Online versus In-Person Services: Effects on Patients and Providers *(Preliminary and incomplete)* *

Amanda Dahlstrand (*Zurich*) Nestor Le Nestour (*Stockholm*)
Guy Michaels (*LSE*)

September 15, 2023

Abstract

Online delivery of one-to-one services offers potential to increase convenience, decrease costs, and reduce inequalities in users' access. But we have relatively little evidence on how the delivery mode impacts both providers and consumers. This paper focuses on online delivery of healthcare services and specifically – primary care doctor consultations. To study this we use new data from Sweden and effectively random assignment of patients to nurses with different propensities to refer patients to online versus in-person doctor consultations. We find that while online consultations are delivered sooner and are shorter, they yield similar in-meeting outcomes, including rates of diagnosis, prescription, specialist referral, patient satisfaction, and 30-day post-consultation avoidable hospitalizations. Online consultations are, however, followed by more visits to emergency departments and in-person primary care consultations. Nevertheless, patients' medium-run outcomes do not differ significantly after online consultations. Adding the costs of increased in-person follow ups, online visits offer modest overall cost savings.

KEYWORDS: Telehealth, Remote work, Online services
JEL CLASSIFICATION: J44, I11, O33

*We thank Ellen Donath for outstanding research assistance, and helpful comments from many colleagues, including seminar participants at Bergen, Chicago Booth, Harvard-MIT, LSE-CEP, LSE Health Policy, and NYU Stern. We thank the healthcare company for allowing us to study their data. We gratefully acknowledge financial support from the British Academy Leverhulme Research Grant SRG2223 230212, the University of Zurich's Research Priority Program "Equality of Opportunity", and the ESRC's Centre for Economic Performance. Institutional Review Board approvals: Stockholm and Swedish ethics review board.

1 Introduction

In a hybrid world, many decisions are made on which services to provide online or in person, and to whom. Such decisions are important for providers, including firms and governments, as well as to workers and users. While online offers potential benefits, a concern is that despite technological advances, the anatomy of one-to-one meetings may still differ through a screen. Consequently, switching services online may have implications for costs, quality, and worker and user experiences. But there is limited evidence from direct head-to-head comparisons of in-person and online customer-facing services. To better inform decision makers, we need to know more about the trade-offs in the mode of service delivery.

The choice of delivery mode is particularly important in healthcare. For providers, online meetings offer potential productivity improvements, which are urgently needed due to rising costs in aging societies, while for patients, online offers convenience, around-the-clock access, and time savings. Online services may also reduce gaps in access to healthcare from rural and urban areas, or rich and poor ones (Dahlstrand, 2023). Key to healthcare delivery in many countries are patient consultations with primary care physicians (PCPs), also known as general practitioners (GPs). In this paper we provide evidence on the consequences of switching doctor consultations from in person to online for a variety of patient and doctor outcomes and costs.

To do so, we assemble new data on individual consultations from Sweden, whose national health insurance covers both public and private providers. The provider whose data we study is Europe's largest digital healthcare firm. Since 2019 it provides patients who chose to register with it all their primary care, including both in-person and online doctor consultations. Our data cover both consultation types and are matched to panel data on the patients themselves, including demographics, socioeconomic characteristics, and health outcomes.

Our analysis focuses on registered patients who request doctor consultations. After these patients briefly meet a nurse online, this nurse decides whether to direct them to a doctor consultation, and if so, whether this consultation should be online or in person. To use the variation arising from the nurses' decision, we begin by estimating OLS regressions of the consequences of nurses directing patients online versus in person, controlling for time and location fixed effects, and in some specifications a rich set of other potential confounders.

Since the sorting of patients across delivery modes (online versus in person) may depend in part on factors that we cannot observe, we also develop an instrumental variable strategy. Our instrument is the share of patients that each nurse referred online among all the patients they referred to doctor consultations when meeting other patients. To demonstrate that the instrument is valid, conditional on our controls, we begin by showing that the first stage is large and precisely estimated. To show that the instrument satisfies the independence assumption, we develop an econometric framework that applies to our setting, and demonstrate that: (i) the characteristics of patients are uncorrelated with the instrument; (ii) nurses' propensity to direct to any doctor consultation is uncorrelated with their propensity to direct online; and (iii) the characteristics of patients who are directed to doctor consultations are also uncorrelated with the instrument. To demonstrate that the instrument satisfies average monotonicity and average exclusion (Frandsen et al., 2023), we show that (i) institutional rules tightly circumscribe nurses' actions, limiting their potential to affect patient outcomes other than by directing them to doctors; (ii) nurses' meetings with patients last on average less than five minutes, leaving little time for anything except a brief exchange to inform the directing of patients; (iii) nurses who refer more online are not differentially likely to make rare mistakes (Chan et al., 2022); and (iv) the first stage is large, positive, and precisely estimated for patients with different demographics and health conditions, suggesting that most patients comply with the instrument.

Our instrumental variable estimates show that compared to in-person doctor consultations, online ones take place sooner after patient's request, are shorter overall, with much shorter patient-facing time but longer administrative time for the doctor to e.g., fill in details, such as prescriptions and notes, after meeting the patient. Despite the difference in duration and timing, online consultations have similar in-meeting outcomes, including rates of diagnosis, prescription (and prescription collection), specialist referral, and patient satisfaction. Looking within 30 days after the consultation, we see similar rates of (very rare) avoidable hospitalizations after online consultations. Overall hospitalizations may be higher after online consultations, although these estimates are imprecise. Two significant differences do arise, however: first, online consultations are more likely to be followed by Emergency Department (ED, also known as accident and emergency - A&E) visits within 30 days; and second, they are more likely to be followed by additional primary care consultations within

30 days, most of which are in-person consultations booked by the doctor. Looking over more than 30 days after the consultation (in some cases up to around a year), we find no significant differences in healthcare outcomes between online and in-person consultations.

Compared to our OLS estimates, our IV estimates generally show lower cost savings from online, suggesting that on average sicker patients tend to be triaged to in-person consultations, and IV overcomes this sorting problem to give a more realistic assessment of respective costs. Nevertheless, our findings still suggest that online consultations offer a modest cost saving to providers or insurers, even when we account for the costs from subsequent ED visits and primary care consultations. Moreover, since the patients we study are city dwellers, they live closer to ED hospitals and primary care clinics than most Swedes.[1] Consequently, the patients we study may be more likely to seek additional in person care than the average patient, and adjusting our estimates for this yields larger cost savings from online consultations. Online consultations also offer time savings for patients on travel and waiting, which we also assess and are part of the societal benefits of online consultations. Other benefits, such as convenience, around-the clock access, and more equitable healthcare delivery across locations and income groups are harder to quantify.

Our findings contribute to the recent literature on the shift to hybrid work, which is an important development in labor markets in the last few years (Bloom et al., 2015; Aksoy et al., 2022; Bloom et al., 2022; Goodman et al., 2019; Ertem et al., 2021). We contribute to this literature by studying in detail the trade-offs between online and in-person 1:1 service provision.[2] Current work on remote or hybrid work has mostly studied settings where the mode (online vs. in person) only changes for the workers (Bloom et al. 2015, Emmanuel and Harrington 2023, Emmanuel, Harrington and Pallais 2023), but there is no change in mode for the customer or client. Also changing the mode for the client can have important implications. The exception is the studies on online teaching, where students have also been switching mode during covid (see e.g. Jack, Halloran, Okun and Oster 2022). However, the forced switching of mode during lockdown may have completely different implications

[1] This is because we study the patients who registered with the firm's first in-person clinics, which opened in cities.

[2] By 1:1 service provision, we mean meetings between one service provider and one user. Examples are banking/financial advice, tutoring, mental health therapy, legal advice, and healthcare consultations.

than other forms of voluntary online services. Moreover, settings where collaboration or peer effects are central (Emmanuel, Harrington and Pallais 2023; Agostinello et al. 2022) will likely find more negative effects from online work than settings in which there is 1:1 service provision. Carlana and La Ferrarra (2021) have a setting with online 1:1 service provision (online remedial tutoring) and find positive effects, but they do not have an in-person comparison group. To the best of our knowledge, we provide the first paper studying the effects of online vs. in person 1:1 services.

There is also a nascent literature on telehealth (Greene, 2022; Hatef et al., 2022; Powell et al., 2017; Kletečka-Pulker et al., 2021; Rodler et al., 2020; Slavova-Azmanova et al., 2021; Ladin et al., 2021; Santarossa et al., 2018; Dingel et al., 2023). We add to this by studying the tradeoffs between online and in-person using a large-scale real-world setting with a rich set of controls and a strategy to overcome selection across delivery modes on based unobserved factors.

Perhaps more closely related to ours are recent studies of changing access to or price of online care (Zeltzer et al., 2023; Ellegård et al., 2021; Rabideau and Eisenberg, 2022). Our paper differ from these by studying a setting where assignment to online versus in person occurs *after* patients sort into care. This allows us to shed more light on the respective effects of these modes of delivery, without concerning ourselves about different patient (or the same patients with different symptoms) sorting into care when they anticipate either online or in-person consultations.

