Time Constraints and Productivity in Health Care

Miguel Alquézar-Yus*

August, 2022

Abstract

This paper explores how reviewing time affects physicians' medical decisions. Insufficient examination time may hamper physicians' care and diagnostic provision, leaving physicians more inclined to over-prescribe medication. I test this prediction using high-frequency data from a Spanish outpatient department and leverage on-the-day cancellations as exogenous time shocks. I find that longer visits lead to more valuable care, measured by the provision of more detailed diagnoses, to higher testing intensity, and to lower drug prescriptions. These effects are driven by junior physicians, who use this extra time to compensate for their more overloaded shifts.

Keywords: Healthcare productivity, quantity-quality trade-off, contracts, decision-making

JEL Classification: H0; I0; J0

*Department of Economics, European University Institute, Via delle Fontanelle, 18, 50014 Fiesole, Italy, miguel.alquezar@eui.eu. I wish to thank Sule Alan, Josep Amer-Mestre, Michèle Belot, Elisa Failache, Christian Fons-Rosen, Ezra Golberstein, Annika Herr, Andrea Ichino, Maria Ptashkina, Ity Shurtz, and Ana Tur-Prats, together with seminar participants at EuHEA PhD conference 2021, WIPE 2022, COPE 2022, SOLE 2022, XV RIDGE Forum, SMYE 2022, ESPE 2022, EALE 2022, Universitat Pompeu Fabra, and EUI, for their helpful comments. A special thanks to the outpatient department for their support throughout the process. Financial support from the Salvador Madariaga-EUI scholarship is gratefully acknowledged. All errors remain our own.

1 Introduction

Working under critical time pressure has become a hallmark of today's economy. According to a survey conducted by Eurofound (2017), 36% of the workers in the European Union work under tight deadlines, while 10% report not having sufficient time to complete their tasks. These structural deficiencies may hinder and undermine decision-making, especially in high-stakes situations, possibly leading to long-lasting social costs.[1]

In this paper, I study the causal relationship between task's completion time and its success rate. More concretely, I examine how the time allocated for reviewing patients affects physicians' medical decisions. This research is especially relevant today, as aging populations and costly technological innovations are increasingly demanding expansions in health care spending (Dieleman et al., 2017). According to Eurofound (2017), 14% of healthcare workers report having insufficient time to do their job, while also being the sector most affected by high emotional demands (22% of the respondents highlighted this issue). In this context, learning how to better manage existing resources, such as human capital, could have sizable welfare-improving implications.

The main empirical challenge to estimating the causal effect of visit length on physicians' decisions is to obtain a relevant source of time that is also exogenous to the patients' characteristics. This is essential as physicians have a full picture of their shift in real time, which allows them to adjust to sudden changes, but also to provide longer visits based on patients' characteristics. On the one hand, physicians may decide to spend more time with those patients found to be more challenging, allowing them to better assess whether further treatment is required. On the other hand, physicians may decide to provide patients with diagnostic inputs as substitute for extra needed time. I address this issue by leveraging on-the-day cancellations as random time shocks to the physicians' schedules, and by focusing on first visits, as new patients have no prior contact with their treating physician.

When a cancellation occurs, I find that a physician generally spends more time with all the visits taking place for the remainder of the shift, but also provides to the very next scheduled visit with an unexpected extra visit length. I focus on such *bonus* time to extract conclusions on how physicians' diagnostic behavior responds to an unexpected increase in consultation time. On-the-day cancellations represent 15% of all visits.

I examine how visit length affects diagnostic intensity and treatment choice, using de-

[1] By comparison, in 1991, only 23% of the workers in the European Union worked under tight deadlines (Eurofound, 1993).

tailed high-frequency data from a Spanish outpatient department. The main specification uses an IV approach, instrumenting the time allocated to review each patient with whether the prior scheduled visit got cancelled. I include physician fixed effects to control for inherent physicians' characteristics, and by construction, by those of their specializations; month-by-year fixed effects to control for seasonality confounders; hour fixed effects to account for different hour-patient compositions; and a wide set of controls for patients' characteristics.

I find that longer visits increase the likelihood of providing a diagnosis, which is the main objective of outpatient departments. For every extra reviewing minute, the likelihood of providing a diagnosis increases by 4.4%. This effect is driven by uncommon diagnoses, while no effect is found on the most common diagnoses, suggesting that longer visits allow physicians to review patients in more detail, and possibly to provide them with a more valuable service. Furthermore, longer first visits increase the diagnostic input utilization, in the form of procedures and laboratory tests, while decreasing drug prescriptions. For every extra minute used to review patients, physicians provide 3.3% more tests, increasing the overall testing cost by 6.3%, and reduce the drug doses prescribed by 20.1%. These results evidence that test ordering, and especially more expensive tests, complement longer visits, while drug prescription is used as a substitute for insufficient reviewing time. Overall, physicians use extra visiting time to assess in more detail the patient's health problems, and in the event of indecision, to request further diagnostic inputs, which ultimately will improve the service provided.

These findings could be, however, hiding an intertemporal input substitution decision. If this were the case, we would expect physicians who were shocked during a given first visit to inversely adjust their input utilization during the corresponding follow-up visit. I find no evidence of an input substitution effect, suggesting that longer first visits have a lasting impact in the clinical process.

I further explore which patient characteristics help explain time utilization. I find that physicians use that extra consultation time to provide a more valuable service, in terms of more testing and diagnostic provision, only to Spanish-born and female patients, while I find no significant effects tied to patient's age. In contrast, physicians use such extra time to provide a more valuable service to urgent patients. These results provide suggestive evidence that physicians use the time created by cancellations to discriminate in favor of patients sharing certain personal characteristics, but also to improve the care provided to those more in need.

Finally, I look at how physicians' contracts influence diagnostic provision. These contracts, based on seniority, provide senior physicians with less overloaded shifts, at the expense of their junior colleagues. I find that longer visits only lead to changes in the input composition and to the provision of a diagnosis when such extra time is provided to junior physicians, while the extra time has no effect on senior physicians. While both junior and senior physicians react to cancellations by providing more time to their subsequent patients, only junior physicians use such *bonus* time to effectively promote a more valuable service. With these results in mind, I provide a back-of-the-envelope calculation for the direct labor cost of increasing diagnosis rates. Policy makers could attempt to improve diagnostic rate by increasing every physician's visiting times across the board. Doing so might prove to be inefficient, however, as it does not take into account that senior physicians' practices are not affected by longer visit length. A tailored approach targeting only junior physicians, however, might help improve health provision while minimizing expenditure.

Understanding the trade-off between time and employee productivity is important from a policy perspective. On the one hand, the provision of longer visits comes at the expense of fewer visits per shift, and in equilibrium, of longer waiting lists to access the outpatient department. On the other hand, longer visits may lead to higher visit quality, which in turn might improve patients' health outcomes and reduce their need for readmission. To the best of my knowledge, this paper is the first one to provide causal evidence that longer reviewing time improves visit quality, and to show that correcting the distortionary incentives created by seniority-based contracts may be welfare-improving.

This paper contributes to different strands of the literature. First, it complements the growing literature on the determinants of physicians' labor supply. Recent literature has looked at the role of financial incentives (Powell-Jackson et al., 2015; Gupta, 2021), co-working (Chan, 2016), peer pressure (Silver, 2021), and scheduling (Chan, 2018).

More specifically, this paper complements the recent literature studying the workload-quality trade-off in the healthcare sector.[2] Mixed evidence has been found on how workload affects physicians' decisions. Shurtz et al. (2019) evaluates how physicians' decisions depend on their daily workload and finds that physicians provide higher diagnostic inputs and lower drug prescriptions on high workload days. Neprash (2016) finds that when physicians fall behind schedule, they spend less time with their subsequent visits, order fewer procedures, and provide fewer diagnoses. Freedman et al. (2021) investigates how primary care providers react to moments of high time pressure, induced by

[2] This issue has also been studied in the service industry (Tan and Netessine, 2014; Bruggen, 2015), banking industry (Xu et al., 2017), and justice system (Coviello et al., 2015), among others.

cancellations and add-ins, finding such pressure pushes physicians to provide fewer diagnostic inputs, more follow-up care, and lower referral rates. This paper is the first one to causally study physicians' direct response to longer reviewing time, as opposed to indirect measures of workload or time pressure. I further contribute to the literature by rejecting that on-the-day workload directly influences physicians' decisions, once visit duration is accounted for.

Second, this work relates to the literature on the impact of reviewing time on output quality. Theoretically, Anand et al. (2011) defines this relationship as an optimal trade-off between the number of customers served and the quality of service, in which congestion and service costs play a major role. Empirically, Song et al. (2022) studies how longer visits affect the probability of readmission in the post-acute care system, finding that there exists a negative relationship between those two variables. Frakes and Wasserman (2017) estimates the causal relationship between the time allocated to review patents and the examiners' effort. They find that lower examination time lead to reductions in the examination scrutiny and to granting patents of weaker-than-average quality. This paper contributes to this literature by causally estimating the relationship between reviewing time and physicians' performance in a setting in which physicians work as single units in a single-stage process. Furthermore, this paper looks at the incentives at play, highlighting that seniority-based contracts might lead to inefficient time-to-input utilization.

The remainder of the paper is organized as follows: Section 2 explains the institutional setting. In Section 3, I present and describe the data used. Section 4 exposes the empirical strategy followed. Section 5 presents the main results, and Section 6 provides a quantification exercise. Finally, Section 7 concludes.

2 Institutional Setting

2.1 Spanish Healthcare System

In Spain, healthcare is universal and free of charge. Its provision is structured around two main actors, Primary Care Centers and Specialized Care Centers, which together form Basic Health Zones (hereafter, BHZ). Essentially, a BHZ is an administrative unit containing several Primary Care Centers mapped into a Specialized Care Center. Individuals are sorted into different BHZs based on their place of residence. Specialized Care Centers cover multiple services such as the intensive care unit, the emergency room, and the outpatient department, all of which are usually located in hospitals.

This study focuses on the outpatient department. Initial access to this department is solely decided by the patients' treating Primary Care Center which, upon analyzing the patients' health conditions, allocates them to outpatient physicians based on the physicians' availability. The referral notification from the primary care center to the outpatient department is provided to patients some days after their visit with the general practitioner, including information on the appointment time and date and the physician's name. This implies that the individuals' place of residence fully determines their outpatient department of reference, disallowing walk-in visits and blocking patients from choosing among clinics. Moreover, the region in which the department of my sample operates, Catalonia, does not allow patients to choose their outpatient physician, minimizing any possible relationship between physicians and their patients prior to a first visit.

2.2 Hospital Management Flow

The hospital manages patients following a production-line approach. Upon arrival, patients register at the main counter, where the administration secretary gives them directions to the waiting rooms, and electronically notifies the physician in charge of the patient's arrival. From the waiting room, the physician calls patients following the appointment schedule, keeping track of who is in the waiting room. After the visit is completed, if a follow-up visit is ordered, patients return to the main counter, where they select the date and time slot of the follow-up visit within the physician's date recommendations. Throughout this process, physicians have access to real-time information on all patients' availability status and health conditions.

Figure 1 presents the type of agenda displayed to physicians. At any given point in time, a physician knows precisely which patients have not yet showed up, which patients have cancelled their visits, and which patients are already in the waiting room. Physicians are also presented with a set of patient characteristics such as the patient's name and place of residence. In our example, a given physician is looking at her schedule at 10:00 a.m. The physician has already seen six patients, while one patient did not show up to his 9:00 a.m. appointment. She has four more visits until the end of the shift, one of which was already cancelled. Moreover, the physician is working ahead of time, as she has already completed the appointment which was scheduled to start at 10.00 a.m. Due to such a comprehensive information system, physicians have a complete picture of their shift, allowing them to react on the spot to changes such as cancellations.

Over the course of a shift, physicians are mandated to provide care to any patient with an appointment, as well as updating patients' medical records. When physicians experience a cancellation, or simply finish a visit faster than expected, they use such extra time not only to catch up on their schedule, but also to fulfil their updating obligations. Additionally, physicians are provided with non-scheduled time for breaks, including a lunch break. Furthermore, physicians have to finish all their visits on their corresponding appointment date, following their pre-booked appointment order.

3 Data

3.1 Hospital Data

I use data from one Spanish medium-sized, contracted hospital covering a wide range of specializations in the metropolitan area of Barcelona (Catalonia). My dataset contains all the 67,530 first visits to outpatient physicians from January 1, 2016 through and June 30, 2018, assigned to 86 physicians covering 19 different specializations. These physicians always operate within their specialization, giving clear-cut advice to patients referred from Primary Care Centers.[3]

This dataset consists of high-frequency visit times and medical treatment information. It includes information on the patient's time of arrival in the hospital, the visit appointment time, the referral date, and the visit starting and ending times. Visit length is measured according to the time that the patient's profile was opened and closed on the physician's terminal. These times are automatically recorded by the terminals used, rather than being self-reported by physicians. Referral and appointment dates are also key for mapping the visits creating the outpatient process, from the first visit to all other follow-up visits. Outpatient processes are used to test whether and how visit lengths result in changes over the whole process. The dataset also covers the treatments provided in each visit, such as imaging and laboratory tests, drugs prescribed, and the testing cost.[4] Table 1 provides descriptive statistics for the variables used in the analysis. For instance, the average patient is a middle-aged Spanish woman living in an area adjacent to the hospital and with public coverage. The average first visit takes 12 minutes, with an average

[3] Physicians are also involved in other minor tasks, not included in the present study, such as night shifts, surgery visits, and rehabilitations.

