INSEAD The Business School for the World®

Working Paper

2022/36/TOM

(Revised version of 2021/32/TOM)

Continuity of Care Increases Physician Productivity in Primary Care

Harshita Kajaria-Montag
University of Cambridge, hk437@cam.ac.uk

Michael Freeman
INSEAD, michael.freeman@insead.edu

Stefan Scholtes
University of Cambridge, s.scholtes@jbs.cam.ac.uk

Continuity of care, defined as an ongoing therapeutic relationship between a patient and a physician, is a defining characteristic of primary care. However, arranging a consultation with one's regular doctor is increasingly difficult as practices face physician shortages. We study the effect of declining care continuity on the productivity of physicians by analyzing data of over 10 million consultations in 381 English primary care practices over a period of 11 years. Specifically, we examine whether a consultation with the patient's regular doctor is more productive than with another doctor in the practice. Using statistical models that account for confounding and selection bias and restricting the sample to consultations with patients who had at least three consultations over the past two years, we find that the time to a patient's next visit is on average 18.1% (95% CI: [16.9%, 19.2%]) longer when the patient sees the doctor they have seen most frequently over the past two years, while there is no operationally meaningful difference in consultation duration. The data shows that the productivity benefit of care continuity is larger for older patients, patients with multiple chronic conditions, and patients with mental health conditions. We estimate that the total consultation demand in our sample could have fallen by up to 5.2% had all practices offered continuity of care at the level of the top decile of practices while prioritizing patients expected to yield the largest productivity benefits. We discuss operational and strategic implications of these findings for primary care practices and for third-party payers.

Keywords: Healthcare; Continuity of Care; Productivity; Primary Care
History: August 23, 2022

Continuity of care increases physician productivity in primary care

Harshita Kajaria-Montag
Judge Business School, University of Cambridge, Cambridge CB2 1AG, United Kingdom, hk437@cam.ac.uk

Michael Freeman
INSEAD, 1 Ayer Rajah Avenue, Singapore 138676, michael.freeman@insead.edu

Stefan Scholtes
Judge Business School, University of Cambridge, Cambridge CB2 1AG, United Kingdom, s.scholtes@jbs.cam.ac.uk

Continuity of care, defined as an ongoing therapeutic relationship between a patient and a physician, is a defining characteristic of primary care. However, arranging a consultation with one's regular doctor is increasingly difficult as practices face physician shortages. We study the effect of declining care continuity on the productivity of physicians by analyzing data of over 10 million consultations in 381 English primary care practices over a period of 11 years. Specifically, we examine whether a consultation with the patient's regular doctor is more productive than with another doctor in the practice. Using statistical models that account for confounding and selection bias and restricting the sample to consultations with patients who had at least three consultations over the past two years, we find that the time to a patient's next visit is on average 18.1% (95% CI: $[16.9\%, 19.2\%]$) longer when the patient sees the doctor they have seen most frequently over the past two years, while there is no operationally meaningful difference in consultation duration. The data shows that the productivity benefit of care continuity is larger for older patients, patients with multiple chronic conditions, and patients with mental health conditions. We estimate that the total consultation demand in our sample could have fallen by up to 5.2% had all practices offered continuity of care at the level of the top decile of practices while prioritizing patients expected to yield the largest productivity benefits. We discuss operational and strategic implications of these findings for primary care practices and for third-party payers.

Key words: healthcare; continuity of care; productivity; primary care
History: August 23, 2022

---

1. Introduction

Primary care practices around the world are experiencing rising demand at a time when their most critical resource – primary care physician time – is becoming more scarce and more expensive. In the UK, the number of primary care physicians per 100,000 population decreased from 67 in 2009 to 60 in 2018, despite increasing demand from an aging population (Palmer 2019). The UK's Nuffield Trust estimates a shortfall of 7,000 general practitioners in the UK by 2023-24 (Beech et al. 2020), while the Association of American Medical Colleges estimates a shortfall of between 21,400 and 55,200 primary care physicians in the US by 2033 (Dall et al. 2020). These projections were made before the COVID-19 pandemic, which will likely exacerbate the shortfall. To match labor shortages with demographically driven demand growth, primary care practices need to increase the productivity of their physicians.

The productivity of primary care physicians has two main dimensions: the number of consultations that they can perform on a day and the extent to which they are able to extend the time to the patient's next consultation. Much emphasis is currently placed on the former – increasing daily throughput per physician – not least because it directly improves on-the-day access for patients. This focus on daily throughput has an unintended but important consequence: It reduces continuity of care as it becomes more difficult for patients to arrange a consultation with their regular doctor (Kajaria-Montag and Freeman 2020). In 2009, 77% of respondents to the UK's annual GP Practice Survey reported being able to see their preferred doctor at least most of the time. Ten years later, in the last survey prior to COVID-19, this proportion had dropped to 45% (Institute for Government 2019). Surprisingly, despite strong evidence of the benefits of continuity of care for patient health, health system utilization, and staff satisfaction (Palmer et al. 2018), there is a dearth of evidence on how declining rates of care continuity may be affecting physicians' productivity. It is important to fill this gap in the literature, as, if continuity of care increases physician productivity, then too much focus on throughput at the expense of care continuity will be counter-productive and can lead to an overall productivity drop. The present paper is to our best knowledge the first large-scale empirical study of the relationship between continuity of care and physician productivity.

We analyze data from over 10 million face-to-face consultations between over 14,000 primary care physicians and 1.8 million patients in 381 English primary care practices over a period of 11 years. For each consultation, we identify the patient's regular doctor as the doctor who had the most consultations with the patient over the past two years, restricting the sample to patients with at least 3 consultations over that period. We then analyze whether a patient's revisit interval (i.e., the time between the focal consultation and the patient's next visit) differed when the consultation was with his regular doctor or with another doctor in the practice. Using a range of empirical methods to control for potential selection and omitted variable bias, we find robust evidence that the revisit interval is extended by an estimated 18.1% (95% CI: $[16.9\%, 19.2\%]$) if the patient sees her regular doctor. At the same time, we find no evidence that the consultations with regular doctors are longer.

Having established the main effect, we then study its heterogeneity. We show that the productivity benefit of care continuity is particularly pronounced for patients with more complex needs, specifically older patients, patients with chronic diseases, and patients with mental health conditions. For the 30.7% consultations with patients over the age of 70, seeing the regular doctor extends the revisit interval by 20.8% (95% CI: [19.8%, 22.0%]) versus 13.5% (95% CI: [12.4%, 14.6%]) for younger patients; for the 80.6% consultations with patients with at least one chronic illness, the revisit interval is extended by 16.9% (95% CI: [15.8%, 18.0%]) versus 10.9% (95% CI: [9.7%, 12.0%]) for patients without chronic illnesses; and for the 27.9% consultations with patients with recorded mental health conditions, the revisit interval is extended by 17.5% (95% CI: [16.4%, 18.6%]) versus 15.2% (95% CI: [14.1%, 16.3%]) for patients without such conditions.

Finally, we demonstrate how the estimation methods deployed in this paper can be used as a scoring tool to enable practice managers to identify patients with the largest productivity benefit of continuity. We apply this scoring method retrospectively to the data to estimate a counterfactual demand reduction. The model suggests that if all practices in the data had offered well-targeted continuity of care for 75% of their consultations (a level achieved by the top decile of practices in the data) then the total consultation demand in the sample would have dropped by up to 5.2%.

2. Literature Review

This paper contributes to the healthcare operations and medical literature on continuity of care. These literature streams have largely focused on the consequences of care continuity on patient outcomes and secondary care utilization, such as emergency visits or hospital admissions, typically for patients with specific conditions. In contrast, we are concerned with the effect of care continuity on physician productivity in primary care practices themselves.

2.1. Continuity of care in healthcare operations

Ahuja et al. (2020b) investigate the effect of providing continuity to patients with diabetes and find that continuity improves three important system utilization metrics: inpatient admissions, hospital length of stay, and readmission rate. In a follow-up study, Ahuja et al. (2020a) partially explain this relationship between continuity and the system utilization metrics by showing that continuity can lead to higher rates of medication adherence and consequently to lower glycemic variability. Senot (2019) also studies the effect of continuity on secondary care usage. Specifically, the study follows the journey of heart failure patients over a one-year period and finds that the continuity of the individual referring provider (along with continuity of the physical location and the accountable care organization) contributes to a reduced risk of hospital readmission for heart failure patients. Queenan et al. (2019) find that providing technology-enabled continuity coupled with increasing patient engagement in their own health reduces hospital readmissions.

The present study complements this research on system utilization effects by focusing on productivity effects within the primary care setting itself. This is important because the direction of the internal productivity effect tells us whether or not practices need to be externally incentivized to provide the continuity of care that will create the documented system utilization benefits.

While the healthcare operations literature has hitherto not engaged much with primary care productivity, Bavafa et al. (2018) is a notable exception. Their paper focuses on the impact of complementing office visits by e-visits on demand in primary care, and the authors show that the introduction of this new communications channel increases demand for office visits. We also address how demand for primary care services changes as a function of how the service provided. Our focus, however, is on the potential demand-inducing effect of reduced care continuity.

Li et al. (2021) focus on telemedicine adoption in an outpatient context. They show that adoption of telemedicine reduces productivity in the short term, by shortening the interval between patient visits, but that the interval between visits increases in the long run. We follow this study in using the revisit interval as a measure of productivity but our focus lies on continuity in primary care.

2.2. Continuity of care in the medical literature

The medical literature differentiates between different types of care continuity, specifically relational continuity, management continuity, and informational continuity (Haggerty et al. 2003). Within a primary care context, the terms relational continuity and continuity of care are often used synonymously and defined as an ongoing therapeutic relationship between a patient and a physician. In this paper, the focus is on this relational component of continuity, as captured by repeat consultations with the same primary care physician.

It is well documented that continuity of care in primary care is valued by patients and doctors alike and surveys highlight various benefits of providing care continuity (Freeman et al. 2010). Specifically, the medical literature has demonstrated various direct health benefits for patients and improved management of health conditions for those who receive care continuity. For instance, studies have shown improvements in quality of life outcomes (Drury et al. 2020, Chen et al. 2017, Ye et al. 2016), blood pressure for diabetic and hypertensive patients (Leniz and Gulliford 2019), mortality (Maarsingh et al. 2016, Cho et al. 2015), adherence to medication plans (Dossa et al. 2017), and the likelihood of filling risky prescriptions (Hallvik et al. 2018).

In terms of system benefits, a meta-analysis by Huntley et al. (2014), involving participants from OECD countries, found that unscheduled secondary care usage is highly influenced by care continuity in the primary care setting. For example, primary care continuity has been associated with reductions in emergency department visits (Pourat et al. 2015) and unplanned hospitalizations of patients with ambulatory care sensitive conditions (Barker et al. 2017). Such advantages have been consistently demonstrated across different patient populations, including patients with serious mental illness (Ride et al. 2019), dementia (Amjad et al. 2016), COPD (Lin et al. 2015), and diabetes (Worrall and Knight 2011, Dossa et al. 2017), as well as older patients (Tammes et al. 2017, Katz et al. 2015, Bayliss et al. 2015, Nyweide et al. 2013). We contribute to this stream of literature by demonstrating that care continuity also affects the need for primary care visits themselves and that this effect is particularly pronounced in older patients and patients with complex conditions, such as chronic diseases or mental illnesses.

In summary, there is rich evidence to show the benefits of relational continuity for both patients and the wider health system in terms of reduced utilization. It is therefore somewhat surprising that the effect of continuity of care on physician productivity within primary care practices has not yet been investigated. This study expands existing knowledge of the effects of care continuity by showing that care continuity not only improves outcomes and system utilization but also enhances the productivity of primary care physicians.

3. The Hypotheses

Before we present and discuss the paper's central hypotheses, we clarify the key variables – care continuity and physician productivity. We focus on relational continuity of care, defined as a sustained therapeutic relationship between a patient and a doctor. In primary care, this relationship is epitomized by the notion of a "patient list" or "patient panel" that many primary care physicians hold, either formally or informally (Wilkin and Metcalfe 1984, Tammes et al. 2017). These are the patients for whose primary care service the doctor takes responsibility over a prolonged period of time. While, in reality, patients will have consultations with their continuity doctor as well as with other doctors in the same practice, they can be expected to see their continuity doctor most frequently. This is how we identify the patient's *regular doctor*. We study productivity effects at the level of individual patient consultations and distinguish between consultations with a patient's regular doctor, who provides continuity of care, and consultations with another doctor in the practice, who provides a more transactional service and who we call a *transactional provider*.

The productivity of a patient consultation consists of two components: it's duration and it's effectiveness, i.e., how well the doctor is able to "get it right the first time". A short but ineffective patient consultation is likely to create the need for additional consultations in the near future, while a longer, more thorough consultation may alleviate this need. Consequently, we need to consider two clinical productivity metrics: (i) the duration of the consultation, as a measure of throughput, and (ii) the revisit interval, i.e., the time to the next consultation, as a proxy measure for the effectiveness of the service.

3.1. Consultation duration

The relationship between continuity of care and consultation duration is determined by two countervailing factors. On the one hand, a regular doctor has a larger incentive to take more time to treat her regular patients thoroughly than a transactional provider. Getting it right the first time will reduce her future workload by preventing revisits, which would likely be her responsibility, while a transactional provider is less likely to see the patient for her next visit (Jeffers and Baker 2016). On the other hand, the regular doctor is more familiar with the patient than a transactional provider and can therefore be more economical in her collection of information (Hill and Freeman 2011, Rosen et al. 2020). It is difficult to argue a priori which of these two factors will prevail and there is therefore no justification for a directional hypothesis.[1] This question needs to be left to an empirical study, which we provide in Section 6.5.

3.2. Revisit interval

We focus in this paper on the second aspect of consultation productivity, the revisit interval. The paper sets out to provide evidence in support of the following two hypotheses.

HYPOTHESIS 1. *If a patient is seen by her regular doctor, there will be a longer time interval to her next consultation than if she is seen by another doctor in the practice.*

HYPOTHESIS 2. *The effect posited in Hypothesis 1 will be larger for (a) patients with multiple comorbidities, (b) older patients, and (c) patients with mental health conditions.*

Before we present the empirical evidence, we propose a number of a prior arguments in support of these hypotheses. Specifically, we draw on four mechanisms that are frequently discussed in the literature on continuity of care (Freeman et al. 2010) – familiarity, security and trust, "collusion of anonymity", and one-stop-shop incentives – and relate these mechanisms to the time to the patient's next consultation.

Familiarity. As the patient's regular doctor sees the patient repeatedly, she becomes familiar with the patient's health trajectory as well as with their preferences, behaviors and personal circumstances. This allows the regular doctor to provide more effective customized advice and treatment and thereby reduces the need for revisits (Hjortdahl and Borchgrevink 1991).

Specifically, because a regular doctor forms an impression of the patient's health trajectory she will be able to respond not only to the patient's health status at the time of the consultation but also to *changes* in their health (Koopman et al. 2003). This longitudinal, and often unrecorded information that a one-off provider does not have allows the regular doctor to make a more accurate assessment of the patient's health needs (O'Connor et al. 1998, Ramanayake and Basnayake 2018). Indeed, the primary care physicians we interviewed for this study confirmed that when one of their regular patients enters the consultation room, they often know at a glance whether or not the patient is seriously ill and that they cannot do this for patients they are not familiar with.

We have also interviewed geriatricians, who have confirmed that the knowledge of a patient's longitudinal health trajectory is particularly relevant for older patients, where first-time impressions of patients can be very misleading. More information is gained from an observed change of the patient's health. Being able to identify these changes and act appropriately allows the patient's regular doctor to reduce the need for future visits and this effect will be stronger for older patients.

[1] In fact, because our study uses UK data, we expect, a priori, little difference in the average duration of patient consultations between the regular doctor and a transactional provider in this data. This is because, although not a formally imposed guideline, 10 minute consultations have become the de facto standard for primary care in the UK (Royal College of General Practitioners 2019).

Trust. Repeated patient interactions over time provides the regular doctor not only with a more comprehensive understanding of the patient's needs and circumstances but also allows her to form a trust-based relationship that may help her communicate more effectively with the patient (Tarrant et al. 2010, Özer et al. 2014, Hill and Freeman 2011). The patient may feel more secure sharing information with her regular doctor and, as a consequence, the doctor will be able to design a more appropriate treatment plan (von Bültzingslöwen et al. 2006). Furthermore, trust will increase adherence to an agreed disease management plan (Ahuja et al. 2020a, Dossa et al. 2017, Brookhart et al. 2007) which is likely to reduce the need for revisits.

Trust goes both ways - the regular doctor also knows and trusts her patients and may therefore be more willing to adopt a wait-and-see approach when this is indicated, to avoid undesirable medicalisation (Hjortdahl and Borchgrevink 1991). They feel both sufficiently familiar with the patient and secure to propose such an approach when appropriate. By contrast a transactional provider is more likely to start the patient on an interventional diagnostic trajectory, to be "on the safe side" and avoid the potential risk of litigation. Unnecessary medicalisation is likely to increase the need for follow-up visits and reduce the time to the next consultation.

We expect the advantage of a trust-based relationship to be particularly important for patients with mental health conditions, such as anxiety, depression and schizophrenia (Biringer et al. 2017). The stigma that is attached to mental health might make patients reluctant to be fully transparent with an unfamiliar physician (Knaak et al. 2017). Moreover, medication compliance is a particular problem for mental health patients, due in part to the negative side effects associated with commonly prescribed medications (e.g. weight gain and fatigue) (Semahegn et al. 2018).

Collusion of anonymity. In the medical literature, collusion of anonymity refers to a situation where a physician has an incentive to do the minimal possible to pass the patient safely on to the next provider (Freeman et al. 2010, Balint 1955). This incentive is larger for a transactional provider, who is less likely to have to take potential follow-up visits. By contrast, the regular doctor has a strong incentive to reduce the likelihood of follow-up consultations, to reduce her future workload. This makes her more likely to diagnose the patient's problem more carefully in an attempt to "get it right the first time" (Koopman et al. 2003). From a time and productivity perspective, such a more careful root-cause diagnosis may not even cost the doctor much time if she is familiar with the patient.

One-stop-shop incentive. A regular doctor has an incentive to leverage the patient encounter to explore the patient's health needs beyond the immediate reason for the consultation, as this may prevent an unnecessary visit in the near future (Hill and Freeman 2011). She may check her notes and proactively deal with multiple illnesses or health issues in a single appointment. Transactional providers lack both the incentive to expand their scope of service beyond the immediate clinical need expressed by the patient and the holistic patient knowledge that facilitates proactive management of the patient's health (Balint 1955). Opportunities for such proactive interventions, outside the scope of the immediate reason for the visit, are particularly salient when the patient has a chronic disease (Koopman et al. 2003, Goodwin et al. 2010). We therefore expect the productivity benefit of seeing a regular doctor to be larger for patients with multiple chronic conditions.

In summary, these four mechanisms lend a priori support to Hypothesis 1 and 2.

4. Clinical Setting, Data and Variables

In this section, we first provide a brief overview of the specifics of the UK primary care context that are relevant for this study. We then describe the dataset in detail and conclude with the description of the dependent variables, independent variables, and controls to be included in the analysis.

4.1. Primary care context

Although the English National Health Service (NHS) is publicly funded through taxation, primary care practices are privately owned businesses, organized as partnerships of primary care physicians. Unlike hospitals, they operate as independent contractors of the NHS and therefore not under its direct control. Instead, the NHS controls their services through standardized contractual arrangements. A typical practice has 8,000-10,000 registered patients, 4-5 full-time equivalent physicians, and a small number of other healthcare workers and administrative staff. Practice income is largely capitation-based, adjusted for demographic and socioeconomic characteristics of the practice population and geography. In 2021, a typical practice with 9,000 registered patients received an income of £1.44M, approximately £160 per registered patient (NHS Digital 2021).

The contract of a general practice in England defines the geographical catchment area for the practice. Patients who live in this area have the right to register with that practice. Patients may apply for registration with any practice but practices have discretion to accept or deny out-of-area patients. Importantly, patients can only register with a single practice in England and are automatically deregistered when they register with a new practice. Since our data is practice-based, we therefore have full visibility of all primary care appointments of patients for their period of registration in the study practices.

Primary care services in England are free at the point of care. Patients can request to see any doctor at their practice, and practices generally try to accomodate this request if the doctor is available. The NHS contract requires that each patient registered at a practice is assigned to a *named doctor*, who is responsible for ensuring that the patient's needs are met (Tammes et al. 2017). However, some practices regard this as a purely administrative requirement, so the patient's named doctor may not be her regular doctor. This study focuses on the patient's regular doctor, the physician who has seen the patient most frequently in the past.

Consultations can be face-to-face, over the phone or video link or, in rare cases, at the patient's home. Appointments are generally booked via the phone with a receptionist. Practices have to accommodate routine and urgent appointments, the latter requiring on-the-day access. Some practices reserve a number of appointment slots for urgent services and, if those slots are booked, refer patients to a hospital emergency department or ask them to call again the next day. Other practices accept all patients who call in before a certain cut-off time or offer unlimited access for acute care throughout the day. These practices have *duty doctors* who are dedicated to serving urgent care patients, typically in a round-robin fashion. The duty doctor role will normally rotate around all doctors in the practice.