Methodologically, we build on the literature using expert propensities as instruments (Kling, 2006; Doyle Jr, 2007; Anwar et al., 2012; Dahl et al., 2014; Aizer and Doyle Jr, 2015; Dobbie et al., 2018; Bhuller et al., 2020; Bakx et al., 2020; Chan et al., 2022; Frandsen et al., 2023). We apply this approach to a different research question, namely assessing the impact of online consultations. Our work is also related to the literature on IV with multi-valued treatments, which are typically either ordered (Angrist and Imbens, 1995; Heckman and Urzua, 2010) or unordered (Lee and Salanié, 2018; Mountjoy, 2022). Our paper differs in focusing on a partially ordered model, where a doctor consultation ranks above no consultation, but the trade-offs between in person and online may differ across patients.

Finally, our paper is related to the literature on the importance of cities as loci of face-to-face interactions even as communication technology improves (e.g., Gaspar and Glaeser (1998) and Michaels et al. (2019)). In our setting at least one-third of online consultations are followed up in person, so centrally located patients retain an advantage in a hybrid world. This advantage also limits the extent to which online consultations can *completely* eliminate spatial inequalities in healthcare access, even though it can reduce them.

The remainder of our paper is organized as follows. Section 2 presents the background, including the institutional setting we study; Section 3 presents our econometric model; Section 4 explains our data sources and our dataset construction; Section 5 presents the research design and empirical findings; and Section 6 concludes.

2 Institutional background

Assessing the impact of online consultations relative to in-person ones involves overcoming two main challenges. First, in many settings doctor consultations are either only in person, as in most countries until a few years ago, or only online, as in some countries during the Covid pandemic.[3] To compare online versus in-person, we need to observe patients across both types of delivery modes. Second, in settings with both online and in-person consultation, sorting of patients is a concern. When faced with a change in the relative price or convenience of online consultations, for example, different patients may schedule appointments for different symptoms, and healthcare providers may also sort patients across delivery modes based on their own criteria. To make progress, we need a setting that allows us to overcome this selection problem.

Our setting is helpful in addressing both of these issues. To observe patients across both consultation modes, we focus on a Swedish firm, which started as a digital service in 2016. Since 2019, this firm also set up clinics for in-person consultations for patients who chose this firm as their primary care provider under the national health insurance, but the online-first model remained the firm's strategy.[4] Here we use novel data on online and in-person doctor consultations for patients who were registered in four clinics: one which opened in Lund in September 2019, and three which opened in the Stockholm area since September 2020.

To resolve the sorting of patients, we use key aspects of the institutional setting in the firm. Figure 1 illustrates the flow of patients who were registered in those

[3]Sweden is an exception for a rich country, as it maintained a combination of both even during the worst phases of the Covid pandemic.

[4]Primary care provision is publicly funded in Sweden with a mix of public and private provision. Patients can choose a clinic, but most patients choose one or default to one which is close to their home. For patients who are registered at a certain clinic, the national health insurance pays mostly through capitation, but in some regions also with some fee for service.

clinics when they requested doctor consultations. The patient's request was submitted via a mobile phone application, at which point an algorithm determined whether a doctor consultation was immediately available, taking into account the symptoms the patient entered and the current waiting time for doctors. In most cases, the algorithm assigned patients directly to an online doctor consultation, but in others the patients were directed to the next available online nurse – these nurses were based anywhere in the country. The online nurse then had to make two quick sequential decisions.[5] First, the nurse had to decide whether to resolve the case without a doctor, or whether a doctor consultation is needed.[6] Second, if the nurse decided that a doctor consultation was needed, the nurse went on to decide whether to book one in person or online. As we explain in Section 3 below, this setting allows us to use variation across nurses in the propensity to direct patients across the two doctor consultation delivery modes.

Before we proceed to the model, however, it is useful to note a few more points about the setting. First, doctors were paid a rate for each shift they work (effectively an hourly rate), and they work from home when online and from clinics when in person. Second, the service is covered by universal health insurance, with small co-pay. Third, the application lets doctor and patient see each other, so it differs from a typical phone conversation without video. Fourth, we study a broad set of patients with a wide range of ages and conditions, who chose primary care provider with an online option; we discuss their representativeness of the broader population below. Fifth, given the data restrictions (our sample ends in December 2020), we focus primarily on short-run outcomes, although we use the fact that we observe some patients, mainly from Lund, for a longer time to study their medium-run outcomes. Finally, the treatment we study is bundled with the identity of individual doctors who chose to work online versus in person. We note that this would have been the case even if we had randomly assigned patients online versus in person, as long as we had not manipulated the work assignment of doctors. At the same time, as we discuss in the data section, almost all the doctors we study worked at least some of the time online, since online remains the firm's core business, and the only option for

[5]The nurses' job is to primarily triage, i.e., sort patients to consultations. If a nurse decides (based on the patient's symptoms) that there is a need for a doctor consultation, the nurse should not give advice but rather direct to a doctor as soon as possible, to minimize unnecessary repetitions.

[6]In some cases, the nurse decides that there is no need for a doctor consultation, as the patient's symptoms are not severe or in need of treatment. In those cases, the nurse can provide some advice about self care and resolve the case.

non-registered patients, who are the majority the firm treats.

3 Data

This section briefly outlines the data sources we use and the way we construct our dataset, leaving the details to the data appendix.

3.1 Data sources

Our starting point is a dataset covering the roughly 1.8 million primary care meetings in a large healthcare provider in Sweden during the 24 months spanning 2019-2020. These include doctor consultations and nurse meetings, both in person and online. We have matched these with data from Statistics Sweden and the Swedish National Board of Health and Welfare from 2013-2020, covering three main data aspects. First, the matched data cover healthcare provision outside the firm, including inpatient and outpatient care as well as prescriptions and their collection. Second, the data contain demographic information, including age, gender, education, and immigration status. Finally, the data contain socioeconomic information, such as earnings and receipt of benefits.

3.2 Dataset construction

As discussed in Section 2, we focus our analysis on doctor consultations booked by online nurses. Within these, we focus on patients who were registered at a center that was open at the time of the nurse meeting, and could have been directed to either an in-person or an online doctor consultation.

To do so, we first restrict our sample of meetings to over 105,000 (primary care) doctor consultations, which were booked rather than direct drop-ins. Appendix Table A1 shows the subsequent sample restrictions that we impose. First, we keep doctor consultations booked by online nurses.[7] Second, we keep consultations with patients who were registered at one the firm's clinics, which had an option for in-person consultations. Third, we exclude patients who called due to Covid (for which there was no treatment) or chlamydia (where typically nurses prescribe a home test) or breastfeeding (where nurses may direct to a breastfeeding consultant). Fourth, we

[7]See the Data appendix for how we constructed the matching between nurse meetings and doctor consultations.

restrict to the time period when the in-person clinic had opened, and thus patients were "at risk" to have an in-person meeting. Finally, we keep only consultations booked by nurses who were observed at least 20 times in the remaining sample.

The remaining sample includes 8907 online meetings of patients with 62 nurses in the sample that we refer to as the "Nurse meeting sample" (or "Nurse sample" in brief), which includes the cases where nurses decide whether to direct to doctor consultations (see Figure 1). Roughly half of these meetings – about 4662 – were directed by nurses to doctor consultations, a sample that we refer to as the "Doctor consultation sample" (or "Doctor sample" in brief). Within this doctors' sample, roughly 57% of doctor consultations were in person and the rest online. These consultations were conducted by 400 doctors, of which 338 appear within our sample as working only online; 38 work both in person and online; and 24 in only in person.[8]

The definitions of the main variables that we use in the paper are shown in Appendix Table A2 and summary key statistics are reported in Table A3.

To help assess the validity of our main (doctor) sample, Table 1 compares the patients we study (column 1) compare to the population in the same municipalities (column 2) and to the Swedish population more generally (column 3). In some respects (including gender and mean income) our sample is quite representative. The patients we study are, however, more likely to be in large cities (a consequence of their being listed in centers open in big cities), a bit younger, better educated, more likely to have an immigrant background, and less likely to be married. These differences notwithstanding, our sample has a broad representation of different segments of Swedish society.

4 Model

This section outlines an econometric model of the assignment of patients to online and in-person doctor consultations. The model illustrates the problem of sorting of patients into consultations on characteristics that are – and are not – observed to the econometrician. It also justifies our use of nurses' propensities to direct to

[8] Since the firm's core business is online provision, however, almost all the doctors who work for it have some online experience. Therefore, of the 24 above-mentioned doctors, at least 17 had worked online at least at some point in 2019-2020. These doctors who consulted patients without nurse redirection or patients who had not registered with this firm as their in-person primary care provider.

online doctors in all but the current meeting as instrumental variables, following the literature on expert propensities, and especially the recent work by Frandsen et al. (2023). We differ from existing work by presenting a semi-ordered model of instrumental variables, where there is an ordered decision (a doctor consultation is a more intensive treatment than having no doctor consultation) and an unordered one (some patients may prefer online consultations and others may prefer in-person consultations). This section begins with an outline of our model, after which we explain how we take the model to the data.

4.1 The model assumptions

As was outlined in Section 2 and Figure 1, we consider patients who request primary care consultations using the firm's application. We focus on patients who are registered with this provider as their main primary care clinic, whose in-person center is open and who, due to temporary congestion, the firm's algorithm directs to online nurse meetings. We assume – and later verify – that this assignment to the next available nurse is conditionally random. We focus on patients (indexed by i), who are assigned to nurses (indexed by j), where the number of patients that nurse j sees (N_j) is relatively large (>20). Each nurse briefly assesses each patient and makes two sequential decisions. First, whether to direct the patient to a doctor consultation, and second, if the nurse does direct the patient to a doctor consultation, whether this doctor consultation should be in person or online.