[4] We cannot retrieve the drug costs as prescriptions are issued based on the drugs' active components. See Law 29/2006.

waiting list of 30 days, and an 8% likelihood of receiving a diagnosis.[5]

3.2 Shift distortions - Cancellations

A standard work shift is from 9 a.m. to 1.30 p.m., with a structure and composition decided ex-ante on a yearly basis, being specific to the physician's specialization. Shifts are characterized by being fully booked (the average waiting time for a first visit is 30 days), and compulsory for the physician (i.e., the physician cannot prioritize or decline visits). The whole dataset available contains 347,277 scheduled visits, divided into first visits (24.43%), follow-up visits (50.10%), and external consultations (23.06%). Figure 2a shows how the encounters spread over the shift by visit type. The period from 9 a.m. to 1.30 p.m. is the busiest, with 86.54% of the visits being concentrated in that period. While in the morning, the hospital facilities are used by the outpatient department, later in the day, they are used for rehabilitations and surgeries, which are not included in the study. First visits spread homogeneously along the shift.

Cancellations represent the main perturbations on the physicians' schedules, providing unexpected free time. I use only those that occur on the visit date, as those happening on a prior date are easily re-booked. In absolute terms, the whole dataset available contains 54,057 cancellations, which comprise those visits withdrawn before their appointed slot (18%) and no-shows for a visit (82%). No patient walk-outs are found in the data. Figure 2b shows the share of cancellations over the shift using all the visits' appointment time at the 30-minutes bin level. Visually, no clear pattern is found to indicate clustered cancellation periods.

I focus on how prior cancellations affect subsequent first visits during a shift. Using the benchmark sample as in Table 1, Figure 2c shows how the number of cancellations before a given visit accumulate over the schedule, with higher variation in the evening shift due to the combination of newly arrived physicians with those who are continuing on from the preceding morning shift. Figure 2d shows how the probability of the previous visit being dropped evolves over the schedule, exhibiting a higher incidence at the beginning of the morning and evening shifts. Hour fixed effects are used in the study to account for such variation in the propensity of receiving a shock.

[5]The hospital managing our outpatient department is considered to be a high-performing center within the Catalan health system. To give a few examples, as of 2017, i) the reported patient satisfaction was 8.2 out of 10 in my outpatient department (7.5 in the region); ii) the probability of readmission within 30 days was 9.22% (9.81% in the region); and iii) the average waiting time to access a first visit was 41 days (121 days in the region). For those reasons, I consider the results presented in this study as a lower bound when compared to other outpatient departments.

Figure 3 exposes how visits are distributed with respect to prior cancellations. Subfigure 3a presents the fraction of visits with no prior cancellation, as well as the fraction of visits with a prior cancellation at higher horizons. In total, 62% of all the first visits had at least one preceding visit cancelled, and 16% had the prior visit dropped. Subfigure 3b presents how the average actual and expected visit lengths evolve with respect to the distance from a prior cancellation.[6] We can appreciate that i) the hospital structurally assigns more time-consuming visits to earlier slots, where there is a lower probability of having any prior cancellation; ii) when shocked with a cancellation, physicians spend more time on their next visit; and iii) for all distances, expected visit length is generally greater than the actual visit length, which shows that the outpatient department provides visits with insufficient time to compensate for overbooking and other administrative duties.

4 Empirical Strategy

In the first empirical exercise, I examine the extent to which medical treatments are influenced by the time spent on a visit, using the following model:

$$Y_{i,j,s} = \beta_0 + \beta_1 Length_{i,j,s} + \theta T_s + \delta_j + \Psi X_{i,s} + \varepsilon_{i,j,s} \quad (1)$$

where Y identifies a given visit outcome, as described in Section 4.1, for a patient i, a physician j and a slot s. The key independent variable, *Length*, identifies how many minutes a physician j spends with patient i in a visit slot s. I control for patient characteristics, $X_{i,s}$, such as gender, age, nationality, insurance coverage, and district distance to the hospital. All regressions include i) physician fixed effects, δ_j, which allows us to account for time-invariant variation across physicians, and by construction, across specializations; ii) month-year fixed effects, T_s, which mitigate the fact that results are confounded by seasonality (e.g. periods in which patients are more prone to suffer from diseases, such as with the seasonal flu, may also lead them to miss their hospital visits more frequently); and iii) hour fixed effects, T_s, which accounts for different hour-patient compositions.

The estimation of Equation 1 using OLS may result in biased estimates for serveral reasons. First, there may be omitted variables not captured by the rich set of controls and fixed effects. These confounding variables may correlate with our measure of visit length

[6] Expected visit length is a measure provided by the outpatient department, which identifies how much time an average visit should take, based on the type of visit, specialization, and administrative duties involved.

and with some unobserved components in the error term. For example, the good/bad mood or health condition of physicians may affect both visit lengths and the medical treatments. Second, given that physicians have full information on all their on-the-day visits, they may decide to allocate visit lengths based on their current and future patient characteristics. Such anticipation may facilitate a simultaneous causation between the time spent with patients and the treatment provided. On the one hand, physicians may decide to spend more time with those patients found to be more challenging, allowing them to better assess if further treatment is required. On the other hand, physicians may decide to provide a patient with a treatment as a substitute for the time spent. Such substitution decision is plausible as reviewing and testing physicians may differ. To tackle these concerns, I only use first visits, as patients accessing these initial visits have no knowledge of the physician's schedule nor do physicians know these patients; and use cancellations as an exogenous variation on the physicians' disposable time.

I use prior cancellations as an instrument to capture exogenous variations on the available time physicians have. Those cancellations are composed of all the on-the-day visit withdraws, including those that happened prior to their appointment time, together with no-shows. I define a first visit to be affected by a cancellation if the visit preceding it was cancelled using their real cancellation time. In practice, $PriorCancel$ is a dummy variable which takes value 1 in the event that the previous scheduled visit was a no-show,[7] or in the event that another visit, which is supposed to happen later in the same day, is cancelled during the current visit. The use of the exact cancellation time is important for the study, as physicians can smooth out cancellations for which they have been notified. I use this approach since it represents a lower bound of a cancellation's impact on the subsequent visits, taking as not treated any other first visit that is not immediately after a cancellation, whether notified or a no-show. Figure 4a displays how cancellations impact the length of subsequent visits. We can see that physicians utilize significantly more time in visits after a cancellation than before it. The figure also highlights that the time used in those first visits right after a cancellation is significantly different from any other first visit. Figure 4b shows that physicians are unable to anticipate cancellations, changing their reviewing time accordingly. For these reasons, the present analysis defines a treated visit as a first visit which takes place immediately after a cancellation; all other first visits

[7] A no-show is a visit for which its patient never showed up. I.e., I do not leverage on the *extra* visiting time provided by those *pending* patients who did not show up on time to their visits, but showed up later in the shift.

are defined as not treated.[8]

The validity of the instrument hinges on various considerations. A first issue relates to the random assignment of cancellation times. Those patients who drop a visit do so with no knowledge of the physician's schedule. Yet, it could be that visits are more frequently cancelled at times when a patient characteristic is more present, such as older patients or those with more chronic problems. Moreover, physicians could also be deciding in practice which patients to take after a visit is cancelled. Table 2 displays the covariate test on the patient characteristics and on the shared physician-patient characteristics. Prior cancellation does not predict any of the patient characteristics, used in the study, and that are visible to physicians. More importantly, physicians do not seem to select patients based on their shared characteristics, namely sex and age. This supports the claim that first visits are indeed randomly affected by prior cancellations.[9]

A second issue pertains to any other utilization of the physician's extra disposable time created by a prior cancellation. Prior to the treated visit, physicians decide how soon to take the new patient, which in turn might reduce their working delay. In turn, the estimates presented would be biased if such a less-rushed environment affects medical treatment directly, and not only via visit length, thus violating the exclusion restriction.[10] This indirect path is mostly attenuated by the use of fixed effects at the hour and the physician level, as the visits with a prior cancellation are compared to adjacent visits with similar levels of time pressure. Nevertheless, it could still be the case that the allocation of such extra time affects the first visit after a cancellation significantly differently, compared to those at different horizons. To test whether a less-rushed environment directly affects the outcomes of interest, I extend Equation 1 to include the variable $Delay$, which represents the difference between the visit start time and the visit appointment time. The average $Delay$ in the sample is 16.2 minutes. Following Neprash (2016), I instrument the variable $Delay$ using a dummy variable to indicate whether the preceding realized visit arrived late to her appointment time, $Prior\,Late$. The variable takes value 1 in the event that the patient appointed before a given visit arrived at the outpatient department

[8]Table A1 shows that prior cancellations lead to extra visiting time to all patients reviewed during the physician's shift. When physicians work overtime, prior cancellations do not lead to any extra reviewing time.

[9]As an exception, some specializations in the outpatient department allow their first visit patients to choose their preferred slot at their corresponding Primary Care Centers. Using only those patients, Table A2 shows that having a prior cancellation is not predictive of either those patients' characteristics nor shared physician-patient characteristics.

[10]The relationship between time pressure and individual decision-making has been broadly studied, finding a negative trade-off between time pressure and output quality (Tversky and Kahneman, 1974; Svenson and Maule, 1993; Maule et al., 2000).

after her scheduled appointment time. When patients arrive late to the outpatient department, physicians await them for some courtesy time, which might lead to higher delays suffered by the following patients.[11] Table A4 evidences no clear link of *Delay* directly affecting visit outcomes. Moreover, when comparing the variable *Length* in Table A4 to the main result provided in Table 3, we can see how including *Delay* does not affect the predictability of our variable of interest. For such reasons, I dismiss the premise that, in the context of this study, changes in time pressure, originating from sudden schedule changes, affect the outcomes of interest, other than through the visit duration.

4.1 Outcomes of interest

I use the instrumental variable framework previously detailed to study how physicians respond to extra time with patients, examining a wide set of outcomes, which can be broadly classified into diagnosis provision and treatment choice.

Regarding diagnosis provision, I investigate whether having longer visits proves to be beneficial in assessing patients' diagnoses. Given that the outpatient department's main objective is to provide a correct assessment of the patient's problems, due to their clear-cut medical knowledge, I use the provision of a diagnosis, *Diagnosis*, as a proxy of the visit successful completion. According to Aranaz et al. (2005), the probability of a diagnostic error in the Spanish healthcare system is 0.13%. It is important to note is that making a diagnosis is not excludable from the provision of other inputs, such as testing, as physicians use tests both to assess and to corroborate diagnoses. Following that logic, I include as outcome a variable identifying whether the current first visit had a follow-up visit in the same hospital, named *Follow-up*.

Referring to the treatment choice, I investigate whether visit length is used as substitute or complement to the provision of tests and drugs during the visit. On the one hand, physicians with extra visiting time are able to examine patients more thoroughly, inspecting their symptoms more carefully, with such an exhaustive examination obviating the need to incur intensive testing. In such a case, we would expect testing to be substitute of visit length. On the other hand, we could think of visit length as a complement of intensive care as physicians presented with such extra visiting time could use it to further deepen their knowledge of the clinical case, and consequently order more tests. Moreover, we expect that extra visit length provides physicians with a clearer idea of the

[11] Table A3 tests whether the instruments *PriorCancel* and *PriorLate* predict observable patient's characteristics, finding no systematic evidence.

patient's needs, thus modifying their drug prescription towards more accurate doses.

The variables used to explore how visit length relates to treatment choices are i) *Tests*, which is a dummy variable measuring whether medical tests, e.g. imaging and laboratory tests, have been ordered, ii) *Num. Tests*, which is a variable identifying the absolute number of tests ordered in a given visit, iii) *Test Cost*, which measures the cost of the tests ordered, iv) *Drugs*, which is a dummy variable measuring whether drugs have been prescribed, and v) *Num. Drugs*, which is a variable measuring the total number of drug doses ordered in a given visit. I compute the testing cost using internal cost information provided by the outpatient department in the sample. As for the number of drugs prescribed, I follow the aggregation method based on the Defined Daily Doses prescribed as proposed by the WHO. A Defined Daily Dose is a measure of drug utilization which stands for the assumed average maintenance dose per day for a drug used for its main indication in adults. I use this measure, as opposed to the number of drugs provided, as it enables aggregation of different drug groups weighted by their relative intensity, avoiding issues related to the drugs' package size and strength.

5 Results

Table 3 reports the estimation results using the 2SLS model previously outlined.[12] Column 1 introduces our first stage estimates using *Prior Cancel* as source of exogenous variation and controls by a wide set of fixed effects. Our first coefficient of interest, *Prior Cancel*, tells us when shocked by a cancellation, physicians decide to spend an average of 1.62 minutes more with the next patient, compared to any other patient with no immediately prior cancellation. Such a significant effect represents an increase of 12.8% over the average visit duration and corresponds to the lower bound effect of cancellations' impact on visit duration, given that visits at higher distances from a notification, used in this study as controls, may also be affected.