A limitation of most primary care medical records, including ours, is that they do not record when an appointment was requested through a receptionist. We can therefore not use the period between appointment request and consultation time to distinguish between urgent and routine appointments. Instead, we use markers, such as antibiotic prescriptions, that are more commonly associated with urgent appointments to help us distinguish the two appointment types.

4.2. Data and sample

In order to understand the effect of a patient seeing her regular doctor in one consultation on the time to her subsequent consultation, we perform a cross-sectional analysis with individual patient consultations as units of observation. We obtained consultation-level data from the UK Clinical Practice Research Datalink (CPRD). This database consists of anonymized electronic medical records covering over 11.3 million patients across 674 practices in the UK; it is representative of the population in terms of age, sex and ethnicity (Herrett et al. 2015). The database encompasses a wealth of information about patients, visits, providers, diagnoses, prescriptions, referrals, treatments, immunization records, and test records. A patient's primary care data can be linked to several other data sources, including secondary care services, resulting in a fairly complete medical record of the patient's health resource usage during their registration period with one of the participating practices.

We obtained data for all English practices that had consented to linkage to secondary care usage data. The restriction to English practices improves the homogeneity of the sample as the national health systems operate differently in the four constituent countries of the UK and there are differences in their standard primary care contracts. The starting data set comprised information on 370,890,526 primary care consultations corresponding to 5,475,342 patients. The analysis sample was derived from this data using the inclusion criteria described in Section EC.1 of the e-companion and summarized in Table 1.

Table 1 Data and sample inclusion criteria

Sample inclusion criteria	Patients	Consultations	Sample reduction
Doctor consultations in 407 primary care practices	5,475,342	370,890,526	–
Face-to-face doctor consultations only	5,335,945	200,312,789	46.0%
Consultations after date at which practice data is of research quality	5,037,650	161,556,335	19.3%
Consultations during a patient's continuous registration period	4,921,208	139,455,412	13.7%
Consultations at which the patient was over 18	3,855,445	86,399,813	38.0%
Consultations with ≥ 3 and ≤ 104 consultations in the preceding 2 years	2,952,445	71,797,380	16.9%
Consultations occurring after a patient's first two years following registration	2,537,781	63,087,124	12.1%
Only consultations with a valid revisit interval	2,410,189	60,894,300	3.5%
Only consultations between January 2007 and December 2017	2,322,773	51,711,037	15.1%
Consultations occurring when the patient's regular doctor is available	2,273,571	45,376,070	12.3%
Random sample of consultations from 381 remaining practices	1,883,626	11,344,065	75.0%

Specifically, the sample for analysis consists of face-to-face clinical consultations between primary care doctors and patients over the age of 18 that took place between 1 January 2007 and 31 December 2017 at times during which (i) a practice's data is confirmed to be of research quality by CPRD, (ii) the patient had sufficient past appointments to identify their regular doctor, and (iii) the patient's regular doctor was available to see them. Data was also filtered to exclude patients visiting more than once per week on average, to only include data from periods in which the patient was continuously registered at a practice, and to exclude any consultations lacking an observable revisit interval. (For more justification, see Section EC.1 of the e-companion.)

The final analysis sample thus consists of 11,344,065 consultations between 1,883,626 patients and 14,123 doctors in 381 practices across 11 years between 2007 and 2017.

4.3. Variable description

4.3.1. Dependent variable. The main dependent variable throughout this study is the patient's revisit interval (RI), which is defined by the time between the focal face-to-face consultation with a doctor and the next face-to-face consultation with a doctor. The RI is measured in days as the data only records the day of the consultation, not the time of the day. In the infrequent event that a patient has multiple face-to-face consultations with doctors on the same day (occurring for 1% of consultations), we set the revisit interval length to 0.5 days. As is usually the case with durations, the distribution of revisit intervals is right-skewed. We therefore transform the variable by taking its natural logarithm. Figure 1 shows the distribution of the log-transformed revisit interval, and Table 2 contains summary statistics.

4.3.2. Independent variables. The main independent variable is a binary variable which indicates whether or not the focal consultation was with the patient's regular doctor. Since this study spans an 11-year time horizon, the patient's regular doctor may change over time. For example, a doctor may retire or leave the practice, or the patient may switch due to a change in

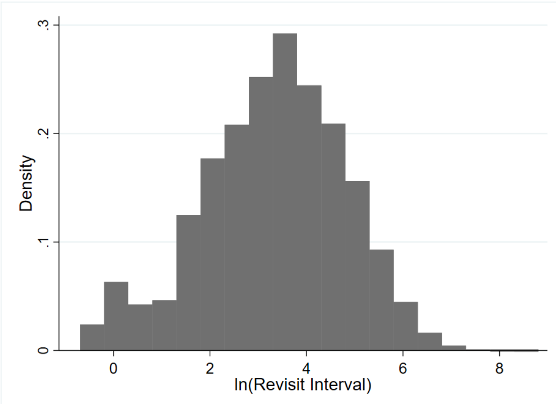


Figure 1 **Distribution of ln(Revisit Interval)**

	Revisit Interval	ln(Revisit Interval)
Mean	68.22	3.29
Median	28.00	3.33
Min	0.50	-0.69
Max	4240.00	8.34
Std. Dev.	116.80	1.48

Table 2 **Descriptive statistics for the dependent variable**

circumstances or a positive experience with another doctor. When we determine a patient's regular doctor we therefore use a dynamic measure.

Specifically, we consider a rolling two-year time window over which we calculate the patient's regular doctor. For consultation i, we define a patient's regular doctor as the doctor with whom the patient had the most face-to-face consultations over the preceding two year. To break ties (occurring in 11% of cases), we choose established over unestablished doctors[2] and, if the tie persists, we choose the doctor who the patient saw most recently. The independent variable RD_i is thus a binary variable that is equal to 1 for consultation i if the consultation is with the patient's regular doctor at the time of the consultation, and 0 otherwise. Overall, 50% of consultations in our analysis sample occur between a patient and their regular doctor. We provide additional information on the independent variable and the patient-level factors that affect it in Section EC.3.

4.3.3. Control variables. The relationship between continuity of care and a patient's revisit interval is likely to be confounded by demographic factors, the patient's consultation and revisit interval history, attributes of the regular doctor, temporal factors, and practice-level factors. To account for these potential confounders, we define a range of control variables for inclusion in the empirical analysis. A summary of the controls is provided in Table 3 and an explanation for each of the controls is given in Section EC.5.

5. Econometric Specifications

The idealized randomized experiment to test Hypothesis 1 is a random assignment of each patient consultation to either the patient's regular doctor or a randomly chosen other doctor from the

[2] The data allow us to distinguish between two types of physicians: *established* physicians, who have a contract with the practice, and *unestablished* physicians, who are not permanent employees of the practice but may be self-employed or employed with an agency. The latter work on an ad-hoc basis, often in multiple practices, and are paid on an hourly basis.

Table 3 Table of controls

Variable	Type	Description	Time-Invariant
Patient demographics			
Age	Categorical (14)	Age of the patient at the time of consultation, split into age bands (18-25, 26-30, 31-35, ..., 81-85, 86+)	No
Number of comorbidities	Categorical (6)	Number of comorbidities at the time of consultation, calculated using the Cambridge Comorbidity Index (CCI), split into bands (0, 1, ..., 4, 5+)	No
Individual comorbidities	Binary (26)	For each of the 26 comorbidities as defined by the CCI, a variable to indicate whether the patient suffers from that comorbidity at the time of the consultation	No
Mental health	Binary	A variable to indicate whether the patient suffers from a mental health condition as defined by the CCI	No
Gender	Binary	Equal to 0 if the patient is female and 1 if the patient is male	Yes
Deprivation Index	Categorical (5)	Index of multiple deprivation assigned to the patient	Yes
Prescriptions	Categorical (10)	The number of repeat prescriptions the patient is prescribed within the 6 months preceding the focal consultation, split into bands (0, 1, 2, 3, 4-5, 6-7, 8-9, 10-12, 13-15, 16+)	No
Tests	Binary	A variable to indicate whether a test was ordered within the 6 months preceding the focal consultation	
Referrals	Binary	A variable to indicate whether there was an outpatient referral within the 6 months preceding the focal consultation	
Patient's past history			
# of past consultations	Categorical (20)	The total number of consultations in the 2 years preceding the patient's focal consultation, split into 20 bands: 3-10 consultations (8 categories of size 1), 11-20 consultations (5 categories of size 2), 21-26 consultations (2 categories of size 3), 27-30 consultations (1 category of size 4), 31-35 consultations (1 category of size 5), 36-55 consultations (2 categories of size 10), 56+ consultations (1 category).	No
Past revisit interval	Continuous	The 2-year past average revisit interval of the patient calculated as described in Section EC.7.4 of the e-companion.	No
Attributes of the regular doctor			
Established doctor	Binary	A variable to indicate if the patient's assigned regular doctor for the focal consultation is an established or unestablished doctor (see Footnote 2)	No
Temporal factors			
Year	Categorical (11)	Year during which the consultation took place (2007-2017)	No
Month of Year	Categorical (12)	Month of the year in which the consultation falls (Jan-Dec)	No
Day of Week	Categorical (7)	Day of the week in which the consultation took place (Mon-Sun)	No
Practice-level factors			
Practice-level demand	Continuous	Total practice demand during the focal week of each consultation, standardized by a weekly average in a 52-week period around the focal week	No
Practice	Categorical (381)	The practice at which the consultation took place	Yes

Notes: If a variable is categorical, the number in $(\cdot)$ in the " Type" column indicates the number of levels.

same practice. We would then measure the time to the next consultation (and the duration of the visit) and report statistical differences. This section describes empirical strategies to identify these statistical differences using observational instead of experimental data.

5.1. Ordinary least squares estimator

We first perform a consultation-level least squares estimation, using the logged revisit interval after consultation i, $\ln(RI_i)$, as the dependent variable and the indicator for the regular doctor, RD_i, as the main explanatory variable, where the index i refers to a consultation. Specifically,

$$\ln(RI_i) = \beta_0 + \beta_1 RD_i + \mathbf{X}_i \boldsymbol{\beta} + \epsilon_i \ , \tag{1}$$

where $\mathbf{X}_i$ specifies the $1 \times n$ vector of controls corresponding to consultation i (as defined in Table 3) and $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ is the error term. We do not assume independence of the observations and cluster standard errors at the patient level. The effect of interest is captured by the coefficient β_1, where Hypothesis 1 posits that $\beta_1 > 0$.

The OLS estimate of β_1 provides an unbiased estimate of the average treatment effect (ATE) when the conditional independence assumption (CIA) holds, i.e., when there is independence between assignment to the treatment (here, seeing the regular doctor or not) and the outcome (here, the revisit interval), conditional on observed variables (here, $\mathbf{X}$). When one is confronted with selection on unobserved variables, the CIA fails to hold and the ATE is biased. The remainder of Section 5 is dedicated to econometric approaches for addressing violations of the CIA.

5.2. Confounding

There are two major confounding sources which can bias the coefficient β_1 in equation (1): Patient acuity and patient selection. They bias the coefficient in different directions.

Patients who present with high acuity needs may be unable to wait for an appointment with their regular doctor and are therefore more likely to see a transactional provider. At the same time, because of the acute nature of their condition, these patients may also require near-term follow-up appointments, leading to a shorter revisit interval. This biases the estimated β_1 in the OLS model (1) upwards. In the extreme case when all acute consultations are with transactional providers and all non-acute consultations with regular doctors, the variable RD_i picks up non-acuity, not continuity of care, and a positive β_i tells us that non-acute patients have longer re-visit intervals, as expected because they require fewer near-term follow-ups.

The second major confounding source is patient selection during the appointments process and as a reflection of patient preferences. Doctors and their patients are more likely to acknowledge the benefits of a continuity of care relationship when the patient has more severe medical needs. It is therefore reasonable to assume that they will make more of an effort to ensure a consultation takes place with the regular doctor. Table 4 provides some descriptive evidence that this is the case. These sicker patients are likely to visit the practice more frequently and will therefore, on average, have shorter revisit intervals. They may also have waited longer for the appointment with

Table 4 Consultation level summary statistics for transactional providers ($RD_i = 0$) and regular doctors ($RD_i = 1$)

	$RD_i = 0$	$RD_i = 1$
Average revisit interval (days)	71.11	68.35
Average age	55.88	60.68
Average number of comorbidities	2.06	2.37
Average number of prescriptions	3.93	4.70
Average % mental health prevalence	0.26	0.30

their doctor and their health may deteriorate in the meantime. Again, this should lead to shorter revisit intervals. This selection effect may not be fully captured by the available control variables and may therefore work in the opposite direction and attenuate the coefficient β_1 in Equation (1).

5.3. Higher acuity subsamples

As we cannot directly measure patient acuity, nor do we observe when the appointment was booked, adding control variables to account for heterogeneity in acuity is not feasible with our data.

To explore the potential for acuity-related confounding to be biasing the results, we first perform exploratory analyses using subsamples of consultations that are associated with patients who are likely to be higher in their level of acuity: (i) the subsample of consultations where the patient was prescribed an antibiotic, and (ii) the subsample of consultations with patients who visited an emergency department (ED) in the seven-day window prior to the focal consultation. If acuity confounds the effect, then we would expect a substantially smaller and perhaps insignificant coefficient β_1 in (1) when estimated on these subsamples. (Analyses conducted on additional acuity subsamples are documented in EC.8 of the e-companion.)

5.4. Instrumental variable estimators

The acuity subsample analyses provide useful prima facie evidence but they do not account for all confounding, and specifically not for the patient selection effect described in Section 5.2. To address confounding more generally, we use two instrumental variable estimation techniques, control functions and two-stage least squares.

5.4.1. Control functions. The control function (CF) approach is based on estimating the following selection and outcome equations

$$RD_i^* = \alpha_0 + \mathbf{X}_i\boldsymbol{\alpha} + \alpha_{n+1} IV_i + \delta_i\,, \;\; RD_i = \mathbb{1}[RD_i^* > 0]\,, \tag{2}$$

$$\ln(RI_i) = \beta_0^{CF} + \beta_1^{CF} RD_i + \mathbf{X}_i\boldsymbol{\beta}^{CF} + \gamma_1^{CF}\widehat{RD}_i + \epsilon_i^{CF}\,, \tag{3}$$

where RD_i^* is a latent variable, $\mathbb{1}[\cdot]$ is the indicator function, $\mathbf{X}_i$ is the $1 \times n$ vector of controls, IV_i is an instrumental variable (to be described shortly), $\widehat{RD}_i$ is the generalized probit residual of observation i, and $\delta_i \sim \mathcal{N}(0,1)$, $\epsilon_i^{CF} \sim \mathcal{N}(0,\sigma^2)$. We do not assume independence of the observations and cluster standard errors at the patient level.

Following the probit estimation of equation (2) as a first stage, $\widehat{RD_i}$ is calculated as

$$\widehat{RD_i} = \frac{\phi(\mathbf{X}_i'\boldsymbol{\alpha}')\,[RD_i - \Phi(\mathbf{X}_i'\boldsymbol{\alpha}')]}{\Phi(\mathbf{X}_i'\boldsymbol{\alpha}')\,[1-\Phi(\mathbf{X}_i'\boldsymbol{\alpha}')]} ,$$

where $\mathbf{X}_i'\boldsymbol{\alpha}' = \hat{\alpha}_0 + \mathbf{X}_i\hat{\boldsymbol{\alpha}} + \hat{\alpha}_{n+1}IV_i$, and $\phi(\cdot)$ and $\Phi(\cdot)$ denote the density and cumulative distribution functions of the standard normal distribution, respectively. The estimation is then completed by estimating the second stage equation (3).

The CF approach is similar to the more common two-stage least square (2SLS) method. However, in contrast to 2SLS, the CF approach estimates a probit model (i.e., equation (2)) in the first stage and then uses the generalized probit residual $\widehat{RD_i}$ as an additional control in the outcome equation (i.e., equation (3)). The addition of the generalized probit residual then adjusts the coefficient β_1^{CF} for unobserved confounders that affect both the endogenous regressor RD_i and the dependent variable $\ln(RI_i)$ (Wooldridge 2015a).[3] The t-statistic of the coefficient γ^{CF} can be used in a straightforward manner to test for endogeneity in the CF model (Wooldridge 2002).[4]

Consistency of the CF approach requires the probit model to be a correct specification for the likelihood of seeing the regular doctor, i.e., $P(RD_i = 1|\mathbf{X}_i, IV_i) = \Phi(\mathbf{X}_i'\boldsymbol{\alpha}')$. In contrast, the 2SLS estimator does not impose strict distributional assumptions on $P(RD_i = 1|\mathbf{X}_i, IV_i)$. However, using the standard 2SLS estimator with a nonlinear model in the first stage renders the estimates inconsistent (Wooldridge (2002) refers to this as the "forbidden regression"). An alternative is a 2SLS approach with a linear probability specification in the first stage, i.e., replacing the probit first stage with an OLS estimate. We use this method as a second IV approach.

5.4.2. The instrumental variable. Both the CF and 2SLS approaches rely on the availability of an instrumental variable (IV) (Wooldridge 2002). The IV must be relevant, i.e., a significant predictor in the first stage equation, and it must be valid, i.e., affect the dependent variable $\ln(RI_i)$ in the second stage equation only through its correlation with the independent variable RD_i. In this study, we use an IV that captures whether the focal patient's regular doctor is relatively more accessible for her regular patients during the week of the focal consultation, compared to her long-run average accessibility. This measure is calculated using the set of patients who share the same regular doctor as the focal patient at the time of the consultation, but it excludes any visits by the focal patient herself. The formal description of the IV calculation is given in Section EC.7.

[3] $\widehat{RD_i}$ is the generalized residual after accounting for observed confounders in the first-stage selection equation, and serves as a proxy for unobserved confounders in the outcome equation. It is used as a control in the outcome equation and if appropriately controlled for, then conditional on observed covariates, RD_i will be exogenous.

[4] We estimate the CF model using Stata's etregress command with the two-step option and bootstrapped standard errors. We refer to Wooldridge (2002, 2015a) for a technically detailed explanation of the CF approach.

We believe this is a relevant IV in our study context. A doctor who is more (or less) accessible than usual to her other regular patients is also likely to be more (or less) accessible to the focal patient (who is also one of her regular patients). Consistent with this intuition, we find a positive correlation between the IV and the patient's propensity to see her regular doctor ($\rho = 0.16$, $p < 0.001$). Formal hypothesis testing for under- and weak identification, reported in Section EC.7.1 of the e-companion, provide strong evidence that the instrument is relevant and the endogenous regressor is not weakly identified.

Turning to the validity condition, there is no reason to believe that the accessibility of the regular doctor for other patients should directly affect the revisit interval of the focal patient. Yet, it is possible that there are unobserved factors that correlate with both the relative accessibility of the regular doctor for other patients and the focal patient's revisit interval (e.g., a flu outbreak that reduces the expected revisit interval of the focal patient and also makes it harder than normal for patients to access their regular doctors). Importantly, however, such factors should affect the revisit interval not only of the focal patient but also of the doctor's other patients. Therefore, it is possible to account for these unobservable factors by adding as a control variable the average of $\ln(RI)$ of other patients who (i) share the same regular doctor as the focal patient and (ii) visit a doctor in the same week as the focal patient. When the average revisit interval of other patients changes (e.g., due to a flu outbreak), then this control variable also adjusts the expected revisit interval of the focal patient in the same direction. This control variable accounts for unobserved factors correlated with both the IV and the outcome, thus strengthening the validity of the IV. A similar approach is used in Freeman et al. (2020) and Bobroske et al. (2021). The full description of this control variable is provided in Section EC.7.4 of the e-companion.

5.5. Propensity score-based estimation

Since the validity of the instrumental variable cannot be formally tested, we add two propensity score-based estimators to further further stress-test the robustness of the results. The first uses matching, which can alleviate some of the impact associated with unobserved heterogeneity (Stuart 2010). The second, the minimum bias estimator (MBE), is an approach designed to reduce endogeneity bias without relying on an instrumental variable (Millimet and Tchernis 2013).

5.5.1. Estimating propensity scores. To implement the propensity score-based approaches, we first reduce the sample to include only one randomly chosen consultation per patient. This ensures that a patient is not matched with themselves when later applying the matching procedure to the data. However, this reduced sample is not representative of the original sample. Every patient is represented only once, independently of the frequency of their use of primary care, and therefore the reduced sample is biased towards consultations with healthier and younger individuals relative to the original sample of consultations. We address this bias by drawing a sample of 25% of the reduced consultation sample using the frequencies of the patients' visits over the entire observation period as probability weights (Jann 2006). This probability weighting re-balances this second sample towards more frequent users of primary care (i.e., the older and less healthy population) and produces a more representative subsample of consultations.