We assume that each person is characterized by their sickness, θ_i, which is causing them to request a doctor consultation; a vector of observable pre-determined characteristics ψ_i; and a measure of how keen they are to consult with a doctor, ϕ_i. We assume that $\phi_i = \theta_i + g(\psi_i) + \eta_i$, where η_i is mean 0 independent and identically distributed (i.i.d.) noise.

We also assume that the patient has a preference τ_i for an in-person (versus online) doctor consultation, such that $\tau_i = 1$ denotes indifference between in person and online. We are agnostic about the relationship between τ_i and the other patient parameters. For example, a sick patient with mobility difficulties may expect a higher payoff from an in-person consultation, but also a higher cost.

We model the patient's utility as

$$U_i = \begin{cases} \phi_i & D_{ij}^0 = 1, D_{ij} = 1 \\ \phi_i \tau_i & D_{ij}^0 = 1, D_{ij} = 0 \\ 0 & D_{ij}^0 = 0, \end{cases} \tag{1}$$

where D_{ij}^0 is an indicator for patient i being directed to any doctor consultation (after meeting nurse j); and D_{ij} is an indicator for patient i being directed to an online doctor consultation, as opposed to an in-person one (after meeting nurse j). We note that D_{ij} is only defined for patients for whom $D_{ij}^0 = 1$.

We also define $D_i(j)$ as the potential potential treatment of patient i meeting nurse j, and $\mathbf{Y}_i(d, j)$ as the vector of potential outcomes of i meeting nurse j, where d is an indicator for online versus in-person. The vector of outcomes for patient i who met nurse j can be written as $\mathbf{Y}_{ij} = \mathbf{Y}_i(1, j) D_{ij} + \mathbf{Y}_i(0, j)(1 - D_{ij})$.

Turning to the nurses, we assume that they all make the decision of whether to direct a patient to a doctor consultation (vs no consultation) based on patient sickness, but nurses differ in their assessment of whether online or in-person doctor consultations are generally preferable. When a nurse decides that a doctor consultation is needed, the nurse balances their preferences for online versus in-person consultations with those of patient. Specifically, we define ρ_j as the preference of nurse j to refer patients online, where $\rho_j > 0$ varies across nurses, so $\rho_j \neq \rho_{j'}$ for some j, j'. We also define θ_{ij} as nurse j's assessment of patient i's sickness, where $\theta_{ij} = \theta_i + \eta_{ij}$, and η_{ij} is mean zero i.i.d. noise.

We define the utility of nurse j, who meets patient i, as

$$\widetilde{U}_j = \begin{cases} 1_{\theta_{ij} > 0} & D_{ij}^0 = 1, D_{ij} = 1 \\ \frac{\tau_i}{\rho_j} 1_{\theta_{ij} > 0} & D_{ij}^0 = 1, D_{ij} = 0 \\ 1_{\theta_{ij} \leq 0} & D_{ij}^0 = 0. \end{cases} \tag{2}$$

Since the nurse decides the outcome of the meeting, the patient's treatment is an online doctor consultation ($D_{ij}^0 = 1, D_{ij} = 1$) when $\theta_{ij} > 0$ and $\tau_i \leq \rho_j$; an in-person doctor consultation ($D_{ij}^0 = 1, D_{ij} = 0$) when $\theta_{ij} > 0$ and $\tau_i > \rho_j$; and no doctor consultation ($D_{ij}^0 = 0$) when $\theta_{ij} \leq 0$.[9]

4.2 Identification in the model

Panel A of appendix Figure A1 illustrates the treatment of patient i when nurses make no mistakes about the sickness of the patient ($Var(\eta_{ij}) = 0$). Here only patients

[9]Without loss of generality, we assume that nurses break ties between online and in person by assigning patients online.

with $\theta_i > 0$ receive a doctor consultation. Among those who consult doctors, there may be three types of patients. Patients with very strong preferences for in person ($\tau_i > \max(\rho_j)$) always consult in person, and those with very strong preferences for online ($\tau_i \leq \min(\rho_j)$) always consult online. Patients whose preferences for online versus in person are intermediate (($\min(\rho_j) < \tau_i \leq \max(\rho_j)$) are compliers – the mode of consultation is determined by the nurse to whom they are (conditionally) randomly assigned. Panel B of appendix Figure A1 shows that when nurses make mistakes ($Var(\eta_{ij}) \neq 0$) the situation is similar, except that some patients who should have consulted a doctor do not, while some who should not, do have a consultation.

Like other papers on expert propensities, we cannot observe ρ_j, so we define nurse j's propensity to refer online (conditional on referring to a doctor consultation) as $\widehat{\pi}_j \equiv \frac{\sum_{i'=1,\dots,N_j} D_{i'j}}{\sum_{i'=1,\dots,N_j} D^0_{i'j}}$. We similarly define the instrument as nurse j's propensity to refer online, leaving out patient i's meeting: $\widehat{\pi}_{ij} \equiv \frac{\sum_{i'=1,\dots,N_j;i'\neq i} D_{i'j}}{\sum_{i'=1,\dots,N_j;i'\neq i} D^0_{i'j}}$.

To use $\widehat{\pi}_{ij}$ as an instrument for D_{ij}, we need to verify the identification assumptions outlined by Frandsen et al. (2023). To satisfy the first stage, we require sufficient variation across nurses in ρ_j.

To satisfy independence in the doctor consultation sample we rely on the (conditional) random assignment of patients to nurses and the orthogonality of nurse errors to nurse propensities to online. This allows us to write:

$$\widehat{\pi}_{ij} \perp \left\{\mathbf{Y}_i(d,j), D_i(j) | D^0_{ij} = 1\right\}$$
$$\widehat{\pi}_{ij} \perp \theta_i, \eta_{ij}, \{\mathbf{Y}_i(d,j), D_i(j)\} \Rightarrow \widehat{\pi}_{ij} \perp \{\mathbf{Y}_i(d,j), D_i(j) | \theta_i + \eta_{ij} > 0\}$$

Average exclusion and average monotonicity (Frandsen et al., 2023) are weaker versions of exclusion and monotonicity, which allow for identification in settings using expert propensities as instruments. Average monotonicity requires that the covariance between each patient's nurse-specific treatment status and nurse overall treatment propensities is weakly positive. The model above assumes strict monotonicity.

Average exclusion assumes that the direct effects of nurses on patient outcomes are uncorrelated with treatment propensity. In our setting, strict exclusion holds if $Y_i(d,j) = Y_i(d)$.

4.3 Taking the model to the data

To verify that the model's assumptions apply in our setting, we begin by testing the first stage assumption by regressing D_{ij} on $\widehat{\pi}_{ij}$ in the doctors' sample ($D^0_{ij} = 1$).

Verifying the applicability of the independence assumption involves three steps. First, to test the (conditional) random assignment of patients to nurses, we regress $\widehat{\pi}_{ij}$ on ψ_i in the nurses' sample (i.e., for all patients in the sample, whether the nurse directs them to a doctor consultation or not). Here and in the regressions below we gradually add a set of controls of fixed effects for: four-hour time blocks; days of week; years x months; and the firm's centers with which the patient is registered. These controls are useful, because they address the possibility that patients (nurses) with different characteristics might be differentially likely to request a doctor consultation (work) at particular times or related to particular clinics. If such sorting occurs, our identification strategy relies on the random matching of patients to nurses conditional on this set of time and location controls. In some specifications, we also control for an indicator for prior patient comorbidity and fixed effects for patients' ICD code groups, as determined by the nurse.[10] To test independence, we report the p-value on a joint F-test for ψ_i.

Second, to test that the instrument is orthogonal to nurses' propensity to refer to any doctor consultations, we regress $\widehat{\pi}_j$ on nurse j's propensity to assign to any doctor, $\frac{1}{N_j}{}_{i=1,\ldots,N_j} D^0_{ij}$.

Finally, to test that the instrument is orthogonal to the characteristics of patients referred in the doctors' sample (i.e., those actually sent to a doctor consultation by the nurse), we regress $\widehat{\pi}_{ij}$ on ψ_i in the doctors' sample ($D^0_{ij} = 1$), and report the p-value on a joint F-test for ψ_i.

To establish average exclusion and average monotonicity, we proceed as follows. First, we discuss the institutional rules, which circumscribe nurses' decisions in our setting; for example, nurses cannot prescribe medications or refer to outside specialists. Second, we report the distribution of duration of the patient-facing time of nurse meetings, to show that these are very short (and much shorter than the doctor consultations). This leaves little time for anything other than a brief conversation about the patient's symptoms and potentially about some self-care if no doctor consultation is deemed necessary. Third, in the spirit of Chan et al. (2022), we check if nurses who direct more online are differentially likely to make rare mistakes. To do so, we define as a mistake an instance where the nurse did not refer a patient to a doctor

[10] These covariates are all fully pre-determined, except the nurse ICD groups. But these ICD codes are more informative of the patient's condition than the results of the patients' self-diagnosis questionnaire.

consultation ($D_{ij}^0 = 0$), but the patient was in any emergency care (A&E/ED and hospitalizations) within ten days of meeting the nurse (thus indeed having a serious medical condition). We then regress $\widehat{\pi}_j$ on share of mistakes of nurse j. Finally, to verify that we observe no first stage reversals, we regress D_{ij} on $\widehat{\pi}_{ij}$ in subsamples of the doctors' sample.