In Column 2, I test whether longer visit duration helps physicians to assess patients' diagnoses. We observe that longer visits have a positive effect on the provision of a diagnosis, implying that for every minute spent with a patient, the probability of providing a diagnosis increases linearly by 0.36 percentage points. In other words, when compared to the average probability of providing a diagnosis, every extra minute spent with a patient

[12] For completeness, I include in the Appendix the benchmark specification without controls (Table A5), the ITT estimation (Table A6), and the OLS estimation (Table A7). They are quantitative and qualitatively similar to the benchmark estimation.

translates into a 4.39% higher chance of providing it. However, the fact that longer visits imply a higher probability of providing a diagnosis could be measuring a more in-depth examination process, driven by longer reviewing time, but also the fact that physicians had enough time to record the diagnosis. I test a such hypothesis by identifying the most repeated diagnosis for each specialization. On the one hand, it could be that physicians' extra time only implies a higher diagnosis rate because they have the time to record the diagnoses. In such a case, both common and uncommon diagnoses would be recorded more frequently as the visit length increases. On the other hand, if physicians do indeed use such extra visiting time to provide a more in-depth examination, we would observe that they provide more uncommon diagnoses, since they are able to observe the patient's symptoms more thoroughly, thus providing a more accurate, and less standard, diagnosis. Table 4 shows that indeed longer reviewing time leads to the provision of more uncommon diagnoses, while no effect is found on the provision of those diagnoses repeated most frequently. In fact, the provision of a common diagnosis takes on average 12.5 minutes, while an uncommon diagnosis requires an average of 13.3 minutes. This provides suggestive evidence that physicians use extra reviewing time to provide a more precise service.

Back to Table 3, we proceed to investigate how visit length affects input choices. In Column 3, I study how visit length causally relates to the probability of ordering tests during a given visit. *Length* significantly predicts that every extra visiting minute accounts for an increase in the probability of ordering tests by 0.65 percentage points. When compared to the average visit ordering pattern, a one-minute increase in the visit length, due to a prior cancellation, implies a 3.6% higher chance of ordering tests. Column 4 broadens the outcome definition by checking whether visit duration affects the total number of tests ordered. As in Column 3, we can see how increases in visit duration significantly predict increases in testing. The estimated effect is low in magnitude, with an increase in the number of tests of 0.0096 per extra minute spent on the consultation. Despite that, when compared to the average number of tests ordered, we can see how an increase of one minute in the visit duration implies a 3.35% increase in the number of tests ordered. These two results together suggest that test ordering is used as a complement of visit duration, meaning that when physicians are exogenously exposed to more time, they employ it in ordering more tests. Due to the test-ordering distribution being right-skewed, as shown in Table 1, the main driver in this relationship is played by the testing extensive margin. Column 5 further checks whether increases in visit duration affect testing cost. We can see how indeed a unit increase in visit duration corresponds to an increase in

Test Cost of €0.8. This means that an extra minute on a consultation translates into a 6.35% increase in the average testing cost. Overall, this implies that visit duration and total testing cost are complementary inputs.

In Columns 6 and 7, I focus on drug prescription. Column 6 shows that that visit duration does not have an effect on the probability of prescribing drugs, however, it does have an effect on the dose prescribed. Column 7 shows how an increase of one minute in a given visit weakly leads to a reduction of 0.4 prescribed doses. Such sizable effect represents a 20.10% reduction in the average dose. These results indicate that the provision of time to physicians help them reduce the overall dose provided to patients. Under the assumption that a longer visit duration helps the physician to have a clearer idea of the patient's problems, the provision of lower doses of drugs could be understood as a convergence to the optimal prescription. This further highlights the role of drug prescription as substitute for reviewing time.[13]

Lastly, Column 8 analyzes whether longer visit duration affects the probability of having a follow-up visit. On the one hand, physicians might decide to provide patients with a follow-up visit at the hospital because longer visit duration might imply further tests to be checked in-situ. On the other hand, the extra visit length might enable to better assess the patient's diagnosis, thus redirecting the patient back to the primary health care center of origin. We can observe how a one-minute increase in visit length increases the probability of a follow-up visit by 0.92 percentage points, which represents a 3.28% increase over its mean. Table A8 shows that indeed a one-minute increase in reviewing a first visit leads to an increment on the total clinical case duration of 1.16 days. This suggests that physicians' complementary use of time and testing implies longer clinical processes.

All these results suggest that visit length is a key factor in understanding input utilization. However, they could actually hide an intertemporal input substitution decision followed by physicians, motivated by the extra time available during their first visits. If this were the case, we would expect that physicians who were *shocked* during a given first visit to inversely adjust their input utilization during the corresponding follow-up visit.[14] Table 5 tests such a hypothesis, using a similar strategy as in Table 3, in a subsample of

[13] The medical literature has found negative correlations between consultation length and medical overprescription, meaning that longer visit help physicians to save time on the patients' education or psychological support (Dugdale et al., 1999; Ventelou et al., 2010; Khori et al., 2012).

[14] I test whether having a prior cancellation predicts any patient characteristic in the sample of follow-up visits. Table A9 in the Appendix shows there is no systematic sample selection based on observable patient characteristics.

first visits with a follow-up visit in our outpatient department. We can see how increases in visit length during the first visit do not actually have any significant impact on the input utilization during the follow-up visit. This reinforces the idea that physicians do not use extra visit length to intertemporally transfer treatments; rather, they provide patients with care that they would not have otherwise received in their medical process.

Column 7 in Table 5 introduces a new variable identifying whether first and follow-up visits were conducted by the same physician. We can see how increasing visit length during a first visit relates to keeping the same treating physician. For every extra minute spent in a first visit, the likelihood that a patient will continue with the same physician increases by 1.05 percentage points. This suggests that longer visits provide physicians with a reason to keep the same patients, possibly due to their more exhaustive knowledge of the patient's case or increased satisfaction from the visit. Table 6 provides further evidence that indeed physicians are the ones pushing for preserving their patients and not the other way around. Physicians achieve this by securing that ordered diagnostic procedures are ready by the time a follow-up visit occurs, thus keeping the same patients over the process. In practice, for every extra reviewing minute spent in a first visit, the probability that physicians cancel the follow-up visit decreases by 12.9%. No effect is found on patient-motivated cancellations.

These results suggests that physicians use extra time to better assess the patient diagnosis, to recommend further intensive care treatments, and to correct drug prescription excess. Nevertheless, how intensely physicians use such time might depend on multiple factors. In the following subsections, I explore whether patients' characteristics are key in understanding time utilization and shed light on the relevance physicians' contracts have on such a relationship.

5.1 Which patients' characteristics are driving these effects?

In this section, I explore the influence patient's and shared patient-physician's characteristics have on time utilization.

I begin by examining whether the patients' gender influences how physicians use extra visiting time. While patients may differ in required treatments along their gender, the exogenous exposure to cancellations allows us to study whether physicians actually discriminate between them. Table 7 shows how visit length affects male and female patients differently. Firstly, we can see that after the realization of a cancellation, physicians employ more time similarly with both male and female patients. Nevertheless, physicians use such extra time only input intensively with female patients, with an increase in the tests ordered and a lower prescription dose. This differential input use is not explained by a systematic difference in their unconditional means, suggesting some limited preferential treatment towards women. I further inspect whether physicians treat patients differently depending on whether they share the same gender as the patient. On the one hand, we could expect that physicians use time more intensively on those patients sharing their gender, following their probable higher *proximity*. On the other hand, given that physicians might be able to screen those patients sharing their gender more quickly, we could expect that extra visiting time could indeed be only used efficiently on those patients from the other gender. Table 8 shows that, when exposed to cancellations, physicians use extra visiting time more intensively only with those patients with a different gender. Putting both results together, they suggest that physicians partially discriminate along gender, providing more intensive care to female patients and to those patients not sharing their gender.

I then look at whether physicians discriminate patients along their nationality of origin. Following the previous approach, Table 9 analyzes whether physician treat native patients patients differently than those born in other countries. While both national and non-national patients get more consultation time after a cancellation, physicians only provide diagnostic inputs and more tests to national patients. Such differential productivity by physicians is not explained by the patients' inherent characteristics. Moreover, the outpatient department considers non-national patients, if anything, more demanding, indicated by providing them longer expected visit lengths, 15.17 minutes, as compared to 14.8 minutes for national patients. The results highlight that, despite the fact that the outpatient department considers non-national patients more demanding, physicians discriminate along nationality, providing a more valuable service only to national patients when given extra visit time.

I next focus on the treatment physicians provide to patients depending on the number of days patients had to wait to access the outpatient department. As we previously explained in Section 2.1, patients are scheduled for a first visit with an outpatient physician at their primary care health centers. At that level, given the hospital scheduling limitations, primary care physicians can decide to speed up patients' first visit with a specialist. In practice, this implies that those patients with worse health conditions will be granted appointments at shorter notice and flagged as urgent to the outpatient physician. Moreover, given that accessing the emergency room is always an option, those patients waiting for a long period will presumably be those with less urgent health issues. Table

10 provides evidence that physicians use extra visiting time differently depending on the patient waiting time. Physicians use longer visits to order more tests, decrease the drug dose prescribed, and provide a diagnosis, but only for those patients whose waiting time was below the mean time for their specialization. These results suggest that physicians internalize the time that patients had to wait for the first visit, providing more urgent patients with a more valuable service.

A remaining question would be whether all specializations in the hospital spend reviewing time in the same fashion. To examine that issue, I classify the specialties into internal medicine or surgical, as these two categories require different input compositions. While surgical specialties use pre-established treatment protocols and surgical procedures in finding and solving the patients' health problems, internal medicine specialties are characterized by more intense use of visiting time and drug prescriptions. Table A10 shows how physicians react to extra visiting time, depending on their specialty. We can observe how those visits to internal medicine specialists are, on average, characterized by more visiting time and drug prescription as compared to the surgical specialties, which emphasize providing more tests. When physicians are notified of a cancellation, those in both categories respond by increasing their visiting time. This extra time is then used by internal medicine physicians to provide patients with more tests but also with a diagnosis (for every extra minute, physicians increase their probability of providing a diagnosis by 7.36% over their average diagnostic probability). On the opposite side, surgical specializations use such increased time to provide patients with tests, and reduce their drug doses, but no impact on the diagnosis rate.[15]

Overall, the way physicians use extra visiting time greatly depends on the patients' inherent characteristics and on the physicians' specialization. These results bring to light that physicians' reaction to the relaxation of their time constraints is not monotonic across subgroups, specially favoring female, Spanish-born and more urgent patients.

5.2 Role of physicians' contracts

In this section, I study how physicians' contracts shape the way extra visiting time is used.

[15] For completion, I include Table A11, which shows there is no distinctive time use along the patients' age profile; and Table A12, which shows that the nature of the shock, namely whether the prior visit was a no-show or a notification, does not differently influence how extra reviewing time is used. Contrarily, Table A13 shows that physicians react significantly more to extra visiting time on overloaded days, in which time is most precious.

The hospital organizes its employees according to the general Spanish healthcare legislation, which determines that contracts are composed by a fixed-wage component and a flexible component, mainly mainly depending on physicians' tenure.[16] These contracts are updated annually on a per-physician basis, including adapting visit workloads according to the physicians' responsibilities and tenure, which might ultimately lead to a differential use of the extra time provided by cancellations.[17]

I use physicians' age as proxy of their tenure, given that i) physicians enter the medical market right after finishing their studies,[18] and ii) the market for physicians enjoys low unemployment.[19] I define physicians to be senior if their age is higher than the median age ($\approx$ 50 years old); otherwise, I define them as junior. As indicated previously, the older physicians are, the more seniority they are likely to have, thus the higher their salary. While the hospital has an incentive to retain these experienced physicians, it is unable to freely raise the physicians' salary, which is publicly regulated. Therefore, senior physicians might be compensated with more advantageous shifts instead. Table 11 shows that senior physicians' schedules include lower numbers of patients per hour and fewer overbooked visits, while expected visit duration is similar to that of junior physicians. Furthermore, Table 12 shows that patients visiting senior outpatient physicians do not differ systematically from those visiting their junior colleagues. These tests show that while seniority affects the physician workload through more relaxed schedules, it does not imply a change in patient composition.[20]

Back to our benchmark specification, Table 13 shows how extra visit duration affects the input utilization depending on whether it is provided to senior or junior physicians. The first insight that we obtain from Columns 1 and 2 is that both senior and junior physicians react to cancellations in a similar manner, by increasing the reviewing time with their subsequent patient. Despite this similar increase, the unconditional visit length for junior physicians is 11.7 minutes, while for their senior colleagues it is 14 minutes.

[16] Such fixed component is similar across physicians as it is based on education attainment, which is, by law, required to be a bachelor's degree in medicine and to have passed a national exam (See Art. 4 in the Royal Decree 127/1984).