Using this subsample of consultations, we then estimate propensity scores, denoted $\hat{P}(\mathbf{X}_i)$. These are generated from a probit regression in which the dependent variable specifies whether or not the patient saw their regular provider (i.e., RD_i) and with regressors including the full set of covariates specified within the control vector $\mathbf{X}_i$ in equation (1).

5.5.2. Matching-based estimator. Using the propensity scores, $\hat{P}(\mathbf{X}_i)$, we next match consultations in the control group (those with a transactional provider for the patient's randomly chosen consultation) to the treated group (those with the patient's regular doctor for their randomly chosen consultation) using nearest neighbor matching without replacement. We also include the condition that the maximum distance between the propensity scores of two observations chosen as potential neighbors (i.e., the caliper) is 0.001. This narrow caliper helps reduce the potential for bias when examining the difference between the average revisit intervals in the two subsamples. The matched sample consists of 494,810 consultations, half of whom saw their regular doctor and half who did not. Summary statistics comparing the covariate profile of the full analysis sample and the matched sample are provided in Section EC.11 of the e-companion.

Using the matched sample, we can estimate the average treatment effect (ATE) in two ways. First, we report the difference in averages of $\ln(RI)$ between the control and the treated groups. Second, we re-estimate the OLS regression specified in equation (1) using the matched sample to control for the effect of other covariates. Generally, regression after matching is not recommended because regression standard errors do not correct for the fact that the matching step has already been performed. However, Austin and Small (2014) and Abadie and Spiess (2021) show that this can be addressed by correcting standard errors by clustering them at the level of the matched pair (i.e., two observations per cluster). We follow this recommendation when reporting effects.

5.5.3. Minimum bias estimation. There is still the potential for confounding from unobservables when estimating the ATE using the matching based approaches described in Section 5.5.2. Moreover, matching methods tend to be inefficient as they discard data. An alternative approach to matching is to use inverse-probability weighting (IPW) to balance the covariates distributions between the control and the treatment groups. When the CIA holds, the normalized IPW estimator of Hirano and Imbens (2001) provides an unbiased estimate of the ATE.

The minimum bias estimator (MBE) approach, described in Millimet and Tchernis (2013), relates closely to the normalized IPW estimator. In particular, when the CIA assumption fails to hold, the MBE aims to adjust the IPW estimate of the ATE to minimize the effect of unobserved bias without relying on instrumental variables and exclusion restrictions. The approach is motivated by the observation that any hidden bias has the biggest impact on the tails of the distribution of selection probabilities (i.e., propensity scores $\hat{P}(\mathbf{X}_i)$ close to 0 or 1). It then follows that the bias in the estimate of the ATE is minimized for observations with propensity scores close to a so-called bias-minimizing propensity score, denoted by P^*. Therefore, instead of estimating the IPW treatment effects using all observations, a subset Ω of observations close to P^* is included so as to minimize the inherent bias in the estimate of the ATE when the CIA fails. The MBE estimates P^* and Ω to minimize the bias by using Heckman's bivariate normal (BVN) selection model (see Millimet and Tchernis (2013) for technical details).[5]

To estimate the MBE, the neighborhood around P^* which determines Ω must be set by the user, who specifies the smallest percent θ of both the treatment and control groups to be contained in Ω. For example, if $\theta = 5\%$, then at least 5% of the control and treatment groups must be in Ω. Smaller values of θ result in a greater reduction in bias at the expense of greater variance in the estimate of the ATE. In our analysis, we report results using a range of values of θ, in particular values of 25%, 10% and 5%. We also follow convention and impose the condition that only observations with corresponding values of $\hat{P}(\mathbf{X}_i) \in [0.02, 0.98]$ can be included in Ω.

It is important to note that since the MBE is estimated on only a subsample of the data, the parameter being estimated will generally differ from the population ATE unless the treatment effect does not vary with $\mathbf{X}$. Therefore, we are less interested in the size of the ATE estimated using the MBE, and more interested in comparing the MBE estimates to the IPW estimate. This allows us to explore the potential direction and size of any endogeneity bias. In particular, when the CIA holds, the MBE provides a consistent, but inefficient, alternative to IPW. However, when the CIA fails, the MBE has a smaller bias than IPW (Millimet and Tchernis 2013).

6. Results

6.1. OLS findings

Table 5 shows the OLS estimates, based on equation (1), for various control structures. Standard errors are clustered at the patient level to account for correlations of error terms when consultations are associated with the same patient. Controls are added step-wise into the model to help us understand how the inclusion of different control categories alters the main effect estimate.

[5] To test the sensitivity of the MBE to potential deviations from the normality assumptions underlying Heckman's BVN model, the MBE can be extended using Edgeworth-expansion (EE) versions of the relevant estimator (Millimet and Tchernis 2013). In our case, we find no significant difference between the BVN and EE estimates.

Table 5 **Sensitivity of OLS coefficient estimates to the inclusion of different control structures**

	OLS: Dependent variable = Natural logarithm of the revisit interval					
$RD_i = 1$	0.074*** [0.072,0.076]	0.060*** [0.058,0.062]	0.127*** [0.125,0.129]	0.158*** [0.156,0.160]	0.158*** [0.156,0.159]	0.157*** [0.155,0.159]
Practice fixed effects	No	Yes	Yes	Yes	Yes	Yes
Temporal factors	No	Yes	Yes	Yes	Yes	Yes
Patient demographics	No	No	Yes	Yes	Yes	Yes
Patient's past history	No	No	No	Yes	Yes	Yes
Attributes of regular doctor	No	No	No	No	Yes	Yes
Practice-level demand	No	No	No	No	No	Yes
Adjusted R^2	0.000	0.044	0.103	0.159	0.159	0.159
Number of observations	11,344,065	11,344,065	11,344,065	11,344,065	11,344,065	11,344,065

Notes: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$; 95% confidence intervals in square brackets, with standard errors clustered at the patient level; Variables included within the categories of controls are specified in Table 3.

The step-wise introduction of factors that are correlated with poorer heath, such as consultation frequency, past average revisit interval, comorbidity controls and age and socioeconomic factors, increase the effect of seeing the regular doctor. This is suggestive of heterogeneity, which we will address directly in Section 6.4.

Examining the results, we find evidence that patients' consultations with their regular doctor are robustly associated with a longer revisit interval ($\beta_1 = 0.157$, 95% CI: $[0.155, 0.159]$, p-value < 0.001 in the fully controlled model). Since the dependent variable has a log-scale, the fully controlled OLS model suggests that the revisit interval increases by 17.0% [16.8%,17.2%], on average, when a patient is seen by their regular doctor.[6]

6.2. Acuity subsamples

Estimating the fully controlled OLS model on the subsample of consultations with antibiotic prescriptions provided a smaller coefficient $\beta_1 = 0.117$ (95% CI: $[0.112, 0.122]$), suggesting that the continuity of care effect may be less pronounced for acute conditions. This is expected, as after an acute visit, a regular doctor is more likely to see the patient again for a follow-up and may have less leeway to extend the revisit interval length. However, the coefficient still remains significantly positive and the effect remains large. This observation is corroborated by the second subsample analysis, consisting of consultations with patients who had visited an ED in the seven-day window prior to the focal consultation. Again, the coefficient is lower than in the full sample ($\beta_1 = 0.120$, 95% CI: $[0.104, 0.135]$) but the effect remains statistically highly significant and the average extension of the revisit interval remains practically significant as well.

[6] Note that since we estimate a log-linear model with a binary independent variable, the reported effect size (17.0%) differs from the β_1 coefficient in Table 5 (15.7%) and is calculated as $(Y(1) - Y(0))/Y(0) = e^{\beta_1} - 1$. All effect sizes and confidence intervals reported in the body of the paper are similarly calculated (Wooldridge 2015b).

Table 6 **Coefficients of RD_i (seeing the regular doctor) on $\ln(RI_i)$ (log revisit interval) for different model specifications**

Dependent variable = Natural logarithm of the revisit interval								
Model	Sample	Method	Coefficient (β_1)	Std. Error	t-statistic	P> $\|t\|$	95% CI	
1a	consultations	OLS	15.7%	0.1%	177.33	0.00	15.5%	15.9%
1b	consultations	CF	16.6%	0.5%	30.57	0.00	15.6%	17.6%
1c	consultations	2SLS	16.2%	0.6%	28.01	0.00	15.1%	17.3%
2a	patients	PSM	15.6%	0.4%	34.59	0.00	14.7%	16.5%
2b	patients	PSM OLS	15.4%	0.4%	36.91	0.00	14.6%	16.2%
3a	patients	IPW	15.9%	–	–	–	15.3%	16.6%
3b	patients	MBE ($\theta = 0.25$)	20.8%	–	–	–	19.6%	22.3%
3c	patients	MBE ($\theta = 0.10$)	20.9%	–	–	–	19.2%	22.7%
3d	patients	MBE ($\theta = 0.05$)	22.3%	–	–	–	19.3%	25.0%

Notes: This table reports the estimated effects (β_1) of seeing the regular doctor on the natural logarithm of the revisit interval for the taxonomy of models specified in Section 5. Standard errors are clustered at the patient level for models 1a-1c and the matched pair level for models 2a-2b. Confidence intervals are bootstrapped for models 3a-3d, and thus standard errors, t-statistics and p-values are not available.

In summary, the subsample analyses suggest that confounding by acuity, if it occurs, is relatively small and does not explain the main effect. The analyses are fully documented in Section EC.8 of the e-companion, alongside analyses corresponding to additional acuity subsamples.

6.3. Alternative model specifications

Table 6 summarizes the results from all models introduced in Section 5. The top panel reports the estimated coefficient associated with seeing the regular doctor on the natural logarithm of the revisit interval using OLS (1a), the CF model (1b), and the 2SLS model (1c). The second panel corresponds to estimates using propensity score matching (PSM). For PSM (2a) we report differences in averages of the revisit interval, $\ln(RI)$, between the control and treated groups; PSM OLS (2b) corresponds to an OLS regression on the matched sample that includes controls (see Section 5.5.2). In the bottom panel, we report the results from the IPW estimator (3a) and the MBE using three different values of θ (3b-3d) (see Section 5.5.3). All PSM, IPW and MBE coefficients are estimated using the reduced sample, where a single consultation was randomly chosen per patient, as explained in Section 5.5.1.

We find that the results are consistent across all modeling techniques employed and confirm the beneficial effect of continuity of care posited in Hypothesis 1. The coefficient of the bias correction term in the CF model is statistically insignificant at the 5% significance level ($\hat{\gamma}^{CF} = -0.006$, 95% CI: $[-0.012, 0.001]$), hence we find that there is little evidence of major confounding. Since the dependent variable has a log-scale, the CF model suggests that the revisit interval increases by 18.1% [16.9%,19.2%], on average, when a patient is seen by their regular doctor. Meanwhile, the PSM and IPW estimates are similar in magnitude to the OLS estimate. Finally, for the MBE estimator, the direction of change is towards higher values of the coefficient for decreasing values of θ, as compared to the coefficient of the IPW estimator. Thus, both the IV-based (CF/2SLS) and the non-IV-based (MBE) estimates suggest that, if anything, endogeneity biases the estimated coefficients downwards. This is consistent with an unobserved selection of less healthy patients for continuity of care (see Section 5.2).

As the CF approach is indicated for use in contexts with a binary first stage selection equation (see Section 5.4.1), we select this approach as the main model specification going forward. The CF method estimates that patients who see their regular provider have a 18.1% longer revisit interval (95% CI: $[16.9\%, 19.2\%]$).

6.4. Moderating effects

In this section, we test the moderating effects of comorbidity, age and mental health as posited in Hypothesis 2. To this end, we re-estimate the CF model (equations (2) and (3)) but now also include interaction terms between the main independent variable RD_i. We also include interactions of the three patient-level moderation variables with the instrumental variable to create three new instruments (Wooldridge 2011). As the moderators are naturally correlated (in particular age and comorbidity), we have included all interaction terms in a single model to estimate their moderation effect net of the correlated effects of the other moderators.

Estimates of the average marginal effects based on the moderation results are reported in Table 7 and a graphical representation of the results is given in Figure EC.10 in the e-companion. Table 7 reports the estimated revisit interval (in natural logarithm) for an average individual (within the segment specified by the first column of the row) assuming they either saw a transactional provider ($RD_i = 0$ columns) or their regular doctor ($RD_i = 1$ columns). For example, an average 18-25 year old who saw a transactional doctor is estimated to have a (natural logarithm of the) revisit interval of 3.31, as compared to 3.35 if they had instead seen their regular doctor. The average marginal effect or, difference between these two values, is equal to 6.8% (95% CI: $[5.5\%, 8.1\%]$) for an average 18-25 year old patient. (The coefficients corresponding to the moderating effect estimates are reported in Table EC.21 in the e-companion.)

6.4.1. Comorbidity. The comorbidity panel in Table 7 confirms that patients who see their regular doctor ($RD_i = 1$ columns) have, on average, longer revisit intervals than those who see a transactional provider ($RD_i = 0$ columns), and this effect is independent of the number of comorbidities. The effect size columns confirms Hypothesis 2(a) by showing that the extension of the revisit interval that the regular doctor achieves is higher for patients with comorbidities. The difference between zero and one comorbidities is statistically significant. Additional differential effects beyond two comorbidities are insignificant.

Table 7 Average marginal effects associated with seeing the regular doctor when using age, comorbidity and mental health as moderators, calculated using the control function model

	$RD_i = 0$		$RD_i = 1$		Marginal Effect	
	$\ln(RI)$	95% CI	$\ln(RI)$	95% CI	M.E.	95% CI
Comorbidity						
0 comorbidites	3.38	[3.38,3.39]	3.51	[3,51,3.52]	12.8%	[11.7%,13.9%]
1 comorbidity	3.29	[3.28,3.29]	3.44	[3.44,3.45]	15.9%	[14.8%,17.0%]
2 comorbidites	3.23	[3.21,3.23]	3.39	[3.38,3.39]	16.7%	[15.6%,17.8%]
3 comorbidites	3.18	[3.17,3.18]	3.35	[3.34,3.36]	17.2%	[16.1%,18.4%]
4 comorbidites	3.16	[3.15,3.16]	3.33	[3.32,3.33]	17.1%	[15.9%,18.3%]
≥ 5 comorbidites	3.15	[3.13,3.15]	3.30	[3.29,3.31]	16.0%	[14.9%,17.2%]
Age band						
18-25yrs	3.33	[3.32,3.33]	3.39	[3.38,3.40]	6.8%	[5.5%,8.1%]
26-30yrs	3.31	[3.30,3.32]	3.40	[3.39,3.41]	8.8%	[7.4%,10.1%]
31-35yrs	3.30	[3.30,3.31]	3.40	[3.39,3.41]	9.9%	[8.6%,11.2%]
36-40yrs	3.32	[3.31,3.33]	3.42	[3.41,3.43]	10.1%	[8.8%,11.3%]
41-15yrs	3.31	[3.31,3.32]	3.43	[3.43,3.44]	12.0%	[10.8%,13.2%]
46-50yrs	3.31	[3.30,3.31]	3.44	[3.43,3.45]	13.3%	[12.1%,14.5%]
51-55yrs	3.30	[3.29,3.30]	3.44	[3.44,3.45]	14.6%	[13.4%,15.8%]
56-60yrs	3.28	[3.28,3.29]	3.44	[3.44,3.45]	16.0%	[14.8%,17.2%]
61-65yrs	3.27	[3.26,3.27]	3.44	[3.43,3.44]	17.2%	[16.0%,18.4%]
66-70yrs	3.24	[3.23,3.25]	3.42	[3.41,3.43]	17.9%	[16.7%,19.1%]
71-75yrs	3.20	[3.20,3.21]	3.40	[3.39,3.40]	19.4%	[18.2%,20.6%]
76-80yrs	3.15	[3.15,3.16]	3.37	[3.36,3.37]	21.1%	[19.9%,22.3%]
81-85yrs	3.10	[3.09,3.11]	3.32	[3.31,3.33]	22.1%	[20.9%,23.3%]
86+yrs	2.99	[2.98,3.00]	3.21	[3.21,3.22]	22.7%	[21.4%,23.9%]
Mental Health						
No	3.24	[3.24,3.25]	3.39	[3.39,3.40]	15.2%	[14.1%,16.3%]
Yes	3.23	[3.23,3.24]	3.41	[3.40,3.41]	17.5%	[16.4%,18.6%]

Notes: '$RD_i = 0$' (resp., '$RD_i = 1$') columns specify the estimated natural logarithm of the revisit interval (ln(RI)) together with 95% confidence intervals (CIs) for a patient who saw a transactional (resp., regular) provider; 'Marginal Effect' columns give the average marginal effect (M.E.), with 95% CIs, associated with a patient seeing a regular doctor versus a transactional provider, implemented using Stata's margins command.

6.4.2. Age. The age band panel in Table 7 shows the effect of patient age. As in the case of comorbidity interactions, the table confirms that patients who see their regular doctor ($RD_i = 1$ columns) have longer revisit intervals than those who see a transactional provider ($RD_i = 0$ columns). The marginal effect columns shows the difference and confirms Hypothesis 2(b): seeing a regular doctor is particularly productive for older patients. The extension of the revisit interval length increases from 6.8% (95% CI: $[5.5\%, 8.1\%]$) for 18-25 year-old patients to 22.7% (95% CI: $[21.4\%, 23.9\%]$) for patients over 86.

6.4.3. Relationship between age and comorbidity Figure 2 combines the effect of providing care continuity to patients in different age groups who have either 0, 1 or 2 or more comorbidities. This is estimated by including the interaction between age, comorbidities and RD_i. The largest productivity-enhancing effect comes from providing care continuity to patients 86+ years of age with one comorbidity (27.6%; 95% CI: $[24.9\%, 30.3\%]$). The productivity gain from providing

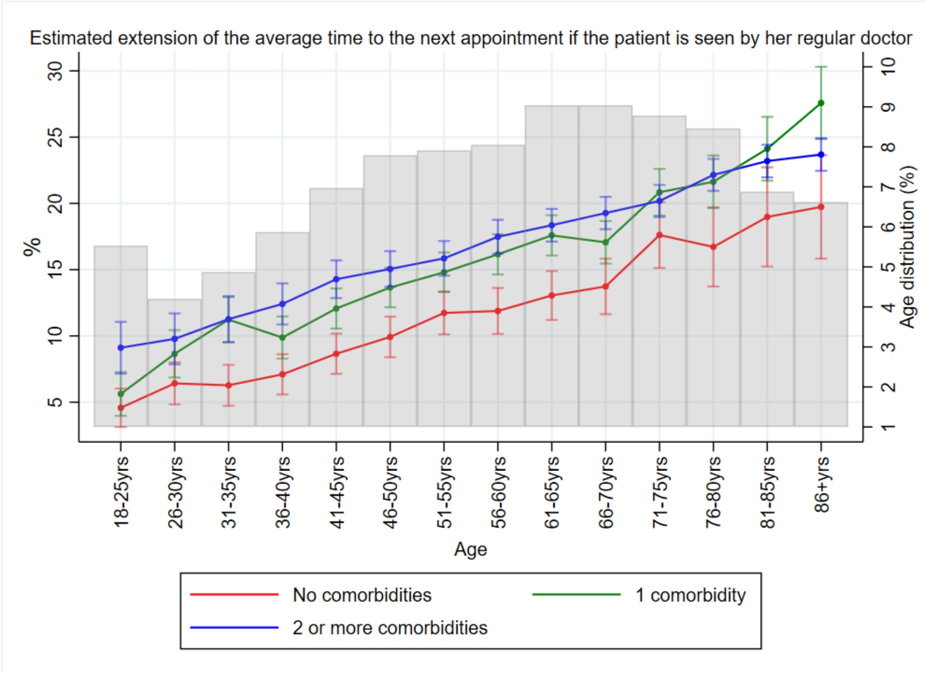


Figure 2 **Differential effect of providing care continuity on the revisit interval for patients across different age groups with 0, 1 or 2 or more comorbidities**

care continuity to younger patients with no comorbidities compared to younger patients with two comorbidities is much higher than for older patients; for older patients, in terms of productivity gains from providing care continuity, it does not make a significant difference whether the patient has 0, 1 or 2 or more comorbidities.

6.4.4. Mental health. Mental health patients are more frequent users of primary care, with a shorter revisit interval, on average, than patients without mental health conditions. Our data supports Hypothesis 2(c) that continuity is more effective for such patients. Specifically, we find that providing continuity extends the revisit intervals of these patients by 17.5% (95% CI: [16.4%, 18.6%]), compared to a 15.2% (95% CI: [14.1%, 16.3%]) improvement for patients without such conditions. This differential effect, albeit significant, is relatively small compared to the effect of comorbidity and age.

6.5. Duration of consultations

As mentioned in Section 3, in addition to the length of the revisit interval, which is the focus of our analysis, consultation length is a second important productivity measure. On the one hand, consultations between a patient and her regular doctor may take longer as the doctor may take more time to avoid a potential revisit that she will have to serve. On the other hand, the regular doctor knows the patient better and may therefore be able to save the time that a transactional provider would need to spend to elicit the necessary information to obtain the same result.

