost triples, for an average last appointment

[3]An appointment scheduled for 15 minutes that actually lasts 17 minutes will have an excess duration of 2 minutes. Conversely, if that appointment had an observed duration of 10 minutes, its excess duration would be -5 minutes.

start time of 28.5 minutes later than the scheduled start time (Panel C of Figure 3).

5 Instrumental Variables Framework

Having presented background and descriptive statistics, I now turn to a discussion of my empirical strategy for identifying the importance of schedule deviations on physician resource utilization and decision-making. In this section, I detail my instrumental variables approach, the underlying identifying assumptions, first stage results, and my outcomes of interest.

A naive approach to this research question might examine the association between appointment schedule disruptions and physician input use or clinical decision-making. There are two main concerns with this approach. First, physicians differ tremendously, both in their deviations from schedule and in their rates of the outcomes of interest. Second, physician schedule disruptions may themselves be endogenous to the treatment choices for a particular patient. Put another way, a physician may allow appointment A to run long, resulting in a late start for appointment B, if the physician knows something about how much time or effort appointment B will require. More generally, any unobserved appointment characteristics that simultaneously affect appointment start time and the outcome of interest render ordinary least squares (OLS) estimation unlikely to provide a causal estimate, even after adopting a within-physician design. To mitigate thes concern, I employ an instrumental variables approach and use physician fixed effects.

5.1 Instrumental Variables Design

I instrument for *current* observed appointment start time relative to scheduled start time with the office arrival time relative to the scheduled appointment start of the *previous* patient. If a late previous patient affects physicians only by generating deviations from the intended appointment schedule, then the arrival time of a physician's previous patient serves as a valid instrument for estimating the effects of current physician lateness on resource use and decision-making.[4] My

[4]A late-arriving patient may affect physician-making in a way that is more related to stress than to time, confounding this identification strategy.

model for estimating the effect of schedule deviations on outcomes is:

$$Y_{ijt} = \beta StartTime_{ijt} + X_{it}'\gamma + A_{it}'\eta + T_{it}'\theta + \delta_j + \epsilon_{ijt} \tag{2}$$

where $StartTime_{ijt}$ is indexed for patient i, seeing physician j, at appointment time t. X_{it} is a vector of patient characteristics including patient age, gender, insurance status, new patient indicators, and chronic condition indicators. A_{it} is a vector of appointment characteristics, including scheduled duration, appointment rank, and an indicator for being a same-day visit. T_{it} are day-of-the-week-season-year fixed effects. δ_j is a physician-practice combination fixed-effect, controlling for time-invariant differences between physicians.

If - as discussed in the previous section - unobserved characteristics affect both the appointment start time and the outcome of interest, OLS estimation of this model is unlikely to provide a causal estimate of β. I therefore predict minutes behind schedule using the arrival time of the physician's previous patient:

$$StartTime_{ijt} = PatientArrivalTime_{ijt-1} + X_{it}'\beta + A_{it}'\eta + T_{it}'\eta + \delta_j + \epsilon_{ijt} \tag{3}$$

where $PatientArrivalTime_{ijt-1}$ is that physician's previous patient's office arrival time, relative to the scheduled start time. This measures the schedule deviation induced by the previous patient and is the excluded instrument in the system of equations specified by Equations 2 and 3. All standard errors are clustered at the physician level.

5.2 Identifying Assumptions

The instrumental variables framework for identifying the causal effect of appointment start time on input use and ordering behavior relies on three main assumptions: a relevance condition, a monotonicity assumption, and an exclusion restriction. I now turn to a discussion of each assumption. The relevance condition requires that the office arrival time of a physician's previous patient affects the current appointment start time. Figure 4 plots the distribution of previous patient arrival times relative to the appointment start time. The average patient in my sample arrives roughly 8 minutes early to the office.[5] Previous patients arrive later than their scheduled appointment time in 21.8% of appointments.

Patient arrival time is likely one of many factors that determines how closely a physician adheres to the intended appointment schedule. In Figure 5, I plot appointment start time (relative to scheduled start time) as a function of the previous patient's office arrival time (again, relative to the scheduled appointment start time), including physician fixed effects, such that all comparisons are within-physician. This shows very little relationship between previous patient arrival time and appointment start time when the previous patient arrived between 10 and 75 minutes early. After that, I observe a strong positive relationship between appointment start time and the office arrival time of the previous patient. This figure suggests that even an "on-time" patient can still delay the appointment start time.

Figure 6 shows estimates of equation 3, including physician fixed effects and clustering standard errors at the physician level, and present results. This specification controls for patient characteristics (age, gender, insurance, new patient status, and chronic condition indicators), appointment characteristics (scheduled duration and indicators for whether the appointment is a same-day visit or double-booked), and time fixed effects (appointment rank within the day, day of the week, season, and year). The office arrival time of a physician's previous patient indeed affects the start time of the current appointment. I find that a 15 minute late previous patient (i.e., roughly 2 standard deviations from the mean arrival time of 8 minutes early) delays the start time of the current appointment by an additional 2 minutes (1.99 s.e. = 0.056) for a 29 minute late start.

A two minute delay resulting from a 15-minute late previous patient may strike the reader as modest, but should be interpreted in light of four characteristics of physician schedules. First, the average appointment begins 27 minutes later than its scheduled start time (Figure 3). This means that the 15-minute lateness of the previous patient could easily be absorbed by the timing of the average visit. Second, appointments generally end early, falling short of their scheduled duration by 2 to 8 minutes (Figure 2). Hence, the average appointment can absorb a certain

[5]Patient arrival time is measured as the start of the checkin stage. My analysis uses patient arrival time relative to the scheduled appointment start time, such that a patient arriving at 2:05 PM for a 2:15 appointment would have an arrival time of -10 minutes.

amount of patient lateness without delaying the next appointment. Third, the intake stage (between checkin and exam) may also absorb some of the schedule disruption. On average, intake lasts nearly 40 minutes, which likely includes measurement of vitals and patient time spent waiting for the physician. Finally, most days have multiple blocks of time that are not scheduled for patient care. These may be designed to absorb schedule disruptions. Given these four schedule characteristics, a 2-minute delay resulting from a 15-minute late previous patient is well within reasonable.

A late-arriving patient may disrupt the start time of multiple subsequent appointments during the physician's day. While my main instrument for appointment start time at time t is the office arrival time of the previous patient $(t-1)$, I test multiple similar instruments and find that a late-arriving patient at time $t-2$, $t-3$, $t-4$, and $t-5$ continues to predict the appointment start time at time t. However, the effect of a late patient decreases with temporal distance, as Figure 6 shows. I return to these possible alternative instruments when I discuss robustness checks.

Finally, I estimate Equation 3 using the current patient's office arrival time to predict the current appointment start time. As expected, this yields the strongest first stage, with a 15 minutes late current patient delaying the start of the current appointment by an additional 8 minutes. That the late patient's appointment seems to be most severely truncated is compelling evidence that physician utility functions do indeed include a patient experience or patient fairness component.

5.2.1 Monotonicity

In addition to the relevance condition, identification of a local average treatment effect within an IV framework is only possible with a monotonicity assumption [Angrist and Imbens, 1994]. In this context, monotonicity assumes that if a 15-minute late previous patient delays a physician's subsequent appointment start time, a 15-minute late previous patient will always do so (rather than causing the physician to run earlier). While monotonicity is fundamentally untestable, I estimate my first stage on a series of distinct subgroups, dividing my sample by patient and appointment characteristics. Figure 7 presents the results of this exercise, splitting the sample by patient gender, insurance, physician relationship (i.e., new vs. established), and age. I find that splitting my sample in these ways yields statistically indistinguishable results, suggesting that late-arriving patients with different characteristics have a similar effect on physicians' subsequent appointment start time.

5.2.2 Exclusion Restriction

Given a strong first stage, my identification strategy depends on the similarly untestable exclusion restriction: the timing of the previous patient's arrival to the office cannot affect the physician's decision-making during the subsequent appointment outside of its effect on the subsequent appointment's start time.

I test for balance on covariates in Table 3, asking whether a given physician sees different patient types immediately following appointments that differed by patient arrival time. I present the difference in conditional means of appointment observables, stratified by quartile of previous patient office arrival time. These estimates are adjusted for physician fixed-effects and are therefore all within-physician comparisons. The covariate of interest is count of chronic conditions, a variable generated using diagnoses from 1-2 years of past claims.[6] Overall, I do not see statistically significant differences in the count of chronic conditions across quartiles of previous patient arrival time. Relative to appointments in the bottom quartile of previous patient office arrival time (i.e., appointments where the previous patient arrived more than 15 minutes early), the difference in chronic condition count (0.002) for appointments in the second quartile (where the previous patient arrived between 15 and 7.5 minutes early) is small and statistically insignificant (p=0.617). This is also true for the difference between the first and third quartile (0.003, p=0.593) and the difference between the first and fourth quartile (-0.006, p=0.191). I find no significant difference between these coefficients (F-statistic = 1.54). I additionally divide appointments by whether they did or did not follow a patient who arrived late to the office. I find that these groups differ by 0.007 chronic conditions (p=0.082), or 0.3% of the mean chronic condition count.

I next rule out current patient arrival time as an instrument for current appointment start

[6] This calculation adheres as much as possible to the algorithm used by the Chronic Condition Warehouse.

time, as it likely violates the exclusion restriction. The last panel of Table 3 shows balance on covariates by quartile of current - rather than previous - patient office arrival time. Here I see that every subsequent quartile of current patient arrival has significantly fewer chronic conditions than the group of patients who arrive earliest to their appointment. Relative to the bottom quartile, the top quartile of appointments by current patient arrival time has 0.23 fewer chronic conditions, or roughly 10% percent of the mean chronic condition count. This suggests the presence of unobserved variables that affect both current patient arrival time and physician decision-making for that patient.

Patient sorting might lead to a violation of the exclusion restriction if physicians systematically schedule particular patients to follow others, based on anticipated arrival time. This is unlikely given the fact that most appointments are scheduled more than a week in advance. However, a physician might choose to reorder her schedule following a late patient or squeeze in a straight-forward same-day appointment. I reestimate Equation 2 in Figure 8, subsetting my main sample to include only non-same-day appointments, which might be systematically and strategically accommodated within a day, based on the complexity or time requirements of other patients. I reestimate using only those appointments following new patients, such that the physician cannot anticipate the previous patient arriving late, based on past experience. I also limit my sample to appointments with physicians in large group practices ($\geq$ 50 physicians), where appointments are most likely to be scheduled by administrative staff, without input on schedule order from the physician. I find that these four coefficients estimated in different samples are statistically indistinguishable from each other. The fact that these sample restrictions result in minimal changes in my first stage result suggests minimal cause for worry about physician selection of patients based on the previous patient's office arrival time.

To summarize, I find considerable degrees of balance in patient characteristics across appointments that vary by office arrival time of the previous patient. Supporting this, I find that omitting certain appointments that might represent the best opportunities for patient sorting yields estimates very similar to my full sample.

5.3 Outcomes

I use the instrumental variables framework detailed in Section 5.1 to examine a range of outcomes along which physicians may respond to schedule disruptions, split broadly into two categories: time-costly physician inputs and follow-up care. Time-costly physician inputs include exam time, procedure count (Current Procedural Terminology [CPT] codes), allowed charges (a measure of visit intensity), and diagnosis count (International Classification of Diseases, 9th edition [ICD-9]). I use diagnosis count as a proxy for how many conditions were discussed, classifying each diagnosis as "new" or "established", based on whether it has previously appeared on any claim for that patient between January 1, 2010 and the appointment date.

I also examine a group of possible time-economizing strategies that physicians may employ, when running behind schedule, including deferred clinical documentation and blow-off behaviors. I create an indicator for deferred documentation (equal to one if time stamps indicate that the physician returned to the EHR after ending the appointment). While I cannot observe blow-off behaviors directly, I do observe subsequent visits made by a patient to either that physician or a hospital (when the PCP practices within a group with an inpatient setting). I create an indicator for whether the patient revisited their PCP or was hospitalized within two weeks.[7] These measure are challenging to interpret, as a revisit or hospitalization could happen for a variety of reasons. The physician may have instructed the patient to book another appointment in the near future or the patient may seek additional care for worsening symptoms. Regardless of the motivation for a follow-up visit or hospitalization, any increase in these measures represents a financial cost to the patient and broader health care system.

Finally, I examine ordering behavior for follow-up care. The conceptual framework predicts that physicians may modify ordering behavior for follow-up care when they are running behind schedule. I observe orders for lab tests, imaging tests, referrals to specialists, and prescription medications - and link these to an appointment using patient-physician-practice-date combinations. For prescriptions, I am able to classify them as "new" or "existing" based on whether

[7] For specifications with hospitalization as the dependent variable, I restrict my sample to appointments at physician offices that are affiliated with a hospital. By doing this, I can plausibly observe hospitalizations, assuming the patient goes to an affiliated hospital.

they have received an order for that medication (from any provider) since January 1, 2010. I can also identify changes to existing prescriptions (e.g., strength or dosage changes) and changes within a therapeutic class (e.g., substitutions between antidepressants).

6 Results

The conceptual framework in Section 3 predicts that physicians respond to a schedule disruption by shortening total appointment duration, reducing other time-costly inputs, and potentially changing their ordering behavior regarding follow-up care for the patient. This section presents results testing these predictions, followed by robustness checks and placebo tests.