[17] For further knowledge on the collective bargaining agreement, please refer to the Resolution EMO/1742/2015 present in the Catalan Regional Bulletin n. 6923.

[18] According to the Spanish Health Ministry, the average age of those physicians entering practice in one of the specialties covered in the sample is 26 years, which corresponds to the age at which students finish their studies (Spanish Health Ministry, 2015).

[19] According to the Spanish Health Ministry, physician's unemployment in 2017 was 2.32% (Spanish Health Ministry, 2019). The unemployment rate in Spain in 2017 was 17.22%.

[20] A total of 13.67% of the visits correspond to 16 physicians who did not want their personal data to be made public. This section does not take them into consideration.

This shows that even if junior physicians utilize more time, it is not enough to compensate for the difference between the average visit length between these two groups. The way contracts are formulated, being physician specific, facilitates less rushed environments for older professionals at the expense of their younger colleagues.

This formulation fully determines how extra visit length is used. In Column 3, we can observe that junior physicians use extra visiting time more effectively by providing more diagnoses. For every extra minute junior physicians spend with a patient, they increase their probability of providing a diagnosis by 0.73 percentage points, i.e., it increases the probability of providing a diagnosis by 9.56% compared to its average. On the opposite extreme, senior physicians, despite spending more time with those patients affected by a prior cancellation, do not use such *bonus* time to modify their diagnosis provision. These results suggest that visit length expansions are not output-efficient on those individuals already enjoying more relaxed schedules. Table 14 shows how such extra visiting time helps only junior physicians to provide a more in-depth diagnosis, measured by more uncommon diagnoses. This provides further evidence that junior physicians effectively use extra time to provide a more valuable service.

Back to Table 13, I display in Column 4 to 8 how consultation time affects input choice. On the one hand, when exposed to extra time, junior physicians provide patients with more tests, both at the intensive and extensive margins, and of higher cost. Quantitatively, for every extra minute junior physicians spend with a patient, the probability of ordering a test increases by 0.68 percentage points (representing a 4% increase over the average ordering probability), the number of tests ordered is increased by 0.013 units (representing a 4.88% increase over their average ordering rate), and testing cost increases by 10.8%. On the other hand, longer visits affect the drug dose level prescribed to patients, as in Table 3, through the intensive margin. For every extra minute junior physicians spend with a patient, they decrease the average dose prescribed by 0.68 daily defined doses (representing a reduction in the prescription dose level by a 28.82% when compared to their average dose prescribed). Similarly, senior physicians decrease the patient's prescriptions by 0.19 doses (reflecting an average reduction of 8.24% in the prescription doses).[21]

Overall, these results highlight that correcting insufficient time per visit might have welfare-improving effects, as in the case of junior physicians. In the case of senior physicians, longer visits do not entail further care expansions, suggesting they are already at

[21] In the same spirit, Table A14 shows that extra visiting time helps least productive physicians catching up in the care provided, while no effects are found on high-performing doctors.

their optimal level of time-to-input utilization. This provides some evidence that defining schedules, based on seniority, might hinder high costs related to a suboptimal utilization of visiting time. In Section 6, I provide a quantification analysis stressing these inefficiencies, and show that time expansions to less experienced physicians might be cost effective.

6 Quantifying the cost of a diagnosis

In this section, I quantify the direct cost of increasing visit length.[22] Suppose that we want to increase the probability of providing a diagnosis by one percentage point ($\approx 12\%$ at the sample average). We can achieve this in two ways: i) by increasing all physicians' visiting times; or ii) by favoring only those physicians with less experience.

6.1 Broad increase

Let us say that we opt to increase the length of all first visits to achieve a one-percentage point increase in the diagnosis rate. This can be achieved with an increase in the average visit length of 2.77 minutes, using the IV-fixed-effects estimates in Column 3 of Table 3.

We calculate the direct costs associated with increasing visit length such that it leads to increasing the diagnosis rate by one percentage point, under the assumption that physicians will optimally utilize their *bonus* visiting time. In our case, using a linear approximation, we have:

$$\hat{\Delta}_{minutes} = 2.77 \times 6.55 \times 102.44 \; = 1,858.62 \text{ minutes per year and physician}$$

where 6.55 refers to the average number of first visit patients per day and physician, and 102.4 the average number of days worked per physician. $\hat{\Delta}_{minutes}$ amounts to about 31 hours extra per year and physician, which represents a 1.8% increase in the physician yearly working hours. We now extrapolate our physician-specific estimates to the general Spanish economy, such that:

$$\hat{\Delta}_{cost} = \hat{\Delta}_{minutes} \times (0.5876 \times (1 + 0.0092 \times 10.55) + 0.8045) \times 76,562 \;\; \approx €206m$$

[22] Throughout the exercise, I assume that the outpatient department's fixed capacities are non-binding along small visit length expansions. Similarly, I do not internalize the positive crowding-out effect longer first visits have on other services, such as the emergency room.

where 0.5876 represents the average physician wage per minute,[23] 0.0092 represents the increased probability of scheduling a follow-up visit as a result of one-minute increase in the first visit duration, and 10.55 the average follow-up visit length. 0.8045 represents the average treatment cost ordered for every extra minute spent with a patient,[24] and 76,562 refers to the total number outpatient physicians in Spain in 2018 (Spanish Health Ministry, 2019). Thus, increasing the diagnosis rate in first visits by one percentage point would have an estimated labor cost of €206m for the general Spanish economy.

6.2 Tailored increase

Suppose we now opt to provide more time per visit only to those physicians who will use it more efficiently. Following the previous procedure, I study how many more minutes junior physicians should have in order to increase their diagnosis rate by one percentage point. This can be achieved by increasing the visit length of junior physicians by 1.37 minutes, using the IV-fixed-effects estimates in Column 3 of Table 13. This change at the visiting intensive margin helps junior physicians assess their patients adequately, while leaving senior physicians' schedules unchanged. Following the same structure as before, we have:

$$\hat{\Delta}_{minutes,junior} = 1.37 \times 6.82 \times 95.45 \ = 891.82 \text{ minutes per year and junior physician}$$

Now we extrapolate these changes to the overall economy, such that:

$$\hat{\Delta}_{cost} = \hat{\Delta}_{minutes,junior} \times \left(0.575 \times (1 + 0.0116 \times 10.09) + 1.294\right) \times 42,863 \ \approx €74m$$

where 0.575 represents the per-minute wage,[25] 0.0116 represents the increased probability of scheduling a follow-up visit as a result of one-minute increase in the first visit duration, and 10.09 the average follow-up visit length. 1.294 represents the average treatment cost ordered for every extra minute spent with a patient, and 42,863 represents the

[23]The average working hours by physician in the Spanish health system is 1,645 hours, regulated by Decree 2/2012 and Royal Decree 20/2012. The average outpatient physician salary in 2018 is €58,000 (Medscape, 2019).

[24]The average treatment cost is calculated using internal information of the sample outpatient department. Both in this and in the next calculations, it is assumed to be representative for the health system as a whole.

[25]The salary used for junior physicians corresponds to a physician with a fixed position, around 40 years old, and 15 years of experience. The annual salary of such a physician is €56,755. For further reference, see OMC (2019).

estimated number of junior physicians.[26]

In sum, comparing this targeted increase to the previous broad increase in reviewing time, it is seen to be more cost-effective in achieving the same result, a one-percentage-point increase in the diagnostic provision. With all due caveats, this exercise highlights how solely exploiting the contracting incentives based on seniority would allow for more efficient diagnostic provision at a reduced cost.

7 Conclusion

This paper estimates and provides evidence of the inefficient allocation of time in the Spanish outpatient system. I leverage on its unique setting and on cancellations as random time shocks to provide a causal interpretation of how the amount of reviewing time shapes physicians' decisions. Conceptually, I compare those first visits affected by an unexpected longer visit time, caused by a prior cancellation, to all other first visits, holding all other parameters in the environment constant.

I find that longer first visits lead to a higher likelihood of providing a diagnosis, the main objective of outpatient departments. The effect is driven by uncommon diagnoses, whose provision require, on average, a more in-depth analysis, while no effect is found for the most common diagnoses. Longer first visits increase the diagnostic input utilization, while decreasing drug dose prescription. These results suggests that physicians use extra visiting time to assess in more detail the patient's health problems, and in the event of indecision, to request further diagnostic inputs, which ultimately will improve the service provided. Moreover, I find no evidence of an input substitution effect between first and follow-up visits, suggesting that longer first visits have a lasting impact in the clinical process.

I shed light on how physicians' responses to extra visiting time depend on patients' characteristics. I find that physicians' reaction to the relaxation of their time constraints is not monotonic across subgroups, specially favoring female, Spanish-born and more urgent patients. This means that physicians use the extra reviewing time to provide a more valuable service, measured by a higher diagnosis and input provision, to these subgroups.

I then look at how relevant working contracts are in shaping physicians' decisions. While the outpatient department has an incentive to retain more experienced physicians, it

[26] I use information from the OECD database - Healthcare Utilization. Given that the number of outpatient physicians is not tabulated by their age, I assume that the distribution of physicians by age is the same for the overall population of physicians and that of outpatient physicians.

is unable to freely raise the physicians' salary, which is publicly regulated, compensating them with more advantageous shifts instead. I find that junior physicians, whose contracts lead to more pressured schedules than those of their senior colleagues, use extra visiting time efficiently, while senior physicians do not. This highlights, as I show quantitatively, that policies increasing all reviewing time across the board might prove to be inefficient.

This avenue of research is extremely important for policy making, as it emphasizes that current promotion incentives might lead to inefficient input utilization. While this paper has focused on one Spanish outpatient department, the message of this study, concerning the effect of remedying insufficient reviewing time on the workers' decisions, is more general. In fact, it applies to all situations in which workers, facing increasing demand, have to speed up to accommodate it. This study indicates that public welfare may be improved by policies aimed at providing additional time to those workers most in need of it.

References

Anand, Krishnan S, M Fazıl Paç, and Senthil Veeraraghavan, "Quality–speed conundrum: Trade-offs in customer-intensive services," *Management Science*, 2011, *57* (1), 40–56.

Aranaz, JM, C Aibar, J Vitaller, and P Ruiz, "Estudio Nacional de Efectos Adversos ligados a la Hospitalización (ENEAS)," 2005.

Bruggen, Alexander, "An empirical investigation of the relationship between workload and performance," *Management Decision*, 2015.

Chan, David C, "Teamwork and moral hazard: evidence from the emergency department," *Journal of Political Economy*, 2016, *124* (3), 734–770.

_ , "The efficiency of slacking off: Evidence from the emergency department," *Econometrica*, 2018, *86* (3), 997–1030.

Coviello, Decio, Andrea Ichino, and Nicola Persico, "The inefficiency of worker time use," *Journal of the European Economic Association*, 2015, *13* (5), 906–947.

Dieleman, Joseph L, Ellen Squires, Anthony L Bui, Madeline Campbell, Abigail Chapin, Hannah Hamavid, Cody Horst, Zhiyin Li, Taylor Matyasz, Alex Reynolds et al., "Factors associated with increases in US health care spending, 1996-2013," *Jama*, 2017, *318* (17), 1668–1678.

Dugdale, David C, Ronald Epstein, and Steven Z Pantilat, "Time and the patient–physician relationship," *Journal of general internal medicine*, 1999, *14* (Suppl 1), S34.

Eurofound, "First European Survey on the work environment 1991-1992," Technical Report, Publications Office of the European Union 1993.

_ , "Sixth European Working Conditions Survey – Overview report," Technical Report, Publications Office of the European Union 2017.

Frakes, Michael D and Melissa F Wasserman, "Is the time allocated to review patent applications inducing examiners to grant invalid patents? Evidence from microlevel application data," *Review of Economics and Statistics*, 2017, *99* (3), 550–563.

Freedman, Seth, Ezra Golberstein, Tsan-Yao Huang, David Satin, and Laura Barrie Smith, "Docs with their eyes on the clock? the effect of time pressures on primary care productivity," *Journal of Health Economics*, 2021, p. 102442.

Gupta, Atul, "Impacts of performance pay for hospitals: The readmissions reduction program," *American Economic Review*, 2021, *111* (4), 1241–83.

Khori, V, S Changizi, E Biuckians, A Keshtkar, AM Alizadeh, AM Mohaghgheghi, and MR Rabie, "Relationship between consultation length and rational prescribing of drugs in Gorgan City, Islamic Republic of Iran," *EMHJ-Eastern Mediterranean Health Journal, 18 (5), 480-486, 2012*, 2012.

Kleibergen, Frank and Richard Paap, "Generalized reduced rank tests using the singular value decomposition," *Journal of econometrics*, 2006, *133* (1), 97–126.

Maule, A John, G Robert J Hockey, and Larissa Bdzola, "Effects of time-pressure on decision-making under uncertainty: changes in affective state and information processing strategy," *Acta psychologica*, 2000, *104* (3), 283–301.

Medscape, "Informe de remuneración económica y satisfacción profesional: España," *www.medscape.com*, 2019.

Neprash, Hannah T, "Better late than never? Physician response to schedule disruptions," 2016.