We explore the duration effect empirically, using the same methods we used for revisit intervals, and report the results in Section EC.12 of the e-companion. Regular doctors spend, on average, less consultation time with their patients than transactional providers. However, the effect size is not large enough to allow a physician to accommodate an additional patient in a typical four-hour clinical session. Nevertheless, these estimations provide strong evidence that the patients' regular doctors do not achieve the extension of the revisit interval at the cost of increasing their consultation time with these patients.

6.6. Further analyses.

To further confirm the robustness of the results outlined in the previous sections, we estimate a number of additional model specifications that: (i) use other continuity measures, drawn from the academic literature; (ii) measure the independent and the instrumental variable over different time frames; (iii) account for doctor-level and patient-level heterogeneity through fixed and random effect models, respectively; (iv) account for potential inaccuracies in the identification of the regular doctor when a patient's regular doctor leaves the practice; (v) apply different inclusion criteria; (vi) use a more granular mapping of mental health conditions; and (vii) add additional time/seasonality controls. The findings, documented in the e-companion, are consistent with those reported here.

The medical literature provides ample evidence of the health benefits of continuity of care in primary care (see Section 2.2). In order to validate these findings and ensure that seeing a regular doctor had health benefits for the patients in our data, we estimated the effect of a consultation with the patient's regular doctor on their propensity to require an ED visit within 1, 3 or 7 days of the focal consultation. Increased ED visit rates would be an indication that regular doctors are, on average, not as thorough in their consultations as transactional providers. We have also estimated whether regular doctors or more reluctant to prescribe medicines, which could also be causing poorer health outcomes. Using the control function model (2, 3), we find the opposite. ED presentation rates are lower and prescription rates are higher after consultations with regular doctors (See e-companion EC.15 for more details).

It is conceivable that a regular doctor achieves the extension of the revisit interval after a face-to-face consultation through more frequent use of follow-up phone consultations, which would add additional work for the regular doctor that we had not accounted for. In our data, 5.51% of face-to-face consultations are followed by a phone consultation before the patient's next face-to-face consultation. Using the main control function model but with a follow-up phone consultation as a binary dependent variable, we find that a phone consultation following a face-to-face visit is 5.50% (95% CI [5.47%,5.52%]) when a patient sees a transactional provider, compared to 5.52% (95% CI [5.51%,5.55%]) when a patient sees her regular doctor. This effect difference is too small to have much explanatory power for the main effect. In fact, adding the binary phone consultation variable as a control to the main model, to assess its mediating effect, shows that the results remain robust (see e-companion EC.16.7 for more details).

7. Counterfactual Analysis: Targeting Continuity of Care

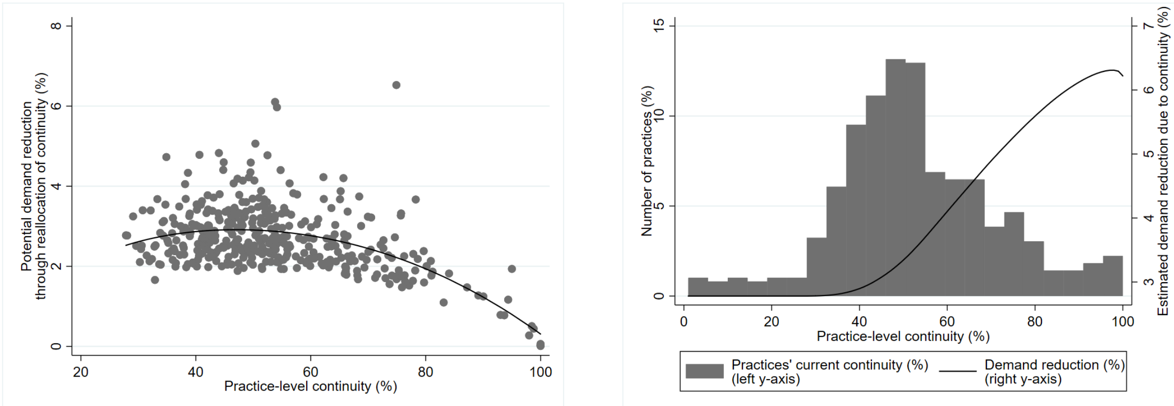
Our results suggest that practices could improve productivity by increasing continuity of care and that they could unlock further productivity gains by reallocating continuity to patients who benefit from it the most. In this section, we conduct two analyses to explore retrospectively what the effect would have been on the consultation demand in our data if practices had followed this recommendation in the past.

Using the insights from Hypothesis 2, we propose a scoring system that can be used by practice managers to prioritize care continuity, targeting those patients for whom it will have the greatest productivity-enhancing effect. The scoring system ranks consultations by the estimated number of days gained if the consultation is offered by the regular doctor rather than a transactional provider. The estimate is obtained as the difference between the predicted return intervals with and without continuity of care. The prediction is based on equation (1), which we augment by including interactions between the regular doctor variable RD_i and every other covariate in the model. Section EC.14 of the e-companion provides more details on this estimation approach.

Our first analysis does not assume that practices changed their proportion of continuity consultations. Instead, we only explore the demand reduction they would have achieved had they *optimized* these consultations by shifting them to the most productivity-enhancing patients, using the above scoring system. Figure 3a shows that such targeting of continuity of care has the potential of unlocking productivity gains, even without changing the overall proportion of continuity consultations. If all practices had retained continuity at the same level but better targeted continuity at the most productivity enhancing patients, then the total consultation demand in the sample would have reduced by 2.7%. In fact, as Figure 3a shows, some practices would have reduced their demand significantly more.

In our second analysis, we consider what would have happened in our sample if practices not only better targeted care continuity but also increased the level of continuity of care provided.[7] To explore this, we first select a target proportion $0\% \leq x \leq 100\%$ of continuity consultations. We then identify all practices that offered less than $x\%$ continuity, and consider the impact on productivity

[7] In this counterfactual analysis, we extrapolate from the individual physician level (the locus of our productivity study) to the practice level. The difference is subtle but important. In particular, we assume that continuity levels can be adjusted within a practice without any negative spillovers onto other drivers of practice productivity. In reality, there may be trade-offs at the practice-level, and thus the findings reported here should be seen as an upper bound on the potential gains that can be achieved from adjusting continuity levels within a practice.



(a) (b)

Figure 3 **(a) Practice-level gains from reallocating consultations with the regular doctor to the most productivity enhancing patients, while keeping the % of the practice's consultations with regular doctors unchanged. (b) System-level reduction in demand if all practices offer a minimum continuity level (specified on the x-axis) to their most productivity-enhancing patients. The underlying histogram shows the current distribution of the continuity levels across practices. If all practices offered continuity at the across-practice 90th percentile (74.1%), the total system-level demand could have reduced by up to 5.2%**

if these practices had offered continuity to x% of their patients instead. (For those practices offering more than x% continuity, we leave their proportion of continuity consultations unchanged.) We then re-allocate, as above, the available proportion of continuity consultations in each practice to the most productivity-enhancing patients. This allows us to estimate the counterfactual demand reduction as continuity of care is increased (i.e., by increasing x) and optimally allocated.

The results are summarized in Figure 3b. If continuity of care levels were increased for those practices with levels below the across-practice median ($x = 51.3\%$ continuity), 75th percentile ($x = 62.3\%$ continuity) and 90th percentile ($x = 74.1\%$ continuity), then total system demand could have reduced by up to 3.4%, 4.3% and 5.2%, respectively. These estimates show that there are significant productivity gains to be realized by increasing the proportion of patients seeing their regular doctors and by better targeting continuity of care at patients who benefit from it most.

8. Managerial and Policy Implications

It is well known that continuity of care in primary care is beneficial for patient health and reduces downstream system utilization. But is it also operationally beneficial for the primary care practice itself? Or do patient health and system benefits accrue at the cost of more up-stream primary care, as primary care physicians spend more time with their continuity of care patients to keep them healthy and out of hospital? If the latter is the case, primary care practices will need to be incentivized to provide continuity of care. Our findings suggest that continuity of care can in fact save primary care resources by making primary care physicians more productive.

We study the physician productivity effect of care continuity by focusing on short-term consultation level duration effects, not on long-term population health effects, of which there is ample evidence in the literature. In other words, the idealized randomized experiment for this study is a coin-flip at the individual consultation level, with a patient assigned to their regular doctor vs another doctor in the practice for that consultation, not a patient level analysis, with a patient provided with a high or low dose of continuity of care over a prolonged period of time.

The study data on over 11M consultations in 381 English practices over a period of 11 years provides evidence that consultations between a patient and a primary care physicians are more productive if the physician is familiar with the patient. Specifically, we find that, after controlling for confounding and patient selection, when patients have a consultation with the doctor they have seen most frequently over the past two years, they have a significantly extended time to their next consultation – by an estimated 18.1% (95% CI: $[16.9\%, 19.2\%]$) – while their consultation duration is, on average, marginally but statistically significantly shorter. We also find that targeting the right patients for continuity of care is important, as the productivity benefit is more pronounced for patients with comorbidities, for older patients, and for patients with mental health conditions.

These findings are of direct relevance to appointments scheduling processes in primary care practices and we have demonstrated how our estimation models can be applied as scoring tools to identify patients for whom continuity of care provides high productivity benefits. In addition, our findings have important strategic implications for practice managers, regulators and payers.

First, as indicated in the introduction, practice managers face a chronic shortage of primary care physicians. The existing workforce has to cope with growing demand and is increasingly stressed, leading to early retirements and part-time working, which further exacerbates the problem. There are two fundamental operational mindsets to address this challenge. The first mindset is to industrialize primary care, to scale it up from a fragmented cottage industry of small shops and organize it so that the number of consultations per physician day is maximized. Fast and convenient access to a primary care consultation, no matter with whom, becomes the operational goal and a physician hour becomes the main operational currency. The second mindset is to recognize the value of continuity of care and double down on relationships between doctors and physicians, to organize the practice so that access to "your primary care physician" is the main goal and the maintenance of the patient-physician relationship becomes the operational currency.

Our findings suggest that practice managers who emphasize daily throughput and fast access should pause and reflect whether this strategy harms continuity and is counter-productive, as it harms the physicians' productivity and generates avoidable demand for future consultations. In fact, doubling down on continuity can be an effective strategy to improve productivity, particularly if a practice operates in a capitation-based funding environment and serves a relatively older or more complex patient population. Practice managers can substantially improve productivity by targeting specific patients for care continuity, as demonstrated in our counterfactual analysis.

Second, our findings are of importance for regulators and third-party payers in relation to the demand for faster and more convenient primary care access, which has led to the emergence of at-scale online providers and has accelerated during the COVID-19 pandemic. Online primary care providers offer a transactional platform for online appointments, matching on-the-spot demand and supply. Much of the advantage comes from scale and the pooling of clinical time. The service is a single consultation between a doctor and a patient, not a long-term relationship. Our findings have several implications for this business model, for regulators and payers who wish to create a sustainable primary care environment, and for primary care practices that need to respond strategically to the threat of the new entrants to the market.

Our findings are consistent with the view that the online business model of primary care is particularly profitable in a fee-for-service environment, where patients or third parties pay for any additional consultations generated by transactional services. By contrast, the model is less compelling in a capitation-based funding environment, where providers are paid a fixed fee per patient per month or year and the risk of excess demand is born by the practice.

Our analysis of moderators – chronic disease, age, and mental health – shows that in a capitation context, online primary care providers will have a significant incentive to design their services to make them less attractive for more demanding or more vulnerable patients. This will minimize the negative productivity effect of the ransactional primary care at the core of their business model. Over time, this dynamic may manifest itself in a segmentation of primary care services, with low-risk patients being served quickly and conveniently by transaction-focused online providers, while high-risk patients are served by local practices that offer the relational continuity of care required for an effective and efficient service for these patients. Most patients will reach times in their lives when they wish to move from one segment to the other. While such a service segmentation may well be beneficial in steady state, regulators must anticipate the destabilizing effect that this transformation will have on traditional practices during a transition period. Specifically, regulators and payers will need to respond with adequate risk-adjusted funding models. The scoring method used in our counterfactual analysis offers a suggestion for how such models could be developed.

The differentiation between relational and transactional primary care services also poses a strategic challenge for local practice managers: Should they provide services for both segments in-house, managing the tensions between the transactional and relational service model, or should they out-source transactional primary care and "double down" on relational services for patients who benefit most from continuity of care and from the close integration of their primary care practice in a local provider network that at-scale online providers will find difficult to replicate? More research is required to advise practices on this decision.

Acknowledgments

This study was approved by the CPRD Independent Scientific Advisory Committee (ISAC) committee. Protocol no.: 19_004R2. We are grateful to Professor Albert Mulley and Dr James Morrow, whose passion for relational medicine greatly influenced our perspective on continuity of care, and to Granta Medical Practices, Cambridgeshire (UK) for serving as a pilot site for the research.

References

Abadie A, Spiess J (2021) Robust post-matching inference. *Journal of the American Statistical Association* 1–13.

Ahuja V, Alvarez CA, Staats BR (2020a) How continuity in service impacts variability: Evidence from a primary care setting, SMU Cox School of Business Research Paper No. 19-13.

Ahuja V, Alvarez CA, Staats BR (2020b) Maintaining continuity in service: An empirical examination of primary care physicians. *Manufacturing & Service Operations Management* .

Amjad H, Carmichael D, Austin AM, Chang CH, Bynum JP (2016) Continuity of care and health care utilization in older adults with dementia in fee-for-service medicare. *JAMA internal medicine* 176(9):1371–1378.

Austin PC, Small DS (2014) The use of bootstrapping when using propensity-score matching without replacement: a simulation study. *Statistics in Medicine* 33(24):4306–4319.

Balint M (1955) The doctor, his patient, and the illness. *The Lancet* 265(6866):683–688.

Barker I, Steventon A, Deeny SR (2017) Association between continuity of care in general practice and hospital admissions for ambulatory care sensitive conditions: cross sectional study of routinely collected, person level data. *Bmj* 356.

Bavafa H, Hitt LM, Terwiesch C (2018) The impact of e-visits on visit frequencies and patient health: Evidence from primary care. *Management Science* 64(12):5461–5480.

Bayliss EA, Ellis JL, Shoup JA, Zeng C, McQuillan DB, Steiner JF (2015) Effect of continuity of care on hospital utilization for seniors with multiple medical conditions in an integrated health care system. *The Annals of Family Medicine* 13(2):123–129.

Beech J, Bottery S, Charlesworth A, Evans H, Gershlick B, Hemmings N, Imison C, Kahtan P, McKenna H, Murray R, Palmer B (2020) Closing the gap: Key areas for action on the health and care workforce — The Nuffield Trust. Technical report, Nuffield Trust.

Biringer E, Hartveit M, Sundfør B, Ruud T, Borg M (2017) Continuity of care as experienced by mental health service users-a qualitative study. *BMC Health Services Research* 17(1):763.

Bobroske K, Freeman M, Huan L, Cattrell A, Scholtes S (2021) Curbing the Opioid Epidemic at its Root: The Effect of Provider Discordance after Opioid Initiation. *Forthcoming in Management Science* .

Brookhart MA, Patrick AR, Schneeweiss S, Avorn J, Dormuth C, Shrank W, van Wijk BL, Cadarette SM, Canning CF, Solomon DH (2007) Physician follow-up and provider continuity are associated with long-term medication adherence: a study of the dynamics of statin use. *Archives of internal medicine* 167(8):847–852.

Chen HM, Tu YH, Chen CM (2017) Effect of continuity of care on quality of life in older adults with chronic diseases: a meta-analysis. *Clinical Nursing Research* 26(3):266–284.

Cho KH, Kim YS, Nam CM, Kim TH, Kim SJ, Han KT, Park EC (2015) The association between continuity of care and all-cause mortality in patients with newly diagnosed obstructive pulmonary disease: a population-based retrospective cohort study, 2005-2012. *PloS one* 10(11).

Dall T, Reynolds R, Chakrabarti R, Jones K, Iacobucci W (2020) The Complexities of Physician Supply and Demand: Projections from 2018 to 2033. Technical Report June, Association of American Medical Colleges.

Dossa AR, Moisan J, Guénette L, Lauzier S, Grégoire JP (2017) Association between interpersonal continuity of care and medication adherence in type 2 diabetes: an observational cohort study. *CMAJ open* 5(2):E359.

Drury A, Payne S, Brady AM (2020) Identifying associations between quality of life outcomes and healthcare-related variables among colorectal cancer survivors: A cross-sectional survey study. *International Journal of Nursing Studies* 101:103434.

Freeman G, Hughes J, et al. (2010) Continuity of care and the patient experience. Technical report, The King's Fund.

Freeman M, Robinson S, Scholtes S (2020) Gatekeeping, fast and slow: An empirical study of referral errors in the emergency department. *Management Science* .

Goodwin N, Curry N, Naylor C, Ross S, Duldig W, et al. (2010) Managing people with long-term conditions. *London: The Kings Fund* .

Haggerty JL, Reid RJ, Freeman GK, Starfield BH, Adair CE, McKendry R (2003) Continuity of care: a multidisciplinary review. *Bmj* 327(7425):1219–1221.

Hallvik SE, Geissert P, Wakeland W, Hildebran C, Carson J, O'kane N, Deyo RA (2018) Opioid-prescribing continuity and risky opioid prescriptions. *The Annals of Family Medicine* 16(5):440–442.

Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, Van Staa T, Smeeth L (2015) Data Resource Profile Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International Journal of Epidemiology* 827–836.

Hill AP, Freeman GK (2011) Promoting continuity of care in general practice. *London: Royal College of General Practitioners* .

Hjortdahl P, Borchgrevink CF (1991) Continuity of care: influence of general practitioners' knowledge about their patients on use of resources in consultations. *British Medical Journal* 303(6811):1181–1184.

Huntley A, Lasserson D, Wye L, Morris R, Checkland K, England H, Salisbury C, Purdy S (2014) Which features of primary care affect unscheduled secondary care use? a systematic review. *BMJ open* 4(5):e004746.

Institute for Government (2019) General practice — The Institute for Government. Available at `https://www.instituteforgovernment.org.uk/publication/performance-tracker-2019/general-practice` (2020/21/07).

Jann B (2006) GSAMPLE: Stata module to draw a random sample. Statistical Software Components, Boston College Department of Economics.

Jeffers H, Baker M (2016) Continuity of care: still important in modern-day general practice. *British Journal of General Practice* 66(649):396–397.

Kajaria-Montag H, Freeman M (2020) Explaining the erosion of relational care continuity: An empirical analysis of primary care in England. *SSRN Electronic Journal* INSEAD Working Paper No. 2020/47/TOM.

Katz DA, McCoy KD, Vaughan-Sarrazin MS (2015) Does greater continuity of veterans administration primary care reduce emergency department visits and hospitalization in older veterans? *Journal of the American Geriatrics Society* 63(12):2510–2518.

Knaak S, Mantler E, Szeto A (2017) Mental illness-related stigma in healthcare: Barriers to access and care and evidence-based solutions. *Healthcare management forum*, 111–116, number 2 (SAGE Publications Sage CA: Los Angeles, CA).

Koopman RJ, Mainous III AG, Baker R, Gill JM, Gilbert GE (2003) Continuity of care and recognition of diabetes, hypertension, and hypercholesterolemia. *Archives of Internal Medicine* 163(11):1357–1361.

Leniz J, Gulliford MC (2019) Continuity of care and delivery of diabetes and hypertensive care among regular users of primary care services in chile: a cross-sectional study. *BMJ Open* 9(10).

Li MM, Nassiri S, Liu X, Ellimoottil C (2021) How does telemedicine shape physician's practice in mental health? *SSRN Electronic Journal* .

Lin IP, Wu SC, Huang ST (2015) Continuity of care and avoidable hospitalizations for chronic obstructive pulmonary disease (copd). *The Journal of the American Board of Family Medicine* 28(2):222–230.

Maarsingh OR, Henry Y, van de Ven PM, Deeg DJ (2016) Continuity of care in primary care and association with survival in older people: a 17-year prospective cohort study. *British Journal of General Practice* 66(649):e531–e539.

MacCallum RC, Zhang S, Preacher KJ, Rucker DD (2002) On the practice of dichotomization of quantitative variables. *Psychological methods* 7(1):19.

Millimet DL, Tchernis R (2013) Estimation of treatment effects without an exclusion restriction: With an application to the analysis of the school breakfast program. *Journal of Applied Econometrics* 28(6):982–1017.

NHS Digital (2021) NHS Payments to General Practice, England 2020/21. Technical Report September, NHS Digital.

Nyweide DJ, Anthony DL, Bynum JP, Strawderman RL, Weeks WB, Casalino LP, Fisher ES (2013) Continuity of care and the risk of preventable hospitalization in older adults. *JAMA Internal Medicine* 173(20):1879–1885.

O'Connor PJ, Desai J, Rush WA, Cherney LM, Solberg LI, Bishop DB (1998) Is having a regular provider of diabetes care related to intensity of care and glycemic control? *Journal of Family Practice* 47:290–297.