After reporting all the checks above, we proceed to use the doctors' sample (i.e., patients who were actually directed to a doctor consultation) to estimate our main specification

$$Y_{ij} = \beta_0 + \beta_1 D_{ij} + \mathbf{Controls}_{ij}^{'} \beta_2 + \epsilon_{ij}, \tag{3}$$

where Y_{ij} are individual outcome components of $\mathbf{Y}_{ij}$ and $\mathbf{Controls}_{ij}$ is a vector, which in the baseline is empty and in some specifications controls for our usual set of fixed effects for: four-hour time blocks; days of week; years x months; and the firm's centers with which the patient is registered. In some specifications, we also control for patient demographics (ψ_i), an indicator for prior patient comorbidity, and fixed effects for nurse-determined patient diagnosis group (ICD group determined by first letter in the code).

Since D_{ij} is potentially endogenous (e.g., if patients with different health problems or other differences that matter for outcomes receive online rather than in-person consultations), we run specifications where we instrument for D_{ij} using $\widehat{\pi}_{ij}$. The differences between the OLS and IV estimates is potentially informative of the direction of this selection.

Following Abadie et al. (2023) we use robust standard errors (s.e.) throughout our regression specifications. To ease comparisons to the existing literature using expert propensities, however, we also report the first stage using standard errors clustered by nurse, which we show are very similar to the robust s.e. in our setting.

5 Empirical findings

We begin this section by discussing evidence on the validity of the model discussed in Section 4. We then discuss our main findings on the similarities and differences between online and in-person in (i) the duration and timing of consultation, (ii) in-meeting outcomes, and (iii) patient outcomes after the consultation. Then we discuss the cost trade-offs for the providers and the patients. Finally, we present evidence on the extent to which patients with different demographics view online consultations as a replacement for in-person consultations.

5.1 Instrument validity

Appendix Figure A2 shows the variation in our instrumental variable, which is the nurse propensity to online leaving out the current meeting. Most of the 62 nurses in our sample direct patients more frequently to in-person consultations, but some direct more often online, resulting in a mean in-person consultation rate of around 56 percent in the doctors' sample.

We use this variation in the instrument to examine the identification assumptions. As Appendix Table A4 shows, the first stage estimate falls slightly from 0.77 without controls to 0.7 when we include the main set of fixed effects (time of day, day of week, month x year, and center), to address the possibility of nurse and patient sorting across times and locations. Reassuringly, when we add further controls (patient demographics, comorbidity indicator, and fixed effects for the nurse-set diagnosis codes, the first stage coefficient remains large (0.67) and stable. The first stage is also precisely estimated when we use either robust s.e. (following Abadie et al. (2023)) or s.e. clustered by nurse (following earlier papers on expert propensities). The F-statistic for the first stage is over 100, so we are not concerned about weak instruments problems, at least for outcomes that we can measure for all or most patients.

Next, we turn to our three tests of the independence assumption. Panel A of Appendix Table A5 shows the balance of the instrument on patient characteristics in the nurses' sample. The p-values are consistent with our assumption that patients and nurses are randomly matched, especially after we include our standard controls for time and location (of the clinic that the patient is registered with). Panel B shows the balance of nurses' propensity to direct online on nurse propensity to direct to *any* doctor (vs. suggesting no doctor visit) in the nurses' sample. The estimates are not only statistically insignificant, but also small in magnitude, when we divide the coefficient by the standard deviation of the main regressor. Finally, Panel C shows the balance of the instrument on patient characteristics in the doctors' sample. Reassuringly, we cannot reject that patient characteristics are balanced among patients referred to doctor consultations, especially with our standard controls for time and location.

In Appendix Table A6 we show results pertaining to the average exclusion assumption (Frandsen et al., 2023). Panel A shows that the mean patient-facing time for nurses is less than five minutes, and the median is four minutes. Such a short meeting may allow a nurse to enquire about the patient's condition and decide whether the patient should consult a doctor and if so – whether this consultation should be online. But it leaves little time for the nurse to affect patient outcomes in other ways than through their assignment to a consultation with a doctor, either in person or online. Panel A also shows that nurse meetings are typically much shorter than doctor consultations: the mean patient-facing time is about four times shorter, and the median is about three times shorter.[11]

Panel B of Appendix Table A6 examines another aspect of the average exclusion assumption, relating to rare mistakes that nurses make. Similarly to Chan et al. (2022), we measure these mistakes as instances when a patient whom a nurse *did not* direct to a doctor consultation is hospitalized within 10 days of meeting the nurse.[12] Even these instances, which are very rare (on average nurses have a mistake share of 10.35%, and 11% of patients experience these events), do not necessarily imply a mistake on the nurse's part, as the health problem may have arisen after the nurse meeting. Nevertheless, our estimates suggest that nurses who refer more online do not significantly differ in the fraction of rare mistakes they made. The estimate is again not only statistically insignificant, but also small in magnitude, when we divide the coefficient by one s.d. of the regressor.

Lastly, we examine the assumption of average monotonicity. Appendix Table A7 follows Frandsen et al. (2023) and Bhuller et al. (2020) by reporting the first stage for different subsamples. As the table shows, the first stage is large and statistically significant when we break down patients by their gender, age, education, income, immigrant status, comorbidity status, whether they specified "general health" in their symptoms form (rather than filling out a specific symptom), and whether they requested the consultation during periods of low or no Covid (versus the first or second Covid wave). This suggests that most patient groups are compliers, responding to the

[11]To ensure comparability of the nurse and doctor meeting durations, Panel A of the table restricts the sample to patients for whom the patient-facing duration is observed for both, although this restriction does not matter much in practice.

[12]Chan et al. (2022) study radiologists' diagnosis of pneumonia, where their decision to diagnose or not is strictly ordered, so less skilled radiologists may be more cautious and over-diagnose. Our setting is different, since we consider the decision to direct to online or in person consultations, which are not necessarily ordered. We separately consider the nurses' decision whether to direct patients to any doctor in our discussion above, and the mistake we measure pertains to that decision rather than to the online vs. in person decision.

nurses' preference for online versus in-person consultations. This is important, since it means that our compliers are broadly representative, at least within the population of patients we study. We return to the point of generalizability below, when we discuss the external validity of our estimates with regards to costs.

5.2 Effects of online versus in-person doctor consultations

Table 2 reports our first set of results, regarding the differences in duration and timing between online and in-person consultations. Panel A shows that the total consultation time is much shorter online according to all the specifications we estimate. This may be one of the reasons why online consultations in themselves are cheaper than in-person (we will investigate post-consultation costs below). Two patterns in these results are worth mentioning, since they recur in many of the other outcomes below. First, the inclusion of different sets of controls makes little difference to the estimates, and second, when OLS and IV differ, and there is a systematic pattern to this difference. Here, the OLS estimates suggest that online meetings are about two-thirds shorter, while the IV estimates suggest that they are only one-third as short. These findings are consistent with patients with (on average) less severe symptoms sorting (or being sorted) into online consultations (assuming that more difficult cases take longer time). Our set of controls, detailed though it is, cannot address this sorting. But the IV estimates overcome this sorting, and they suggest smaller cost savings online than the OLS results would suggest – in this case in time saved. As we discuss below, several of our other findings are also consistent with this interpretation.

Panels B and C break down the total doctor consultation time into patient-facing and administrative parts. Online meetings have significantly shorter patient-facing time but longer administrative time. A possible interpretation of this finding is that when meeting a patient in person, the doctor fills in any notes or forms while the patient is in the room, whereas online meetings end sooner and the doctor fills in some forms by themselves after the meeting. Another possible interpretation (they are not mutually exclusive) is that doctors need some time to consult notes and/or recuperate after consulting patients. Online, this is recorded separately as administrative time, whereas in person this time may in some cases be bundled with the patient-facing time.

In Panel D of Table 2 we show one clear advantage of online consultations: they take place much sooner after the patient's request – typically on the same day. In contrast, in-person consultations are typically held 2-3 days after the nurse meeting, reflecting the need to find availability among the smaller set of doctors working in the nearby clinic, as well as the need to schedule for travelling.

Table 3 examines key within-consultation outcomes. It shows that online and in-person doctor consultations are broadly similar across a range of immediate outcomes. The OLS estimates in Panel A show that the rate of meaningful diagnosis is higher online, while the IV estimates show more negative but imprecise estimates.[13] Panel B shows that online consultations are either more likely (OLS) or equally likely (IV) to yield a prescription.

Panel C of Table 3 shows that the rates of patient prescription collection are similar for online and in person consultations. This measure is interesting, since it can be seen as a measure of patient adherence (Neiman et al., 2018), but is often difficult to measure in data. Nonetheless, these estimates should be taken with caution, since they rely on a subsample of patients (those who receive a prescription), which means we are both conditioning on an outcome and relying on a weaker first stage.

Panel D shows that specialist referrals are either less common online (OLS) or equally common (IV), again consistent with the above-mentioned sorting pattern.[14]

Panel E shows the satisfaction of patients following online consultations. These estimates are available only for patients who scored the meeting, which is more commonly done online (see Appendix Table A8), most likely because patients are more systematically reminded to score consultations online than in person. Consequently, the estimates in this panel (like those in Panel C) condition on an outcome, and should therefore be treated with caution. Nevertheless, the estimates here are again consistent with the similarity of in-person and online in terms of their in-consultation outcomes, we see no significant difference. In sum, meaningful diagnosis, prescription, adherence, referral, and patient satisfaction are all similar.