6.1 Do Physicians Speed Up When They Run Late?

I begin by examining the effect of appointment start time on exam duration. I instrument for appointment start time at time t with the office arrival time of the physician's previous patient at time $t-1$. The first panel of Table 5 reports results of 2SLS estimation of observed appointment duration as a function of predicted appointment start time, using the full analytic sample. In this and all subsequent specifications, I control for patient characteristics (age, gender, insurance, new patient status, and chronic condition indicators), appointment characteristics (scheduled duration and indicators for whether the appointment is a same-day visit or double-booked), and time fixed effects (appointment rank within the day, day of the week, season, and year). The coefficient on minutes behind is negative and significant, indicating that physicians speed up as they run increasingly behind schedule. For every additional 2 minutes of appointment start time (relative to the scheduled start time), physicians shorten the appointment by an average of 0.35 minutes (s.e. = 0.016). Given that the average appointment in my sample lasts 4.4 minutes less than its scheduled amount of time, this represents an 8% decrease in observed duration relative to scheduled duration.

In the remainder of Table 5, I examine heterogeneity in physician response to a late appointment start. First, I limit my sample to appointments occurring in the second half of physician days. With fewer appointments remaining to "catch up" during, one might expect physicians to speed up more in response to a schedule disruption during the second half of their day. As predicted, Panel 2 of Table 5 shows that physicians speed up more in response to a schedule disruption when they have fewer remaining appointments.

I also examine heterogeneity by physician scheduling patterns. The analytic sample includes 10, 15, 20, and 30-minute appointments, but physicians use very different mixes of these possible durations. Roughly 10% of physicians in my sample schedule appointments of a single duration (15 minutes is the most common single duration), while the majority of physicians schedule a mix of 15- and 30-minute appointments. Another $\approx$ 15% of physicians schedule appointments of all four durations (10, 15, 20, and 30 minutes). Physicians who schedule multiple appointment durations may do so to maximize expected net patient benefit over a day. These physicians may respond differently to a schedule disruption, knowing that a complicated patient later in the day has a 30-minute appointment, rather than the same 15-minute appointment as his uncomplicated peers. Panels 3 and 4 of Table 5 report the effect of running behind schedule on excess appointment duration, splitting the sample according to physician scheduling patterns. I find that physicians who only schedule one appointment duration reduce excess duration more that physicians with all four appointment duration templates, in response to the same 2 minute delayed appointment start time perturbation (0.37 minutes [s.e. = 0.072] vs 0.31 minutes [s.e. = 0.044].

6.2 How Do Physicians Speed Up?

Having established that physicians reduce total appointment duration in response to a schedule disruption, I now examine how physicians' use of time-costly inputs varies as a function of minutes behind schedule. I find that physicians reduce both time spent with a patient and use of time-costly inputs in response to a schedule disruption. Table 6 shows that physicians significantly reduce billed procedures when they run late. In response to a two minute delay in appointment start time (relative to scheduled start), physicians reduce their procedure billings by 0.02 CPT codes (s.e. = 0.003), which is a 1% decrease relative to the average of 1.85 CPT codes per appointment. I also find a negative coefficient on spending, but it is not statistically significant. This is not surprising, as evaluation & management (E&M) codes are the bulk of physician reimbursement; the average E&M CPT code reimbursement is $96, compared to an average of $31 for non-E&M codes.

Given the number and heterogeneity of CPT codes (i.e., CPT codes can indicate anything from an office visit to a blood draw to procedures like a joint injection), I also estimate the likelihood of specific, high-frequency CPT codes as a function of appointment start time. Table 6 reports findings from each individual model. Like the overall relationship between CPT count and appointment start time, I find that a delayed appointment start reduces the likelihood that a physician will perform certain common procedures. For every two minutes added to the appointment start time, the likelihood of a blood draw decreases by 0.2 percentage points (s.e. = 0.001). The same perturbation results in a 0.2 percentage point decrease (s.e. = 0.0004) in the likelihood of a lipid panel, a 0.1 percentage point decrease (s.e. = 0.0003; significant at the $p < 0.1$ level) in electrocardiograms, and a statistically insignificant decrease in vaccine administration of 0.1 percentage points (s.e. = 0.001). The absence of a significant decrease in vaccine administration may be explained by the fact that this may be the expressly stated reason for a patient's visit, this may be a responsibility the physician can easily transfer to a nurse practitioner or other mid-level provider, or this may require relatively little time spent in discussion.

The count of recorded diagnoses (ICD-9 codes) also decreases as a function of appointment start time, with a 2 minute delayed appointment start resulting in 0.01 fewer diagnoses (s.e. = 0.002) on a sample mean of 2.94 ICD-9 codes recorded. If each diagnosis is a condition, this finding indicates that physicians speed up by addressing fewer discrete conditions - a result consistent with previous research - in addition to shortening the time devoted to each condition [Tai-Seale and McGuire, 2011]. However, it may also indicate that physicians simply document fewer conditions, despite having discussed them.

I find evidence that physicians are more likely to multitask - seeing multiple patients at once and going between rooms - when running late. Reincorporating overlapping appointments,[8] I find that a 2 minute delay in appointment start time increases the likelihood that the ap-

[8] My main sample excludes appointments with overlapping time stamps because I cannot accurately observe how much time the physician spent with either patient.

pointment will overlap with another by 3.7 percentage points (s.e.=0.001), or 7% on a base of 50%.

In addition to doing less during the appointment and multitasking, physicians seem to push some time-costly inputs to a later time or date. In response to a schedule disruption, PCPs defer documentation to a time outside of and after the appointment. For a 2 minute delay in appointment start time, the likelihood of post-appointment documentation increases by 0.3 percentage points (s.e. = 0.01), or 0.5% on a base rate of 51%. Appointment documentation may be a task with particular flexibility on timing, given that it only requires physician effort and does not rely on labor supplied by any other mid-level provider or administrative staff. However, this outcome raises particular concern regarding physician burnout, given research showing that documentation consumes nearly twice as much time as patient care and is a major source of job dissatisfaction [Sinsky et al., 2016, Shanafelt et al., 2012, Shanafelt et al., 2015].

I also find suggestive evidence that physicians postpone patient care when running late. Table 6 reports the likelihood that a patient revisits that physician within two weeks. A 2 minute delay in appointment start time results in a 0.1 percentage point increase (s.e. = 0.006), or 1% on a base rate of 10%. This finding has multiple possible interpretations: the physician could be directly encouraging the patient to reschedule another appointment soon or the patient may return earlier because a condition that wasn't addressed during the initial appointment has worsened. To provide context for this figure, a back-of-the-envelope calculation suggests that a 1% increase in annual primary care office visits is roughly 5 million additional appointments, at a cost of more than $500 million.[9]

A weakness of my data is the inability to track patients when they see a provider in a non-office care setting. However, a subsample of the data provider's clients have an emergency room or inpatient hospital setting. If the patient receives care there, I will observe it in my data. The final panel of Table 6 reports results for physician appointments occurring within a practice affiliated with an inpatient setting. I construct a binary indicator for the existence of an inpatient hospital admission for a chronic "ambulatory care sensitive condi-

[9] To arrive at this spending estimate, I apply the average appointment-level spending in my sample ($111.25) to 1% of a rough estimate of the number of annual primary care physician office visits: $\$111.25 \times 507,015,000 \times 0.01 = \$564$ million.

tion" within two weeks of a primary care appointment in my sample. Ambulatory care sensitive conditions are those for which outpatient care could potentially prevent the need for hospitalization (e.g., a patient with diabetes may be hospitalized for diabetic complications if inadequately monitored or educated in self-management). The rate of these potentially preventable hospitalizations is a frequently-used quality measure at the provider or market level [Agency for Healthcare Research and Quality, 2002]. I do not find any evidence that a patient is more likely to be hospitalized for an ambulatory care sensitive condition after seeing a physician running behind schedule.

Table 6 also presents reduced form results, which can be interpreted as incorporating all ways in which a late patient affects physician decision making during the subsequent visit. While the primary effect of a late patient is to delay the physician, it is possible that stress or other factors generated by patient lateness may affect physician decision-making. Reduced form estimates are smaller in magnitude, but similar in direction and statistical significance to the two-stage least squares estimates.

6.3 Does Speeding Up Affect Follow-Up Care Decisions for a Patient?

Having examined how input use changes in response to an unexpected delay in appointment start time, I now turn to the effect a schedule disruption has on physician decision-making regarding appropriate follow-up care for the patient.

6.3.1 New Conditions

The conceptual framework in Section 3 yields different predictions regarding schedule disruption driven changes in follow-up care based on the type of condition addressed. For new or acute conditions, I predict the likelihood of orders for follow-up care to increase as a risk averse physician speeds up and is less likely to observe the patient's true health state. Appointments are likely to include discussion of multiple conditions (based on the average of 1.33 ICD-9 diagnoses per appointment) and, while I can match orders to an appointment, I cannot necessarily match orders to the specific condition they address. For this reason, it is necessary to subset my sample by patient type, to focus on certain types of follow-up care relevant to particular conditions.

I begin by focusing on the effect of running behind schedule when a physician sees a new patient.[10] For this subset of patients, all conditions discussed during the appointment are new (though some conditions may be chronic and this is simply the first discussion between this patient-physician pair) and any orders - labs, prescriptions, imaging, or specialist referrals - placed in this context are more likely to indicate a change in the patient's treatment course. The first panel of Table 7 reports results for new patient visits to office-based PCPs. I find no significant change in overall ordering behavior, nor in the likelihood of lab, imaging, or prescription medication orders.

Unlike other order types, the likelihood that a new patient receives a specialist referral increases considerably as their physician falls behind schedule. A 2 minute delay in appointment start results in a 0.5 percentage point increase (s.e. = 0.003) in referrals to a specialist. Relative to a base rate of 13% of appointments resulting in a specialist referral, this is a 4% increase in referral likelihood. To provide context for this figure, a back-of-the-envelope calculation suggests that a 4% increase in annual specialist visits among new patients is roughly 2 million additional appointments, at a cost of more than $350 million.[11]

Next, I examine patients with a first-time diagnosis of a painful condition, including arthropathies, spinal disorders, and rheumatism.[12] For this subset of patients, opioid painkillers are a relevant prescription drug order. By focusing on appointments where the patient receives a first-time diagnosis of a painful condition, I limit my sample to appointments where I likely observe the initiation of opioid treatment. The second panel of Table 7 shows that the likelihood of receiving an opioid prescription during an appointment where a new painful condition is recorded increases as a function of appointment start time. A 2 minute delay in appointment start results in a 0.2 percentage point (s.e. = 0.0009) increase in the likelihood of a narcotic painkiller prescription. This is an increase of 2.5% relative to the base rate of 7% of appointments in this subset that result in an opioid prescription. I also find an increase in non-opioid painkiller prescribing as a

[10] I define a new patient as any patient that the physician has not submitted a claim for since January 1, 2010.

[11] To arrive at this spending estimate, I apply the average appointment-level spending in the full athenahealth dataset for physicians with a non-primary care specialty ($161.76) to 4% of a rough estimate of the number of annual specialist visits, scaled by the proportion of patients in my analytic sample who are new (12.9%): $\$161.76 \times 421,584,000 \times 0.04 \times 0.129 = \$352$ million.

[12] I define a new diagnosis as anything that has not previously been recorded on any claim for that patient.

function of minutes behind, but this is not statistically significant.

6.3.2 Acute Conditions

The predicted response of a physician to a schedule disruption when treating a patient with an acute condition is similar to that for a new condition. I now examine the subset of appointments for which the physician records an upper respiratory infection (URI) diagnosis, which is likely to be an acute condition. For this subset of patients, antibiotics are a possible prescription drug order - and not always an appropriate one. Public health experts have long been concerned about overprescription of antibiotics for URIs [Centers for Disease Control and Prevention, 2015]. I examine the likelihood of an antibiotic prescription as a function of schedule disruptions in the final panel of Table 7. I find that a 2 minute delay in appointment start time increases the likelihood of a patient receiving an antibiotic prescription by 1.9 percentage points (s.e. = 0.007), or 3.4% relative to the sample mean of 57.8%.

6.3.3 Established Conditions

The conceptual framework in Section 3 suggests that physicians may respond differently to a schedule disruption when deciding on follow-up care for a chronic condition. In Table 8, I look at the likelihood of modifying an existing prescription as a function of appointment start time.[13] I find that a 2 minute delay in appointment start time results in a 0.1 percentage point (s.e. = 0.0006) reduction in the likelihood of any change to an existing prescription. In my sample, roughly half of the changes to existing predictions are changes in medication strength, while the other half are brand name changes within a therapeutic class. The decrease in overall prescription modifications as a function of minutes behind schedule is driven by a significant decrease in the likelihood of a brand name change within therapeutic class. The same 2 minute delay in appointment start results in a 0.1 percentage point (s.e. = 0.0005) reduction in the likelihood of a brand name change, which is roughly an 2% decrease relative to the sample mean of 5%. If switching brand name medications is a more drastic change than adjusting the

[13] I identify existing prescriptions as any prescription submitted by any physician on a date prior to the current appointment. This includes prescriptions ordered during or outside of an appointment.

strength of the same medication, this result suggests that physicians are particularly likely to avoid making major treatment course changes when running late.

6.4 Robustness Checks

As discussed in Section 5, the start time of appointment t may be a function of patient arrival time prior to $t-1$ (my primary instrument). In Figure 6, I show that a late-arriving patient at time $t-2$, $t-3$, $t-4$, and $t-5$ continues to significantly delay the appointment start time at time t. However, the effect of a late patient decreases with temporal distance. I repeat my main regressions using each of the four possible alternative instruments and present these results in Appendix Table 3. I find that a delayed start caused by a late patient at time $t-2$ through $t-5$ results in a shortened appointment at time t. Changes in observed appointment duration and documentation deferral remain significant through multiple sequential instruments, while other coefficients are directionally similar, but insignificant.