Özer Ö, Zheng Y, Ren Y (2014) Trust, trustworthiness, and information sharing in supply chains bridging china and the united states. *Management Science* 60(10):2435–2460.

Palmer W (2019) Is the number of GPs falling across the UK? Technical report, Nuffield Trust, UK.

Palmer W, Hemmings N, Rosen R, Keeble E, Williams S, Imison C (2018) Improving access and continuity in general practice. *Research Summary* .

Pollack CE, Hussey PS, Rudin RS, Fox DS, Lai J, Schneider EC (2016) Measuring care continuity: a comparison of claims-based methods. *Medical care* 54(5):e30.

Pourat N, Davis AC, Chen X, Vrungos S, Kominski GF (2015) In california, primary care continuity was associated with reduced emergency department use and fewer hospitalizations. *Health Affairs* 34(7):1113–1120.

Queenan C, Cameron K, Snell A, Smalley J, Joglekar N (2019) Patient heal thyself: reducing hospital readmissions with technology-enabled continuity of care and patient activation. *Production and Operations Management* 28(11):2841–2853.

Ramanayake R, Basnayake B (2018) Evaluation of red flags minimizes missing serious diseases in primary care. *Journal of Family Medicine and Primary Care* 7(2):315.

Ride J, Kasteridis P, Gutacker N, Doran T, Rice N, Gravelle H, Kendrick T, Mason A, Goddard M, Siddiqi N, et al. (2019) Impact of family practice continuity of care on unplanned hospital use for people with serious mental illness. *Health Services Research* 54(6):1316–1325.

Rosen R, Massey Y, Abbas S, Hufflett T (2020) Relational continuity for general practice patients with new and changing symptoms. Technical report, Valentine Health Partnership, The Health Foundation.

Royal College of General Practitioners (2019) Fit for the future: a vision for general practice.

Semahegn A, Torpey K, Manu A, Assefa N, Tesfaye G, Ankomah A (2018) Psychotropic medication non-adherence and associated factors among adult patients with major psychiatric disorders: a protocol for a systematic review. *Systematic Reviews* 7(1):1–5.

Senot C (2019) Continuity of care and risk of readmission: An investigation into the healthcare journey of heart failure patients. *Production and Operations Management* 28(8):2008–2030.

Stuart EA (2010) Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics* 25(1):1.

Tammes P, Purdy S, Salisbury C, MacKichan F, Lasserson D, Morris RW (2017) Continuity of primary care and emergency hospital admissions among older patients in england. *The Annals of Family Medicine* 15(6):515–522.

Tarrant C, Dixon-Woods M, Colman AM, Stokes T (2010) Continuity and trust in primary care: a qualitative study informed by game theory. *The Annals of Family Medicine* 8(5):440–446.

von Bültzingslöwen I, Eliasson G, Sarvimäki A, Mattsson B, Hjortdahl P (2006) Patients' views on interpersonal continuity in primary care: a sense of security based on four core foundations. *Family practice* 23(2):210–219.

Wilkin D, Metcalfe D (1984) List size and patient contact in general medical practice. *British Medical Journal (Clinical Research Ed.)* 289(6457):1501–1505.

Wooldridge JM (2002) *Econometric analysis of cross section and panel data* (MIT Press).

Wooldridge JM (2011) *Econometric analysis of cross section and panel data: Second Edition*, chapter 6 (MIT Press).

Wooldridge JM (2015a) Control function methods in applied econometrics. *Journal of Human Resources* 50(2):420–445.

Wooldridge JM (2015b) *Introductory econometrics: A modern approach* (Cengage learning).

Worrall G, Knight J (2011) Continuity of care is good for elderly people with diabetes: retrospective cohort study of mortality and hospitalization. *Canadian Family Physician* 57(1):e16–e20.

Yang D, Dalton JE (2012) A unified approach to measuring the effect size between two groups using sas®. *SAS global forum*, volume 335, 1–6.

Ye T, Sun X, Tang W, Miao Y, Zhang Y, Zhang L (2016) Effect of continuity of care on health-related quality of life in adult patients with hypertension: a cohort study in china. *BMC Health Services Research* 16(1):674.

e-companion to: "Continuity of care increases physician productivity in primary care"

Outline of the e-companion

EC.1. Sample inclusion criteria Page ec2

EC.2. Patient pathways Page ec3

EC.3. What predicts continuity of care? Validation of the independent variable Page ec4

EC.4. Alternative measures of continuity of care Page ec7

EC.5. Description of control variables Page ec8

EC.6. Patient visits Page ec11

EC.7. Instrumental variable construction Page ec12

EC.8. Confounding: Subsample analysis Page ec16

EC.9. Doctor and patient heterogeneity Page ec19

EC.10. Percentage of past visits with regular doctor Page ec20

EC.11. Summary statistics for the matched sample Page ec20

EC.12. Results for the duration of consultations Page ec23

EC.13. Moderation results Page ec23

EC.14. Calculation of the scoring index for targeting continuity of care Page ec24

EC.15. Continuity of care and other outcomes Page ec27

EC.16. Additional robustness checks Page ec28

EC.1. Sample inclusion criteria

Below we provide justification for the various inclusion criteria that were used to transform the initial data set into the final sample for analysis. Each of these points is also summarized in Table 1 in Section 4.2 of the main paper.

- At most practices, patients are able to see a range of staff, such as physicians, nurses, and nurse practitioners, through various channels including face-to-face and phone consultations. In accordance with the literature on continuity of care, we include only consultations with physicians; there is currently only weak evidence of advantages of continuity between patients and non-physician care providers (Tammes et al. 2017, Barker et al. 2017). Moreover, in the English primary care setting, physicians are owners of the practices and hold patient lists of patients who they are responsible for. The arguments that we have put forward in the hypotheses development, specifically the ownership of the patient based on incentives of the regular doctor, only hold for physicians and not nurses. Hence, in our analysis we exclude visits to non-physicians. Advanced nurse practitioners do not hold patient lists and hence are not accountable for specific patients. Moreover, they are almost always used for acute conditions only. We also include only face-to-face consultations, which were the most common way patients interacted with their physicians.
- To ensure high data quality, CPRD determines when the data from a practice is considered to be of sufficient quality for research purposes. We discard any observations made before the date at which a practice's data is considered to be of research quality.
- Patients change registrations between practices over time, for example when they change their place of residency. To ensure that registration gaps with a practice do not affect the analysis, we only include the latest continuous period of registration of a patient with a practice.
- To increase the homogeneity of the sample, we have excluded babies and children as they have different medical needs to adults. The sample thus includes only consultations associated with adult patients over the age of 18.
- In line with previous literature, we exclude consultations with patients who had fewer than three consultations in the two-year window prior to the consultation since an accurate identification of a regular doctor is not possible with very few consultations (Ahuja et al. 2020b). We also exclude consultations that were preceded by more than 104 consultations over a two-year window, an average of one consultation per week. Such patients are likely to have very special needs or are on a specific care management plan that requires frequent planned visits. The relationship between care continuity and the revisit interval is less meaningful for this group of patients.
- Since we use a two-year period to calculate the patient's regular doctor, we exclude consultations that took place in the first two years following the patient's registration date at the practice. During this period, we do not have an accurate estimate of the patient's regular doctor.
- We exclude a patient's last recorded consultation as there is no subsequent consultation with which to calculate the revisit interval. We refer to the remaining consultations as consultations with a valid revisit interval.

- Our data consists of the complete medical record of all patients described above who had a contact with a primary care practice between January 2007 and December 2017[8]. However, we also have partial records on some patients outside of this date range, which we remove from the analysis.
- We wish to estimate the effect of care continuity at times when the practice is able to schedule an appointment with the patient's regular doctor. We therefore exclude consultations that take place during weeks when the focal patient's regular doctor is on leave. Our results should therefore be interpreted as being conditional on the patient's regular doctor being available in the week during which the consultation occurred.
- Finally, due to computational issues related to the large sample size, all of the models are estimated on a random sample of 25% of the remaining 45,376,070 consultations. The random samples are drawn from each of the 381 primary care practices remaining in the data and then merged, thus ensuring that each practice continues to be represented.

EC.2. Patient pathways

To provide clarity on the unit of analysis and the structure of our data, we provide an illustration of the pathway for two different patients. Each row in the tables EC.1 and EC.2 represents a patient-consultation and forms the unit of analysis in our data set. The first column is the date the consultation took place, the second column represents how many consultations the patient had in the 2 years preceding the focal consultation, column 3 is the regular doctor calculated as described in section 4.3.2, the actual doctor is the doctor the patient sees during the consultation and the revisit interval is the time elapsed between the focal consultation and the consultation following the focal consultation as described in Section 4.3.1. We use letters A, B, C, D, E, F, G, H, J, K, X, Y, Z to denote different doctors.

[8] This restriction is applied by CPRD, our data provider

Consultation date	# of consultations in past 2 years	Regular doctor	Actual doctor	Revisit Interval
31-Oct-08	0	NA	X	20
20-Nov-08	1	NA	X	133
02-Apr-09	2	NA	X	169
18-Sep-09	3	X	X	402
25-Oct-10	4	X	A	17
11-Nov-10	4	X	Z	151
11-Apr-11	3	X	Y	36
17-May-11	4	X	Y	378
29-May-12	4	Y	B	143
19-Oct-12	5	Y	Z	42
30-Nov-12	4	Y	C	152
01-May-13	4	Z	D	43
13-Jun-13	4	Z	Z	96
17-Sep-13	5	Z	Y	62
18-Nov-13	6	Z	Y	81
07-Feb-14	7	Z	Y	244
09-Oct-14	7	Y	E	204
01-May-15	6	Y	F	203
20-Nov-15	3	Y	Y	NA

Table EC.1 Patient 1 has 19 consultations between 31 October 2007 and 20 November 2015, and sees doctors A,B,C,D,E,F,X,Y,Z during that period

Consultation date	# of consultations in past 2 years	Regular doctor	Actual doctor	Revisit Interval
09-Mar-07	0	NA	H	7
16-Mar-07	1	NA	H	1
17-Mar-07	2	NA	J	6
23-Mar-07	3	H	H	7
30-Mar-07	4	H	H	6
05-Apr-07	5	H	H	8
13-Apr-07	6	H	K	14
27-Apr-07	7	H	H	14
11-May-07	8	H	H	7
18-May-07	9	H	H	7
25-May-07	10	H	H	14
08-Jun-07	11	H	H	7
15-Jun-07	12	H	H	118
11-Oct-07	13	H	H	NA

Table EC.2 Patient 2 has 14 consultations between 9 March 2007 and 11 October 2007, and sees doctors H,J and K during that period

EC.3. What predicts continuity of care? Validation of the independent variable

To validate that we are capturing continuity correctly in our independent variable, we wish to understand the relationship between patient-level factors and the probability that the patient will see their regular doctor. In table EC.4 we include summary statistics for several factors broken down by consultations that take place

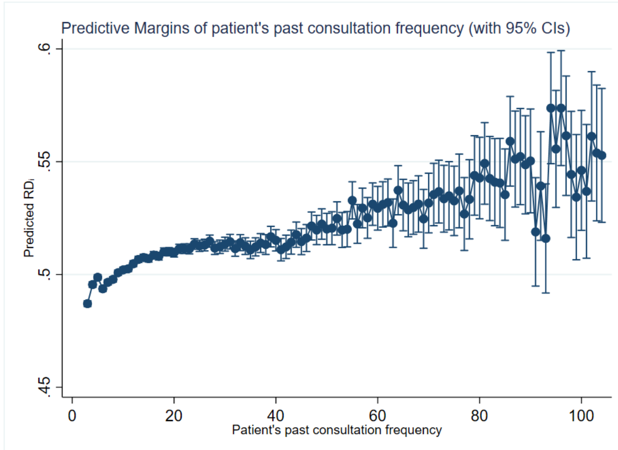


Figure EC.1 Non-linear effect of the patient's past consultation frequency on whether the patient saw her regular doctor

with a patient's regular doctor and consultations that do not. According to the literature (as outlined in Section 2.2 of the main paper), older patients and patients with comorbidities benefit most from continuity of care, therefore we would expect that the patient's past frequency of consultations (heavy users of primary care) and the patient's complexity (age, number of comorbidities, etc.) will be the most important predictors. To investigate this, we estimate a linear probability model of the form $RD_i = \beta_0 + \boldsymbol{\beta}_1\mathbf{X}$, where X is a vector of control variables described in Section 4.3.3 and RD_i is a binary variable, capturing whether the patient sees her regular doctor. The results of the model are given in Table EC.3.

We find that the older the patient is, the higher the probability that they will see their regular doctor. Similarly, the more complex the patient is (higher comorbidities and higher number of prescriptions), the higher the probability that they will see their regular doctor. We also find an increasing relationship with the patient's past frequency of consultation – the probability of seeing the regular doctor increases with the increase in past frequency. However, further analysis shows us that this increase is active only up to a certain point (60 consultations), after which the probability of seeing the regular doctor stabilizes.

As the observed effects in Table EC.3 are in line with expectations, we have confidence that the binary variable RD_i is appropriately identifying the doctor who is more likely to be a patients regular doctor.

	$RD_i = 0$	$RD_i = 1$
Average Revisit interval	71.11	68.35
Average log(Revisit interval)	3.28	3.35
Average consultation frequency	15.19	16.64
Average age	55.88	60.68
Average number of comorbidities	2.06	2.37
Average % mental health prevalence	0.26	0.30
Average % female patients	64.87%	60.43%
Average deprivation level	2.84	2.82
Average number of prescriptions	3.93	4.70

Table EC.4 Summary statistics for several factors broken down by consultations that take place with a patient's regular doctor and consultations that do not

	Dependent Variable: RD_i	
$ln(ExpectedRI)$	-0.026***	[-0.026,-0.025]
IMD=2	-0.007***	[-0.009,-0.006]
IMD=3	-0.008***	[-0.009,-0.006]
IMD=4	-0.017***	[-0.018,-0.015]
IMD=5	-0.020***	[-0.022,-0.019]
Demand	-0.025***	[-0.027,-0.024]
Female	-0.025***	[-0.026,-0.024]
26-30 yrs	0.017***	[0.015,0.019]
31-35 yrs	0.034***	[0.032,0.036]
36-40 yrs	0.050***	[0.048,0.052]
41-15 yrs	0.070***	[0.069,0.072]
46-50 yrs	0.086***	[0.084,0.088]
51-55 yrs	0.101***	[0.099,0.103]
56-60 yrs	0.115***	[0.113,0.117]
61-65 yrs	0.127***	[0.125,0.129]
66-70 yrs	0.140***	[0.138,0.142]
71-75 yrs	0.150***	[0.148,0.152]
76-80 yrs	0.155***	[0.152,0.157]
81-85 yrs	0.150***	[0.147,0.152]
86+ yrs	0.132***	[0.130,0.135]
1 comorbidity	0.028***	[0.026,0.029]
2 comorbidites	0.037***	[0.035,0.039]
3 comorbidites	0.040***	[0.037,0.043]
4 comorbidites	0.044***	[0.040,0.048]
≥ 5 comorbidites	0.049***	[0.044,0.055]
1 prescription	0.022***	[0.021,0.023]
2 prescriptions	0.033***	[0.032,0.035]
3 prescriptions	0.043***	[0.041,0.044]
4-5 prescriptions	0.049***	[0.048,0.051]
6-7 prescriptions	0.052***	[0.051,0.054]
8-9 prescriptions	0.053***	[0.051,0.055]
10-12 prescriptions	0.052***	[0.050,0.054]
13-15 prescriptions	0.053***	[0.050,0.055]
16+ prescriptions	0.059***	[0.055,0.062]
4 consultations	0.008***	[0.007,0.010]
5 consultations	0.012***	[0.010,0.013]
6 consultations	0.007***	[0.005,0.008]
7 consultations	0.009***	[0.008,0.011]
8 consultations	0.011***	[0.009,0.012]
9 consultations	0.014***	[0.012,0.015]
10 consultations	0.015***	[0.013,0.017]
11-12 consultations	0.016***	[0.015,0.018]
13-14 consultations	0.020***	[0.018,0.022]
15-16 consultations	0.021***	[0.019,0.022]
17-18 consultations	0.022***	[0.020,0.024]
19-20 consultations	0.023***	[0.021,0.025]
21-23 consultations	0.024***	[0.022,0.026]
24-26 consultations	0.026***	[0.024,0.028]
27-30 consultations	0.026***	[0.024,0.028]
31-35 consultations	0.026***	[0.023,0.028]
36-45 consultations	0.027***	[0.024,0.029]
46-55 consultations	0.034***	[0.030,0.038]
56+ consultations	0.046***	[0.042,0.051]
Established GP=1	0.023***	[0.021,0.024]
Mental Health=1	0.020***	[0.019,0.021]
Constant	0.616***	[0.607,0.624]
Observations	11,344,065	
All other controls	Yes	

95% confidence intervals in brackets

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table EC.3 Linear probability model to validate the independent variable

EC.4. Alternative measures of continuity of care

We use different measures of continuity of care as specified in the literature to investigate the effect on the revisit intervals and to ensure that our results are not driven by the use of regular doctor to measure continuity of care. First, we estimate the Bice-Boxerman Continuity of Care index (COCI) which represents the dispersion of visits between providers or the degree of coordination needed between different providers (Pollack et al. 2016). This measure is suitable for longitudinal data and provides comparability across patients. The COCI as calculated as follows:

$$COCI = \frac{(\sum_{i=1}^{p} n_i^2) - n}{n(n-1)}$$

where n_i is the number of visits to provider i and n is the total number of visits. For our setting, we calculate the COCI for every consultation by using all consultations that take place in the two-year interval preceding the focal consultation.

The second measure we use is the Herfindahl–Hirschman index (HHI) which is conceptually similar to the COCI but measures the extent to which the patient's visits are concentrated with a single or group of providers (Maarsingh et al. 2016, Pollack et al. 2016). The HHI is most commonly used to measure market concentration in economic analyses but is also now widely used in the medical literature. The HHI is calculated as follows:

$$HHI = \sum_{i=1}^{p} (\frac{n_i}{n})^2$$

where n_i is the number of visits to provider i and n is the total number of visits. For our setting, we calculate the HHI for every consultation by using all consultations that take place in the two-year interval preceding the focal consultation.

The third measure we use is the Usual Provider of Care index (UPC) which measures the density or the concentration of visits to a single usual provider. For each consultation, the UPC is the number of visits to the regular provider divided by the total number of visits (Maarsingh et al. 2016).

Table EC.5 below shows the effect of COCI, HHI and UPC on the revisit intervals. The results suggest that a higher COCI (lower dispersion), higher HHI (lower dispersion) and a higher UPC (higher density with one provider), all lead to a longer revisit interval.

Table EC.5 Coefficient of RD_i, UPC, HHI and COCI.

Dependent Variable = Natural logarithm of the revisit interval				
$RD_i = 1$	0.140*** [0.138,0.141]			
UPC		0.296*** [0.290,0.298]		
HHI			0.149*** [0.145,0.153]	
COCI				0.122*** [0.118,0.126]
Observations	11,344,065	11,344,065	11,344,065	11,344,065
Adjusted R^2	0.144	0.144	0.142	0.142

95% confidence intervals in brackets
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

EC.5. Description of control variables

To account for various factors that may confound the relationship between continuity of care and revisit intervals, we include different control variables.

Patient demographics

First, differences in patient demographics may affect the patient's preference for seeing her regular doctor as well as her primary care consultation frequency and hence the revisit interval. To account for this, we control for patient age, as older patients are likely to be sicker and make more frequent contact with primary care, and they are also more likely to prefer to see their regular doctor. We include patient age at the time of consultation in categories as shown in Figure EC.2 (Left).

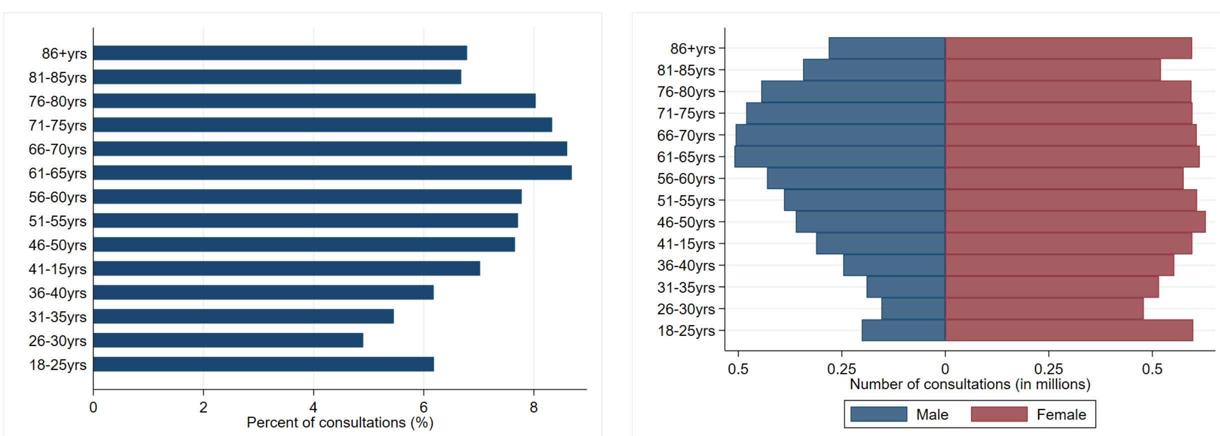


Figure EC.2 (Left) Categories of patient age at the time of consultation; (Right) Breakdown of consultations by gender and age.