In contrast to the similarity of in-consultation outcomes for in person and online, we see some differences between the two delivery modes in the short-run post-consultation outcomes. We study these mostly by looking at patient outcomes within 30 days of the doctor consultation. We focus on extensive-margin short-run outcomes, since if a person has used a follow-up service, any subsequent outcomes may in part

[13] A meaningful diagnosis is one that does not fall in the symptomatic or procedure categories of the Swedish ICD 10 diagnosis classification (which are R or Z diagnoses).

[14] Due to differences across regions, patients were only referred to specialists in Stockholm and not in Lund.

be the result of that follow-up.

This difference between online and in person is not present when it comes to the rare (and negative) outcome of avoidable hospitalizations, where patients are hospitalized for reasons that primary care could plausibly have treated or prevented, but did not do or succeed in (u.s. Agency for Healthcare Research and Quality, 2023). Both the OLS and IV estimates in Panel A of Table 4 show no significant difference in this measure between online and in person, although we note that this outcome is very rare, so the confidence intervals of the estimates are wide compared to the mean of the outcome, suggesting that we may be under-powered to detect significant differences.

Panel B examines overall hospitalizations. All the estimates are statistically insignificant, although the IV estimates are large compared to the mean of the outcome, and in some cases marginally significant. As discussed above, and as we shall also see in the next outcomes, the difference between OLS and IV is consistent with the latter solving a selection problem.

In Panel C we see significant differences between online and in person. Online consultations are more likely to result in Emergency Department (ED or A&E) visits in both OLS and IV, and the IV estimates are large. Our interpretation is that an online consultation is more likely to result in the patient or the doctor – or both – concluding that the patient should see another doctor in person, at least as a precaution. In some cases, ED (A&E) could be the quickest/easiest way to achieve this, given the above-mentioned lag of two-three days until an in-person consultation.

This interpretation is also consistent with the results in Panel D. Here, both OLS and IV show that following an online consultation, the patient is more likely to have another primary care consultation within 30 days (still within the same provider, where patients are registered and should receive their primary care). The estimates are all large and statistically significant, and the IV estimates suggest a higher post-first-consultation cost to online: about one in three online consultations (compared to about one in ten in-person consultations) are followed by another primary care visit within 30 days.

In Appendix Table A9 we repeat the analysis in Table 4, but this time starting the 30-day count from the nurse meeting, since this avoids a gap in observing patients between the nurse and doctor meeting, which (as we discuss above) is larger for in person consultations. The estimates in Table A8 are, nonetheless, broadly similar to those in Table 4.

Appendix Table A10 looks more closely at then increased primary care follow-ups after online consultations. The table shows that this is mostly due to more revisits that are initiated by a doctor (not the patient), and that these are essentially all in-person consultations and not online ones. The table also suggests that there may also be a slightly higher probability of a patient-initiated primary care follow up, although the estimates are smaller and imprecise. Taken together the results in this table suggest that doctors working online are often cautious, and book a follow-up in-person consultation. At the same time, it is possible that some of the follow-up visits reflect patient requests for doctors to inspect unrelated health issues.

In Appendix Table A11 we re-estimate the regressions reported in Table 4, except that we consider "medium run" outcomes - those from over 30 days until the end of our sample. The period over which we observe patients varies, with some (who met a nurse in late 2019) observed for over a year while others (who met a nurse in late 2020) are observed for a much shorter duration. Still, our results suggest no significant differences in these medium run outcomes between in person and online.

5.3 Doctor productivity and the sorting of doctors to online

So far we have seen that doctors working online held shorter meetings, and only part of this productivity advantage was due to sorting of patients. To further investigate differences in doctor productivity online vs in person (in terms of consultations per hour), we study doctors' shifts in these two different delivery modes. Here we use the sample of doctor "shifts" which encompasses the much larger sample of non-registered patients as well as registered ones.

Columns (1) and (2) of Table 5 shows that when we account for the full shift duration, doctors working online are roughly twice as productive as those working in person, although admittedly in this setting we cannot control for the sorting of patients. The shift data does, however, afford sufficient variation to study doctor sorting. As the difference between columns (2) and (3) shows, more productive doctors do indeed sort online, but only about 13 percent of the online productivity gain are explained by doctor sorting. Columns (4)-(6) repeat the analysis excluding any breaks between patients, and the results are broadly similar.

5.4 Cost analysis

We now consider the difference in costs for providers and patients between online and in-person. As Table 6 shows, when we ignore follow-ups, online meetings are almost four times cheaper than in person. This large cost advantage likely reflects the productivity differences discussed above, as well overhead costs from operating clinics and other staff costs. These large cost advantage of online consultations is much reduced, however, when we account for the much higher incidence of follow-ups in primary care and the Emergency Department (ED). Once those follow-up rates are accounted for, online is only about 15 percent cheaper.

The primary care provider company is paid through capitation for the patients studied in this paper in Region Scania, and so face a cost from additional primary care follow ups within the service. In Region Stockholm, they are paid through a combination of capitation and some fee for service, so it is more unclear what the incentives are for additional primary care visits. In Region Stockholm, the primary care provider faces a penalty if a large share of their patients have Emergency Department visits, while they get a bonus if a low share have ED visits. In Region Scania, that was not the case in the study period, but was started only in 2022.

A similar result applies to patient costs. When we account for patient co-pay, the time costs of the duration of the consultations and travel and waiting costs, as well as the travel costs – without follow-ups – online is about 2.4 times cheaper than in person. But accounting for the same costs related to the higher share of follow-ups in primary care and and ED almost erases the cost advantage of online. Nevertheless, as discussed above, online patients still benefit from seeing doctors sooner. They may also benefit from not having to travel and wait when sickest, from scheduling convenience, and from being able to see doctors almost 24 hours a day, every day of the week – much more than in person.

Still, to achieve greater cost savings for both providers and patients, reducing revisits without sacrificing healthcare quality remains an important challenge.

5.5 Patient heterogeneity in viewing online as substitute for in person

Finally, we consider the extent to which patients with different characteristics view online as a substitute for in-person consultations. Table 7 shows this for sample of non-registered (drop-in) patients, which is much larger than that of registered patients, and allows for heterogeneity analysis. Here we use a question that was asked only of online patients – did they consider their online consultation a replacement for an in-person consultation? An important caveat here is that just under half the patients who were asked answered this question, and it is plausible that those who answered were more positively disposed towards online. Still, as the table shows, about 95 percent of those who answered said that online was a substitute for in-person. Those who were less likely to consider online a replacement were predominantly older patients in their 80s or 70s and to a lesser extent immigrants who were neither from the EU15 nor from Scandinavia.

6 Conclusion

Online delivery is now possible for many services, such as banking/financial advice, tutoring/teaching, therapy, and healthcare. Within healthcare, healthcare systems around the world are struggling to find the right mix of online and in person consultations after the pandemic. This decision can be related to the sticker price of consultations which is often lower than that of in-person consultations, which is important in a situation with tighter finances for healthcare providers and systems. It can also be related to preferences about work mode among doctors, and patient preferences for having a more convenient consultation and not travelling when sick. But it is crucial that we also understand the health outcomes and the following healthcare costs that result from online compared to in-person consultations. To the best of our knowledge, we provide the first paper studying the effects of online vs. in person 1:1 services. Moreover, this is the first paper to compare online healthcare with in person healthcare in a setting where patients have already sorted to care.

Despite online meetings being shorter than in-person meetings, we find that the in-meeting outcomes such as diagnosis, prescription, specialist referral and patient satisfaction are largely similar to in-person meetings. This can help explain why the sticker price of online meetings is cheaper than in-person, on top of reasons such as lower overhead due to less office space when doctors work from home.

Yet, it is crucial to measure what happens after the meeting. We find that online meetings are followed (in the month following the consultation) by considerably more in-person follow ups, both in the same primary care service and in the Emergency

Department. The larger share of in-person follow ups in primary care are mostly initiated by the doctor, and hence seems to reflect cautiousness among doctors about fully completing a case without an in-person consultation. Other outcomes during the 30 days after the initial consultation are not different between online and in person. No outcomes in the longer run, post 30 days after the consultations, differ between online and in person, so it seems that there are no differential medium term health or healthcare consequences from online visits.

Given the higher share of follow-ups, the initial 2.4 times total cost difference between online and in person consultations is mostly erased, leaving online meetings (counting additional follow ups) just a little cheaper than in-person consultations.

We have also been able to study doctor sorting between in person and online work, and find that slightly more productive (in terms of patients per hour) doctors sort to online. This is one of the first measures of such sorting, and interestingly goes the opposite direction as contemporaneous work by Emmanuel and Harrington (2023) who study sorting of call center workers to remote vs in-person work.

A pattern that we see in several results is that OLS suggests that online meetings are even shorter or have lower specialist referrals, compared to IV. We interpret this as evidence of the sorting problem, that nurses may sort more severe cases to in person – which is what our IV method is designed to deal with. Future work may investigate whether nurses' current sorting is optimal, or whether it can be improved so that less patients who need in-person follow ups are sorted to online.

References

Abadie, A., Athey, S., Imbens, G. and Wooldridge, J. (2023), 'When should you adjust standard errors for clustering?', *The Quarterly Journal of Economics* **138**(1), 1–35.