I also construct an instrument for appointment start time that is a binary indicator of whether the physician's previous patient arrived at the office after their scheduled appointment start time - rather than a continuous measure of patient arrival time at $t-1$. Like the continuous instrument, this binary instrument has a strong first stage, with a late previous patient adding an additional 1.35 minutes (s.e.=0.06) to the current appointment's predicted start time. Again, physicians respond to a delayed appointment start time by truncating the appointment duration, billing fewer procedures, recording fewer diagnoses, and deferring appointment documentation (Panel 1, Appendix Table 4).

Finally, I construct an instrument for appointment start time that is the log of the previous patient's arrival time. I find that a 10% increase in previous patient arrival time delays the start of the current appointment by 0.2% (s.e. = 0.02%). A 10% increase in appointment start time reduces the observed appointment duration by 4.9% (s.e.=0.5%), procedure use by 0.03 CPT codes (s.e. = 0.009), spending by 0.5% (0.003%), and diagnosis count by 0.02 ICD-9 codes (s.e. = 0.007). The likelihood of deferred appointment documentation increases by 0.07 percentage points (s.e. = 0.002) and a revisit within 2 weeks by 0.002 percentage points (s.e. = 0.0002) (Panel 2, Appendix Table 4).

6.4.1 Placebo Tests

To confirm that my results are not spurious, I conduct the following placebo tests. I instrument for appointment start time using the office arrival time of a physician's *subsequent* patient. I present results of this placebo tests in Appendix Table 5. As anticipated, I find that the subsequent patient's office arrival times (for appointments $t+1$ and $t+2$) are not predictive of the current patient's appointment start time.

To test my finding regarding increased likelihood of patient revisit to the PCP within two weeks, I create an indicator variable for whether the patient revisited the PCP within two weeks for an appointment that had already been scheduled prior to the current visit. Schedule disruptions should not affect the likelihood that a patient keeps an already-scheduled appointment in the future. I present results in Appendix Table 6. As anticipated, I find that a 2 minute delay in appointment start time results in a statistically insignificant change in the likelihood of a *previously scheduled* patient revisit to the PCP within two weeks.

7 Implications and Conclusion

At the individual appointment level, the magnitudes of my findings are modest. A 15-minute late previous patient delays the current patient's appointment by an additional 2 minutes. In response to the unexpected schedule disruption, physicians reduce input use by 1 to 4%, depending on the input (i.e., procedure use, time). They also modify ordering behavior for certain relevant patient populations (e.g., new patients, patients with a new pain diagnosis or upper respiratory infection) by 2.5 to 4%.

To appreciate the broader impact of schedule disruptions, consider that nearly four out of five physician-days in my sample contain one or more late-arriving patients. More than 15% of the days in my sample include a patient who arrived at least 15 minutes late, resulting in changes in input use and ordering behavior for subsequent appointments. One in seven of all appointments followed a 15-minute late patient closely enough to delay the start time and affect the physician's decision-making.

My identification strategy relies on one source of plausibly exogenous variation in appointment start time, but many other factors likely contribute to a late appointment start. As physicians' time becomes an increasingly scarce resource, appointment schedules may adapt in a few possible ways: more appointments, denser appointment schedules, and/or shorter appointments. I examine the association between these three possible changes and schedule disruptions, specifically how accumulated daily "lateness" (i.e., how late the day's last appointment started - how late the day's first appointment started) responds. I find that all three of the possible schedule changes are associated with an increase in accumulated daily lateness.

First, physicians may add appointments to the day. In my analytic sample, I find that each additional appointment in a physician's day adds almost half a minute, on average, to accumulated daily lateness. Second, appointment schedule density may increase, with more double-booked appointments or fewer non-scheduled blocks of time. In my analytic sample, I find that a 10 percentage point increase in the number of double-booked appointments adds roughly 2 minutes to a physician's accumulated daily lateness. The same 10 percentage point increase in the share of in-office time devoted to patient care adds 3 minutes to a physician's accumulated daily lateness. Finally, appointment schedules may respond by shortening the time allotted to each patient. I examine the association between the daily share of short appointments (i.e., those scheduled for 10 or 15 minutes rather than 20 or 30 minutes) and the accumulated daily lateness. I find that a 10 percentage point increase in the share of short appointments adds 1 minute to a physician's accumulated daily lateness.

The association between possible appointment schedule changes (i.e., more appointments, denser appointment schedules, and/or shorter appointments) and schedule disruptions suggests two implications. First, more time-series data on primary care appointment scheduling are needed to understand how physicians respond to increased demands on their time. Ideally, these measures will *not* be derived from physician or patient surveys. Instead, these measures should be derived from utilization data with time-stamps, a type of data that is quite new to the health economics community and may require a reorientation towards proprietary data sources. Second, if physicians respond to increased time pressure in any of the ways described above, this will likely have implications for spending and quality of care.

7.1 Conclusion

In this paper, I show that schedule disruptions (specifically, running behind schedule) affect the input use and ordering behavior of office-based primary care physicians. In response to a delayed appointment start, PCPs spend less time with the patient, bill fewer procedures, record fewer diagnoses, and defer appointment documentation to after the appointment. A patient whose physician is running late is more likely to revisit that physician within the next two weeks, however likelihood of a potentially preventable hospitalization does not change. I find some evidence that a schedule disruption induces physicians to order more follow-up care for a new condition or complaint (i.e., specialist referrals for a new patient, antibiotics for a patient with an upper respiratory infection, opioid painkillers for a patient with a new diagnosis of back pain), but to decrease orders that change the existing course of treatment for an established condition.

My findings are relevant to multiple actors at various levels of the health care system. For policymakers and private payers, my findings suggest an unintended consequence of steadily squeezing primary care physician's appointment schedules. Whether the increasing scarcity of physician time is driven by increased documentation and quality reporting requirements or major insurance expansions, my results suggest that more schedule disruptions and shorter effective appointment times may have implications for health care spending and quality. Specifically, more frequent return visits to a PCP and specialist referrals place upward pressure on health care spending. Care quality may suffer if physicians increase use of contraindicated care (e.g., antibiotics for certain conditions) in response to a schedule disruption. Additionally, schedule deviations may be one factor that contributes to decisions of particular concern to policymakers, like opioid prescribing. To the extent that policymakers and payers are concerned about these outcomes, it may be appropriate to take legislative or regulatory action to modify the administrative requirements placed on physicians.

My results may also have bearing on discussions regarding expansion of the primary care workforce. I find that physicians respond to schedule disruptions by deferring appointment documentation. Current research estimates that physicians do nearly two hours of EHR and desk work for every one hour of direct clinical face time with patients. Given the clear relationship between administrative work and job dissatisfaction, increasing the time spent on such tasks poses the threat of exacerbating the already accelerating primary care physician burnout rates.

Finally, my results may have implications at the physician practice level. I find that certain scheduling patterns seem to weather the schedule disruption storm better than others. For example, physicians who schedule single-duration appointments (e.g., only 15-minute appointments) shorten appointment duration more than their multi-duration peers as they fall behind schedule. This finding may suggest a role for targeted clinical decision support, deployed in clinically questionable areas (e.g., antibiotic prescriptions for upper respiratory infections) as physicians run increasingly behind schedule. Additionally, there may be a welfare-improving role for appointment schedule innovations that incorporate existing data on patients into future scheduling decisions.

References

[Abbo et al., 2008] Abbo, E. D., Zhang, Q., Zelder, M., and Huang, E. S. (2008). The Increasing Number of Clinical Items Addressed During the Time of Adult Primary Care Visits. *Journal of General Internal Medicine*, 23(12):2058–2065.

[Agency for Healthcare Research and Quality, 2002] Agency for Healthcare Research and Quality (2002). Guide to Prevention Quality Indicators. pages 1–115.

[Angrist and Imbens, 1994] Angrist, J. D. and Imbens, G. W. (1994). Identification and Estimation of Local Average Tretment Effects. *Econometrica*, 62(2):467–475.

[Arrow, 1963] Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review*, 53(5):941–973.

[Blumenthal et al., 2015] Blumenthal, D., Malphrus, E., and MacGinnis, J. M. (2015). *Vital Signs: Core Metrics for Health and Health Care Progress*. National Academies Press, Washington (DC).

[Bodenheimer and Bauer, 2016] Bodenheimer, T. and Bauer, L. (2016). Rethinking the Primary Care Workforce - An Expanded Role for Nurses. *New England Journal of Medicine*, 375(11):1015–1017.

[Bodenheimer and Pham, 2010] Bodenheimer, T. and Pham, H. H. (2010). Primary Care: Current Problems And Proposed Solutions. *Health Affairs*, 29(5):799–805.

[Brandenburg et al., 2015] Brandenburg, L., Gabow, P., Steel, G., Toussaint, J., and Tyson, B. J. (2015). Innovation and Best Practices in Health Care Scheduling. *Institute of Medicine*, pages 1–24.

[Brownlee, 2012] Brownlee, S. (2012). Why Your Doctor Has No Time to See You.

[Buchmueller et al., 2014] Buchmueller, T. C., Miller, S., and Vujicic, M. (2014). How Do Providers Respond to Public Health Insurance Expansions? Evidence from Adult Medicaid Dental Benefits. *National Bureau of Economic Research Working Paper Series*, pages 1–58.

[Casalino et al., 2016] Casalino, L. P., Gans, D., Weber, R., Cea, M., Tuchovsky, A., Bishop, T. F., Miranda, Y., Frankel, B. A., Ziehler, K. B., Wong, M. M., and Evenson, T. B. (2016). US Physician Practices Spend More Than 15.4 Billion Annually To Report Quality Measures. *Health Affairs*, 35(3):401–406.

[Centers for Disease Control and Prevention, 2015] Centers for Disease Control and Prevention (2015). Adult Treatment Recommendations. Technical report.

[Centers for Disease Control and Prevention, 2016] Centers for Disease Control and Prevention (2016). 1 in 3 Antibiotic Prescriptions Unnecessary. Centers for Disease Control and Prevention.

[Chan, 2015a] Chan, D. (2015a). The Efficiency of Slacking Off: Evidence from the Emergency Department. *NBER Working Paper Series*, pages 1–66.

[Chan, 2015b] Chan, D. C. (2015b). Teamwork and Moral Hazard: Evidence from the Emergency Department. *Working Paper*, pages 1–42.

[Chen, 2013] Chen, P. W. (2013). For New Doctors, 8 Minutes Per Patient.

[Clynch and Kellett, 2015] Clynch, N. and Kellett, J. (2015). Medical documentation: Part of the solution, or part of the problem? A narrative review of the literature on the time spent on and value of medical documentation. *International Journal of Medical Informatics*, 84(4):221–228.

[Coviello et al., 2010] Coviello, D., Ichino, A., and Persico, N. (2010). Don't Spread Yourself Too Thin: The Impact of Task Juggling on Workers' Speed of Job Completion. *NBER Working Paper*, pages 1–45.

[Cutler et al., 2013] Cutler, D., Skinner, J., Stern, A. D., and Wennberg, D. E. (2013). Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending. *NBER Working Paper Series*, pages 1–50.

[Finkelstein et al., 2016] Finkelstein, A., Gentzkow, M., and Williams, H. (2016). Sources of Geographic Variation in Health Care: Evidence from Patient Migration. *Working Paper*, pages 1–50.

[Fiscella and Epstein, 2008] Fiscella, K. and Epstein, R. M. (2008). So much to do, so little time: care for the socially disadvantaged and the 15-minute visit. *Archives of internal medicine*, 168(17):1843–1852.

[Garthwaite, 2012] Garthwaite, C. L. (2012). The Doctor Might See You Now: The Supply Side Effects of Public Health Insurance Expansions. *American Economic Journal: Economic Policy*, 4(3):190–215.

[Health Resources and Services Administration, 2013] Health Resources and Services Administration (2013). Projecting the Supply and Demand for Primary Care Practitioners Through 2020. Technical report, Rockville, MD.

[Higgins et al., 2013] Higgins, A., Veselovskiy, G., and McKown, L. (2013). Provider Performance Measures In Private And Public Programs: Achieving Meaningful Alignment With Flexibility To Innovate. *Health Affairs*, 32(8):1453–1461.

[IHS, 2015] IHS (2015). The Complexities of Physician Supply and Demand: Projections from 2013 to 2025. Technical report.

[Kaiser Family Foundation, 2016] Kaiser Family Foundation (2016). Total Professionally Active Physicians. Technical report.

[Konrad et al., 2010] Konrad, T. R., Link, C. L., Shackelton, R. J., Marceau, L. D., von dem Knesebeck, O., Siegrist, J., Arber, S., Adams, A., and McKinlay, J. B. (2010). Its About Time. *Medical care*, 48(2):95–100.

[Mas and Moretti, 2009] Mas, A. and Moretti, E. (2009). Peers at Work. *American Economic Review*, 99(1):112–145.

[Medicare Payment Advisory Commission, 2016] Medicare Payment Advisory Commission (2016). Physician and Other Health Professional Services. In *Report to the Congress: Medicare Payment Policy*, pages 1–30. Medicare Payment Advisory Commission, Washington.

[Merritt Hawkins, 2014] Merritt Hawkins (2014). Physician Appointment Wait Times and Medicaid and Medicare Acceptance Rates. Technical report.

[Molitor, 2016] Molitor, D. (2016). The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration. *Working Paper*, pages 1–66.

[National Center for Health Statistics, 2014] National Center for Health Statistics (2014). Summary Health Statistics for U.S. Adults: National Health Interview Survey, 2012. pages 1–171.

[National Center for Health Statistics, 2015] National Center for Health Statistics (2015). National Ambulatory Medical Care Survey: 2012 State and National Summary Tables. pages 1–40.

[National Center for Health Statistics, 2016] National Center for Health Statistics (2016). Health, United States, 2015. pages 1–461.

[National Resident Matching Program, 2016] National Resident Matching Program (2016). Results and Data: 2016 Main Residency Match. Technical report, Washington, DC.