Olea, José Luis Montiel and Carolin Pflueger, "A robust test for weak instruments," *Journal of Business & Economic Statistics*, 2013, *31* (3), 358–369.

OMC, "Estudio comparativo de las Retribuciones de los Médicos de Atención Primaria en España en 2018 y diferencias desde 2009," 2019.

Powell-Jackson, Timothy, Sumit Mazumdar, and Anne Mills, "Financial incentives in health: New evidence from India's Janani Suraksha Yojana," *Journal of health economics*, 2015, *43*, 154–169.

Shurtz, Ity, Alon Eizenberg, Adi Alkalay, and Amnon Lahad, "Physician workload and treatment choice: the case of primary care," 2019, p. 70.

Silver, David, "Haste or waste? Peer pressure and productivity in the emergency department," *The Review of Economic Studies*, 2021, *88* (3), 1385–1417.

Song, Hummy, Elena Andreyeva, and Guy David, "Time is the wisest counselor of all: The value of provider–patient engagement length in home healthcare," *Management Science*, 2022, *68* (1), 420–441.

Spanish Health Ministry, "Tercer Informe Descriptivo MIR-R1. Convocatoria 2013-2014," 2015.

_ , "Estimación de la oferta y demanda de médicos. España 2018-2030," 2019.

Svenson, Ola and A John Maule, *Time pressure and stress in human judgment and decision making*, Springer Science & Business Media, 1993.

Tan, Tom Fangyun and Serguei Netessine, "When does the devil make work? An empirical study of the impact of workload on worker productivity," *Management Science*, 2014, *60* (6), 1574–1593.

Tversky, Amos and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty.," *science*, 1974, *185* (4157), 1124–1131.

Ventelou, Bruno, Sophie Rolland, and Pierre Verger, "A case study on the substitution effect between the length of GP consultation and drug prescribing practices," *Healthcare policy*, 2010, *5* (4), 58.

Xu, Yuqian, Tom Tan, and Serguei Netessine, "The Impact of Workload on Operational Risk: Evidence from a Commercial Bank," *SMU Cox School of Business Research Paper*, 2017, (18-13).

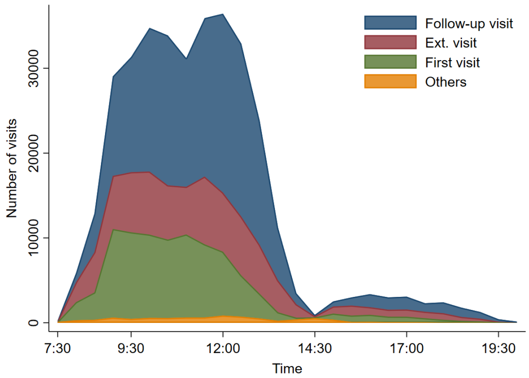
Figures and Tables

Figure 1: Daily physician's schedule viewed at 10:00 am.

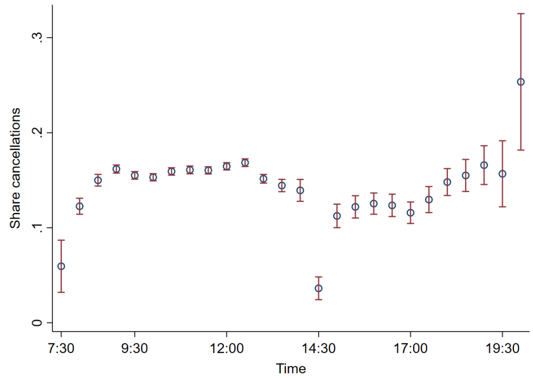
Appointment Time	Patient ID	Patient	Basic Health Zone	Status	Arrival time	Visit type
8:30	1	Antonio García Gracia	Barcelona 2-B	Completed	8:25	Follow-up
9:00	2	Jordi Bosch Fernández	Barcelona 3-A	Not present	-	Follow-up
9:10	3	Montserrat Muñoz Sánchez	Barcelona 4-D	Completed	9:05	First Visit
9:15	4	María del Carmen González Serra	Barcelona 5-D	Completed	9:00	First Visit
9:30	5	Anna Solé Pérez	Barcelona 1-C	Completed	9:10	Follow-up
9:40	6	José Giménez Sánchez	Barcelona 2-E	Completed	9:00	Long Cure
10:00	7	Wei Wang	Barcelona 8-B	Completed	9:40	Injection
10:15	8	María José Pérez Iglesias	Barcelona 4-C	Pending	9:45	First Visit
10:25	9	Montserrat Batlle Figueres	Barcelona 5-C	Pending	-	Follow-up
10:43	10	María del Mar Cardel Pérez	Barcelona 3-E	Cancelled	-	First Visit
11:00	11	Mohammed Alaoui	Barcelona 5-A	Pending	-	Follow-up

Notes: The figure shows how the schedules used in the outpatient department look like, using fictitious information. *Appointment Time* refers to the time at which a patient is appointed to start her visit. *Status* refers to the visiting status, which can be "Completed" if the visit finished already, "Not Present" if the visit was supposed to happen but the patient was not present, "Pending" if the visit will happen later, and "Cancelled" if the visit was appointed for a later time but cancelled earlier on the day. *Arrival time* refers to their arrival time to the outpatient department. If *arrival time* is not displayed (e.g. −), it means the patient has not registered yet at the outpatient department. *Visit type* highlights broadly the type of visit, which can be "First Visit", "Follow-up", "Long Cure", or "Injection".

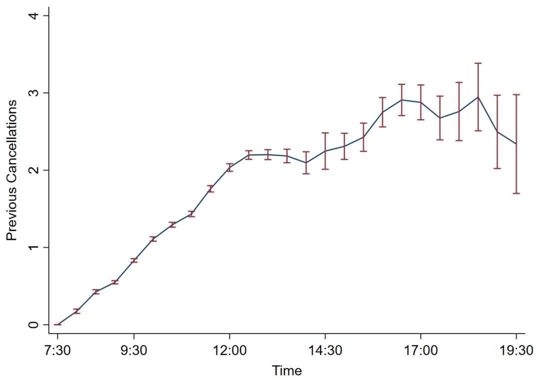
Figure 2: Distribution of visits over the day



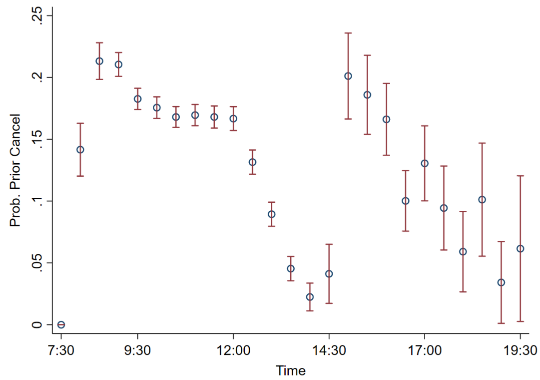
(a) Distribution by visit type



(b) Share of cancellations (Appointment)



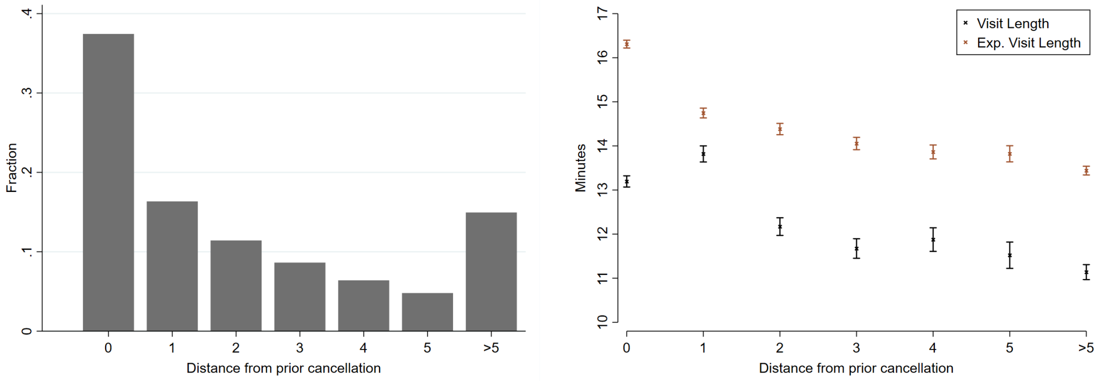
(c) Previous cancellations (First visits)



(d) Prob. Prior cancel (First visits)

Notes: The figure reports how visits and cancellations span over the schedule. Subfigures 2a and 2b use the sample including all the visits, namely first, external, and follow-up visits, cancelled or not, while subfigures 2c and 2d only use our final sample of first visits. Subfigure 2b displays the share of cancellations as to when those visits were appointed. Subfigures 2c and 2d use the real notification time of those cancellations as in our main analysis. Prior cancel identifies those visits that had their prior visit slot cancelled using their real cancellation time. All subfigures use 30-minutes bin sizes.

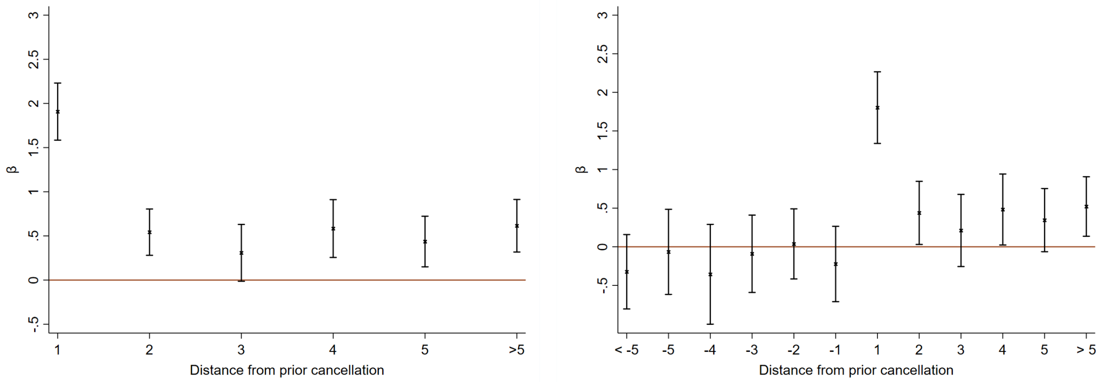
Figure 3: Distances to prior cancellation



(a) Distribution (b) Unconditional means

Notes: The figure reports the proportion of visits by distance to a cancellation and their visit lengths. The sample use corresponds to the final sample as exposed in Table 1. Subfigure 3a shows the proportion of visits which had no previous cancellation (distance 0), a cancellation in the previous visit (distance 1), and so forth. Subfigure 3b displays the unconditional mean of both visit length and expected visit length by the distance to a preceding cancellation.

Figure 4: First stage at multiple distances



(a) All shifts (b) Shifts with one or no cancellation

Notes: The figure reports how cancellations impact surrounding visits. Subfigure 4a uses the final sample as exposed in Table 1, and shows graphically the first stage results using dummy variables identifying those visits at 1, 2, 3, 4, 5, or more than 5 visits from a cancellation. Subfigure 4b uses only those shifts with one or no cancellations, and shows graphically the first stage results using dummy variables identifying those visits at 1, 2, 3, 4, 5, or more than 5 visits from a cancellation, both prior and posterior to a cancellation. The results presented in both figures include all the fixed effects and controls as in our benchmark specification (see Table 3). Confidence intervals at the 95%.

Table 1: Summary Statistics

	Mean	SD	Min	Max	N
Patient characteristics					
Male	0.45	0.50	0	1	67530
Age	58.85	19.55	0	106	67530
Reference BHZ	0.60	0.49	0	1	67530
Distance from hospital (km)	4.37	12.87	0	1979	67530
Born in Spain	0.68	0.47	0	1	67530
Public coverage	0.98	0.12	0	1	67530
Chronic condition	0.06	0.23	0	1	67530
Physician characteristics					
Physician: Male	0.59	0.49	0	1	66350
Physician: Age	49.78	9.32	32	65	58301
Visit characteristics					
Visit length (mins)	12.58	9.59	1	120	67530
Follow-up visit	0.28	0.45	0	1	67530
Out of agenda	0.15	0.35	0	1	67530
Internal referral	0.11	0.32	0	1	67530
Waiting list (days)	29.73	52.02	0	770	67530
Waiting room (mins)	27.22	32.62	0	545	67530
Tests	0.29	0.75	0	15	67530
Test cost	12.67	50.58	0	2019	67530
Drugs	2.04	27.34	0	2600	67530
Diagnosis	0.08	0.27	0	1	67530

Notes: The table provides a summary statistics for our sample of interest. Reference BHZ is an indicator variable that identifies whether the patient comes from a Basic Health Zone covered by the outpatient department. Distance from hospital is a variable that measures how many kilometers apart is the patient's Basic Health Zone centroid from the hospital using a linear distance algorithm. Public coverage is an indicator variable that identifies whether the treated patient is covered by the general public health insurance. Chronic condition is an indicator variable that identifies if the patient previously was been diagnosed any chronic condition. Visit length identifies how long a visit is using the patient's profile opening and closure in the physician's terminal. Out of agenda identifies whether the visit was placed in a slot not covered by the physician's agenda (visit schedule). Internal referral identifies if the visit was appointed by another hospital physician as opposed to a general practitioner. Waiting room is a variable that measures how many minutes has the patient been waiting prior to the visit start. Test cost indicates the testing cost per visit in euros. The variable Drugs captures the number of drugs prescribed measured using the Defined Daily Dose (DDD) definition. Diagnosis is an indicator variable identifying if a visit led to the definition of a precise diagnosis. Physician related variables such as age or sex have missing observations as some physicians preferred not disclosing such information. All other variables are self-explanatory.