Aizer, A. and Doyle Jr, J. (2015), 'Juvenile incarceration, human capital and future crime: Evidence from randomly-assigned judges', *The Quarterly Journal of Economics* **130**(2), 759–803.

Aksoy, C. G., Barrero, J., Bloom, N., Davis, S. J., Dolls, M. and Zarate, P. (2022), Working from home around the world, Working paper no. w30446, National Bureau of Economic Research.

Angrist, J. and Imbens, G. (1995), 'Two-stage least squares estimation of average causal effects in models with variable treatment intensity', *Journal of the American Statistical Association* **90**(430), 431–442.

Anwar, S., Bayer, P. and Hjalmarsson, R. (2012), 'The impact of jury race in criminal trials', *The Quarterly Journal of Economics* **127**(2), 1017–1055.

Bakx, P., Wouterse, B., Van Doorslaer, E. and Wong, A. (2020), 'Better off at home? Effects of nursing home eligibility on costs, hospitalizations and survival', *Journal of Health Economics* **73**, 102354.

Bhuller, M., Dahl, G., Løken, K. and Mogstad, M. (2020), 'Incarceration, recidivism, and employment', *Journal of Political Economy* **128**(4), 1269–1324.

Bloom, N., Han, R. and Liang, J. (2022), How hybrid working from home works out, Working paper no. w30292, National Bureau of Economic Research.

Bloom, N., Liang, J., Roberts, J. and Ying, Z. (2015), 'Does working from home work? Evidence from a chinese experiment', *The Quarterly Journal of Economics* **130**(1), 165–218.

Chan, D., Gentzkow, M. and Yu, C. (2022), 'Selection with variation in diagnostic skill: Evidence from radiologists', *The Quarterly Journal of Economics* **137**(2), 729–783.

Dahl, G., Kostøl, A. and Mogstad, M. (2014), 'Family welfare cultures', *The Quarterly Journal of Economics* **129**(4), 1711–1752.

Dahlstrand, A. (2023), 'Defying distance? The provision of services in the digital age'.

Dingel, J., Gottlieb, J., Lozinski, M. and Mourot, P. (2023), Market size and trade in medical services, Working paper no. w31030, National Bureau of Economic Research.

Dobbie, W., Goldin, J. and Yang, C. (2018), 'The effects of pre-trial detention on conviction, future crime, and employment: Evidence from randomly assigned judges', *American Economic Review* **108**(2), 201–240.

Doyle Jr, J. (2007), 'Child protection and child outcomes: Measuring the effects of foster care', *American Economic Review* **97**(5), 1583–1610.

Ellegård, L., Kjellsson, G. and Mattisson, L. (2021), An app call a day keeps the patient away? substitution of online and in-person doctor consultations among young adults, Working Papers in Economics 808, University of Gothenburg.

Ertem, Z., Schechter-Perkins, E., Oster, E., van den Berg, P., Epshtein, I., Chaiyakunapruk, N., Wilson, F., Perencevich, E., Pettey, W., Branch-Elliman, W. and Nelson, R. (2021), 'The impact of school opening model on SARS-CoV-2 community incidence and mortality', *Nature Medicine* **27**(12), 2120–2126.

Frandsen, B., Lefgren, L. and Leslie, E. (2023), 'Judging judge fixed effects', *American Economic Review* **113**(1), 253–277.

Gaspar, J. and Glaeser, E. (1998), 'Information technology and the future of cities', *Journal of Urban Economics* **43**(1), 136–156.

Goodman, J., Melkers, J. and Pallais, A. (2019), 'Can online delivery increase access to education?', *Journal of Labor Economics* **37**(1), 1–34.

Greene, J. (2022), *The doctor who wasn't there: Technology, history, and the limits of telehealth*, The University of Chicago Press.

Hatef, E., Lans, D., Bandeian, S., Lasser, E., Goldsack, J. and Weiner, J. (2022), 'Outcomes of in-person and telehealth ambulatory encounters during COVID-19 within a large commercially insured cohort', *JAMA Network Open* **5**(4), e228954–e228954.

Heckman, J. and Urzua, S. (2010), ‘Comparing IV with structural models: What simple IV can and cannot identify’, *Journal of Econometrics* **156**(1), 27–37.

Kletečka-Pulker, M., Völkl-Kernstock, S., Fassl, A., Klager, E., Willschke, H., Klomfar, S., Wochele-Thoma, T., Schaden, E. and Atanasov, A. (2021), ‘Telehealth in times of COVID-19: Spotlight on Austria’, *Healthcare (Basel)* **9**(3), 280.

Kling, J. (2006), ‘Incarceration length, employment, and earnings’, *American Economic Review* **96**(3), 863–876.

Ladin, K., Porteny, T., Perugini, J., Gonzales, K., Aufort, K., Levine, S., Wong, J., Isakova, T., Rifkin, D., Gordon, E. and Rossi, A. (2021), ‘Perceptions of telehealth vs in-person visits among older adults with advanced kidney disease, care partners, and clinicians’, *JAMA network open* **4**(12), e2137193–e2137193.

Lee, S. and Salanié, B. (2018), ‘Identifying effects of multivalued treatments’, *Econometrica* **86**(6), 1939–1963.

Michaels, G., Rauch, F. and Redding, S. J. (2019), ‘Task specialization in US cities from 1880 to 2000’, *Journal of the European Economic Association* **17**(3), 754–798.

Mountjoy, J. (2022), ‘Community colleges and upward mobility’, *American Economic Review* **112**(8), 2580–2630.

Neiman, A. B., Ruppar, T., Ho, M., Garber, L., Weidle, P. J., Hong, Y., George, M. G. and Thorpe, P. G. (2018), ‘Cdc grand rounds: Improving medication adherence for chronic disease management—innovations and opportunities’, *American Journal of Transplantation* **18**(2), 514–517.

Powell, R. E., Henstenburg, J. M., Cooper, G., Hollander, J. E. and Rising, K. L. (2017), ‘Patient perceptions of telehealth primary care video visits’, *The Annals of Family Medicine* **15**(3), 225–229.

Rabideau, B. and Eisenberg, M. D. (2022), ‘The effects of telemedicine on the treatment of mental illness: Evidence from changes in health plan benefits’, *Available at SSRN 4065120* .

Rodler, S., Apfelbeck, M., Schulz, G. B., Ivanova, T., Buchner, A., Staehler, M., Heinemann, V., Stief, C. and Casuscelli, J. (2020), ‘Telehealth in uro-oncology beyond the pandemic: Toll or lifesaver?’, *European Urology Focus* **6**(5), 1097–1103.

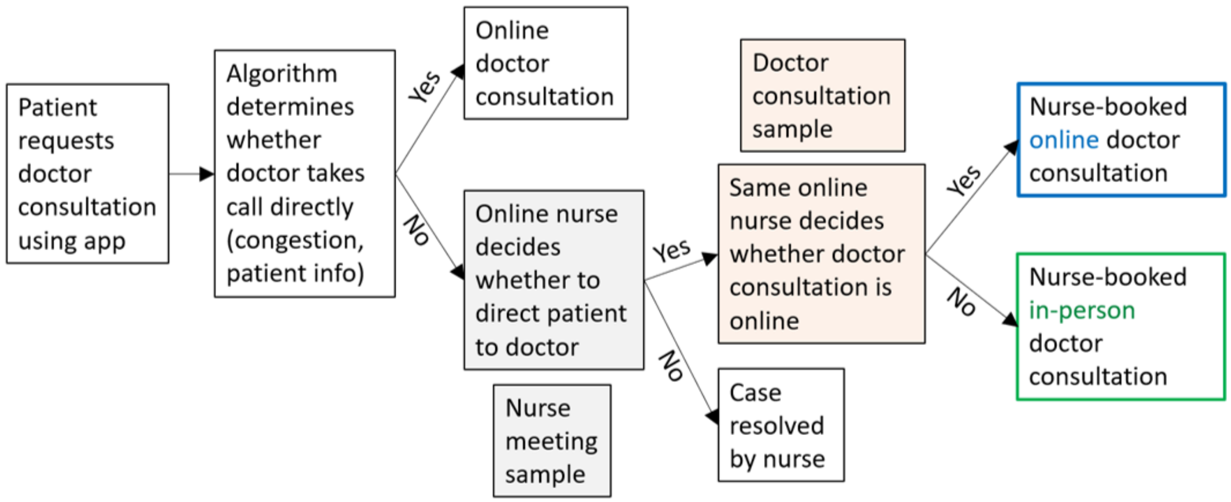
Santarossa, S., Kane, D., Senn, C. Y. and Woodruff, S. J. (2018), 'Exploring the role of in-person components for online health behavior change interventions: Can a digital person-to-person component suffice?', *Journal of medical Internet research* **20**(4), e144.

Slavova-Azmanova, N. S., Millar, L., Ives, A. and Saunders, C. M. (2021), "nothing beats the doctor's face to impart trust in their judgement': The role of telehealth in cancer care', *Australian Health Review* **45**(4), 521–522.

u.s. Agency for Healthcare Research and Quality (2023), 'Potentially avoidable hospitalizations'. Last accessed September 2023.
URL: *https://www.ahrq.gov/research/findings/nhqrdr/chartbooks/carecoordination/measure3.html*

Zeltzer, D., Einav, L., Rashba, J. and Balicer, R. D. (2023), 'The impact of increased access to telemedicine', *Journal of the European Economic Association* .