[Oxentenko et al., 2010] Oxentenko, A. S., West, C. P., Popkave, C., Weinberger, S. E., and Kolars, J. C. (2010). Time spent on clinical documentation: a survey of internal medicine residents and program directors. *Archives of internal medicine*, 170(4):377–380.

[Peckham, 2015] Peckham, C. (2015). Medscape Physician Compensation Report 2015 . Technical report.

[Peckham, 2016] Peckham, C. (2016). Medscape Physician Compensation Report 2016. Technical report.

[Petterson et al., 2012] Petterson, S. M., Liaw, W. R., Phillips, R. L., Rabin, D. L., Meyers, D. S., and Bazemore, A. W. (2012). Projecting US Primary Care Physician Workforce Needs: 2010-2025. *The Annals of Family Medicine*, 10(6):503–509.

[Rabin, 2014a] Rabin, R. C. (2014a). 15-Minute Visits Take A Toll On The Doctor-Patient Relationship.

[Rabin, 2014b] Rabin, R. C. (2014b). You're on the Clock: Doctors Rush Patients out the Door.

[Shanafelt et al., 2015] Shanafelt, T., Hasan, O., Dyrbye, L., Sinsky, C., Satele, D., Sloan, J., and West, C. (2015). Changes in Burnout and Satisfaction WithWork-Life Balance in Physicians and theGeneral US Working Population Between2011 and 2014. *Mayo Clinic Proceedings*, 90(12):1600–1613.

[Shanafelt et al., 2012] Shanafelt, T. D., Boone, S., Tan, L., Dyrbye, L. N., Sotile, W., Satele, D., West, C. P., Sloan, J., and Oreskovich, M. R. (2012). Burnout and Satisfaction With Work-Life Balance Among US Physicians Relative to the General US Population. *Archives of internal medicine*, 172(18):1377.

[Shaw et al., 2014] Shaw, M. K., Davis, S. A., Fleischer, A. B., and Feldman, S. R. (2014). The duration of office visits in the United States, 1993 to 2010. *The American journal of managed care*, 20(10):820–826.

[Shunko et al., 2015] Shunko, M., Niederhoff, J., and Rosokha, Y. (2015). Humans Are Not Machines: The Behavioral Impact of Queueing Design on Service Time. pages 1–33.

[Silver, 2016] Silver, D. (2016). Haste or Waste? Peer Pressure and the Distribution of Marginal Returns to Health Care. *Job Market Paper*, pages 1–95.

[Sinsky et al., 2016] Sinsky, C., Colligan, L., Li, L., Prgomet, M., Reynolds, S., Goeders, L., Westbrook, J., Tutty, M., and Blike, G. (2016). Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. *Annals of internal medicine*.

[Song et al., 2016] Song, H., Tucker, A. L., Murrell, K. L., and Vinson, D. R. (2016). Public Relative Performance Feedback:Improving Worker Productivity through Adoption of Coworkers' Best Practices. *Working Paper*, pages 1–38.

[Starfield, 1992] Starfield, B. (1992). *Primary Care: Concept, Evaluation, and Policy*. Oxford University Press, New York, NY.

[Tai-Seale and McGuire, 2011] Tai-Seale, M. and McGuire, T. (2011). Time is up: increasing shadow price of time in primary-care office visits. *Health Economics*, 21(4):457–476.

[Tai-Seale et al., 2007] Tai-Seale, M., McGuire, T. G., and Zhang, W. (2007). Time Allocation in Primary Care Office Visits. *Health Services Research*, 42(5):1871–1894.

[US Department of Health and Human Services, 2016] US Department of Health and Human Services (2016). The Opioid Epidemic: By the Numbers. Technical report.

[White House, 2015] White House (2015). National Action Plan for Combating Antibiotic-Resistant Bacteria. Technical report.

[Zismer et al., 2015] Zismer, D. K., Christianson, J., Marr, T., and Cummings, D. (2015). An Examination of the Professional Services Productivity for Physicians and Licensed, Advance Practice Professionals Across Six Specialties in Independent and Integrated Clinical Practice. Technical report.

8 Tables

Table 1: Appointment Descriptives

	Full Dataset	Analytic Sample
Age		
0-9	0.095	-
10-19	0.081	-
20-29	0.099	0.061
30-39	0.108	0.090
40-49	0.118	0.135
50-59	0.153	0.196
60-69	0.155	0.223
70+	0.192	0.268
Female	0.588	0.577
Insurance		
Commercial	0.583	0.602
Medicare	0.235	0.322
Medicaid	0.094	0.048
No Insurance	0.048	0.023
Workers' Compensation	0.014	0.003
Chronic Condition Count		
0	0.787	0.470
1	0.098	0.134
2	0.041	0.129
3+	0.073	0.267
Geography		
Northeast	0.222	0.263
Midwest	0.194	0.194
South	0.482	0.427
West	0.107	0.124
New Patient	0.194	0.137
Season		
Winter (Jan-March)	0.223	0.237
Spring (Apr-June)	0.242	0.247
Summer (July-Sept)	0.256	0.246
Fall (Oct-Dec)	0.279	0.270
2014	0.558	0.571
Total Charges	$147.97	$111.25

Note: This table presents descriptive statistics on all appointments within the full dataset, compared to the main analytic sample of appointments with office-based primary care physicians.

Table 2: Physician and Practice Descriptives

	Full Dataset	Full Dataset (PCPs only)	Analytic Sample
Physician Characteristics			
Volume			
Weekly appointment count: Average physician	41.8	45.0	68.0
Weekly appointment count: 10th percentile physician	2.2	2.8	31.3
Weekly appointment count: 90th percentile physician	93.0	97.2	108.8
Panel size: Average physician	1103.1	1089.9	1855.5
Panel size: 10th percentile physician	12.0	15.0	759.0
Panel size: 90th percentile physician	2764.0	2600.0	3146.0
Days active, per week	3.3	3.5	4.6
Appointment Time			
Scheduled Duration: Average physician	18.5	18.8	17.1
Scheduled Duration: 10th percentile physician	11.1	13.9	14.1
Scheduled Duration: 90th percentile physician	29.9	29.9	22.4
Practice Characteristics			
Physician Count: Average Practice	11.7	22.5	21.0
Nurse Practitioner Count: Average Practice	3.2	4.5	4.5
Physician Assistant Count: Average Practice	1.4	2.5	2.2

Note: This table presents descriptive statistics on physicians and practices within the full dataset, primary care physicians in the full dataset, and the analytic sample used in this paper. Panel size is defined as the number of distinct patients who see any given physician or practice during the study period. This likely underestimates true panel size, as some patients may not seek an appointment with their physician, despite having an established relationship. Days active is defined as the average weekly count of days during which a physician billed for patient care. While the analytic sample only includes primary care physicians, this table presents the average number of mid-level providers (nurse practitioners and physician assistants) billing within the same practice. A practice is defined as a context ID, which is the primary unit of contracting.

Table 3: Balance on Patient Chronic Conditions Count

	Difference	Std. Error	P-Value
Quartile of Previous Patient Lateness			
2	0.002	0.005	0.617
3	0.003	0.005	0.593
4	-0.006	0.005	0.191
Previous Patient Late: Yes or No			
Yes	-0.007	0.004	0.082
Quartile of Current Patient Lateness			
2	-0.0417	0.007	0.000
3	-0.122	0.008	0.000
4	-0.233	0.009	0.000

Note: This table presents the difference in conditional means of observables for appointments by quartile of previous patient arrival time to the office. Estimates are adjusted for physician fixed effects, so the comparisons are all within-physician.

Table 4: Common Orders, by Category

Order Type	% of Orders	Order Type	% of Orders
Prescription Drugs		Imaging Tests	
Analgesic Narcotic Agonist Combinations	6.9%	Electrocardiogram	18.2%
Antihyperlipidemics	6.4%	Mammogram	9.6%
Antidepressants	5.3%	X-Ray, Chest	8.5%
Antisecretory Agents	3.6%	Bone Density Study	3.5%
Anti anxiety Agents	3.2%	Echocardiogram	3.4%
ACE Inhibitors	3.1%	X-Ray, Knee	2.2%
Thyroid Hormones	3.0%	CT, Abdomen and Pelvis	1.5%
Glucocorticoids	2.9%	X-Ray, Lumbar Spine	1.4%
Beta Blockers	2.7%	X-Ray, Shoulder	1.3%
Anticonvulsants	2.5%	X-Ray, Foot	1.3%
Lab Tests		Specialist Referrals	
Complete Blood Count	9.5%	Gastroenterology	8.3%
Comprehensive Metabolic Panel	8.4%	Physical Therapy	8.2%
Thyroid Stimulating Hormone	5.5%	Orthopaedic	6.9%
Lipid Panel	5.3%	Cardiology	5.1%
Hemoglobin A1c	4.9%	Dermatology	4.8%
Urinalysis	2.7%	General Surgery	3.9%
Basic Metabolic Panel	1.9%	Neurology	3.8%
Prothrombin Time	1.2%	ENT	3.6%
Free T4	1.1%	Urology	3.5%
Vitamin D, 25-Hydroxy	1.1%	Colonoscopy	3.4%

Note: This table presents the most common physician orders placed during appointments in the athenahealth analytic sample. Orders are categorized into prescription drugs, imaging tests, lab tests, and specialist referrals.

Table 5: Physicians Speed Up When Running Late

	(1) 2SLS	(2) Reduced Form	(3) Mean
Analytic Sample			
Observed Duration	-0.1764***	-0.0151***	13.7
	(0.0078)	(0.0006)	
Observations	1,194,617	1,194,647	
2nd Half of Day			
Observed Duration	-0.2213***	-0.0141***	13.6
	(0.0149)	(0.0008)	
Observations	559,936	559,992	
Single Duration Physicians			
Observed Duration	-0.1861***	-0.0172***	13.6
	(0.0359)	(0.0030)	
Observations	43,657	43,660	
All Duration Physicians			
Observed Duration	-0.1571***	-0.011***	12.7
	(0.0218)	(0.0014)	
Observations	191,891	191,896	

Note: Full controls are used, including patient characteristics (age, gender, insurance, new patient status, and chronic condition indicators), appointment characteristics (scheduled duration and indicators for whether the appointment is a same-day visit or double-booked), and time fixed effects (appointment rank within the day, day of the week, season, and year). Standard errors are clustered at the physician level. * denotes significance at 10% level, ** denotes significance at 5% level, and *** denotes significance at 1% level.

Table 6: Physicians Speed Up by Doing Less During the Appointment

	(1) 2SLS	(2) Reduced Form	(3) Mean
Analytic Sample			
Procedure Count	-0.0077***	-0.0007***	1.85
	(0.0016)	(0.0001)	
Spending	-0.1489	-0.0128	$111.25
	(0.0985)	(0.0084)	
Diagnosis Count	-0.0050***	-0.0004***	1.33
	(0.0011)	(0.0001)	
Post-Visit Documentation	0.0013***	0.0001***	0.51
	(0.0004)	(0.0000)	
Revisit within 2 Weeks	0.0006**	0.0001***	0.10
	(0.0003)	(0.0000)	
Specific Procedures			
Blood Draw	-0.0012***	-0.0001***	0.09
	(0.0003)	(0.00003)	
Electrocardiogram	-0.0003*	-0.00003*	0.06
	(0.0002)	(0.00001)	
Vaccine Administration	-0.0003	-0.00003	0.05
	(0.0003)	(0.00002)	
Lipid Panel	-0.0009***	-0.0001***	0.03
	(0.0002)	(0.0000)	
Observations	1,194,617	1,194,647	
Overlapping Appointment	0.0183***	0.0015***	0.50
	(0.0006)	(0.0000)	
Observations	2,389,877	2,389,878	
Practices with Inpatient Setting			
Ambulatory Care Sensitive Condition	-0.00006	-.000001	0.0002
Hospitalization within 2 weeks	(0.00004)	(0.00001)	
Observations	453,059	453,142	

Note: Full controls are used. Standard errors are clustered at the physician level. To examine multitasking as an outcome, I have reincorporated overlapping appointments into the sample. An appointment is considered to have an available inpatient setting when the physician works at a practice with inpatient hospital or emergency room claims. * denotes significance at 10% level, ** denotes significance at 5% level, and *** denotes significance at 1% level. Full regression results presented in Appendix Table 1.

Table 7: Response of Ordering Behavior to a Schedule Disruption

	(1) 2SLS	(2) Reduced Form	(3) Mean
New Patients			
Prescription Drug	-0.0005	-0.0000	0.49
	(0.0018)	(0.0001)	
Lab Test	-0.0018	-0.0001	0.37
	(0.0019)	(0.0000)	
Imaging Test	-0.0013	-0.0001	0.18
	(0.0017)	(0.0001)	
Specialist Referral	0.0027**	0.0002**	0.13
	(0.0012)	(0.0001)	
All Orders	-0.0005	-0.0000	0.71
	(0.0014)	(0.0001)	
Observations	160,472	160,641	
New Pain Diagnoses			
Opioid Painkiller	0.0009**	0.0004**	0.07
	(0.0004)	(0.0002)	
Non-Opioid Painkiller	0.0008	0.0005	0.13
	(0.0005)	(0.0004)	
Observations	528,325	528,409	
URI Diagnoses			
Antibiotic Order	0.0097***	0.0075***	0.58
	(0.0034)	(0.0025)	
Observations	295,615	295,750	

Note: Full controls are used. URI is upper respiratory infection. New pain diagnoses include any new diagnosis of arthropathies, spinal disorders, or rheumatism. Standard errors are clustered at the physician level. * denotes significance at 10% level, ** denotes significance at 5% level, and *** denotes significance at 1% level.