Table 2: Covariate test

	(1) Male	(2) Age	(3) Ref. BHZ	(4) Dist. BHZ	(5) Chronic	(6) Pub. Cov	(7) Spanish	(8) Waiting list	(9) Same sex	(10) Same age
Prior Cancel	-0.0039	0.2213	0.0057	-0.0434	0.0032	0.0003	-0.0052	1.0568	-0.0030	0.0046
	(0.0045)	(0.1812)	(0.0056)	(0.1492)	(0.0023)	(0.0012)	(0.0047)	(0.7029)	(0.0046)	(0.0043)
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530	67530	66350	58301
Dep. Var. Mean	0.447	58.85	0.598	4.365	0.0582	0.984	0.677	29.73	0.517	0.152

Notes: The table tests whether having a prior cancellation predicts the patient and the shared physician-patient characteristics. Ref. BHZ is an indicator variable that identifies if the patient comes from a Basic Health Zone covered by the hospital. Dist. BHZ measures how many kilometers apart is the patient's Basic Health zone from the hospital using a linear distance algorithm. Chronic is an indicator variable that identifies if the patient previously had any chronic condition. Pub. Cov. identifies if a visit was covered by the public insurance scheme. Spanish identifies those patients born in Spain. Waiting list measures the days that patients wait to access a first visit from their corresponding Primary Care center. Same sex identifies if both physician and patient share the same sex. Same age identifies if both physician and patient have a similar age, measured using a 10 years window. All other variables are self-explanatory. Standard errors are clustered at the physician level. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Effect of Visit Length on Visit Outcomes - Main Analysis

	(1) Length	(2) Diagnosis	(3) Test	(4) Num. Tests	(5) Testing cost	(6) Drug	(7) Num. Drugs	(8) Follow-up
Length		0.0036**	0.0065***	0.0096**	0.8045**	-0.0010	-0.4106*	0.0092***
		(0.0018)	(0.0023)	(0.0042)	(0.3470)	(0.0011)	(0.2166)	(0.0032)
Prior Cancel	1.6222***							
	(0.1598)							
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530	67530
Dep. Var. Mean	12.58	0.0819	0.181	0.286	12.67	0.0333	2.043	0.280
F - Stat	–	104.3	104.3	104.3	104.3	104.3	104.3	104.3

Notes: The reported regressions correspond to the 1st Stage (Col. 1), and the 2nd Stage with multiple outcome variables (Col. 2-8). For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Olea and Pflueger (2013). *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Effect of Visit Length on Diagnosis Provision

	(1) Diagnosis	(2) Common	(3) Uncommon
Length	0.0036**	0.0001	0.0034**
	(0.0018)	(0.0008)	(0.0014)
Month-Year FE	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Observations	67530	67530	67530
Dep. Var. Mean	0.0819	0.0123	0.0695
F - Stat	104.3	104.3	104.3

Notes: The reported regressions correspond to the 2nd Stage with the following outcomes: i) the probability of a diagnosis (Col. 1), ii) the probability of a common diagnosis (Col. 2), and iii) the probability of an uncommon diagnosis (Col. 3). Diagnoses are classified as common identify those diagnoses most repeated in a given specialization, while uncommon represent any other non modal diagnosis. See Table 3 for further reference on the controls used. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Olea and Pflueger (2013). *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Effect of Current Visit Length on the Next Visit Outcomes

	(1) Length	(2) F. Length	(3) F. Tests	(4) F. Num. Tests	(5) F. Drugs	(6) F. Num. Drugs	(7) Same Physician
Length		0.0953	0.0008	-0.0013	0.0008	0.5331	0.0105***
		(0.1848)	(0.0034)	(0.0055)	(0.0008)	(0.4414)	(0.0037)
Prior Cancel	1.8596***						
	(0.2439)						
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14350	14350	14350	14350	14350	14350	14350
Dep. Var. Mean	14.39	11.19	0.143	0.195	0.00613	0.552	0.656
F - Stat	–	58.82	58.82	58.82	58.82	58.82	58.82

Notes: The reported regressions correspond to the 1st Stage (Col. 1), and the 2nd Stage with multiple outcome variables (Col. 2-7). The sample used refers to all those visits that had a follow-up visit appointed on that same first visit. The outcomes used refer to the follow-up visit. For information on the outcome variables, please refer to Section 4.1. *Same Physician* is a dummy variable that takes value one if the visit was conducted by the same physician that conducted the first one, and zero otherwise. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. All the controls used are measured as in the first visit. See Table 2 for further reference. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Olea and Pflueger (2013). *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Effect of Current Visit Length on the Next Visit Cancellation

		Next visit cancelled		
	Length	All	By patient	By physician
	(1)	(2)	(3)	(4)
Length		-0.0001	0.0036	-0.0037**
		(0.0048)	(0.0043)	(0.0018)
Prior Cancel	1.8328***			
	(0.1947)			
Month-Year FE	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Dep. Var. Mean	14.34	0.240	0.211	0.0287
F - Stat	–	89.65	89.65	89.65

Notes: The reported regressions correspond to the 1st Stage (Col. 1), and the 2nd Stage with multiple outcome variables (Col. 2-7). The sample used refers to all those visit that had a follow-up visit appointed on that same visit. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. All the controls used are measured as they were during the first visit. See Table 2 for further reference. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Olea and Pflueger (2013). *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Effect of Visit Length on Visit Outcomes - By Patient sex

	(1) Length	(2) Length Male	(3) Diagnosis	(4) Test	(5) Num. Tests	(6) Test Cost	(7) Drug	(8) Num. Drugs
Length			0.0033	0.0081**	0.0107	1.4358***	-0.0019	-0.5883**
			(0.0022)	(0.0034)	(0.0067)	(0.4917)	(0.0019)	(0.2329)
Length × Male			0.0005	-0.0033	-0.0024	-1.3088*	0.0020	0.3682
			(0.0028)	(0.0044)	(0.0088)	(0.7509)	(0.0025)	(0.2928)
Male	-0.0163	12.3993***	-0.0052	0.0315	0.0138	16.6131*	-0.0232	-4.0772
	(0.1029)	(0.5845)	(0.0352)	(0.0565)	(0.1139)	(9.6802)	(0.0313)	(3.6727)
Prior Cancel	1.5313***	0.0901						
	(0.1754)	(0.0844)						
Prior Cancel × Male	0.2067	1.5743***						
	(0.1656)	(0.2835)						
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530	67530
Joint Length p-value	–	–	0.0949	0.106	0.120	0.810	0.995	0.446
Dep. Var. Mean	–	–	0.0819	0.181	0.286	12.67	0.0333	2.043
F - Stat	–	–	19.15	19.15	19.15	19.15	19.15	19.15

Notes: The reported regressions correspond to the 1st Stage regression (Col. 1-2) and the 2nd Stage with multiple outcome variables (Col. 3-8). The table presents the interaction of *Length* and the patient's sex (*Male*). For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Joint Length p-value is the joint p-value of both *Length* and *Length* × *Male*. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Kleibergen and Paap (2006). *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Effect of Visit Length on Visit Outcomes - By Patient-Physician sex

	(1) Length	(2) Length Male	(3) Diagnosis	(4) Test	(5) Num. Tests	(6) Test Cost	(7) Drug	(8) Num. Drugs
Length			0.0046**	0.0103***	0.0220***	1.4593***	0.0008	-0.3642
			(0.0021)	(0.0036)	(0.0064)	(0.4992)	(0.0016)	(0.2773)
Length × Same sex			-0.0021	-0.0073*	-0.0242***	-1.2207*	-0.0036	-0.0894
			(0.0027)	(0.0044)	(0.0091)	(0.7327)	(0.0026)	(0.2728)
Same sex	-0.0772	12.3465***	0.0292	0.0940*	0.3127***	16.4118*	0.0428	0.6551
	(0.1067)	(0.5926)	(0.0332)	(0.0561)	(0.1161)	(9.2885)	(0.0319)	(3.4255)
Prior Cancel	1.6723***	0.0824						
	(0.1621)	(0.1057)						
Prior Cancel × Same sex	-0.0264	1.4853***						
	(0.1727)	(0.2761)						
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	66350	66350	66350	66350	66350	66350	66350	66350
Joint Length p-value	–	–	0.261	0.288	0.718	0.639	0.116	0.0506
Dep. Var. Mean	–	–	0.0819	0.181	0.286	12.67	0.0333	2.043
F - Stat	–	–	17.44	17.44	17.44	17.44	17.44	17.44

Notes: The reported regressions correspond to the 1st Stage regression (Col. 1-2) and the 2nd Stage with multiple outcome variables (Col. 3-8). The table presents the interaction of *Length* and whether the patient and physician have the sex (*Same sex*). For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Joint Length p-value is the joint p-value of both *Length* and *Length* × *Same sex*. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Kleibergen and Paap (2006). *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Effect of Visit Length on Visit Outcomes - By Nationality

	(1) Length	(2) Length Spanish	(3) Diagnosis	(4) Test	(5) Num. Tests	(6) Test Cost	(7) Drug	(8) Num. Drugs
Length			0.0044**	0.0054*	0.0095*	0.9784**	-0.0000	-0.5015
			(0.0021)	(0.0030)	(0.0057)	(0.4856)	(0.0014)	(0.3364)
Length × Non-Spanish			-0.0024	0.0031	0.0001	-0.5006	-0.0028	0.2617
			(0.0035)	(0.0059)	(0.0096)	(0.8946)	(0.0042)	(0.4364)
Non-Spanish	0.1674	12.2274***	0.0243	-0.0395	-0.0079	5.6797	0.0300	-3.5730
	(0.1450)	(0.5984)	(0.0465)	(0.0754)	(0.1228)	(11.1152)	(0.0487)	(5.5017)
Prior Cancel	1.6390***	0.0538						
	(0.1748)	(0.0677)						
Prior Cancel × Non-Spanish	-0.0507	1.5372***						
	(0.2392)	(0.2804)						
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530	67530
Joint Length p-value	–	–	0.501	0.0629	0.178	0.468	0.392	0.269
Dep. Var. Mean	–	–	0.0819	0.181	0.286	12.67	0.0333	2.043
F - Stat	–	–	20.59	20.59	20.59	20.59	20.59	20.59

Notes: The reported regressions correspond to the 1st Stage (Col. 1), and the 2nd Stage with multiple outcome variables (Col. 2-8). The table presents the interaction of *Length* and whether the patient was born in Spain (*Spanish*). For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Joint Length p-value refers to the joint significance of *Length* and *Length* × *Spanish*. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Olea and Pflueger (2013). *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Effect of Visit Length on Visit Outcomes - By Waiting List

	(1) Length	(2) Length WaitLong	(3) Diagnosis	(4) Test	(5) Num. Tests	(6) Test Cost	(7) Drug	(8) Num. Drugs
Length			0.0061**	0.0109***	0.0153***	0.7824*	-0.0023	-0.4661**
			(0.0025)	(0.0033)	(0.0058)	(0.4064)	(0.0016)	(0.2311)
Length × WaitLong			-0.0091**	-0.0149**	-0.0191**	0.0976	0.0042	0.1840
			(0.0037)	(0.0061)	(0.0097)	(0.5968)	(0.0035)	(0.2255)
WaitLong	-0.7258**	11.6228***	0.1012**	0.1604**	0.1836	-4.6967	-0.0540	-2.7691
	(0.3084)	(0.5002)	(0.0494)	(0.0768)	(0.1196)	(7.3844)	(0.0408)	(2.8035)
Prior Cancel	1.6671***	0.0187						
	(0.1834)	(0.0498)						
Prior Cancel × WaitLong	-0.1346	1.3745***						
	(0.2117)	(0.2154)						
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530	67530
Joint Length p-value	–	–	0.221	0.370	0.593	0.0860	0.432	0.274
Dep. Var. Mean	–	–	0.0819	0.181	0.286	12.67	0.0333	2.043
F - Stat	–	–	25.17	25.17	25.17	25.17	25.17	25.17

Notes: The reported regressions correspond to the 1st Stage regression (Col. 1-2) and the 2nd Stage with multiple outcome variables (Col. 3-9). The table presents the interaction of *Length* and whether the patient had to wait more than the average service waiting list (*WaitLong*). For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. Joint Length p-value refers to the joint p-value of *Length* and *Length* × *WaitLong*. See Table 2 for further reference. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Kleibergen and Paap (2006). *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Visit characteristics by Senior Physicians

	(1) Exp. Visit Length	(2) Overbook	(3) Visits/hour	(4) Overloaded day
Senior Physician	-0.2446	-0.0367***	-0.2967***	-0.0829***
	(0.1983)	(0.0115)	(0.0955)	(0.0268)
Month-Year FE	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes
Specialty FE	Yes	Yes	Yes	Yes
Observations	58301	58301	58301	58301
Dep. Var. Mean	14.93	0.212	4.255	0.240

Notes: The table tests whether senior physicians, measured as those above the median age in the outpatient department, have different type of visits. Exp. Visit Length is a hospital-provided variable that measures how long a given visit should be. Overbook is an indicator variable that identifies those visits that were appointed on the time slot of a prior visit. Overloaded day is an indicator variable that identifies those days in which the total expected visiting time a physician has, exceeds the time he/she is at the outpatient department. All other variables are self-explanatory. Standard errors are clustered at the physician level. *** p<0.01, ** p<0.05, * p<0.1.