Figure 1. Assignment of registered patients to in-person versus online



Note: This paper shows the flows of patients registered with Kry. The cases in the box with grey background are defined as the "Nurse meeting sample" (or "Nurse sample" for short). The cases in the box with orange background are defined as the "Doctor meeting sample" (or "Doctor sample" for short).

Table 1. Summary statistics

	Doctor sample mean	Municipality mean	National mean
Female	0.49	0.50	0.50
Age	35.0	39.3	41.3
University education	0.58	0.49	0.39
Married	0.30	0.40	0.42
Immigrant background	0.39	0.35	0.26
Big city municipality	0.85	0.85	0.32
Income	334.4	353.7	328.9

Note: Panel A of this table comparing patients in our sample to the full Swedish population. The "Sample mean" column reports unweighted means of the (4662) doctor sample observations for which each variable is defined. The "Municipality mean" takes municipality-level means in 2019 and averages them using the share of each of the 96 municipalities in the doctor sample as weights. The "National mean" is the mean for Sweden in 2019. "University education" is reported for people over the age of 23; "Married" is reported for people over the age of 18; "Immigrant background" is an indicator for people who were either born outside Sweden or whose parents were both born outside Sweden. "Big city municipality" is an indicator for municipalities with big cities, including Stockholm and Lund. "Income" includes annual earnings from wages and self-employment in thousands of SEK and is reported for people over the age of 20.

Table 2. Timing and duration of doctor consultations

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Days between nurse meetings and doctor consultation								
Consultation was online	-2.30	-2.28	-2.30	-2.35	-3.13	-2.73	-2.73	-2.77
	(0.065)	(0.066)	(0.068)	(0.069)	(0.34)	(0.40)	(0.42)	(0.43)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4662	4662	4529	4516	4662	4662	4529	4516
First-stage K-P F-statistic					197	145	138	133
Baseline mean	2.4	2.4	2.5	2.5	2.4	2.4	2.5	2.5
B: Total consultation duration (in minutes)								
Consultation was online	-25.8	-25.7	-25.7	-26.0	-12.6	-14.1	-15.1	-15.0
	(0.62)	(0.62)	(0.63)	(0.66)	(3.53)	(4.01)	(3.99)	(4.13)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4510	4510	4382	4369	4510	4510	4382	4369
First-stage K-P F-statistic					193	140	133	130
Baseline mean	39.8	39.8	39.7	39.7	39.8	39.8	39.7	39.7
C: Patient-facing part of the consultation (in minutes)								
Consultation was online	-26.8	-26.7	-26.8	-27.0	-22.6	-23.2	-22.8	-22.7
	(0.44)	(0.44)	(0.46)	(0.48)	(2.16)	(2.42)	(2.48)	(2.56)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4342	4342	4222	4210	4342	4342	4222	4210
First-stage K-P F-statistic					199	147	137	134
Baseline mean	32.2	32.2	32.2	32.2	32.2	32.2	32.2	32.2
D: Administrative part of the consultation (in minutes)								
Consultation was online	1.32	1.38	1.42	1.31	6.32	6.51	6.44	6.32
	(0.34)	(0.34)	(0.35)	(0.35)	(1.94)	(2.21)	(2.27)	(2.34)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4331	4331	4212	4200	4331	4331	4212	4200
First-stage K-P F-statistic					197	145	136	133
Baseline mean	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2

Note: This table reports regressions using the doctor sample (see text for discussion). The instrument in the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. Fixed effects include Year*Month, 4 hour blocks, day of the week and where patient was listed. For a description of the control variables we use, please see main text. Robust standard errors are in parentheses.

Table 3. Within-consultation outcomes

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Doctor set an informative diagnosis								
Consultation was online	0.036	0.037	0.038	0.028	-0.12	-0.14	-0.15	-0.12
	(0.014)	(0.014)	(0.014)	(0.014)	(0.076)	(0.087)	(0.090)	(0.087)
Observations	4662	4662	4529	4516	4662	4662	4529	4516
First-stage K-P F-statistic					197	145	138	133
Baseline mean	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68
B: Patient received a prescription								
Consultation was online	0.15	0.15	0.15	0.15	-0.013	0.016	0.025	0.061
	(0.013)	(0.013)	(0.014)	(0.014)	(0.070)	(0.080)	(0.082)	(0.083)
Observations	4662	4662	4529	4516	4662	4662	4529	4516
First-stage K-P F-statistic					197	145	138	133
Baseline mean	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21
C: Patient collected prescription within 30 days (conditional on getting a prescription)								
Consultation was online	-0.0019	-0.0023	0.0022	0.0047	-0.024	-0.0033	0.032	0.044
	(0.017)	(0.020)	(0.020)	(0.021)	(0.13)	(0.19)	(0.18)	(0.19)
Observations	1074	1074	1043	1040	1074	1074	1043	1040
First-stage K-P F-statistic					27	16	16	17
Baseline mean	0.92	0.92	0.91	0.91	0.92	0.92	0.91	0.91
D: Doctor gave a specialist referral (Stockholm only)								
Consultation was online	-0.092	-0.093	-0.093	-0.096	-0.035	-0.016	-0.0034	-0.018
	(0.010)	(0.010)	(0.011)	(0.011)	(0.057)	(0.068)	(0.070)	(0.068)
Observations	2415	2415	2332	2323	2415	2415	2332	2323
First-stage K-P F-statistic					82	60	58	64
Baseline mean	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
E: Patient satisfaction score (5 is best)								
Consultation was online	-0.012	-0.022	-0.023	-0.038	-0.076	-0.21	-0.21	-0.24
	(0.042)	(0.044)	(0.046)	(0.049)	(0.26)	(0.32)	(0.34)	(0.33)
Observations	1466	1466	1430	1424	1466	1466	1430	1424
First-stage K-P F-statistic					52	33	29	31
Baseline mean	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7

Note: This table reports coefficients from regressions using the doctor sample (see text for discussion). The instrument in the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. For a description of the variables, please see the main text and appendix. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Table 4. Patient outcomes in the 30 days after the doctor consultation

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Avoidable hospitalizations within 30 days								
Consultation was online	-0.00011	-0.00017	-0.00031	-0.00041	0.0021	0.0019	0.0018	0.0023
	(0.0011)	(0.0010)	(0.0011)	(0.0011)	(0.0035)	(0.0054)	(0.0059)	(0.0055)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
B: Any hospitalization within 30 days								
Consultation was online	0.0024	0.0024	0.0022	0.0022	0.035	0.040	0.044	0.047
	(0.0031)	(0.0030)	(0.0032)	(0.0033)	(0.018)	(0.022)	(0.023)	(0.024)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008
C: Any Emergency Department (A&E) visit within 30 days								
Consultation was online	0.017	0.014	0.014	0.016	0.12	0.11	0.13	0.13
	(0.0070)	(0.0070)	(0.0072)	(0.0078)	(0.044)	(0.054)	(0.057)	(0.059)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.040	0.040	0.041	0.041	0.040	0.040	0.041	0.041
D: New visit to primary care provider within 30 days								
Consultation was online	0.082	0.086	0.092	0.098	0.13	0.20	0.24	0.25
	(0.016)	(0.016)	(0.016)	(0.017)	(0.090)	(0.11)	(0.11)	(0.11)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37

Note: This table reports coefficients from regressions using the doctor sample (see text for discussion). The instrument in the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. For a description of the variables, please see the main text and appendix. The baseline mean is the mean of the dependent variable for in-person doctor consultations. There are 37 hospitalizations, 5 avoidable hospitalizations, and 189 Emergency Department visits in the doctor sample. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Table 5. Doctor consultations per hour

	Shift incl. all breaks			Shift excl. all breaks		
	(1)	(2)	(3)	(4)	(5)	(6)
Shift was online	1.88	1.94	1.68	2.34	2.47	2.00
	(0.078)	(0.078)	(0.12)	(0.084)	(0.084)	(0.15)
Time fixed effects		✓	✓		✓	✓
Doctor fixed effects			✓			✓
Observations	78413	78413	78413	78413	78413	78413
Baseline mean	1.81	1.81	1.81	2.88	2.88	2.88

Note: This table reports coefficients from regressions using the doctor shift sample. This sample consists of registered and non-registered patient meetings (excluding prescription renewals, tests ordered, psychologist consultations and nurse meetings) collapsed to doctor*day level. A shift starts with the start of the first consultation and ends with the end of the last consultation. Breaks are times in between the consultations. For the construction of the shift variables, please see the data appendix. Time fixed effects include Year*Month and day of the week fixed effects. The baseline mean is the mean of the dependent variable for in-person doctor shifts. Robust standard errors are in parentheses.