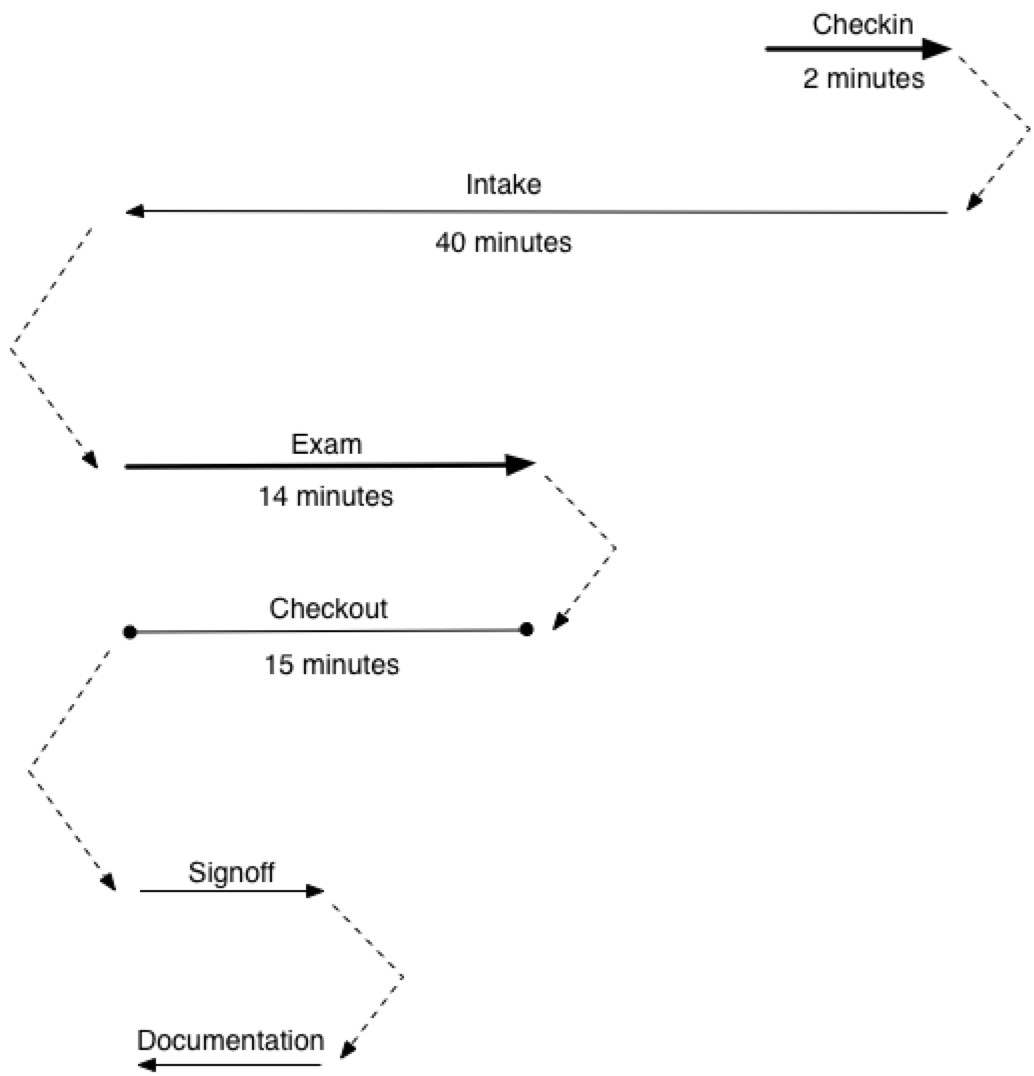
Table 8: Likelihood of a Prescription Modification

	(1) 2SLS	(2) Reduced Form	(3) Mean
Any Change to Prescription	-0.0007**	-0.0001**	0.9
	(0.0003)	(0.0000)	
Strength Change	-0.0003	-0.0000	0.05
	(0.0002)	(0.0000)	
Brand Name Switch	-0.0005**	-0.0001**	0.05
	(0.0002)	(0.0000)	
Observations	1,194,617	1,194,674	

Note: Full controls are used. Standard errors are clustered at the physician level. * denotes significance at 10% level, ** denotes significance at 5% level, and *** denotes significance at 1% level.

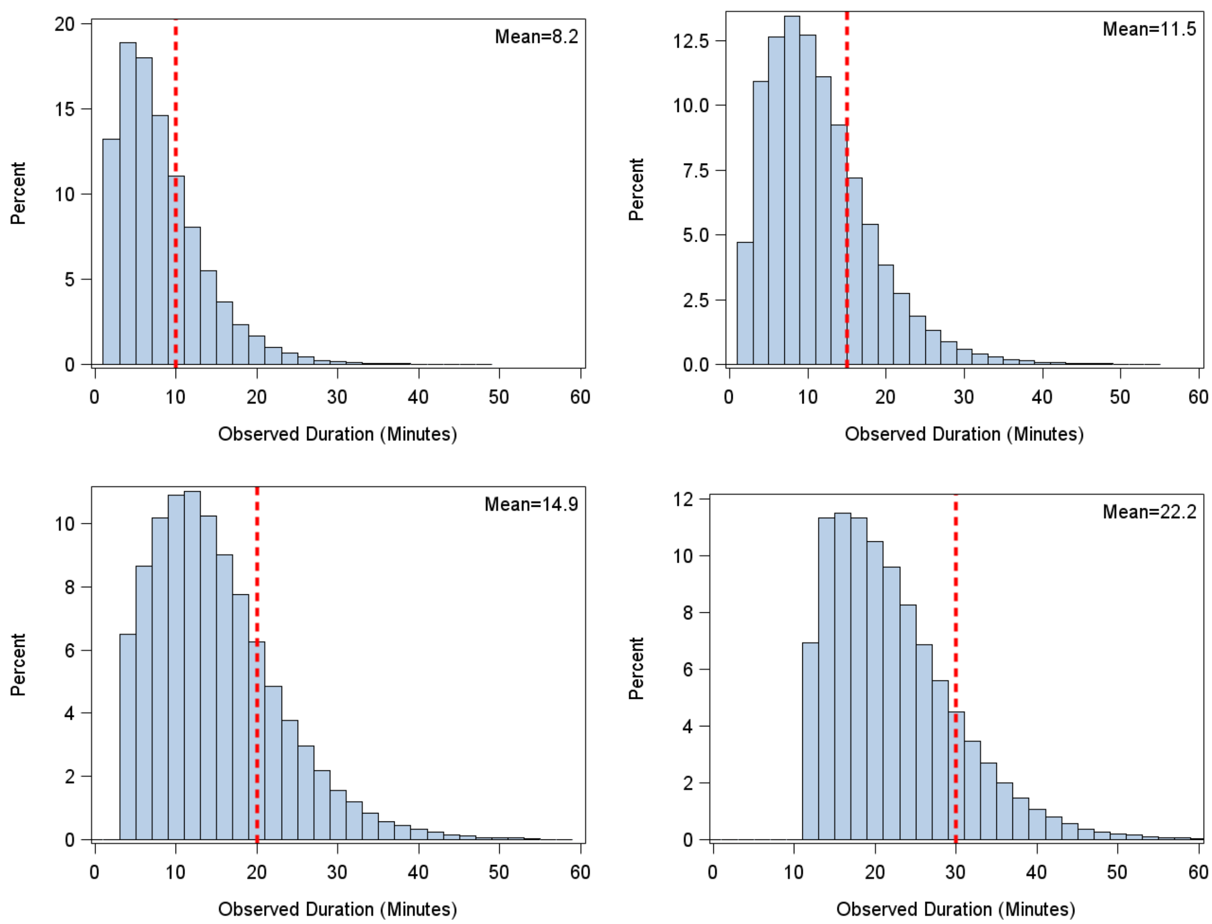
9 Figures

Figure 1: Appointment Stages in the athenahealth, Inc., Electronic Health Record



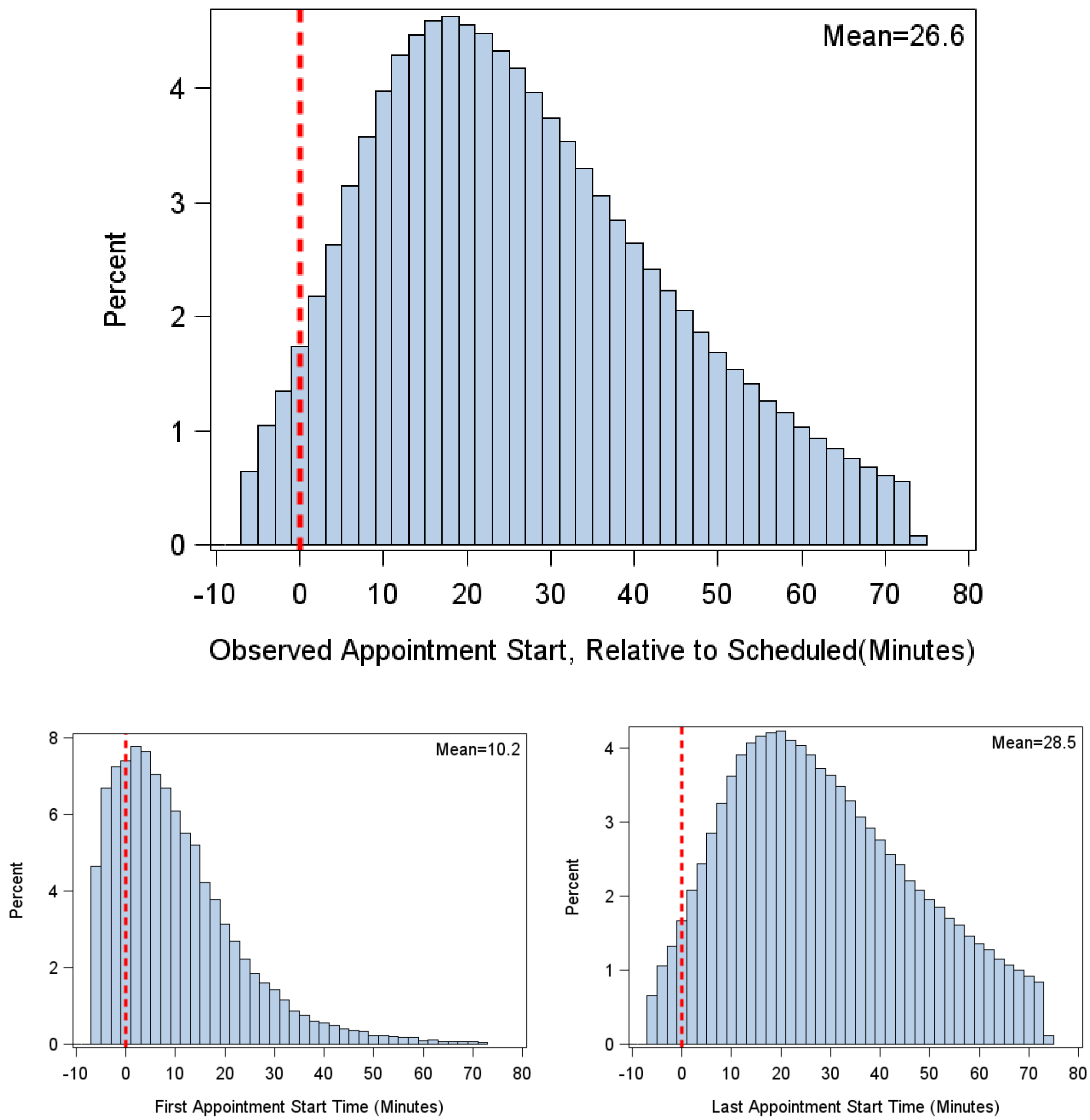
Note: This figure shows a graphical depiction of sequential appointment stages within the athenahealth electronic health record. Arrows indicate the flow through an appointment. Dotted lines indicate that there may be time between each stage. Each stage with recorded time stamps has a time estimate, calculated as the average observed duration within my analytic sample. Bold lines indicate the stages (checkin and exam) where I rely on time stamps for my analysis.

Figure 2: Observed Exam Duration by Scheduled Duration



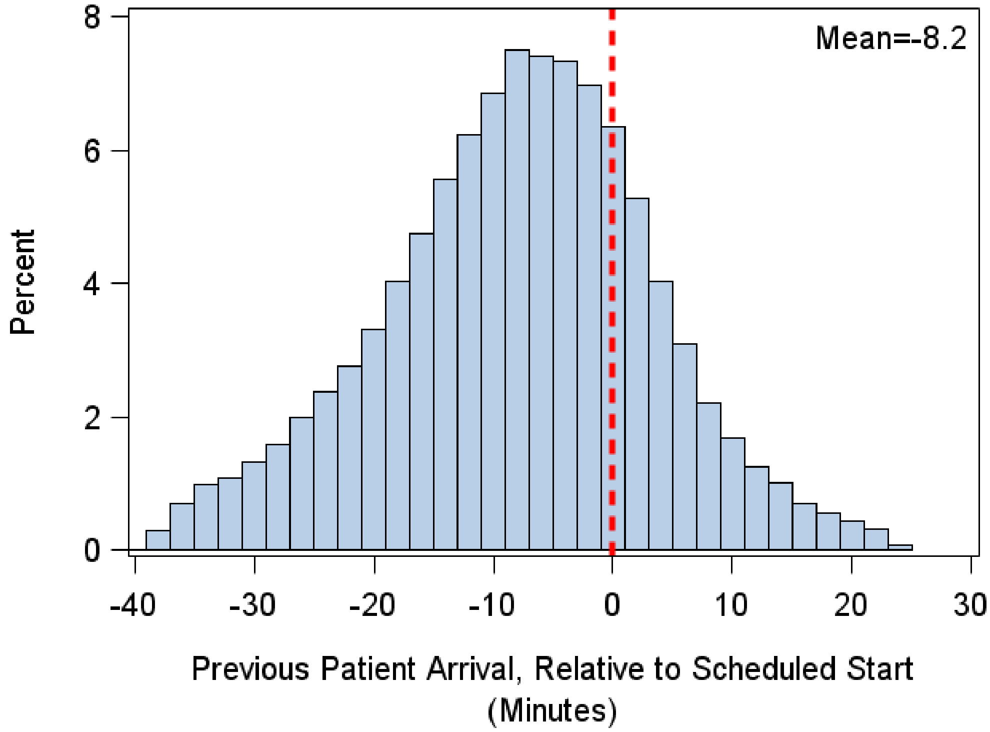
Note: This figure plots the distribution of observed appointment duration for appointments of scheduled durations including 10, 15, 20, and 30 minutes. I trim the top and bottom 5% of appointments, to reduce the influence of outliers on my findings. Dotted red lines display the scheduled appointment duration for each distribution of observed appointment duration.