Table 12: Patient characteristics by Senior Physicians

	(1) Male	(2) Age	(3) Ref. BHZ	(4) Dist. BHZ	(5) Chronic	(6) Pub. Cov	(7) Spanish	(8) Waiting list
Senior Physician	-0.0115	0.5690	0.0243	-0.0957	-0.0007	-0.0237	0.0099	-3.8768*
	(0.0093)	(0.3451)	(0.0306)	(0.5525)	(0.0027)	(0.0157)	(0.0168)	(2.0910)
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Specialty FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	58301	58301	58301	58301	58301	58301	58301	58301
Dep. Var. Mean	0.447	58.85	0.598	4.365	0.0582	0.984	0.677	29.73

Notes: The table tests whether senior physicians, measured as those above the median age in the outpatient department, have different type of patients compared to their junior colleagues. Ref. BHZ is an indicator variable that identifies if the patient comes from a Basic Health Zone covered by the hospital. Dist. BHZ measures how many kilometers apart is the patient's Basic Health zone from the hospital using a linear distance algorithm. Chronic is an indicator variable that identifies if the patient previously had any chronic condition. Pub. Cov. identifies if a visit was covered by the public insurance scheme. Spanish identifies those patients born in Spain. Waiting list measures the days that patients wait to access a first visit from their corresponding Primary Care center. All other variables are self-explanatory. Standard errors are clustered at the physician level. *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Effect of Visit Length on Visit Outcomes - By Seniority

	(1) Length	(2) Length Senior	(3) Diagnosis	(4) Test	(5) Num. Tests	(6) Test Cost	(7) Drug	(8) Num. Drugs	(9) Follow-up
Length			0.0073**	0.0068*	0.0130***	1.2940**	-0.0017	-0.6812*	0.0116***
			(0.0030)	(0.0036)	(0.0049)	(0.5066)	(0.0016)	(0.4093)	(0.0044)
Length × Senior			-0.0075*	-0.0029	-0.0071	-1.1112*	0.0009	0.4865	-0.0064
			(0.0042)	(0.0046)	(0.0091)	(0.6618)	(0.0021)	(0.3965)	(0.0060)
Senior	-0.7348***	11.1466***	0.0473	0.0473	0.1047	17.6521**	-0.0135	-6.2473	0.1078
	(0.2682)	(1.0392)	(0.0523)	(0.0558)	(0.1189)	(7.9696)	(0.0239)	(4.9516)	(0.0729)
Prior Cancel	1.7592***	-0.0574**							
	(0.2560)	(0.0274)							
Prior Cancel × Senior	-0.1712	1.7390***							
	(0.3677)	(0.2621)							
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	58301	58301	58301	58301	58301	58301	58301	58301	58301
Joint Length p-value			0.949	0.223	0.443	0.680	0.598	0.0720	0.245
Dep. Var. Mean			0.0763	0.169	0.266	11.97	0.0384	2.363	0.283
F - Stat			22.05	22.05	22.05	22.05	22.05	22.05	22.05

Notes: The reported regressions correspond to the 1st Stages (Col. 1-2), and to the 2nd Stage with multiple outcome variables (Col. 3-8) and visit length interacted by the physician's seniority. For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Joint Length p-value is the joint p-value of both *Length* and *Length × Senior*. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Kleibergen and Paap (2006). *** p<0.01, ** p<0.05, * p<0.1.

Table 14: Effect of Visit Length on Diagnosis Provision - By Seniority

	(1) Length	(2) Length Senior	(3) Diagnosis	(4) Common	(5) Uncommon
Length			0.0073**	0.0009	0.0064***
			(0.0030)	(0.0014)	(0.0021)
Length × Senior			-0.0075*	-0.0010	-0.0065**
			(0.0042)	(0.0016)	(0.0033)
Senior	-0.7348***	11.1466***	0.0473	-0.0004	0.0476
	(0.2682)	(1.0392)	(0.0523)	(0.0199)	(0.0415)
Prior Cancel	1.7592***	-0.0574**			
	(0.2560)	(0.0274)			
Prior Cancel × Senior	-0.1712	1.7390***			
	(0.3677)	(0.2621)			
Month-Year FE	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Observations	58301	58301	58301	58301	58301
Joint Length p-value	–	–	0.949	0.945	0.959
Dep. Var. Mean	–	–	0.0763	0.0110	0.0653
F - Stat	–	–	22.05	22.05	22.05

Notes: The reported regressions correspond to the 1st Stages (Col. 1-2), and to the 2nd Stage with the following outcomes: i) the probability of a diagnosis (Col. 3), ii) the probability of a common diagnosis (Col. 4), and iii) the probability of an uncommon diagnosis (Col. 5). Diagnoses are classified as common identify those diagnoses most repeated in a given specialization, while uncommon represent any other non modal diagnosis. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Joint Length p-value is the joint p-value of both *Length* and *Length* × *Senior*. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Kleibergen and Paap (2006). *** p<0.01, ** p<0.05, * p<0.1.

Appendix

A Tables

Table A1: Effect of a Prior Cancellation on Visit Length - Time to End Shift

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	6 Hours	5 Hours	4 Hours	3 Hours	2 Hours	Last Hour	Overtime
	Length	Length	Length	Length	Length	Length	Length
Prior Cancel	2.5271***	1.9414***	1.6238***	1.6337***	1.7079***	0.6215**	-0.3333
	(0.6106)	(0.2728)	(0.2345)	(0.2409)	(0.2304)	(0.2858)	(1.1706)
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2014	6558	12031	14323	14687	11964	3862
Dep. Var. Mean	13.44	13.29	12.64	12.64	12.18	12.42	10.92

Notes: The reported regressions correspond to the first stage estimation using *Prior Cancel* as the main regressor. The ending time in a given shift is calculated using appointment times. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Standard errors are clustered at the physician level. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Covariate test - Patient choice specializations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Male	Age	Ref. BHZ	Dist. BHZ	Chronic	Pub. Cov	Spanish	Waiting list	Same sex	Same age
Prior Cancel	-0.0037	0.2639	0.0108	-0.2057	0.0038	-0.0003	-0.0051	1.2624	0.0010	0.0088*
	(0.0054)	(0.2226)	(0.0066)	(0.1277)	(0.0032)	(0.0010)	(0.0055)	(0.9336)	(0.0056)	(0.0052)
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	47832	47832	47832	47832	47832	47832	47832	47832	46652	40682
Dep. Var. Mean	0.465	60.28	0.645	3.812	0.0610	0.991	0.703	31.63	0.526	0.143

Notes: The table tests whether having a prior cancellation predicts the patient and shared physician-patient characteristics, on those specializations in which patients can choose their preferred slot. Ref. BHZ is an indicator variable that identifies if the patient comes from Basic Health Zone covered by the hospital. Dist. BHZ measures how many kilometers apart is the patient's Basic Health zone from the hospital using a linear distance algorithm. Chronic is an indicator variable that identifies if the patient previously had any chronic condition. Pub. Cov. identifies if a visit was covered by the public insurance scheme. Spanish identifies those patients born in Spain. Waiting list measures the days that patients wait to access a first visit from their corresponding Primary Care center. Same sex identifies if both physician and patient share the same sex. Same age identifies if both physician and patient have a similar age, measured using a 10 years window. All other variables are self-explanatory. Standard errors are clustered at the physician level. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Covariate test - Late prior patient

	(1) Male	(2) Age	(3) Ref. BHZ	(4) Dist. BHZ	(5) Chronic	(6) Pub. Cov	(7) Spanish	(8) Waiting list	(9) Same sex	(10) Same age
Prior Cancel	-0.0039	0.2141	0.0060	-0.0497	0.0032	0.0002	-0.0052	1.0555	-0.0032	0.0046
	(0.0045)	(0.1806)	(0.0056)	(0.1495)	(0.0023)	(0.0012)	(0.0047)	(0.7017)	(0.0046)	(0.0042)
Prior Late	0.0009	-0.3156	0.0112	-0.2361***	0.0003	-0.0046**	-0.0016	-0.2962	-0.0096	-0.0011
	(0.0062)	(0.2046)	(0.0068)	(0.0824)	(0.0024)	(0.0020)	(0.0042)	(0.8013)	(0.0059)	(0.0031)
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530	67530	66350	58301
Dep. Var. Mean	0.447	58.85	0.598	4.365	0.0582	0.984	0.677	29.73	0.517	0.152

Notes: The table tests whether a late arrival of the previous patient predicts the current patient and shared physician-patient characteristics. Ref. BHZ is an indicator variable that identifies if the patient comes from Basic Health Zone covered by the hospital. Dist. BHZ measures how many kilometers apart is the patient's Basic Health zone from the hospital using a linear distance algorithm. Chronic is an indicator variable that identifies if the patient previously had any chronic condition. Pub. Cov. identifies if a visit was covered by the public insurance scheme. Spanish identifies those patients born in Spain. Waiting list measures the days that patients wait to access a first visit from their corresponding Primary Care center. Same sex identifies if both physician and patient share the same sex. Same age identifies if both physician and patient have a similar age, measured using a 10 years window. All other variables are self-explanatory. Standard errors are clustered at the physician level. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Effect of Visit Length and Delay on Visit Outcomes

	(1) Length	(2) Delay	(3) Diagnosis	(4) Tests	(5) Num. Tests	(6) Test Cost	(7) Drugs	(8) Num. Drugs	(9) Follow-up
Length			0.0036**	0.0062***	0.0095**	0.7484**	-0.0005	-0.3482**	0.0105***
			(0.0017)	(0.0024)	(0.0045)	(0.3344)	(0.0009)	(0.1726)	(0.0032)
Delay			-0.0000	-0.0003	-0.0001	-0.0686	0.0006*	0.0764	0.0016*
			(0.0006)	(0.0006)	(0.0010)	(0.0802)	(0.0003)	(0.0686)	(0.0008)
Prior Cancel	1.6256***	-1.1688**							
	(0.1595)	(0.4783)							
Prior Late	0.1356	6.1898***							
	(0.1120)	(0.6744)							
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530	67530	67530
Dep. Var. Mean	12.58	16.20	0.0819	0.181	0.286	12.67	0.0333	2.043	0.280
F - Stat	–	–	42.39	42.39	42.39	42.39	42.39	42.39	42.39

Notes: The reported regressions correspond to the two 1st Stages (Col. 1-2), and the 2nd Stage with multiple outcome variables (Col. 3-9). Prior Late is an indicator variable that identifies whether the previous patient arrived to the hospital after her scheduled visit time. For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table A3 for the corresponding instrument covariate test. F-Stat corresponds to the first-stage joint F-statistics measure proposed by Kleibergen and Paap (2006). Standard errors are clustered at the physician level. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Effect of Visit Length on Visit Outcomes - No controls

	(1) Length	(2) Diagnosis	(3) Test	(4) Num. Tests	(5) Test Cost	(6) Drug	(7) Num. Drugs	(8) Follow-up
Length		0.0021	0.0066***	0.0098**	0.7940**	-0.0012	-0.4193*	0.0093***
		(0.0022)	(0.0023)	(0.0043)	(0.3484)	(0.0011)	(0.2182)	(0.0032)
Prior Cancel	1.6205***							
	(0.1585)							
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No	No	No
Observations	67530	67530	67530	67530	67530	67530	67530	67530
Dep. Var. Mean	12.58	0.0819	0.181	0.286	12.67	0.0333	2.043	0.280
F - Stat	–	105.8	105.8	105.8	105.8	105.8	105.8	105.8

Notes: The reported regressions correspond to the 1st Stage (Col. 1), and the 2nd Stage with multiple outcome variables (Col. 2-8). For information on the outcome variables, please refer to Section 4.1. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Olea and Pflueger (2013). *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Effect of a Prior Cancellation on Visit Outcomes - ITT

	(1) Diagnosis	(2) Test	(3) Num. Tests	(4) Test Cost	(5) Drug	(6) Num. Drugs	(7) Follow-up
Prior Cancel	0.0058*	0.0105**	0.0155**	1.3050**	-0.0016	-0.6661*	0.0149***
	(0.0029)	(0.0041)	(0.0072)	(0.5999)	(0.0017)	(0.3433)	(0.0052)
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530
Dep. Var. Mean	0.0819	0.181	0.286	12.67	0.0333	2.043	0.280