Similarly, using the Cambridge Comorbidity Index (CCI), we control for (i) the total number of chronic conditions the patient suffers from at the time of consultation, (ii) a binary variable for each of the 26 individual comorbidities in the CCI and (iii) whether the patient suffers from a mental health condition (described as anxiety, depression or schizophrenia)[9]. The categories used for the total number of chronic conditions the patient suffers from at the time of consultation and the prevalence of the comorbidities in our dataset are represented visually in Figure EC.3.

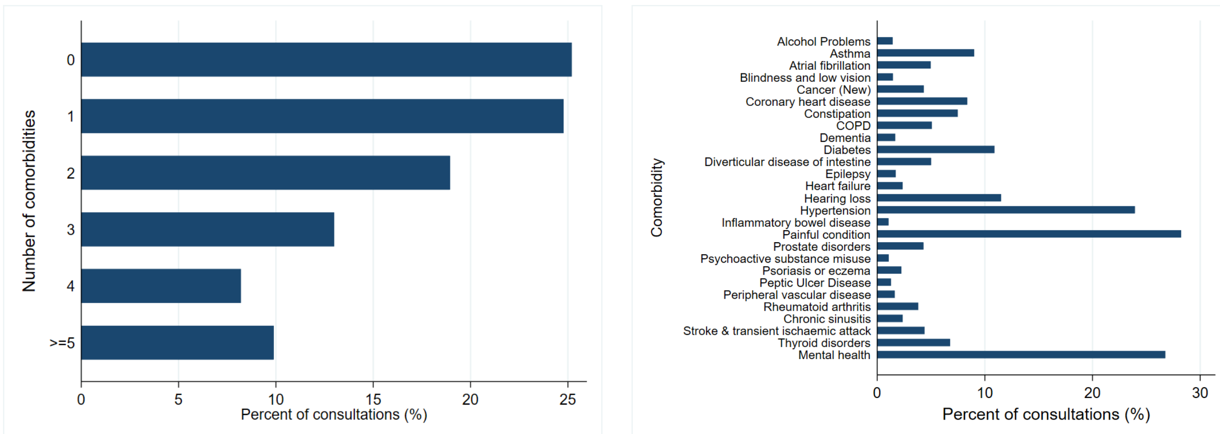


Figure EC.3 (Left) Categories of the number of active comorbidities the patient suffers from at the time of consultation; (Right) Prevalence of each individual comorbidity in the data.

We control for gender as it might be a potential confounder – for example, pregnant women or women who just gave birth are likely to visit a primary care provider at short intervals and also more likely to want to see a doctor they trust. The breakdown of gender by age group is shown in Figure EC.2 (Right) and suggests that for all age groups, women are heavier users of primary care than men.

To further capture the patient's severity and complexity, we calculate the total number of unique repeat medications the patient has been prescribed in the 6-month window preceding the focal consultation. The categories used for the total number of prescriptions is represented visually in Figure EC.4.

Since granular laboratory results are difficult to include meaningfully in a full-patient sample, i.e. across conditions, to control for labs in an aggregate fashion, we include a binary variable that indicates whether hematology-related results are been requested for the patient within the 6 months preceding the focal consultation.

Since, in the UK setting, imaging is not conducted in primary care practices but instead in hospitals, through outpatient referrals, detailed information on imaging is not included in our data. To account for utilization of imaging services, we include a binary control variable that indicates whether the patient has an outpatient referral within the 6 months prior to the focal consultation.

Finally, we control for the socioeconomic status of the patient using the patient-level index of multiple deprivation (IMD), which is provided by CPRD in quintiles) and is a relative characterization of poverty or the socioeconomic situation of the area where the patient resides. The categories of the IMD are shown in Figure EC.5.

[9] There are a total of 38 comorbdities specified in the Cambridge Comorbidity Index, out of which we group 3 comorbidities as mental health condition, and only include those remaining conditions that have a higher than 1% prevalence in our dataset. The leaves us with 26 conditions that we include in the analysis (Payne et al. 2020).

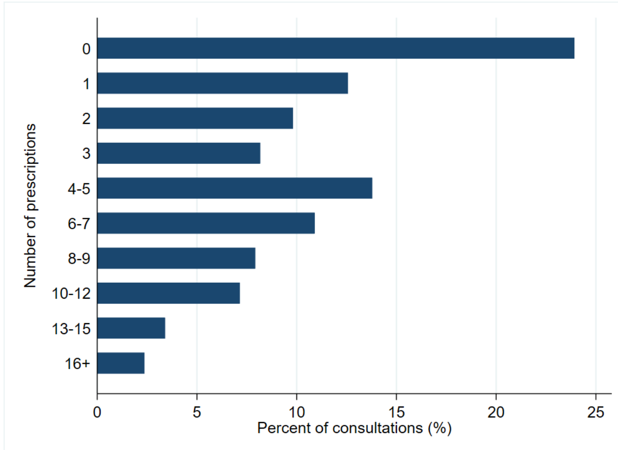


Figure EC.4 **Categories of the number of different active repeat prescriptions the patient is prescribed within a 6-month window preceding the consultation**

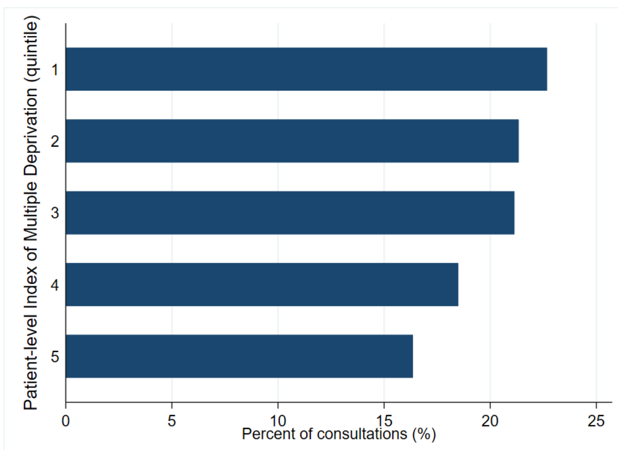


Figure EC.5 **Patient's Index of Multiple Deprivation (IMD) as quintiles (1=least deprived quintile)**

Patient visit history

An important confounding factor that need to be accounted for is the patient's visit history. The more frequently a patient visited in the past, the more likely they are to also visit frequently in the future. Hence, we expect that such a patient will have a shorter revisit interval. At the same time, a patient's likelihood of a consultation with a regular doctor is affected by her consultation frequency (see Figure EC.1). Since we anticipate nonlinear effects (see Figure EC.1), we include both the patient's past consultation frequency as a categorical control as well as her average past revisit interval as a linear continuous control (see Table 3 of the main paper for details). The categories are shown in Figure EC.6.

Doctor type

Third, we note that practices are staffed with two types of GPs:

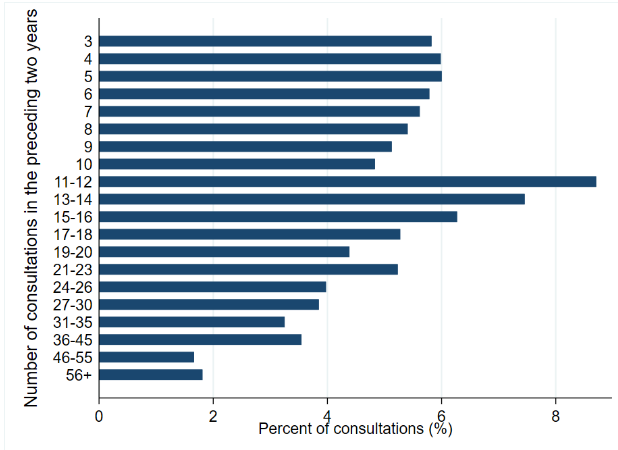


Figure EC.6 Categories of patient's consultations in the two years preceding the focal consultation

- **Established:** Partners, who are co-owners of the practice, and salaried doctors, who are permanent employees of the practice. Usually, established doctors are list-holding doctors who are accountable for the health of their patient list.
- **Unestablished:** Locums or temporary doctors who have temporary contracts with the practice and work ad-hoc shifts. Unestablished doctors are either self-employed or employed at a locum agency, and they are paid on an hourly basis.

Even though the two categories of doctors receive similar training, the incentives for the established doctors to "get it right the first time" might be stronger, which would translate to a higher quality and productivity benefit if the patient's regular provider is an established doctor. According to our regular provider assignment algorithm, established doctors are regular providers of the patient 93% of the time, whereas unestablished doctors are regular providers the rest of the time.

Practice and time-related controls

Fourth, we include practice fixed effects to capture unobserved time-invariant heterogeneity across practices. For practice-level observed time variant heterogeneity, we include a dynamic control that measures practice-level variation in demand. We calculate this as the total practice demand during the focal week of each consultation as compared to a weekly average in a 52-week period around the focal week. This measure captures fluctuations in demand, which may be correlated with the probability that the patient will see her regular doctor and with the scheduling capacity of the practice.

Lastly, time fixed effects are included to account for any factors that change over time and have a common effect on all practices. Controls are included for year to account for trends, month of year to account for seasonality, and day of the week effects to account for differences in demand and supply across different days of the week that may affect both a patient's seeing her regular doctor and the revisit interval.

EC.6. Patient visits

In Figure EC.7 we show the histogram of patient visits in our data. For each patient, we calculate the total number of visits in the data divided by the total number of years registered in the data. We find that the

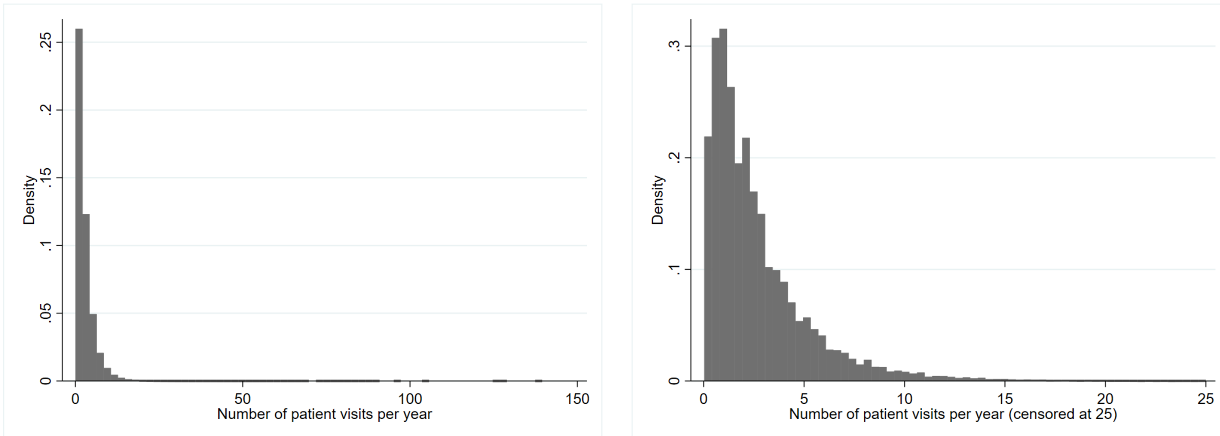


Figure EC.7 **(Left) Histogram of number of patient visits per year; (Right) Histogram of number of patient visits per year (censored at 25 visits)**

average number of visits per year is 2.7, the median number of visits is 2.0, with a standard deviation of 2.8 visits.

EC.7. Instrumental variable construction

To calculate the instrumental variable, let C_{gt} denote the set of consultations in week t for which doctor g is specified as patients' regular doctor. In other words, $|C_{gt}|$ denotes the total number of consultations in week t made by patients for whom doctor g is their regular doctor, where $|\cdot|$ specifies the cardinality of the set. Let g_c denote the actual doctor who the patient saw during consultation c. Then we define the accessibility of doctor g during week t for patients who consider doctor g as their regular doctor to be

$$WkAccessibility_{gt} = \frac{\sum_{c\in C_{gt}} I\left[g_c = g\right]}{|C_{gt}|}$$

where $I\left[\cdot\right]$ is an indicator function that takes value 1 when the condition inside the brackets is satisfied, and 0 otherwise.

Note that $WkAccessibility_{gt}$ also includes any visits by patient i, which may produce a mechanical relationship between the instrumental variable and the dependent variable in the selection equation (i.e., whether or not the patient saw their regular doctor). To prevent this, let C_{igt} be the set of consultations made by patient i in week t for which doctor g was specified as their regular doctor. Then we can define accessibility, excluding patient i, as follows:

$$WkAccessibility_{igt} = \frac{\sum_{c\in\left\{C_{gt}\setminus C_{igt}\right\}} I\left[g_c = g\right]}{|\{C_{gt}\setminus C_{igt}\}|}$$

Next, we standardize $WkAccessibility_{igt}$ by calculating the same measure over a 52-week preceding week t. Specifically, let $T_t = \{t-52,\ldots,t-2,t-1\}$ denote the set of 52 weeks prior to week t, excluding week t itself. Then the average accessibility over this 52-week period can be written as

$$YrAccessibility_{igt} = \frac{\sum_{t\in T_t}\sum_{c\in\left\{C_{gt}\setminus C_{igt}\right\}} I\left[g_c = g\right]}{\sum_{t\in T_t}|\{C_{gt}\setminus C_{igt}\}|}$$

Finally, we define the instrumental variable, IV, as:

$$IV_{igt} = \frac{WkAccessibility_{igt}}{YrAccessibility_{igt}}$$

EC.7.1. Statistical tests of the instrumental variable

To validate the instrumental variable, we perform formal hypothesis test for under- and weak identification. The underidentification test is a Lagrange Multiplier (LM) test that tests the rank of the matrix to determine whether the equation is identified i.e. whether the excluded instrument is "relevant" or correlated with the endogenous regressor in the first stage selection equation. Weak identification is when the excluded instrument is weakly correlated with the potentially endogenous regressor. Weak instruments can lead to poor performance of estimators, specifically, estimators might be inconsistent, confidence intervals can be incorrect and the tests for significance of coefficients might lead to wrong conclusions.

Though these tests are designed for a continuous endogenous variable, we proceed with testing by treating our binary endogenous variable as continuous. While this means that the critical values of the tests and the significance levels might be slightly different from those reported here, we note that the CF (endogeneous variable treated as binary) and 2SLS (endogeneous variable treated as continuous) results are nearly identical (see Table 6 in the main paper).

For the underidentification test, we use the Sanderson and Windermeijer χ^2 Wald statistic as reported using the `ivreg2` routine in Stata 16 (Sanderson and Windmeijer 2016, Baum et al. 2002). Under the null, the equation is underidentified. For our model, the SW χ^2 statistic takes a value of 150,000 with 1 d.f., which has corresponding p-value = 0.00. Hence, there is strong evidence to reject the null hypothesis of underidentification, and we conclude that the excluded instrument is "relevant".

For weak identification, we use the Sanderson and Windermeijer first stage F-statistic which is the F version of the SW χ^2 statistic. It is used as a diagnostic for whether the endogenus regressor is "weakly identified". The F-statistic from the model is compared against the critical values of the Kleibergen-Paap statistic (for cluster robust standard errors) reported in Stock et al. (2005) to determine whether the instruments are weak. The null hypothesis of the test is that the equation is weakly identified and the maximum bias of the IV estimator relative to the bias of OLS is some specified value such as 10%. For a single endogenous regressor with cluster robust standard errors, the Stock-Yogo critical values are 16.38, 8.96, 6.66 and 5.53 for maximum bias of 10%, 15%, 20% and 25%, respectively. If the estimated F-statistic is less than a particular critical value then the interpretation is that the instruments are weak for that level of bias. In our model, the estimated SW F-statistic is equal to 150,000, indicating a maximal bias of (significantly) less than 10%. Hence, this suggests that that there is no evidence to suspect that our models are affected by the problem of weak instruments.

EC.7.2. Sensitivity of the instrumental variable

Due to its construction, the instrumental variable is sensitive to low values of $\{C_{gt} \backslash C_{igt}\}$ and $\sum_{t \in T_t} \sum_{c \in \{C_{gt} \backslash C_{igt}\}}$. This leads to long tails in the distribution, though we note that 99% of the values are in the range 0 and 2, with a mean of 1.08 and a median of 1.08. Given the long tails in the IV distribution, we have also tested two alternative characterizations of the IV as robustness:

1. We use a dichotomized version of the instrumental variable following MacCallum et al. (2002), specifically, $BinaryIV = 0$ if $WkAccessibility_{igt} < YrAccessibility_{igt}$ and $BinaryIV = 1$ otherwise. The interpretation is that when $BinaryIV$ is 0, the doctor is less accessible to his patients during that week

compared to his yearly average, and if $BinaryIV$ takes a value of 1, the doctor is more accessible to his patients than the yearly average.

2. We censor the instrumental variable at 2, referring to it as the $CappedIV$, and include an additional dummy variable as a control to indicate those observations for which the instrumental variable took a value of greater than 2 (1% of the observations). Similar to the interpretation of $BinaryIV$, when $CappedIV$ is less than 1, the doctor is less accessible to her patients during that week compared to her yearly average, and if the value is greater than 1, the doctor is more accessible to her patients than the yearly average.

Summary statistics for the original IV and for the two variations on the instrumental variable described above are given in Table EC.6. Using $BinaryIV$ and $CappedIV$ in our control function model, the estimated effect sizes take value 18.1% (95% CI: [16.9%, 19.2%]) and 17.6% (95% CI: [16.6%, 18.6%]), respectively. Thus, the results are insensitive to the choice of instrument.

Table EC.6 Descriptive statistics for the instrumental variable

	Mean	Median	Min	Max	SD
(1)$WkAccessibility_{igt}$	0.49	0.49	0.00	1.00	0.20
(2)$YrAccessibility_{igt}$	0.48	0.46	0.00	1.00	0.16
(3) Instrumental Variable	1.08	1.08	0.00	252.00	0.64
(4) Binary version of the IV ($BinaryIV$)	0.66	1.00	0.00	1.00	0.47
(5) Capped version of the IV ($CappedIV$)	1.07	1.08	0.00	2.00	0.32

EC.7.3. Different time frames for calculation of the IV

In the calculation of the instrumental variable, we use 'week' as the measurement of time for the calculation of the instrumental variable as it is more granular. As the time-period the IV is calculated over increases, the variation will reduce. Therefore, there is a trade-off between having a weak instrument versus an instrument that is very noisy. Choosing 'day' as the time-period might lead to a noisy measurement of the IV as we can expect large fluctuations; in contrast, a 'month' might have insufficient variation. Additionally, in this section we also consider alternative time-frames such as fortnight and month to ensure that our results are not driven by the choice of week as the measurement of time to calculate the IV.

Table EC.7 Coefficient of RD_i across two different measurements to calculate the IV (i) fortnight, (ii) month.

Dependent Variable = Natural logarithm of the revisit interval		
	Fortnight	Month
RD_i=1	0.146***	0.130***
	[0.136,0.157]	[0.117,0.143]
Observations	11,344,065	11,344,065

Notes: 95% confidence intervals in square brackets

EC.7.4. IV control

One concern is that there may be omitted variables that threaten the validity of the instrumental variable, particularly if there are unobserved factors that correlate with both the relative accessibility of the regular doctor for other patients (instrument) and the focal patient's revisit interval (dependent variable). For example, if there is a flu outbreak, the focal patient's expected revisit interval might decrease, and at the same time other patients would also find it harder than normal to access their regular provider.

However, such omitted factors should also be expected to affect the revisit interval of the doctor's other patients. We use this insight to construct a control variable to improve the validity of the instrumental variable. Specifically, we add a control that takes the focal patient's expected $\ln(RI)$ and adjusts this for the average of $\ln(RI)$ of other patients who (i) share the same regular doctor as the focal patient and (ii) visit a doctor in the same week as the focal patient. Intuitively, when the average revisit interval of these other patients change, then the expected $\ln(RI)$ of the focal patient should change too.

More specifically, we calculate the control described above for each consultation c'_i that occurred on day t in week w for patient i. To do so, first let $C_{i[t-730,t-1]}$ denote the set of consultations that occurred between day $t-1$ and $t-730$ for patient i. The time between each consultation c_i made by patient i in that interval is then given by RC_{c_i}. Thus, the past average revisit interval (RI) for consultations that occurred over the two years $(t-730, t-1)$ prior to consultation c'_i is as follows:

$$PastAvgRI_{c'_i} = \frac{\sum_{c_i \in C_{i[t-730,t-1]}} RI_{c_i}}{\left|C_{i[t-730,t-1]}\right|}$$

where $|\cdot|$ specifies the cardinality of the set.