Figure 3: Distribution of Observed Appointment Start Time



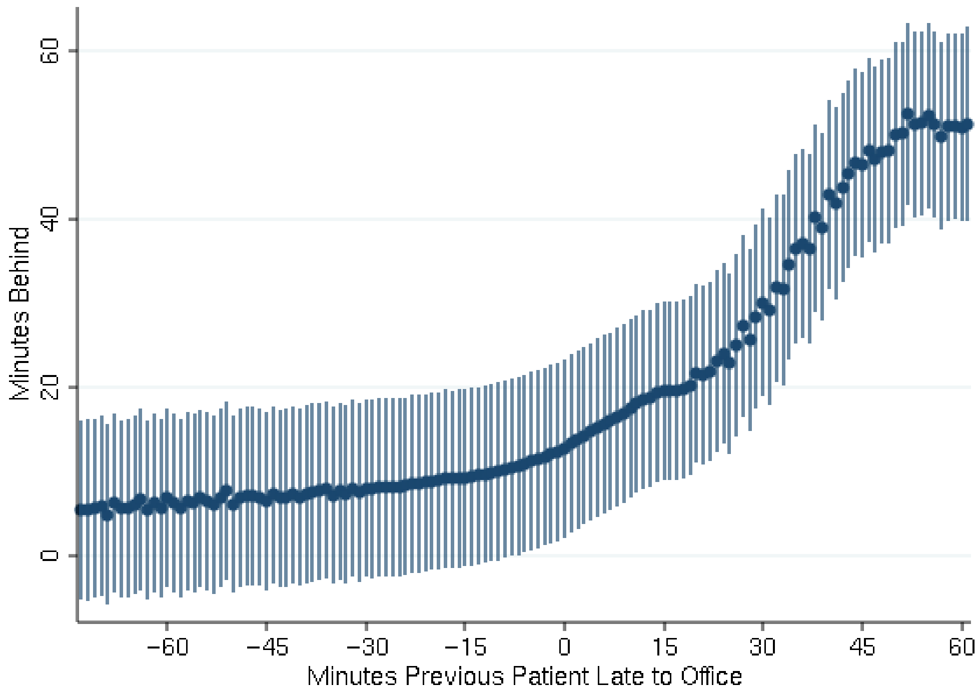
Note: This figure plots the distribution of observed appointment start time, relative to the scheduled start time. The first panel plots the distribution for the entire analytic sample. The second and third panels plot the distribution for only first or last appointments, respectively. Dotted red lines display an on-time appointment start (observed appointment start time minus scheduled start time is equal to zero).

Figure 4: Distribution of Previous Patient Arrival Time



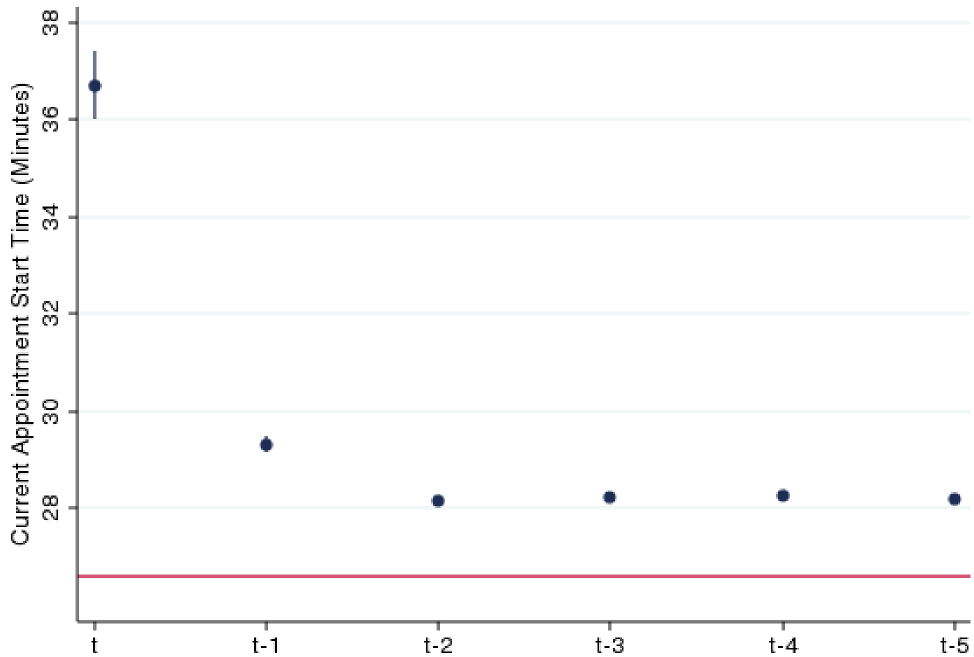
Note: This figure plots the distribution of previous patient office arrival time, relative to the scheduled appointment start time. Dotted red lines display an on-time office arrival time (previous patient office arrival time minus scheduled start time is equal to zero).

Figure 5: Appointment Start Time as a Function of Previous Patient Office Arrival Time



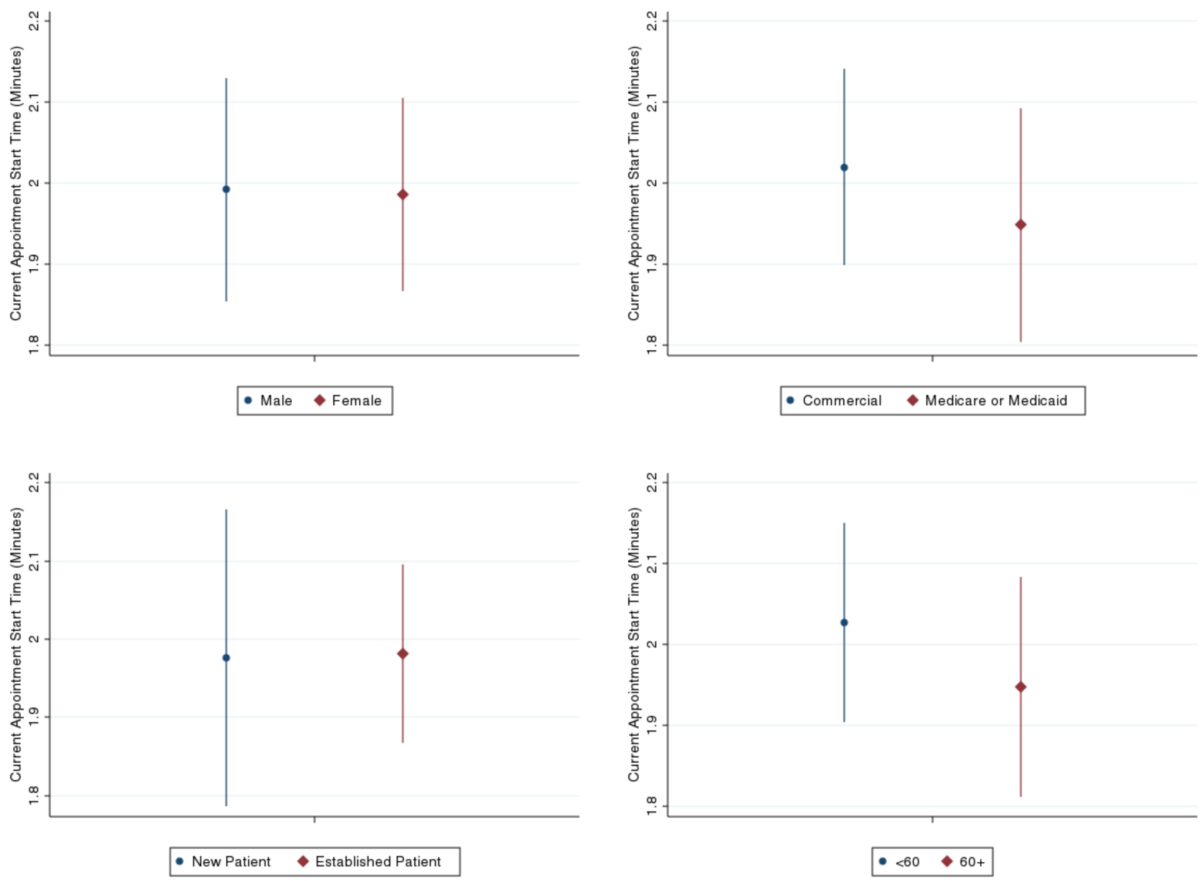
Note: This figure plots average minutes behind schedule as a function of previous patient arrival time, with standard error bars. Estimates are adjusted for physician fixed effects, so all comparisons are within-physician.

Figure 6: Effect of a 15-Minute Late Patient Arrival at Time t through $t-5$ on Appointment Start at Time t



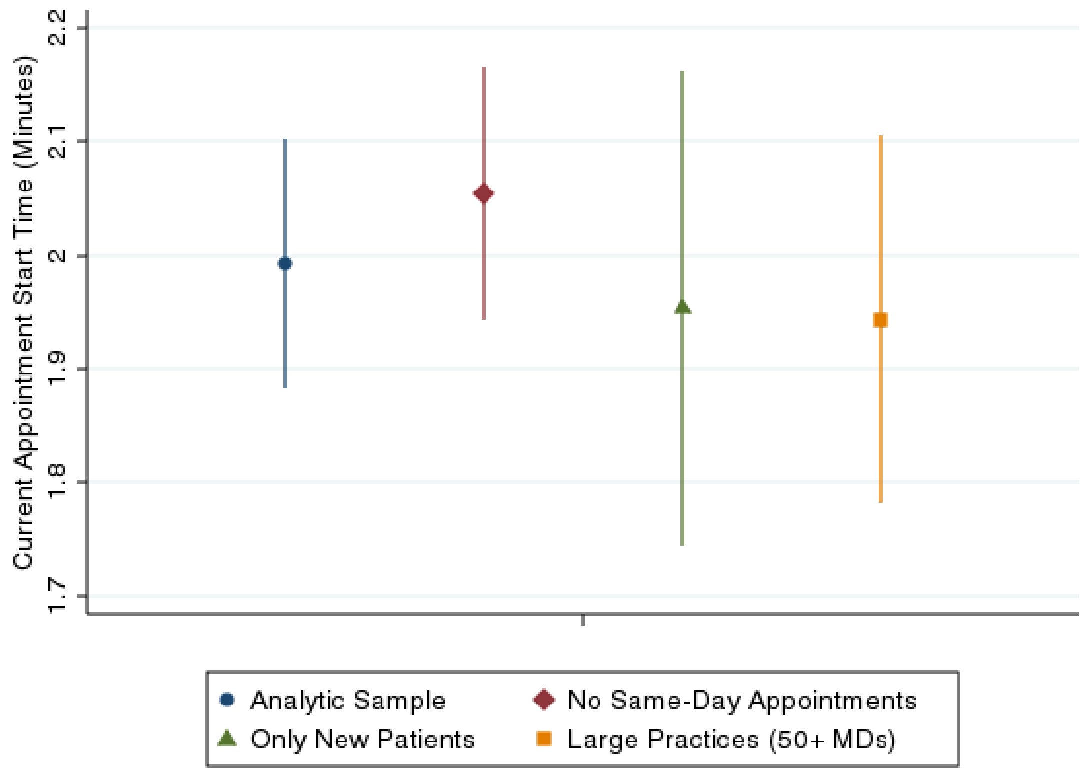
Note: This figure plots the appointment start time at time t as a function of a 15-minute late patient arrival at time t through $t-5$. Full controls are used, including patient characteristics (age, gender, insurance, new patient status, and chronic condition indicators), appointment characteristics (scheduled duration and indicators for whether the appointment is a same-day visit or double-booked), and time fixed effects (appointment rank within the day, day of the week, season, and year). Estimates are adjusted for physician fixed effects, so all comparisons are within-physician. Vertical bars display 95% confidence intervals. The horizontal line plots the average appointment start time (relative to the scheduled start time).

Figure 7: First Stage Estimates, by Mutually Exclusive Patient and Appointment Characteristics



Note: This figure presents first stage results of Equation 3, predicting current appointment start time as a function of the previous patient's office arrival time, using distinct samples that vary on patient and appointment characteristics. Each point depicts the change in appointment start time resulting from each additional minute of previous patient office arrival time, with bars showing 95% confidence intervals. Full controls are used. Standard errors are clustered at the physician level.

Figure 8: First Stage Estimates, Excluding Appointments Likely to Exhibit Selection



Note: This figure presents first stage results of Equation 3, predicting current appointment start time as a function of the previous patient's office arrival time, excluding appointments likely to exhibit selection due to certain patient or appointment characteristics. Each point depicts the change in appointment start time resulting from a 15-minute late previous patient, with bars showing 95% confidence intervals. The left-most point presents first stage regression results using my main sample of consecutive appointments. Point 2 narrows the sample to all appointments scheduled before the day of the appointment. Point 3 includes only appointments following new patients. The right-most point includes only appointments for physicians who work in a large group practice (≥ 50 physicians). Full controls are used. Standard errors are clustered at the physician level.