Notes: The reported regressions correspond to the ITT estimation using *Prior Cancel* as the main regressor. For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Standard errors are clustered at the physician level. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Effect of Visit Length on Visit Outcomes - OLS

	(1) Diagnosis	(2) Test	(3) Num. Tests	(4) Test cost	(5) Drug	(6) Num. Drugs	(7) Follow-up
Length	0.0011***	0.0012**	0.0029***	0.2017***	0.0003***	0.0210**	0.0034***
	(0.0002)	(0.0005)	(0.0011)	(0.0569)	(0.0001)	(0.0100)	(0.0006)
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530
Dep. Var. Mean	0.0819	0.181	0.286	12.67	0.0333	2.043	0.280

Notes: The reported regressions correspond to the OLS estimation using *Length* as the main regressor. For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Standard errors are clustered at the physician level. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Effect of Visit Length on Clinical Process Duration

	(1) Length	(2) Case duration	(3) Length	(4) Case duration
Length		1.1678**		0.6699
		(0.5564)		(0.9142)
Prior Cancel	1.6222***		1.8414***	
	(0.1598)		(0.1961)	
Month-Year FE	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	67530	67530	18846	18846
Dep. Var. Mean	12.58	30.13	14.36	108
F - Stat	–	104.3	–	89.24

Notes: The reported regressions correspond to the 1st Stage (Col. 1 & 3), and the 2nd Stage (Col. 2 & 4). The variable Case duration measures the number of days, after a first visit, that has taken a clinical process to end. Columns 1 and 2 use the whole sample and provide a value 0 to those first visits that had no follow-up, and Columns 3 and 4 use only those first visits that scheduled a follow-up visit. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Olea and Pflueger (2013). *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Covariate test - Follow-up visits

	(1) Male	(2) Age	(3) Ref. BHZ	(4) Dist. BHZ	(5) Chronic	(6) Pub. Cov	(7) Spanish	(8) Waiting list	(9) Same sex	(10) Same age
Prior Cancel	-0.0071	0.2368	0.0158	0.2433	0.0045	0.0031	-0.0226**	1.9820	0.0043	0.0072
	(0.0096)	(0.4141)	(0.0113)	(0.3040)	(0.0051)	(0.0023)	(0.0108)	(1.2038)	(0.0095)	(0.0087)
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14350	14350	14350	14350	14350	14350	14350	14350	14266	12530
Dep. Var. Mean	0.432	62.16	0.663	4.078	0.0702	0.977	0.696	27.09	0.523	0.134

Notes: The table tests whether having a prior cancellation predicts the patient and the shared physician-patient characteristics in a sample of visits with a follow-up appointments. Ref. BHZ is an indicator variable that identifies if the patient comes from a Basic Health Zone covered by the hospital. Dist. BHZ measures how many kilometers apart is the patient's Basic Health zone from the hospital using a linear distance algorithm. Chronic is an indicator variable that identifies if the patient previously had any chronic condition. Pub. Cov. identifies if a visit was covered by the public insurance scheme. Spanish identifies those patients born in Spain. Waiting list measures the days that patients wait to access a first visit from their corresponding Primary Care center. Same sex identifies if both physician and patient share the same sex. Same age identifies if both physician and patient have a similar age, measured using a 10 years window. All other variables are self-explanatory. Standard errors are clustered at the physician level. *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Effect of Visit Length on Visit Outcomes - By Specialization type

	(1) Length	(2) Diagnosis	(3) Tests	(4) Num. Tests	(5) Test Cost	(6) Drugs	(7) Num. Drugs	(8) Follow-up
Panel A: Internal medicine specialization								
Length		0.0055***	0.0060	0.0110**	1.1976**	0.0003	-0.6096	0.0136***
		(0.0015)	(0.0039)	(0.0055)	(0.5922)	(0.0011)	(0.5049)	(0.0034)
Prior Cancel	2.2361***							
	(0.3030)							
Observations	23339	23339	23339	23339	23339	23339	23339	23339
Dep. Var. Mean	15.65	0.0747	0.197	0.278	14.07	0.0295	2.840	0.343
F - Stat	–	55.86	55.86	55.86	55.86	55.86	55.86	55.86
Panel B: Surgical specialization								
Length		0.0022	0.0063**	0.0087	0.4616	-0.0020	-0.2617**	0.0065
		(0.0028)	(0.0028)	(0.0061)	(0.3715)	(0.0016)	(0.1070)	(0.0048)
Prior Cancel	1.3514***							
	(0.1691)							
Observations	44141	44141	44141	44141	44141	44141	44141	44141
Dep. Var. Mean	10.95	0.0857	0.172	0.291	11.94	0.0353	1.624	0.248
F - Stat	–	65.23	65.23	65.23	65.23	65.23	65.23	65.23
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The reported regressions correspond to the 1st Stage (Col. 1), and the 2nd Stage with multiple outcome variables (Col. 2-8). Panel A includes all observations covering visits that happened in an internal medicine specialization, while Panel B includes those that happened at a surgical specialization. For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. The specialties classified as internal medicine are: Allergology, Cardiology, Dermatology, Endocrinology, Internal Medicine, Neurology, Oncology, Pain pathologies, Pulmonology, and Rheumatology; while those specialties classified as surgical are: Cardiovascular surgery, General surgery, Maxillofacial surgery, Ophthalmology, Orthopedics, Otolaryngology, and Urology. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Olea and Pflueger (2013). *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Effect of Visit Length on Visit Outcomes - By Retired Patients

	(1) Length	(2) Length Retired	(3) Diagnosis	(4) Test	(5) Num. Tests	(6) Test Cost	(7) Drug	(8) Num. Drugs
Length			0.0035	0.0043	0.0107	0.7382	-0.0010	-0.4595*
			(0.0029)	(0.0031)	(0.0066)	(0.4982)	(0.0013)	(0.2470)
Length × Retired			0.0000	0.0051	-0.0019	0.1852	-0.0001	0.1144
			(0.0043)	(0.0053)	(0.0095)	(0.7283)	(0.0016)	(0.2279)
Retired	0.2615	12.6102***	-0.0057	-0.0886	-0.0202	-5.4413	0.0006	-1.4926
	(0.1708)	(0.5866)	(0.0569)	(0.0661)	(0.1227)	(9.4129)	(0.0207)	(3.0157)
Prior Cancel	1.5211***	-0.0264						
	(0.1618)	(0.0627)						
Prior Cancel × Retired	0.2585	1.7836***						
	(0.1876)	(0.2516)						
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530	67530
Joint Length p-value	–	–	0.177	0.0180	0.144	0.0683	0.413	0.142
Dep. Var. Mean	–	–	0.0819	0.181	0.286	12.67	0.0333	2.043
F - Stat	–	–	27.58	27.58	27.58	27.58	27.58	27.58

Notes: The reported regressions correspond to the 1st Stage regression (Col. 1-2) and the 2nd Stage with multiple outcome variables (Col. 3-8). The table presents the interaction of *Length* and whether the patient's age is over 65 (*Retired*). For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Joint Length p-value is the joint p-value of both *Length* and *Length* × *Retired*. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Kleibergen and Paap (2006). *** p<0.01, ** p<0.05, * p<0.1.

Table A12: Effect of Visit Length on Visit Outcomes - By Shock type

	(1) Length	(2) Diagnosis	(3) Tests	(4) Num. Tests	(5) Test Cost	(6) Drugs	(7) Num. Drugs	(8) Follow-up
Panel A: No-Show								
Length		0.0036**	0.0078***	0.0111**	0.7637*	-0.0010	-0.3939*	0.0084***
		(0.0018)	(0.0027)	(0.0047)	(0.3910)	(0.0010)	(0.2106)	(0.0032)
Prior No-Show	1.6082***							
	(0.1594)							
Observations	66320	66320	66320	66320	66320	66320	66320	66320
Dep. Var. Mean	12.53	0.0817	0.181	0.286	12.63	0.0332	2.055	0.280
F - Stat	–	103	103	103	103	103	103	103
Panel B: Notification								
Length		0.0037	-0.0036	-0.0022	1.1875	-0.0011	-0.5477	0.0149
		(0.0046)	(0.0058)	(0.0097)	(0.8695)	(0.0027)	(0.3340)	(0.0091)
Prior Notification	1.7260***							
	(0.2885)							
Observations	57702	57702	57702	57702	57702	57702	57702	57702
Dep. Var. Mean	12.39	0.0816	0.180	0.286	12.59	0.0319	2.055	0.279
F - Stat	–	36.20	36.20	36.20	36.20	36.20	36.20	36.20
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The reported regressions correspond to the 1st Stage (Col. 1), and the 2nd Stage with multiple outcome variables (Col. 2-8). Panel A includes all observations, but those with a prior withdrawal, while Panel B includes all observations, but those with a prior no show up. For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Olea and Pflueger (2013). *** p<0.01, ** p<0.05, * p<0.1.

Table A13: Effect of Visit Length on Visit Outcomes - By Overloaded Days

	(1) Length	(2) Length Non Overload	(3) Diagnosis	(4) Test	(5) Num. Tests	(6) Test Cost	(7) Drug	(8) Num. Drugs
Length			0.0043	0.0099*	0.0132*	0.9296	-0.0015	-0.5138*
			(0.0027)	(0.0054)	(0.0079)	(0.5892)	(0.0030)	(0.2643)
Length × Non Overload			-0.0011	-0.0052	-0.0054	-0.1730	0.0007	0.1523
			(0.0037)	(0.0059)	(0.0099)	(0.7140)	(0.0033)	(0.2747)
Non Overload	0.6932***	12.0272***	0.0042	0.0458	0.0549	2.9035	-0.0115	-1.6484
	(0.1776)	(0.5285)	(0.0461)	(0.0715)	(0.1144)	(8.5425)	(0.0393)	(3.3703)
Prior Cancel	1.9344***	0.0763						
	(0.2555)	(0.0887)						
Prior Cancel × Non Overload	-0.4097	1.4005***						
	(0.2570)	(0.2099)						
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530	67530
Joint Length p-value	–	–	0.184	0.0406	0.140	0.0699	0.434	0.136
Dep. Var. Mean	–	–	0.0819	0.181	0.286	12.67	0.0333	2.043
F - Stat	–	–	27.58	27.58	27.58	27.58	27.58	27.58

Notes: The reported regressions correspond to the 1st Stage regression (Col. 1-2) and the 2nd Stage with multiple outcome variables (Col. 3-8). The table presents the interaction of *Length* and whether the physician had a non-pressing day (*Non Overload*). The variable *Non Overload* identifies those days in which the total expected visit length exceeds the physician's daily schedule. For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Joint Length p-value is the joint p-value of both *Length* and *Length* × *Retired*. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Kleibergen and Paap (2006). *** p<0.01, ** p<0.05, * p<0.1.

Table A14: Effect of Visit Length on Visit Outcomes - By High-Performing physicians

	(1) Length	(2) Length High-Performing	(3) Diagnosis	(4) Test	(5) Num. Tests	(6) Test Cost	(7) Drug	(8) Num. Drugs
Length			0.0023	0.0056**	0.0116**	0.7769**	-0.0003	-0.3945
			(0.0016)	(0.0025)	(0.0053)	(0.3669)	(0.0008)	(0.2833)
Length × High-Performing			0.0041	0.0026	-0.0064	0.0846	-0.0021	-0.0510
			(0.0048)	(0.0052)	(0.0084)	(0.7778)	(0.0026)	(0.3128)
High-Performing	-1.2899	9.3470***	-0.1316**	-0.1999**	-0.2265	-15.6842	0.0448	1.6781
	(1.6542)	(1.8655)	(0.0540)	(0.0800)	(0.1508)	(9.5695)	(0.0435)	(4.8503)
Prior Cancel	1.8436***	-0.0519***						
	(0.2197)	(0.0195)						
Prior Cancel × High-Performing	-0.5610*	1.4276***						
	(0.3258)	(0.2391)						
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67530	67530	67530	67530	67530	67530	67530	67530
Joint Length p-value	–	–	0.156	0.0810	0.434	0.225	0.350	0.0280
Dep. Var. Mean	–	–	0.0819	0.181	0.286	12.67	0.0333	2.043
F - Stat	–	–	18.13	18.13	18.13	18.13	18.13	18.13

Notes: The reported regressions correspond to the 1st Stage regression (Col. 1-2) and the 2nd Stage with multiple outcome variables (Col. 3-8). The table presents the interaction of *Length* and whether the physician's average time used to provide a diagnosis is lower than the average time used in her specialization (*High-Performing*). For information on the outcome variables, please refer to Section 4.1. All regressions include the following controls: Patient sex, age, square age, whether the patient is from the reference BHZ, the distance from the patient BHZ to the hospital, whether the patient is a chronic, whether the patient was born in Spain, whether the patient is covered by the public insurance, the days and squared days passed since the first visit referral, whether the visit was forced into the agenda, and whether the visit was referred by a colleague. See Table 2 for further reference. Joint Length p-value is the joint p-value of both *Length* and *Length* × *Retired*. Standard errors are clustered at the physician level. F-Stat corresponds to the first-stage F-statistics measure proposed by Kleibergen and Paap (2006). *** p<0.01, ** p<0.05, * p<0.1.