Next, let $g_{c'_i}$ denote the regular doctor associated with patient i at the time of consultation c'_i. Further, let $J_{c'_i w}$ specify the set of all consultations by *other* patients for which the regular doctor is the same as for consultation c'_i (i.e., is $g_{c'_i}$) and that took place in the same week w in which consultation c'_i occurred. Using the same method as described above, we can then calculate the past average revisit interval for those consultations $c'_j \in J_{c'_i w}$, where j specifies the patient associated with consultation c'_j. These revisit intervals are given by $PastAvgRI_{c'_j}$.

Next, we create a multiplier that captures the difference between the current and past revisit intervals for those consultations $c'_j \in J_{c'_i w}$, which is equal to $MultRI_{c'_j} = RI_{c'_j}/PastAvgRI_{c'_j}$. Averaging the multiplier over all consultations $c'_j \in J_{c'_i w}$ we get

$$\overline{MultRI}_{c'_i} = \frac{\sum_{c'_j \in J_{c'_i w}} MultRI_{c'_j}}{|J_{c'_i w}|}.$$

Notice that when $\overline{MultRI}_{c'_i} < 1$ it suggests that revisit intervals were shorter than normal for *other* patients who visited during the same week w in which consultation c'_i took place and who shared the same regular doctor $g_{c'_i}$. Meanwhile, when $\overline{MultRI}_{c'_i} > 1$ it suggests that revisit intervals were longer than normal.

The multiplier thus allows us to account for unobserved factors that are correlated with both the IV and the outcomes. For example, if there is a flu outbreak, we might expect the multiplier to take a value less than 1, reflecting the fact that parents are expected to return faster on average. Using this insight, we can then adjust the expected revisit interval of the focal patient by taking their past revisit interval and multiplying by this multiplier, giving us:

$$ExpectedRI_{c'_i} = PastAvgRI_{c'_i} \times \overline{MultRI}_{c'_i}$$

This variable thus gives us the expected time to the next visit of patient i based on their past revisit interval and the multiplier calculated from other patients.

We present a simple numerical example to illustrate how this control works. Say, patient I has 3 consultations in the 2 years preceding the focal consultation, with an average revisit interval of 60 days. Patient I has a regular doctor X at the focal consultation. Doctor X also has 3 other patients A,B and C who have consultations during that week and also consider him their regular doctor. Patients A, B and C have a past average revisit interval of 30, 60 and 90 respectively. After the focal consultation, their next consultation is after 70, 80 and 30 days respectively. In this case, the multiplier or $MultRI$ is 30/70, 60/80 and 90/30 or 0.43, 0.75 and 3 respectively. Averaging this number gives us 1.4. Since the $\overline{MultRI}$ is greater than 1, the RI's are longer than usual for the regular doctor's other patients for that week. Hence, for the focal consultation, the $ExpectedRI$ would be 60 times 1.4 which is equal to 84.

Adding $\ln(ExpectedRI_{c'_i})$ as an additional control variable in our model allows us to account for various unobserved factors that might be correlated with both the availability of the regular doctor for other patients (the instrument) and also the expected revisit interval of the focal patient (the dependent variable).[10] Summary statistics for this control variable are given in Table EC.8.

Table EC.8 Descriptive statistics for the control for the instrumental variable

	Mean	Median	Min	Max	SD
(1) $\ln(ExpectedRI)$	4.08	4.10	-3.88	10.25	0.77

EC.8. Confounding: Subsample analysis

As outlined in Section 5.4, patient acuity is the most obvious source of omitted variable bias. Specifically, higher acuity patients may be unwilling to wait to see their regular provider, and hence are more likely to see their non-regular doctor. Moreover, these acute patients are more likely to return sooner after the focal consultation for a follow-up visit after the index consultation. This would lead to a shorter revisit interval and could provide an alternative, non-causal explanation for β_i in the OLS model.

One way we address endogeneity concerns is by subsampling the data and comparing the effect sizes associated with seeing the regular doctor across subsamples that indicate higher acuity patients. In selecting subsamples, we aim to identify potentially "acute" consultations for which patients have a higher likelihood of seeing a non-regular doctor and also a shorter revisit interval. We wish to establish that our results are consistent across acute and non-acute consultations, so we perform the following analyses.

[10] We use the natural logarithm of the control to match with the units of the dependent variable (natural logarithm of the revisit interval).

Antibiotic prescriptions

We first consider samples based on prescriptions. Specifically, we identify (i) the sample of consultations when no medication was prescribed and (ii) the sample of consultations when a new antibiotic was prescribed. Antibiotic prescriptions are identified using Gulliford et al. (2020). The rationale is that patients who are prescribed a new antibiotic during their consultation are more likely than other patients to have an acute problem. We estimate the same model as in equation (1) (re-written below), using the same dependent variable and the same set of control variables:

$$\ln(RI_i) = \beta_0 + \beta_1 RD_i + \mathbf{X}_i \boldsymbol{\beta} + \epsilon_i\,, \tag{EC.1}$$

where the vector $\mathbf{X}_i$ specifies the set of controls corresponding to observation i, as defined in Section 4.3.3, and $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ is the error term.

If acuity is in fact a major confounder, we expect that the acute consultations would have a smaller value of β_1 in equation (EC.1). The results (reported in Table EC.9) show that the effect size is indeed slightly smaller, suggesting that the effect is less pronounced for acute conditions. However, the effect remains large and significantly positive.

Table EC.9 Coefficient of RD_i across two different samples: the sample of consultations when (i) no medication was prescribed, (ii) when a new antibiotic was prescribed.

Dependent Variable = Natural logarithm of the revisit interval		
Subsample:	No New Prescription	New Antibiotic Prescription
RD_i=1	0.158***	0.117***
	[0.155,0.161]	[0.112,0.122]
Observations	6,310,933	1,354,977

Notes: 95% confidence intervals in square brackets; The total number of observations in this table is not equal to the total sample size of 11,344,065 observations, as there is a category of observations at which a non-antibiotic medicine was prescribed, which we do not include in this analysis.

Antiviral drug prescriptions

Additionally, we identify the sample of consultations when an antiviral drug was prescribed. The rationale is similar to antibiotics in that patients who are prescribed a new antiviral drug during their consultation are more likely than other patients to have an acute problem. We estimate the same model as in equation (1) (re-written below), using the same dependent variable and the same set of control variables:

$$\ln(RI_i) = \beta_0 + \beta_1 RD_i + \mathbf{X}_i \boldsymbol{\beta} + \epsilon_i\,, \tag{EC.2}$$

where the vector $\mathbf{X}_i$ specifies the set of controls corresponding to observation i, as defined in Section 4.3.3, and $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ is the error term.

If acuity is in fact a major confounder, we expect that the acute consultations would have a smaller value of β_1 in equation (EC.1). The results (reported in Table EC.10) show that the effect size is indeed slightly smaller, suggesting that the effect is less pronounced for acute conditions. However, the effect remains large and significantly positive.

Table EC.10 Coefficient of RD_i across a sample of consultations when an antiviral drug was prescribed.

Dependent Variable = Natural logarithm of the revisit interval	
	Antiviral Prescription
RD_i=1	0.090*** [0.048,0.131]
Observations	20,100

Notes: 95% confidence intervals in square brackets;

Testing

We also identify two additional samples based on testing data (i) the sample of consultations when any type of test was recorded (ii) the sample of consultations when a blood test was recorded[11]. The rationale is that tests are more likely to be ordered for patients with an acute problem. We estimate the same model as in equation (1) (re-written below), using the same dependent variable and the same set of control variables:

$$\ln(RI_i) = \beta_0 + \beta_1 RD_i + \mathbf{X}_i\boldsymbol{\beta} + \epsilon_i\,, \tag{EC.3}$$

where the vector $\mathbf{X}_i$ specifies the set of controls corresponding to observation i, as defined in Section 4.3.3, and $\epsilon_i \sim \mathcal{N}(0,\sigma^2)$ is the error term.

If acuity is in fact a major confounder, we expect that the acute consultations would have a smaller value of β_1 in equation (EC.1). The results (reported in Table EC.11) show that the effect size is indeed slightly smaller, suggesting that the effect is less pronounced for acute conditions. However, the effect remains large and significantly positive.

Table EC.11 Coefficient of RD_i across two different samples: the sample of consultations when (i) any type of test was recorded, (ii) when a haematology-related test was recorded.

Dependent Variable = Natural logarithm of the revisit interval		
Subsample:	Any test	Haematology-related test
RD_i=1	0.158*** [0.149,0.167]	0.119*** [0.070,0.170]
Observations	360,819	12,307

Notes: 95% confidence intervals in square brackets

Prior emergency department visit

Consultations are more likely to be acute if they are preceded by an emergency department (ED) visit. For example, a patient visiting the ED for an acute reason might be asked by the hospital doctors to follow up with a doctor at their primary care practice. Thus, we class those consultations that are preceded by an ED visit in the seven days prior to the consultation as more acute. We compare the effect of seeing the regular doctor on the revisit interval on two different subsamples: (i) when the index consultation was preceded by

[11] Blood tests are identified using the haematology read codes which start with '42'

an ED visit within seven days and (ii) the remainder of the sample, which had no ED visits within seven days before the index consultation.

Again, we re-estimate the same model as in equation (1). If the results are driven by acute consultations only, we would expect that the effect size of the subsample where the consultation was preceded by an ED visit would have a much larger effect size. However, as we report in Table EC.12, the effect size is large and significant in both cases.

Table EC.12 Coefficient of RD_i across two different samples: the sample of consultations that were (i) preceded by an ED visit in the 7 days prior to the consultation, and (ii) those that were not.

Dependent Variable = Natural logarithm of the revisit interval		
Subsample:	ED visit 7 days prior	No ED visit 7 days prior
RD_i=1	0.120***	0.152***
	[0.104,0.135]	[0.150,0.153]
Observations	175,123	11,168,942

Notes: 95% confidence intervals in square brackets;

EC.9. Doctor and patient heterogeneity

EC.9.1. Patient-level heterogeneity

Patients might have their own unique circumstances that might affect their decision to revisit and when they are able to actually visit. To account for such patient level unobserved time-varying heterogeneity, we include patient random effects to our main model specification. We find that the results remain qualitatively similar.

Table EC.13 Coefficient of RD_i for a model with patient-level random effects.

Dependent Variable = Natural logarithm of the revisit interval	
	Patient-RE
RD_i=1	0.131***
	[0.130,0.134]
Observations	11,344,065

Notes: 95% confidence intervals in square brackets;

EC.9.2. Physician specific effects

According to Schwartz et al. (1999), physicians receive very little guidance on visit scheduling times and hence widely vary in their recommendations of visit intervals. To capture physician specific variation and scheduling preferences, we include physician fixed effects (instead of practice fixed effects) to our main model. Additionally, we also believe that physician roles can be a confounder. For example, partners co-own the practice, hence have greater responsibility towards their patients, and have financial liability as owners of the practice. Salaried physicians, on the other hand, do not have a stake in the ownership and receive a fixed salary. Hence, partners may be more confident in scheduling a longer revisit interval. We find that the results remain qualitatively similar for both specifications.

Table EC.14 Coefficient of RD_i for a model with physician-level fixed effects.

Dependent Variable = Natural logarithm of the revisit interval	
	Physician-FE
RD_i=1	0.146***
	[0.144,0.148]
Observations	11,344,065

Notes: 95% confidence intervals in square brackets;

Table EC.15 Coefficient of RD_i for a model with physician roles.

Dependent Variable = Natural logarithm of the revisit interval	
	Physician role
RD_i=1	0.139***
	[0.137,0.141]
Observations	11,344,065

Notes: 95% confidence intervals in square brackets;

EC.9.3. Practice−year fixed effects

To account for the changes in practice scheduling policy over time, we additionally include practice−year fixed effects and find that the results remain similar.

Table EC.16 Coefficient of RD_i for a model with practice−year fixed effects.

Dependent Variable = Natural logarithm of the revisit interval	
	Practice−Year fixed effects
RD_i=1	0.154***
	[0.152,0.156]
Observations	11,344,065

Notes: 95% confidence intervals in square brackets;

EC.10. Percentage of past visits with regular doctor

To differentiate between patients who have had a larger share of their past consultations with their regular doctor and patients who have had a smaller share of their consultations, we include $VisitPercentage$ as an additional control variable which is calculated at each consultation as the number of visits with the regular doctor in the 2 years preceding the focal consultation as a proportion of the total number of visits in the 2 years preceding the focal consultation.

EC.11. Summary statistics for the matched sample

In Table EC.18 we compare the proportions of the control variables for the full sample with those corresponding to the control group and the treatment group. (Specifically, in this table we compare the patient demographics across the different groups. Comparison across other factors (not shown) also suggests that there is no significant different between the groups.) We report p-values corresponding to Pearson χ^2 tests

Table EC.17 Coefficient of RD_i for a model with $VisitPercentage$ control.

Dependent Variable = Natural logarithm of the revisit interval	
	Including $VisitPercentage$
RD_i=1	0.153***
	[0.151,0.154]
Observations	11,344,065

Notes: 95% confidence intervals in square brackets;

for differences in the proportions between the control and treatment groups in the matched sample. As can be seen, the proportions remain the same across the three columns of Table EC.18, as desired. We also report standardized differences that measure effect sizes between different groups and are recommended for comparing baseline covariates in propensity matched studies (Yang and Dalton 2012). For each of the covariates, the low standardized differences do not exceed the imbalance threshold of 0.2.

Table EC.18 Comparing the balance of the full sample with the control group and treatment group of the matched sample

	Full sample (%)	Control group (%)	Treated group (%)
Age band			
18-25yrs	5.53	5.03	4.88
26-30yrs	4.20	3.95	3.91
31-35yrs	4.87	4.68	4.72
36-40yrs	5.87	5.92	6.00
41-15yrs	6.97	7.12	7.14
46-50yrs	7.79	8.00	8.00
51-55yrs	7.91	8.18	8.19
56-60yrs	8.05	8.03	8.31
61-65yrs	9.04	8.28	9.34
66-70yrs	9.04	9.30	9.08
71-75yrs	8.88	9.03	8.71
76-80yrs	8.46	8.73	8.35
81-85yrs	6.88	6.79	6.74
86+yrs	6.62	6.59	6.62
		p-value for Pearson test (χ^2)=0.789	
		Standardized difference: 0.009	
Comorbidity			
0 comorbidites	19.39	17.94	17.75
1 comorbidity	21.65	22.07	22.18
2 comorbidites	18.62	19.07	19.19
3 comorbidites	14.26	14.45	14.48
4 comorbidites	10.16	10.30	10.25
≥ 5 comorbidites	15.92	16.18	16.14
		p-value for Pearson test (χ^2)=0.568	
		Standardized difference: 0.006	
Mental Health			
No	72.07	72.46	72.85
Yes	27.93	27.55	27.51
		p-value for Pearson test (χ^2)=0.774	
		Standardized difference: 0.001	
Gender			
Male	37.84	37.83	37.88
Female	62.16	62.17	62.12
		p-value for Pearson test (χ^2)=0.771	
		Standardized difference: 0.001	
Index of Multiple Deprivation			
1	23.00	23.05	23.02
2	21.50	21.80	21.80
3	21.17	20.73	20.62
4	18.35	18.24	18.32
5	16.00	16.18	16.23
		p-value for Pearson test (χ^2)=0.866	
		Standardized difference: 0.004	
Prescriptions			
0	22.80	21.94	21.79
1	12.44	12.68	12.72
2	9.89	10.10	10.13
3	8.33	8.59	8.61
4-5	14.16	14.17	14.24
6-7	11.18	11.18	11.24
8-9	8.10	8.13	8.12
10-12	7.27	7.33	7.27
13-15	3.44	3.53	3.51
16+	2.29	2.36	2.36
		p-value for Pearson test (χ^2)=0.977	
		Standardized difference: 0.005	

Notes: The null hypothesis for the Pearson χ^2 test is that there is no difference between the two proportions. In each case, we find that we cannot reject the null at the 5% significance level; standardized differences (calculated using Yang and Dalton (2012)) = difference in means or proportions divided by standard error; imbalance defined as absolute value greater than 0.20 (small effect size)

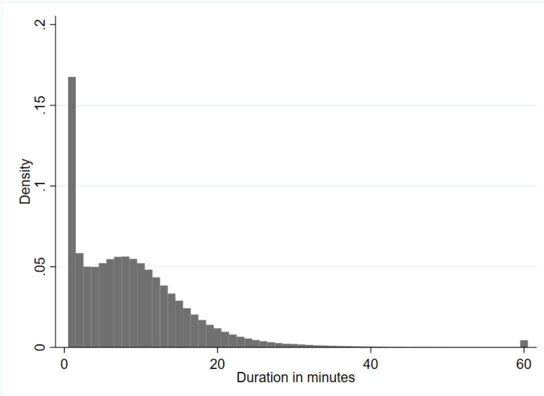


Figure EC.8 **Distribution of Duration**

	Duration
Mean	9.05
Median	8.00
Min	0.50
Max	60.00
Std. Dev.	8.11

Table EC.19 **Descriptive statistics for duration**

EC.12. Results for the duration of consultations

We first report the summary statistics and the histogram of the distribution of duration of consultations, the change in average duration over time, and the results from the range of models.

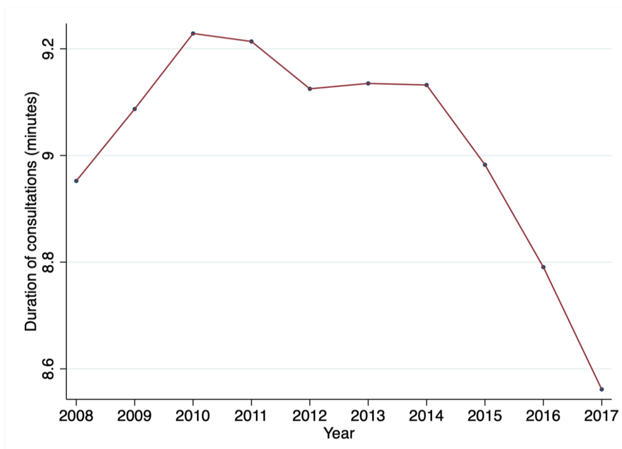


Figure EC.9 **Change in duration over time**

Table EC.20 summarises the results of all the models introduced in Section 5 but using the length of consultation as the dependent variable. The results show that the regular doctors spend less time on average with their patients than transactional providers. Hence, providing care continuity will not come at a cost of reducing daily clinical throughput in primary care practices. However, the effect size is not large enough to be actionable. Physicians typically work in four-hour shifts and will be unable to accommodate an additional patient.

EC.13. Moderation results

Additionally, Figure EC.10 shows a graphical representation of the estimates of the average marginal effects based on the moderation results corresponding to Table 7 in the main paper.

Table EC.20 Coefficient of RD_i across all modeling techniques

	Dependent variable = Length of the consultation (minutes)							
	Sample	Model	Coefficient	Std. Error	t-statistic	P> $\|t\|$	95% CI	
1a	consultations	OLS	-0.37	0.01	-66.0	0.00	-0.38	-0.36
1b	consultations	CF-IV	-0.15	0.03	-5.01	0.00	-0.20	-0.09
1c	consultations	IV-2SLS	-0.11	0.03	-3.67	0.00	-0.17	-0.05
2a	patients	PSM	-0.32	0.02	-13.86	0.00	-0.37	-0.27
2b	patients	PSM OLS	-0.32	0.02	-14.47	0.00	-0.37	-0.28
3a	patients	IPW	-0.29	–	–	–	-0.33	-0.25
3b	patients	MBE ($\theta = 0.25$)	-0.28	–	–	–	-0.38	-0.16
3c	patients	MBE ($\theta = 0.10$)	-0.26	–	–	–	-0.55	-0.13
3d	patients	MBE ($\theta = 0.05$)	-0.19	–	–	–	-0.66	-0.07

Notes: Standard errors clustered at the patient level for models 1a-1c, the matched pair level for models 2a-2b, and bootstrapped standard errors for models 3a-3d.
This table reports the coefficient estimates of RD_i (seeing the regular doctor) for the taxonomy of models specified in Section 5. OLS regression refers to the ordinary least squares regression equation 1. CF refers to the control function approach specified in Section 5.4.1 and 2SLS refers to the two-stage least squares method using the same IV as used for the CF approach. PSM corresponds to the matching-based effect estimation where we report differences in averages of $\ln(RI)$ (log revisit interval) between the control and treated groups, using nearest neighbor matching without replacement and a caliper of 0.001 (Section 5.5.2). PSM OLS corresponds to an OLS regression on the matched sample that includes the covariates (Section 5.5.2). Models 3a correspond to the IPW estimator whereas models 3b-3d correspond to the minimum bias estimators with $\theta = 25\%, 10\%, 5\%$ respectively. (Section 5.5.3).