10 Appendix A

Table 1: Analytic Sample versus National Ambulatory Medicare Care Survey Estimates

	(1) CPT Count	(2) Spending ($)	(3) Diagnosis Count	(4) Post-Visit Documentation	(5) Revisit w/in 2 Weeks
Observed Start Time	-0.0077***	-0.1489	-0.0050***	0.0013***	0.0006**
	(0.0016)	(0.0985)	(0.0011)	(0.0004)	(0.0003)
30-39	0.0176*	0.8490**	0.1607***	0.0060***	-0.0006
	(0.0093)	(0.3879)	(0.0069)	(0.0017)	(0.0016)
40-49	0.0720***	5.0442***	0.3286***	0.0135***	0.0034*
	(0.0123)	(0.5671)	(0.0084)	(0.0018)	(0.0018)
50-59	0.1089***	6.4033***	0.4538***	0.0186***	0.0055***
	(0.0152)	(0.5845)	(0.0100)	(0.0020)	(0.0018)
60-69	0.1797***	4.0146***	0.5194***	0.0196***	0.0112***
	(0.0174)	(0.7463)	(0.0112)	(0.0022)	(0.0020)
70+	0.1654***	-4.7927***	0.5457***	0.0179***	0.0207***
	(0.0201)	(1.0967)	(0.0134)	(0.0025)	(0.0021)
Male	0.0179***	1.6331***	0.0062*	0.0016**	0.0003
	(0.0052)	(0.3100)	(0.0037)	(0.0008)	(0.0007)
Medicare	0.2271***	25.7797***	-0.0621***	-0.0056**	-0.0156***
	(0.0177)	(1.1074)	(0.0097)	(0.0026)	(0.0019)
Other	-0.2389***	-18.9272***	-0.2937***	-0.0061	0.0064
	(0.0609)	(2.6941)	(0.0380)	(0.0093)	(0.0071)
Commercial	0.0785***	27.3001***	-0.1357***	-0.0068***	-0.0218***
	(0.0164)	(1.2239)	(0.0102)	(0.0026)	(0.0018)
Uninsured	-0.2304***	-85.2876***	-0.3115***	0.0099**	-0.0551***
	(0.0265)	(1.3465)	(0.0160)	(0.0044)	(0.0028)
Worker's Comp	-0.2945***	10.5423***	-0.6044***	0.0032	0.0666***
	(0.0381)	(2.3433)	(0.0441)	(0.0070)	(0.0102)
15 Minutes	0.2574***	11.3098***	0.2037***	0.0008	-0.0110**
	(0.0793)	(2.2521)	(0.0437)	(0.0158)	(0.0048)
20 Minutes	0.3487***	18.4650***	0.3759***	0.0007	-0.0086**
	(0.0514)	(1.9302)	(0.0392)	(0.0097)	(0.0041)
30 Minutes	0.9121***	49.2970***	0.7880***	0.0242	-0.0009
	(0.0845)	(2.3494)	(0.0469)	(0.0151)	(0.0048)
Same-Day Appt	-0.0429***	-8.2005***	-0.6023***	-0.0100***	0.0592***
	(0.0148)	(0.5380)	(0.0116)	(0.0016)	(0.0018)
New Patient	0.0443**	15.8062***	-0.1161***	0.0120***	0.0060***
	(0.0192)	(0.9557)	(0.0121)	(0.0025)	(0.0019)
Double-Booked	0.0279***	0.4784	-0.0413***	0.0015	0.0148***
	(0.0108)	(1.1447)	(0.0078)	(0.0020)	(0.0018)

Appt Rank 3	0.0059	0.5430	-0.0012	0.0042*	0.0010
	(0.0111)	(0.5424)	(0.0072)	(0.0023)	(0.0017)
Appt Rank 4	-0.0103	-0.2056	-0.0134	0.0087***	-0.0006
	(0.0144)	(0.7101)	(0.0090)	(0.0030)	(0.0022)
Appt Rank 5	-0.0114	-0.2905	-0.0208**	0.0139***	-0.0002
	(0.0166)	(1.2922)	(0.0102)	(0.0035)	(0.0027)
Appt Rank 6	-0.0256	-0.8893	-0.0301***	0.0146***	0.0004
	(0.0179)	(1.1385)	(0.0112)	(0.0040)	(0.0029)
Appt Rank 7	-0.0445**	-2.0786*	-0.0456***	0.0182***	-0.0001
	(0.0195)	(1.2462)	(0.0120)	(0.0043)	(0.0031)
Appt Rank 8	-0.0715***	-2.7754**	-0.0518***	0.0228***	0.0011
	(0.0199)	(1.3595)	(0.0124)	(0.0044)	(0.0033)
Appt Rank 9	-0.0947***	-2.9268**	-0.0735***	0.0225***	0.0000
	(0.0196)	(1.2135)	(0.0125)	(0.0046)	(0.0033)
Appt Rank 10	-0.1140***	-3.5941***	-0.0811***	0.0187***	0.0019
	(0.0191)	(1.0602)	(0.0124)	(0.0044)	(0.0032)
Appt Rank 11	-0.1325***	-4.3871***	-0.0920***	0.0222***	0.0037
	(0.0189)	(1.0905)	(0.0118)	(0.0044)	(0.0032)
Appt Rank 12	-0.1577***	-4.2086***	-0.1052***	0.0240***	0.0036
	(0.0189)	(1.1173)	(0.0119)	(0.0043)	(0.0032)
Appt Rank 13	-0.1568***	-5.5931***	-0.1128***	0.0214***	0.0006
	(0.0198)	(0.9942)	(0.0121)	(0.0042)	(0.0030)
Appt Rank 14	-0.1781***	-5.5925***	-0.1094***	0.0226***	0.0023
	(0.0198)	(1.3606)	(0.0124)	(0.0044)	(0.0032)
Appt Rank 15	-0.1849***	-6.2993***	-0.1235***	0.0246***	0.0007
	(0.0202)	(1.0883)	(0.0126)	(0.0045)	(0.0031)
Appt Rank 16	-0.1982***	-6.7291***	-0.1374***	0.0256***	-0.0011
	(0.0205)	(1.3765)	(0.0133)	(0.0045)	(0.0033)
Appt Rank 17	-0.2007***	-6.9051***	-0.1317***	0.0239***	0.0006
	(0.0215)	(1.3590)	(0.0135)	(0.0047)	(0.0033)
Appt Rank 18	-0.2130***	-7.5993***	-0.1527***	0.0248***	-0.0032
	(0.0233)	(1.5476)	(0.0144)	(0.0050)	(0.0035)
Appt Rank 19	-0.2342***	-9.1742***	-0.1479***	0.0265***	-0.0004
	(0.0241)	(2.1446)	(0.0146)	(0.0050)	(0.0035)
Appt Rank 20	-0.2401***	-8.4437***	-0.1685***	0.0270***	-0.0025
	(0.0256)	(1.7152)	(0.0159)	(0.0053)	(0.0039)
Appt Rank 21	-0.2824***	-8.4357***	-0.1654***	0.0275***	-0.0019
	(0.0277)	(1.7583)	(0.0174)	(0.0054)	(0.0041)
Appt Rank 22	-0.2775***	-9.3848***	-0.1827***	0.0217***	-0.0044
	(0.0297)	(1.9949)	(0.0174)	(0.0059)	(0.0041)
Appt Rank 23	-0.2868***	-11.6984***	-0.1884***	0.0261***	-0.0040
	(0.0324)	(3.5395)	(0.0191)	(0.0060)	(0.0047)
Appt Rank 24	-0.3136***	-9.9714***	-0.2063***	0.0292***	0.0044
	(0.0327)	(1.9658)	(0.0203)	(0.0066)	(0.0053)
Appt Rank 25	-0.3035***	-12.5510***	-0.1810***	0.0257***	-0.0056
	(0.0347)	(3.0392)	(0.0202)	(0.0069)	(0.0051)

Appt Rank 26	-0.2932***	-13.1632***	-0.2216***	0.0285***	-0.0022
	(0.0371)	(3.2687)	(0.0222)	(0.0081)	(0.0056)
Appt Rank 27	-0.3350***	-8.1362	-0.1999***	0.0290***	0.0018
	(0.0366)	(11.1420)	(0.0230)	(0.0084)	(0.0060)
Appt Rank 28	-0.3574***	-14.0445***	-0.2193***	0.0195**	-0.0009
	(0.0400)	(3.4735)	(0.0262)	(0.0077)	(0.0065)
Appt Rank 29	-0.3417***	-16.0214***	-0.2340***	0.0344***	0.0001
	(0.0474)	(5.3034)	(0.0297)	(0.0098)	(0.0065)
Appt Rank 30	-0.2520***	-9.2051***	-0.2110***	0.0301***	-0.0042
	(0.0506)	(2.2071)	(0.0337)	(0.0091)	(0.0079)
Appt Rank 31	-0.3950***	-13.6030***	-0.2489***	0.0222**	0.0167*
	(0.0537)	(2.6236)	(0.0350)	(0.0102)	(0.0094)
Appt Rank 32	-0.3797***	-9.6954***	-0.2760***	0.0187	-0.0219*
	(0.0739)	(2.2400)	(0.0305)	(0.0160)	(0.0117)
Appt Rank 33	-0.3615***	-14.0704***	-0.2644***	0.0285**	-0.0191*
	(0.0648)	(2.1978)	(0.0422)	(0.0119)	(0.0103)
Appt Rank 34	-0.3771***	-15.6563***	-0.1812***	0.0143	-0.0127
	(0.0964)	(4.2813)	(0.0375)	(0.0121)	(0.0134)
Appt Rank 35	-0.5342***	-16.3697***	-0.3005***	0.0472*	-0.0126
	(0.0901)	(4.6851)	(0.0373)	(0.0264)	(0.0157)
Appt Rank 36	-0.5130***	-14.2701***	-0.3008***	0.0302*	-0.0020
	(0.1064)	(3.0361)	(0.0623)	(0.0172)	(0.0164)
Appt Rank 37	-0.4348***	-16.4210**	-0.2582***	0.0406*	-0.0270*
	(0.0941)	(8.0996)	(0.0639)	(0.0212)	(0.0164)
Appt Rank 38	-0.5057***	-12.3148***	-0.2755***	0.0361	-0.0303
	(0.1295)	(4.6050)	(0.0841)	(0.0245)	(0.0235)
Appt Rank 39	-0.3879***	-9.2831**	-0.1601	0.0018	-0.0061
	(0.1121)	(3.6651)	(0.0994)	(0.0229)	(0.0210)
Appt Rank 40	-0.4766***	-17.6386***	-0.3519***	-0.0051	-0.0233
	(0.1324)	(5.7062)	(0.0763)	(0.0373)	(0.0290)
Tuesday	0.0111*	0.7785	-0.0251***	0.0029**	0.0043***
	(0.0064)	(0.5441)	(0.0072)	(0.0012)	(0.0010)
Wednesday	-0.0020	1.1977	-0.0239***	0.0062***	-0.0065***
	(0.0057)	(0.9590)	(0.0045)	(0.0013)	(0.0010)
Thursday	0.0004	1.2689	-0.0244***	0.0074***	0.0003
	(0.0067)	(1.0155)	(0.0041)	(0.0014)	(0.0010)
Friday	-0.0084	-0.1790	-0.0559***	0.0065***	0.0060***
	(0.0059)	(0.5797)	(0.0046)	(0.0014)	(0.0011)
Spring	-0.2474***	-5.2422***	-0.0715***	-0.0004	-0.0498***
	(0.0105)	(0.2768)	(0.0060)	(0.0030)	(0.0013)
Summer	-0.1521***	-2.8713***	-0.0378***	-0.0060***	-0.0512***
	(0.0082)	(0.4084)	(0.0052)	(0.0023)	(0.0013)
Winter	-0.2100***	-4.9523***	-0.0626***	0.0038	-0.0461***
	(0.0102)	(0.2934)	(0.0062)	(0.0035)	(0.0014)
2014	-0.1263***	2.7928***	-0.0109	-0.0049	0.0227***
	(0.0118)	(0.2967)	(0.0073)	(0.0052)	(0.0011)

Hypothyroidism	0.0172**	-0.0622	0.0998***	0.0015	-0.0025**
	(0.0082)	(0.4547)	(0.0053)	(0.0012)	(0.0011)
AMI	0.0065	2.3921	0.0342	0.0066	0.0041
	(0.0346)	(1.6942)	(0.0258)	(0.0087)	(0.0071)
Alzheimers	0.0050	-0.3778	0.0082	-0.0010	-0.0047
	(0.0333)	(1.1772)	(0.0238)	(0.0063)	(0.0065)
_Icc_alzheia1	-0.0252	0.2162	0.0401***	-0.0016	-0.0084**
	(0.0176)	(0.8585)	(0.0138)	(0.0038)	(0.0036)
Anemia	0.0737***	4.7793	0.0590***	0.0068***	0.0130***
	(0.0129)	(3.5511)	(0.0073)	(0.0016)	(0.0018)
Asthma	-0.0198**	-0.5568	0.0497***	0.0009	0.0131***
	(0.0090)	(0.8049)	(0.0068)	(0.0017)	(0.0016)
Atrial Fibrillation	0.0435***	-3.0281***	0.0873***	0.0026	0.0294***
	(0.0101)	(0.9927)	(0.0079)	(0.0019)	(0.0023)
Benign Prostatic Hyperplasia	-0.0161	-2.2818***	-0.0250***	-0.0037	-0.0056***
	(0.0136)	(0.6238)	(0.0096)	(0.0023)	(0.0021)
Cataract	0.0151	-5.0795***	-0.0224	0.0016	0.0061*
	(0.0203)	(1.6751)	(0.0162)	(0.0038)	(0.0034)
CKD	-0.0106	-0.6475	0.0798***	0.0042*	0.0094***
	(0.0152)	(1.0724)	(0.0109)	(0.0022)	(0.0017)
Pulmonary Disease	-0.0118	-2.0767***	0.0664***	0.0001	0.0184***
	(0.0099)	(0.7307)	(0.0073)	(0.0018)	(0.0016)
Depression	-0.0697***	-3.0879***	0.1103***	0.0068***	0.0064***
	(0.0084)	(0.5809)	(0.0065)	(0.0014)	(0.0012)
Diabetes	0.1699***	2.5050***	0.2247***	0.0082***	0.0072***
	(0.0128)	(0.3788)	(0.0067)	(0.0011)	(0.0009)
Glaucoma	0.0376	-8.2569*	-0.0573**	0.0010	0.0016
	(0.0375)	(4.4173)	(0.0273)	(0.0069)	(0.0061)
Heart Failure	-0.0263**	-1.7084**	0.0784***	0.0060***	0.0146***
	(0.0112)	(0.8557)	(0.0096)	(0.0020)	(0.0023)
Hip or Pelvic Fracture	-0.0294	1.5744	-0.0071	-0.0026	-0.0022
	(0.0408)	(5.3950)	(0.0305)	(0.0089)	(0.0093)
Hyperlipidemia	0.0103	-1.6527**	0.0750***	-0.0040***	-0.0097***
	(0.0087)	(0.7086)	(0.0061)	(0.0013)	(0.0009)
Hypertension	-0.0223***	-0.1907	0.1424***	-0.0018	-0.0025***
	(0.0066)	(0.6523)	(0.0050)	(0.0015)	(0.0009)
Ischemic Heart Disease	-0.0272***	-2.1328**	0.1160***	0.0034**	0.0042***
	(0.0077)	(0.9407)	(0.0076)	(0.0016)	(0.0014)
Osteoporosis	0.0319*	1.6615	-0.0176	0.0001	-0.0022
	(0.0187)	(1.7578)	(0.0111)	(0.0022)	(0.0017)
Arthritis	0.0056	-0.3428	0.0319***	-0.0007	0.0056***
	(0.0121)	(0.4944)	(0.0090)	(0.0016)	(0.0012)
Stroke	-0.0177	-0.7515	0.0477***	0.0010	0.0091***
	(0.0136)	(0.6818)	(0.0105)	(0.0029)	(0.0028)
Breast Cancer	-0.0367	16.9149	0.0025	0.0036	-0.0021
	(0.0247)	(11.0277)	(0.0139)	(0.0037)	(0.0040)

Colorectal Cancer	0.1285	23.7523	0.0636***	0.0004	-0.0071
	(0.0984)	(15.8122)	(0.0223)	(0.0055)	(0.0068)
Prostate Cancer	0.0065	2.5787	0.0236	-0.0033	-0.0063*
	(0.0253)	(3.2682)	(0.0152)	(0.0041)	(0.0036)
Lung Cancer	0.0443	41.1054	0.0838**	0.0054	0.0173*
	(0.0673)	(35.1478)	(0.0351)	(0.0064)	(0.0092)
Endometrial Cancer	-0.1091	-32.4998	-0.0226	-0.0020	0.0244
	(0.0766)	(30.9788)	(0.0631)	(0.0168)	(0.0219)
Observations	1,194,617	1,194,617	1,194,617	1,194,617	1,194,617

Note: This presents full regression results for Panel 1 of Table 6.

Table 2: Analytic Sample versus National Ambulatory Medicare Care Survey Estimates

	NAMCS	Analytic Sample
Age		
Category 1 (Youngest)	0.091	0.061
Category 2	0.239	0.225
Category 3	0.352	0.419
Category 4 (Oldest)	0.317	0.268
Female	0.582	0.577
Insurance		
Commercial	0.602	0.602
Medicare	0.249	0.322
Medicaid	0.127	0.048
No Insurance	0.048	0.023
Workers' Compensation	0.014	0.003
Chronic Condition Count		
0	0.449	0.470
1	0.236	0.134
2	0.133	0.129
3+	0.149	0.267
Geography		
Northeast	0.199	0.263
Midwest	0.184	0.194
South	0.396	0.427
West	0.222	0.124
New Patient	0.159	0.137
Practice Size		
Solo	0.340	0.104
2	0.095	0.058
3-5	0.263	0.069
6-10	0.174	0.072
11+	0.123	0.697

Note: This table presents appointment-level descriptive statistics on a nationally representative sample of office visits, compared to the analytic sample of 2013-2014 athenahealth appointments with office-based primary care physicians. NAMCS is the National Ambulatory Medical Care Survey. NAMCS insurance categories are not mutually exclusive and therefore do not sum to one. NAMCS age categories do not match those in athenahealth and are therefore assigned to relative categories. Age category 1 is 15-24 in NAMCS and 20-29 in the analytic sample. Category 2 is 24-44 and 30-49. Category 3 is 45-64 and 50-69. Category 4 is 65+ and 60+.
Source: Authors' analyses and National Ambulatory Medical Care Survey (NAMCS) Summary Tables from the 2012 NAMCS.

Table 3: Robustness Check: Alternative Instruments

	(1) $t-1$	(2) $t-2$	(3) $t-3$	(4) $t-4$	(5) $t-5$
Observed Duration	-0.1764***	-0.0599***	-0.0488***	-0.0400***	-0.0254***
	(0.0078)	(0.0077)	(0.0094)	(0.0103)	(0.0130)
Procedure Count	-0.0077***	-0.0019	0.0009	0.0034	0.0097**
	(0.0016)	(0.0022)	(0.0022)	(0.0037)	(0.0048)
Spending	-0.1489	-0.2568*	0.1931	-0.1467	1.9321
	(0.0985)	(0.1451)	(0.2496)	(0.1585)	(1.2361)
Diagnosis Count	-0.0050***	-0.0017	0.0002	-0.0002	0.0043*
	(0.0011)	(0.0013)	(0.0015)	(0.0019)	(0.0026)
Post-Visit Documentation	0.0013***	0.0007*	0.0011**	0.0004	0.0009
	(0.0004)	(0.0004)	(0.0005)	(0.0006)	(0.0008)
Revisit within 2 Weeks	0.0006**	0.0000	-0.0003	-0.0002	0.0007
	(0.0003)	(0.0003)	(0.0004)	(0.0005)	(0.0007)
Observations	1,194,617	1.090,822	984,954	884,421	790,754

Note: This replicates elements of Tables 5 and 6, using alternative instruments of patient office arrival time at $t-2$ through $t-5$. Full controls are used. Standard errors are clustered at the physician level. * denotes significance at 10% level, ** denotes significance at 5% level, and *** denotes significance at 1% level.

Table 4: Robustness Check: Binary and Log Instrument

	2SLS (Instrument: Previous Patient Late Y/N)
Observed Duration	-0.2211***
	(0.0140)
Procedure Count	-0.0086***
	(0.0025)
Spending	0.1076
	(0.2030)
Diagnosis Count	-0.0052***
	(0.0019)
Post-Visit Documentation	0.0019***
	(0.0007)
Revisit within 2 Weeks	0.0006
	(0.0005)
First Stage	
Appointment Start Time	1.3466***
	(0.0612)

	2SLS (Instrument: Log of Previous Patient Arrival Time)
Log Observed Duration	-0.4888***
	(0.0512)
Procedure Count	-0.3121***
	(0.0936)
Log Spending	-0.0522**
	(0.0261)
Diagnosis Count	-0.1546***
	(0.0653)
Post-Visit Documentation	0.0707***
	(0.0222)
Revisit within 2 Weeks	0.0176***
	(0.0017)
First Stage	
Appointment Start Time	0.0209***
	(0.0020)

Note: This replicates elements of Tables 5 and 6, instrumenting for current appointment start time using a) a binary indicator for a previous patient arriving to the office after his or her scheduled appointment start time and b) the log of the previous patient's arrival time. Full controls are used. Standard errors are clustered at the physician level. * denotes significance at 10% level, ** denotes significance at 5% level, and *** denotes significance at 1% level.

Table 5: Placebo Test: Subsequent Patient Arrival Instruments

	Current Appointment (t) Start Time (Minutes Behind)
Primary Instrument	
Patient Arrival for Appointment $t-1$	0.0858***
(Previous Appt)	(0.0024)
Alternative Instruments	
Patient Arrival for Appointment $t+1$	-0.2336
	(0.7476
Patient Arrival for Appointment $t+2$	-0.3162
	(0.7726)

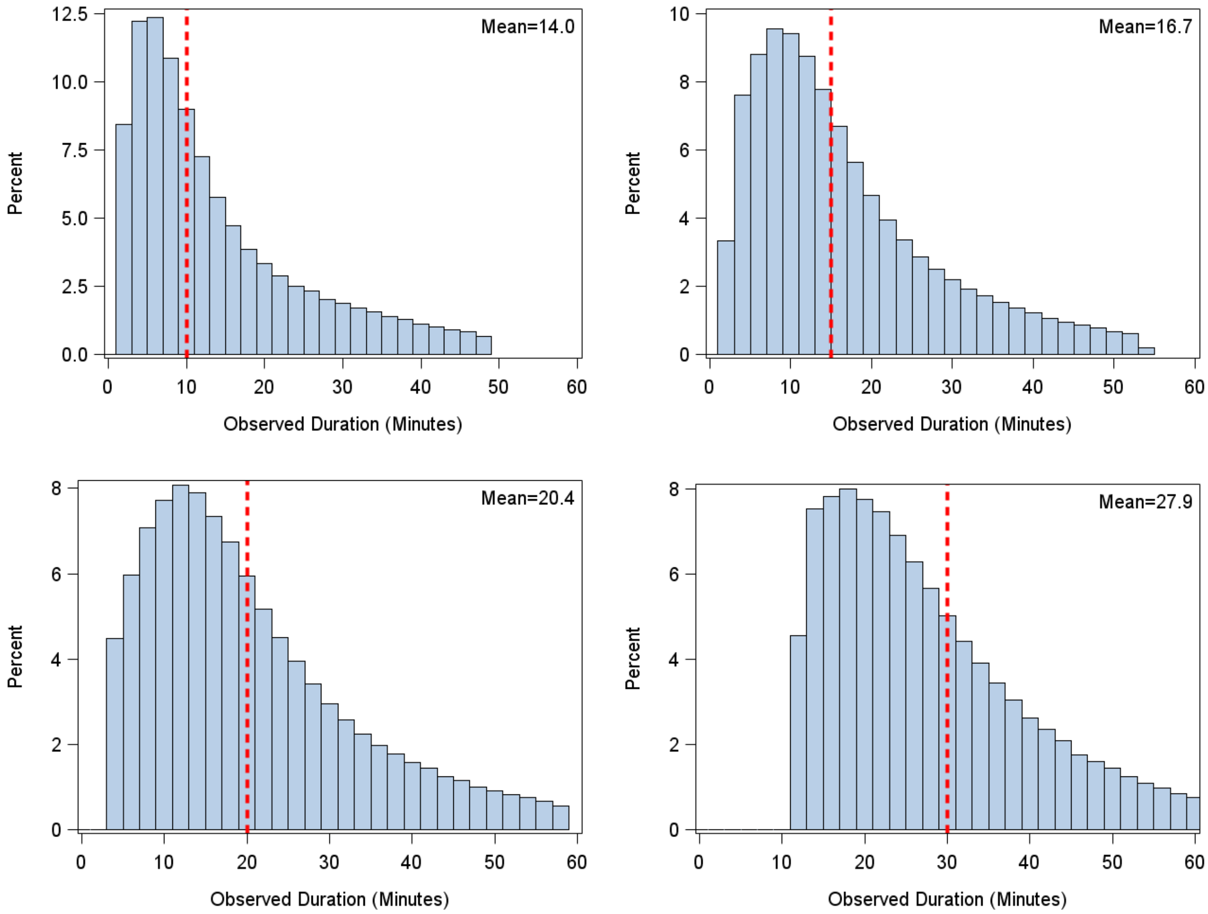
Note: This table presents first stage results of my instrumental variables estimating equation, predicting current appointment start time (minutes behind) as a function of the previous patient's office arrival time ($t-1$), the arrival time of the patients 1 and 2 appointments later. Full controls are used. Standard errors are clustered at the physician level. * denotes significance at 10% level, ** denotes significance at 5% level, and *** denotes significance at 1% level.

Table 6: Placebo Test: Revisit Within Two Weeks

	(1) 2SLS
Revisit Within 2 Weeks	0.0006**
	(0.0003)
Previously Scheduled Revisit Within 2 Weeks	0.0002
	(0.0002)

Note: Full controls are used. Standard errors are clustered at the physician level. * denotes significance at 10% level, ** denotes significance at 5% level, and *** denotes significance at 1% level.

Figure 1: Observed Exam Duration by Scheduled Duration



Note: This figure replicates Figure 2, without excluding appointments with overlapping timestamps.

11 Appendix B: Sample Selection

Table 1: Sample Definition

Sample Description or Step	n =
Raw time stamp data	72,674,234
Restrict to office or HOPD-based PCPs	19,866,414
Exclude weekend appointments, those with durations other than 10, 15, 20, or 30 minutes, pediatric appointments, those on days with < 7 or > 40, those that start before 8am or after 7pm	11,135,787
Exclude appointments with anomalous time stamps	5,695,661
Exclude non-consecutive appointments and those with overlapping time stamps	1,194,617

Note: This table describes each step in sample construction. HOPD is hospital outpatient department. PCP is primary care physician. Primary care physicians are defined as those listing internal medicine, general practice, family practice as their primary specialty.