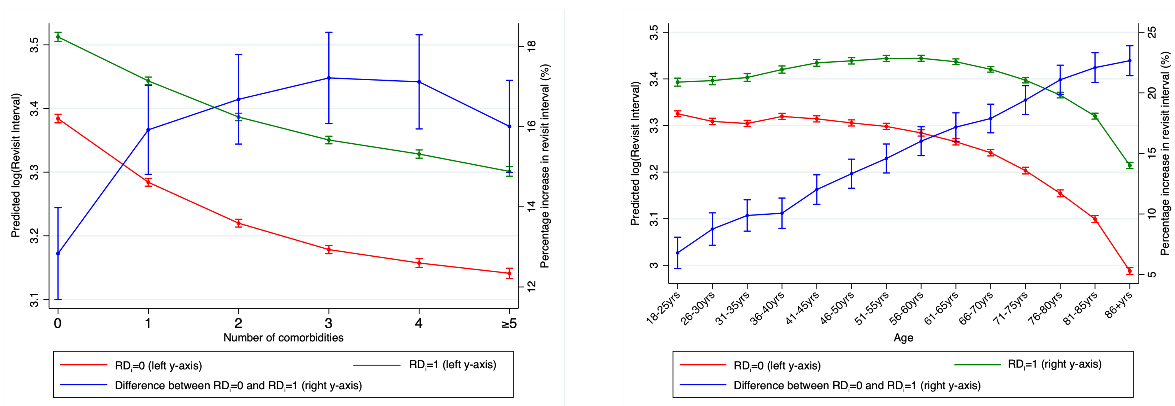


Figure EC.10 Percent increase in ln(Revisit Interval) when the patient sees a transactional provider $(RD_i = 0)$ compared to when the patient sees her regular doctor $(RD_i = 1)$, (left) for patients with multiple chronic diseases (H(a)) and (right) across different age groups (H(b))

EC.14. Calculation of the scoring index for targeting continuity of care

Our results suggest that practices should increase their continuity offering to improve productivity. However, we do not recommend that practices offer increased continuity on a random basis, but take into account our results to design the reallocation of continuity to patients who benefit from it the most. Using the insights from Hypothesis 2, we propose a scoring system that can be used by practice managers to prioritize those patients for whom care continuity has the most productivity-enhancing effect and target them for more relational services. Given the current lack of tools to monitor continuity or to measure how well practices have been providing continuity of care to patients who benefit from it, we believe this methodology is an important practical contribution in its own right (Palmer et al. 2018, Hill and Freeman 2011).

Table EC.21 Moderating effects of age, comorbidities and mental health, using the control function approach

CF: Dependent variable = Natural logarithm of the revisit interval				
	$RD_i = 0$		Interaction	
	Coefficient	95% CI	Coefficient	95% CI
Baseline ($RD_i = 1$)	–	–	0.032	[0.019,0.045]
26-30yrs	-0.016***	[-0.023,-0.009]	0.020*	[0.010,0.030]
31-35yrs	-0.021***	[-0.027,-0.015]	0.031***	[0.021,0.041]
36-40yrs	-0.006	[-0.012,0.000]	0.033***	[0.023,0.042]
41-45yrs	-0.011***	[-0.017,-0.005]	0.052***	[0.043,0.062]
46-50yrs	-0.020***	[-0.026,-0.014]	0.065***	[0.056,0.074]
51-55yrs	-0.027***	[-0.033,-0.021]	0.078***	[0.069,0.087]
56-60yrs	-0.041***	[-0.047,-0.035]	0.092***	[0.083,0.101]
61-65yrs	-0.060***	[-0.066,-0.054]	0.104***	[0.095,0.113]
66-70yrs	-0.083***	[-0.090,-0.077]	0.111***	[0.102,0.120]
71-75yrs	-0.122***	[-0.128,-0.108]	0.126***	[0.117,0.136]
76-80yrs	-0.171***	[-0.178,-0.164]	0.143***	[0.133,0.152]
81-85yrs	-0.226***	[-0.233,-0.219]	0.153***	[0.143,0.163]
86+yrs	-0.337***	[-0.345,-0.330]	0.159***	[0.149,0.169]
1 comorbidity	-0.100***	[-0.104,-0.096]	0.031***	[0.026,0.036]
2 comorbidities	-0.164***	[-0.169,-0.159]	0.038***	[0.033,0.044]
3 comorbidities	-0.206***	[-0.212,-0.199]	0.044***	[0.037,0.050]
4 comorbidities	-0.227***	[-0.234,-0.219]	0.043***	[0.036,0.050]
≥ 5 comorbidities	-0.243***	[-0.252,-0.234]	0.032***	[0.025,0.039]
Mental health condition	-0.008***	[-0.011,-0.005]	0.023***	[0.019,0.028]

Notes: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$; 95% confidence intervals in square brackets, with standard errors clustered at the patient level; 'Baseline' row gives the effect of seeing the regular doctor for an 18-25 year old with no comorbidities and no mental health condition; The '$RD_i = 0$' column specifies the effect of age, comorbidity and mental health for a patient who did not see their regular doctor; The effect of age, comorbidity and mental health for a patient who saw their regular doctor can be determined by taking the baseline effect, and adding to this the sum of the coefficients from the '$RD_i = 0$' and 'Interaction' columns.

The proposed scoring system ranks consultations by the estimated number of days gained if the consultation is offered by the patient's regular doctor as compared to another doctor. To determine the revisit interval days gained, we begin by estimating an interaction model in which we include interactions between the binary indicator for a consultation with the patient's regular doctor (i.e., RD_i) and every other independent variable, using the same method as in equation (2).[12] For each consultation i, we then use this model to predict the revisit interval under two scenarios: (i) assuming the patient receives care continuity ($RD_i = 1$), which we denote $\ln(\hat{RI}_1)_i$, and (ii) assuming the patient does not receive care continuity ($RD_i = 0$), which we denote $\ln(\hat{RI}_0)_i$. In order to calculate the productivity gain in number of days rather than in logarithmic units, we transform $\ln(\hat{RI}_0)_i$ and $\ln(\hat{RI}_1)_i$ using Duan's smearing transformation to get $R\hat{I}_{0i}$ and $R\hat{I}_{1i}$ respectively, in days.[13] For each patient-consultation, we define $days\_gained_i = R\hat{I}_{1i} - R\hat{I}_{0i}$, which is the expected days gained or "score" in days from providing care continuity for consultation i.

[12] This model gives us the best in-sample fit for the predicted revisit interval.

[13] In a model $\ln(y_i) = \mathbf{x}\beta + \epsilon_i$, if the errors are normally distributed as $N(0, \sigma_\epsilon^2)$, the estimates can be recovered as follows: $E[y|\mathbf{x}] = exp(\mathbf{x}\hat{\beta} + 0.5\hat{\sigma}_\epsilon^2)$ where $\hat{\sigma}$ is the usual unbiased estimator of σ. But if the errors are not normally distributed, as it is in our case according to the Jarque-Bera test for normality, then Duan's smearing estimator is used. These estimates are recovered as: $E[y|\mathbf{x}] = exp(\mathbf{x}\hat{\beta}) * (N^{-1}\sum exp(\epsilon_i))$ where ϵ_i is the i^{th} residual from the model and $N^{-1}\sum_i exp(\epsilon)$ is the smearing factor (Duan 1983).

Once we have calculated the scores, we can calculate the maximal productivity advantage that a practice would have gained from care continuity if it had used a specific proportion of its overall consultations as its care continuity consultation budget. The maximal number $total\_days\_gained_p$ that practice p can gain, relative to offering no continuity of care, is the sum of consultation-level $days\_gained_i$ achieved by allocating regular doctors to consultations in rank order of $days\_gained_i$, from highest to lowest, until the practice runs out of its continuity of care consultation budget.

To apply this to our data, let C_p denote the set of recorded consultations i that took place in practice p, and let N_p be the total number of these consultations. Then for all consultations $i_p \in C_p$, we can sort the $days\_gained_{i_p}$ in descending order and denote this $\text{sort}(days\_gained_{i_p})^{(a)}$ where a corresponds to the rank once the consultations are sorted using this sorting scheme. Next, let $x_p \in [0\%, 100\%]$ specify the percent of continuity provided in practice p. Then the total days gained by offering $x_p\%$ continuity to the most productivity-enhancing patients at practice p can be calculated as

$$total\_days\_gained_p^{[x_p]} = \sum_{a, a\in[0, N_p x_p]} \text{sort}(days\_gained_{i_p})^{(a)}.$$

Change in practice demand by re-allocating continuity

First, note that an extension of the average revisit interval by a factor r leads to a demand reduction of $r/(1+r)$.[14] Therefore, to estimate the effect of continuity on practice demand we can equivalently focus on estimating the impact on the revisit intervals (i.e., estimating the value of r).

To identify r, we start by observing that $\sum_{i_p\in C_p} R\hat{I}_{0i_p}$ gives the sum over the predicted revisit intervals of all consultations C_p that occurred at practice p under the assumption that *no* patient saw their regular doctor. Based on the results in the main paper, we should expect intuitively that as more patients see their regular doctor these revisit intervals should increase. For example, $\sum_{i_p\in C_p} R\hat{I}_{0i_p} + total\_days\_gained_p^{[x_p]}$ gives the sum of predicted revisit intervals assuming instead that the $x_p\%$ most productivity-enhancing patients at practice p saw their regular doctor, while all other patients saw a transactional provider. Meanwhile, $\sum_{i_p\in C_p} RI_{i_p}$ (i.e., the sum of the revisit intervals as given in the raw data) gives the sum of the revisit intervals under the status quo, i.e., assuming no change in the percentage of consultations between patients and their regular doctor and no change in the allocation of continuity to consultations. Combining these two measures, we can estimate the percentage increase in the revisit interval if continuity at a practice were retained at the same level, which we denote x_p^*, but if continuity were instead re-allocated from those patients who actually received it to those patients who stood to benefit from it the most. This is given by the following expression:

$$\frac{\sum_{i_p\in C_p} R\hat{I}_{0i_p} + total\_days\_gained_p^{[x_p^*]}}{\sum_{i_p\in C_p} RI_{i_p}}$$

Subtracting 1 from this expression gives us the factor r, i.e., the increase in the revisit interval by better targeting care continuity. Observe that we should expect numerator to be greater than the denominator

[14] Assuming that each patient i has a demand for x_i consultations over a period $[0, T]$, an extension of all revisit intervals by a factor r extends this demand to x_i consultations over a period $[0, (1+r)T]$ or, equivalently, to a demand of $x_i/(1+r)$ consultations over the period $[0, T]$. Demand is therefore reduced by a factor $r/(1+r)$ to the fraction $1/(1+r)$ of the original demand.

(i.e., $r > 0$), since in both the numerator and denominator we offer continuity to the same proportion of consultations ($x_p^*\%$) but the consultations in the numerator are allocated to the most productivity-enhancing patients.

Using r as calculated above, we can then estimate for each practice the reduction in demand that would have been expected from better allocating continuity while keeping continuity-levels unchanged. The productivity-gains for each practice are plotted in Figure 3a in the main paper.

Change in system demand by re-allocating and increasing continuity

While in the above analysis we have focused on re-allocating continuity within a practice, we can also ask what would happen if practices increased the level of continuity provided *above* their current level x_p^*. To do this, we start by estimating the total reduction in system demand if all practices had re-allocated continuity to the most productivity-enhancing patients. This is given by the expression:

$$\frac{\sum_p \left(\sum_{i_p \in C_p} R\hat{I}_{0 i_p} + total\_days\_gained_p^{[x_p^*]}\right)}{\sum_p \sum_{i_p \in C_p} RI_{i_p}}$$

This is the baseline result if all practices kept continuity at their current levels x_p^*. Now let P_x denote the set of practices with current continuity levels below $x\%$, i.e., with $x_p^* < x$. Then we ask what would have happened to demand if all practices with $x_p^* < x$ had increased continuity levels to x *and* had re-allocated continuity to the $x\%$ most productivity-enhancing patients. (Note that we will assume that all practices with $x_p^* \geq x$ keep continuity at the current level x_p^* but also improve targeting of continuity). For a given minimum continuity level x, the total reduction in system demand is then given by the expression

$$\frac{\sum_{p \in P_x} \left(\sum_{i_p \in C_p} R\hat{I}_{0 i_p} + total\_days\_gained_p^{[x]}\right) + \sum_{p \notin P_x} \left(\sum_{i_p \in C_p} R\hat{I}_{0 i_p} + total\_days\_gained_p^{[x_p^*]}\right)}{\sum_p \sum_{i_p \in C_p} RI_{i_p}}$$

We can estimate this for all values $x \in [0\%, 100\%]$. This is shown in Figure 3b in the main paper.

EC.15. Continuity of care and other outcomes

There are various studies that look at the effect of continuity of care on healthcare utilization, such as prescription rates, referral rates and ED visits as described in Section 2.2. To complement this line of research and to ensure that a push for productivity improvement is not at the cost of health outcomes, we investigate the effect of seeing the patient's regular provider on such health care resources for this cohort of patients. Specifically, we look at (1) the probability of ED admission within 1 day, 3 days and 7 days from the focal consultation (columns 2,3 and 4 in Table EC.22) (2) prescription rates (column 5 in Table EC.22) and (3) referral rates (column 6 in Table EC.22). For each of these models, we use the different measures as dependent variable in a control function model as specified in Equation 2.

For all the above outcomes, we find that the effect sizes are very small, but providing continuity leads to a reduction in ED visits rates and referral rates, and a slight increase in prescription rates. While ED referrals and prescription rates are likely to be markers of better health outcomes, the reduced referral rate is more difficult to interpret. We believe a regular doctors who is more familiar with the patient is less likely to feel the need to practice defensive medicine, which leads to over-referrals. However, we do not have the data on the outcomes of the outpatient appointments to validate this conjecture.

Table EC.22 Coefficient of RD_i across various outcomes.

Dependent variable=	ED visit rate (1-day)	ED visit rate (3-days)	ED visit rate (7-days)	Prescription rate	Referral rate
$RD_i = 1$	-0.003***	-0.005***	-0.005***	0.016***	-0.008***
	[-0.004,-0.003]	[-0.006,-0.004]	[-0.006,-0.004]	[0.012,0.019]	[-0.010,-0.007]
Observations	11,344,065	11,344,065	11,344,065	11,344,065	11,344,065

Notes: 95% confidence intervals in square brackets;

EC.16. Additional robustness checks

EC.16.1. Doctors who may leave or retire from the practice

We also run additional analysis to account for the fact that doctors may leave or retire from the practice during the study period. Specifically, we first identify if and when a doctor retires or leaves the practice by looking at the date the doctor was last present in the data. Next, for all the patients who had such doctors as their regular doctor, we remove all the consultations of such patients that take place in the two years following the leaving or retirement of their regular doctor. Lastly, we reintroduce the patient into the sample after this two year period so that there has been enough time for the patient to readjust her preferences. The results are reported in Table EC.23 and we find that the results are qualitatively similar.

Table EC.23 Coefficient of RD_i for a model accounting for physicians leaving the practice.

Dependent Variable = Natural logarithm of the revisit interval	
	Exclusion of consultations by doctors who leave the practice
RD_i=1	0.154***
	[0.153,0.156]
Observations	11,186,978

Notes: 95% confidence intervals in square brackets;

EC.16.2. Different time frames to establish a regular provider

We establish robustness of our results by considering a 1 and 3-year dynamic rolling window for the identification of the regular doctor rather than a 2-year dynamic window.

Table EC.24 Coefficient of RD_i across two different time periods to establish the regular provider (i) one year , (ii) three years.

Dependent Variable = Natural logarithm of the revisit interval		
	One year	Three years
RD_i=1	0.156***	0.157***
	[0.155,0.158]	[0.155,0.158]
Observations	9,817,679	11,452,414

Notes: 95% confidence intervals in square brackets;

EC.16.3. Analysis with only established providers

We establish robustness of our results by considering only consultations with established doctors.The coefficient of $RD_i = 1$ remains positive and significant.

Table EC.25 Coefficient of RD_i for a model with established providers only.

Dependent Variable = Natural logarithm of the revisit interval	
	Established providers only
RD_i=1	0.134*** [0.132,0.136]
Observations	10,400,162

Notes: 95% confidence intervals in square brackets;

EC.16.4. Exclusion of weekends from the data

As a robustness test, we exclude weekends from the data as they might indicate anomalies, and find that the results remain qualitatively similar.

Table EC.26 Coefficient of RD_i for a model without weekends.

Dependent Variable = Natural logarithm of the revisit interval	
	Excluding weekends
RD_i=1	0.130*** [0.128,0.131]
Observations	10,946,888

Notes: 95% confidence intervals in square brackets;

EC.16.5. Non-linear effect of mental health comorbidities

In order to capture potential non-linear effects of the number of comorbidities, we substitute the binary mental health indicator with the number of mental health comorbidities. This ranges from 0 to 3 (depression, anxiety and schizophrenia). In table EC.27, we report results for the mental health number as a (i) linear control, (ii) quadratic control and (iii) moderator. We find that even though the coefficient of the mental health comorbidities is significant, the effect size is small, and therefore the coefficient of $RD_i = 1$ remains qualitatively similar for all three models.

Table EC.27 Coefficient of RD_i across mental health comorbidity as a (i) linear control, (ii) quadratic control and (iii) moderator.

Dependent Variable = Natural logarithm of the revisit interval			
	Linear control	Quadratic control	Moderator
RD_i=1	0.127*** [0.126,0.129]	0.127*** [0.126,0.129]	0.121*** [0.120,0.124]
Number of MH comorbidities	-0.003*** [0-.005,-0.001]	0.011*** [0.007,0.016]	-0.0126*** [-0.015,-0.010]
Number of MH comorbidities^2		-0.009*** [-0.011,-0.007]	
RD_i*Number of MH comorbidities			0.0180*** [0.0150,0.020]
Observations	11,344,065	11,344,065	11,344,065

Notes: 95% confidence intervals in square brackets;

EC.16.6. Additional controls for seasonality

Instead of using fixed effects for the month of the year, we substitute these with season fixed effects as a robustness test (Spring: March, April, May; Summer: June, July, August; Autumn: September, October, November; Winter: December, January, February). The results are reported in Table EC.28 and we find that the coefficient of $RD_i = 1$ remains qualitatively similar.

Table EC.28 Coefficient of RD_i for a model with season fixed effects.

Dependent Variable = Natural logarithm of the revisit interval	
	Season-FE
RD_i=1	0.127*** [0.125,0.129]
Observations	11,344,065

Notes: 95% confidence intervals in square brackets;

EC.16.7. Effect of phone consultations on the revisit interval

We conduct additional analysis to check whether non-face to face visits in between consultations drive the extension of the revisit interval. We first define a binary variable $PhoneConsult$ which takes a value of 1 if a phone consultation took place between two face to face visits, and 0 if not.

Next, we estimate a model with $PhoneConsult$ as the dependent variable and RD_i as the main independent variable, along with all the controls as defined in the previous models. Results show that the probability of a phone consultation following a face-to-face visit is 5.50% (95% CI [5.47%,5.52%]) when a patient does not see his regular doctor, compared to 5.52% (95% CI [5.51%,5.55%]) when a patient sees his regular doctor, suggesting that seeing the regular doctor could increase the probability of a phone consultation, but the effect is very small.

To ensure that the phone consultation is not driving the extension of the revisit interval, we estimate the main model as specified in Equation 1 along with $PhoneConsult$ as an additional control variable. As expected, we find that effect size of RD_i does not change, suggesting that the phone consultations following the face to face visit is not driving the extension of the revisit interval.

Table EC.29 Coefficient of RD_i across two models using (i) $PhoneConsult$ as the dependent variable and (ii) revisit interval as the dependent variable with $PhoneConsult$ as an additional control variable.

Dependent Variable =	PhoneConsult	Natural logarithm of the revisit interval
RD_i=1	0.000* [0.000,0.000]	0.157*** [0.155,0.159]
Observations	11,344,065	11,344,065

Notes: 95% confidence intervals in square brackets

References for the e-companion

Baum CF, Schaffer ME, Stillman S (2002) IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation. Statistical Software Components, Boston College Department of Economics.

Duan N (1983) Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association* 78(383):605–610.

Gulliford MC, Sun X, Anjuman T, Yelland E, Murray-Thomas T (2020) Comparison of antibiotic prescribing records in two uk primary care electronic health record systems: cohort study using cprd gold and cprd aurum databases. *BMJ open* 10(6):e038767.

Hill AP, Freeman GK (2011) Promoting continuity of care in general practice. *London: Royal College of General Practitioners* .

Palmer W, Hemmings N, Rosen R, Keeble E, Williams S, Imison C (2018) Improving access and continuity in general practice. *Research Summary* .

Payne RA, Mendonca SC, Elliott MN, Saunders CL, Edwards DA, Marshall M, Roland M (2020) Development and validation of the cambridge multimorbidity score. *CMAJ* 192(5):E107–E114.

Sanderson E, Windmeijer F (2016) A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of econometrics* 190(2):212–221.

Schwartz LM, Woloshin S, Wasson JH, Renfrew RA, Welch HG (1999) Setting the revisit interval in primary care. *Journal of General Internal Medicine* 14(4):230–235.

Stock JH, Yogo M, et al. (2005) Testing for weak instruments in linear iv regression. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* 80(4.2):1.