Table 6. Cost table for providers and patients (in SEK)

	In-person	Online	Table
A. Provider cost			
Cost of doctor consultation w/o in-person follow-up	1905	500	
Follow-up cost of in-person primary care	133	629	Table A10, Panel B
Follow-up cost of in-person ED	182	760	Table 4, Panel C
Total provider cost incl. follow-up cost times fraction of follow-ups	2220	1889	
B. Patient cost			
Co-pay/Patient fee in primary care (Average)	159	159	
Patient-facing consultation time	156	46	Table 2, Panel C
Waiting time for the doctor	145	74	
Two-way commuting costs to the GP (Travel/Parking time/Fuel costs)	220	0	
Parking fee (Primary care, during the day)	5	0	Table 2, Panel C
Public transport fee (Single ticket, One-way)	7	0	
Patient cost without in-person follow-up	692	279	
Follow-up cost of in-person primary care	48	228	Table A10, Panel B
Follow-up cost of in-person ED	68	284	Table 4, Panel C
Total patient cost incl. follow-up cost times fraction of follow-ups	808	791	

Note: This table reports cost estimates in SEK. Estimates listed under the "Table" column are coefficients taken from column 8 of the corresponding table. First visits are either in person or online primary care consultations. Follow-ups are either in-person revisits to primary care or Emergency departments (ED) visits within 30 days. The provider costs are weighted by the probability that the treatment happens. The patient time cost estimates are the product of patient time spent in (or getting to and from) a consultation, multiplied by the mean hourly wage of white-collar workers in Sweden (290 SEK/hour in March 2023). For both online and in-person consultations, the fee for paying patients is 225 (the mean of 250 SEK for Stockholm and 200 SEK for Scania), and the fee per patient is multiplied by the fraction of paying patients (70.58% of our sample). The mean patient-facing consultation time for in-person GP visits is 32.2 minutes; for online 9.5 minutes. The waiting time in the doctor's office for in-person is an educated guess of 30 minutes (Ekman 2018) and 15.31 minutes for online (based on our data). The commuting costs to a GP include travel time weighted by transport/parking time both ways – to the doctor and back - multiplied by the average hourly wage. Transport includes commuting by car (incl. fuel costs), public transport (incl. ticket), biking and walking. We calculate the costs based on the probability of commuting type (Rosberg & Enström 2019). The average time to a GP is 23.42 minutes two-way after including frequencies of commuting. We assume 5 minutes of parking/walking to the doctor's office before and after the consultation. For follow-ups, we multiply the costs by the probability that the follow-up occurs. The ED fees for Stockholm and Scania are 400 SEK for both and were multiplied by the fraction of paying patients. We assume that patients only drive by car to an ED. The commuting costs to an ED take the average travel time to an ED multiplied by the average hourly wage and include parking time and fuel costs. The mean travel time by car to an ED is 31.06 minutes two-way. The median stay time of a patient in an ED is 3.18 hours. The two-way commuting costs are around 420 SEK, and the ED time costs are 957 SEK.

Table 7. Drop-ins: Replacement for in-person answer

Has a comorbidity	-0.004					-0.003
	(0.0008)					(0.0008)
In employment (ages 16-74)		0.02				
		(0.001)				
Age 10-19			-0.004			0.0007
			(0.0010)			(0.001)
Age 20-29			-0.004			0.004
			(0.0009)			(0.002)
Age 30-39			-0.002			0.009
			(0.0009)			(0.002)
Age 40-49			-0.003			0.007
			(0.001)			(0.002)
Age 50-59			-0.01			-0.0002
			(0.002)			(0.002)
Age 60-69			-0.02			-0.01
			(0.003)			(0.003)
Age 70-79			-0.04			-0.03
			(0.006)			(0.006)
Age 80+			-0.07			-0.07
			(0.02)			(0.02)
In education				-0.0002		0.002
				(0.0008)		(0.002)
Primary school education				-0.01		-0.008
				(0.001)		(0.001)
Short high-school				-0.02		-0.01
				(0.002)		(0.002)
Uni $<$ 3 years				-0.004		-0.003
				(0.001)		(0.001)
Uni $\geq$ 3 years				-0.005		-0.005
				(0.001)		(0.001)
Second gen. immigrant					-0.01	-0.01
					(0.001)	(0.001)
Immigrant (EU15/Scandi.)					-0.01	-0.009
					(0.002)	(0.002)
Immigrant (non-EU15/Scandi.)					-0.03	-0.03
					(0.001)	(0.001)
Observations	456498	262893	437297	434144	437023	433878
Baseline mean	0.95	0.95	0.95	0.95	0.95	0.95

Note: This table is based on the non-registered online drop-in sample. It shows estimates of a survey, which asked patients whether the online consultation replaced an in-person consultation. Positive answers to the question are coded as 1, Don't know" responses as 0.5," and negative responses as 0. Consultations related to Chlamydia or Covid-19 are dropped. The baseline for the age bins is children aged 0 - 9, and for the education variables, the baseline is high-school education. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Online Appendix

A Data appendix

A.1 Data sources

Our analysis is primarily based on consultation-level data between 2019-2020 from a Swedish private healthcare provider offering both in-person and online medical consultations. For all individuals observed in this consultation level data, who will be referred to as patients from now on, we have administrative individual-level panel data from Statistics Sweden's Integrated Database for Labour Market Research (LISA) from 2013 - 2020. Furthermore, for all patients, we have specialist care panel data on the visit level (e.g., a hospital stay or specialist visit) from 2013 - 2020. The specialist care data includes in-patient and out-patient care data and is provided by the National Board of Health and Welfare (NBHW / Socialstyrelsen). Finally, for all patients, we have prescription data on the drug level, i.e., each prescribed drug makes an observation line over 2013-2020, again provided by Socialstyrelsen.

We note that all datasets are proprietary and confidential and were accessed after applications to the Stockholm Regional Ethics Council (update Amanda's text) had been approved. Additionally, Statistics Sweden and the other entities carried out their own confidentiality assessments before approving the sharing of data. Statistics Sweden anonymized the personal identifiers and matched the identifiers with all datasets, and then shared only an anonymized version of the data with us.

A.1.1 Primary care provider data

The consultation level data from 2019 to 2020 from a large private Swedish primary care provider provides the backbone for our analysis. The primary care provider started in 2016 as a digital-only healthcare provider. However, since 2019, it has extended its offering to include in-person doctors' consultations. In-person services were rolled out at different times for different locations, with services first offered in Lund (City in the Scania region) and then expanded; see table ?? for the observed opening dates. The observed opening date is the date of the first logged in-person doctors consultation at the clinic.

For each consultation, we have data on when it occurred (minute precision), the form it took, if it was in-person or online, and the fee the patient paid. We do not directly know where the meeting took place; however, for all patients registered with a Kry in-person clinic, we do know which one. All Swedish citizens are registered with some primary care provider and can change at will without any fee. Therefore, for patients registered with Kry, we have information about the (likely?) location of the in-person consultations.

Moreover, we have data on the duration of the consultation, including a breakdown of the patient's and clinician's consultation duration, where the latter also encompasses administrative work related to the consultation. We also have data on the provider's internal code for the symptom the patient provides when seeking care through the provider's mobile app. The providers' mobile app is the primary channel for seeking care and the only one relevant to our study. In the app, the patients start by filling in what symptoms they are seeking care for before being matched with a clinician. Finally, we also know the consultations "type"; this is an internal categorization of consultations depending on whom the patient met (e.g., "nurse meeting" or "psychology meeting"), whether it was booked ahead of time (e.g., "drop-in" or "doctor booked revisit"), or the purpose (e.g., "prescription renewal" or "test ordered"). We are primarily concerned with a sequence of consultations starting with a "nurse meeting" and resulting in a "doctor booked revisit" (booked by the nurse); see section A.2 for more details.

Regarding the outcome of the consultation, we have data on the clinician's diagnosis and whether the patient was prescribed anything. The patient diagnosis is in the form of an ICD-10-SE code with 4-5 characters of precision. We do not have data from the provider on what was prescribed. However, we have prescription data from Socialstyrelsen to fill in the gap. We also have information on whether the doctor referred the patient to a specialist for the Stockholm-based clinics.

For the clinicians and patients, we know their age and gender; however, with a very high share of missing values for the clinicians. Furthermore, for the clinicians, we know their specialization and, from that, their seniority level.

A.1.2 Demographics and socioeconomic data

To complement the primary care data, we have demographic and socioeconomic micro-data on patients from SCB drawn from the Integrated Database for Labour Market Research (LISA). This panel data provides information on individual income, educational attainment, municipality, immigration background, and marriage status. The variables are provided at the patient level, with yearly measurements. The income measurement is a summation for the year, the education attainment is measured at the end of the year's spring semester, immigration background is constant, and municipality and marriage status are measured either 31st of December or the 1st of January of the year after. Therefore we exclusively use the 2018 values for the demographic and socioeconomic controls employed throughout the paper. This decision ensures that all variables were measured before the start of our sample in 2019.

A.1.3 Specialist care data

The National Board of Health and Welfare (NBHW / Socialstyrelsen) provided the specialist care data. This data, covering 2013-2020, may be divided into in-patient and out-patient care data. Both are forms of specialist care, but in-patient care requires the patient to be admitted overnight. Out-patient care includes emergency department visits and other non-primary care visits to clinics and clinicians to patients' homes. The in-patient and out-patient datasets contain up to 30 ICD-10 diagnostics codes with three characters of precision. The datasets also include external causes codes for applicable cases, classifying events such as falls and bites. Both datasets provide the visit date (and discharge date for in-patient visits). However, the out-patient data also provides the exact admission, assessment, and discharge time for emergency visits. Finally, the in-patient data also provides data on the mode of contact, for instance, a home visit, a visit with a team of clinicians at a clinic, or a visit with just one clinician at a clinic.

A.1.4 Prescription data

The prescription data was also provided by the NBHW (Socialstyrelsen). This data encompasses all prescriptions patients picked up between 2013 and 2020. Each distinct drug picked up takes up an observation line, so a single trip to the pharmacy may be represented by multiple observations. We know the date of when the drugs were picked up, but not whether they were picked up at the same pharmacy or during the same visit to the pharmacy. The data includes some information about the prescriber. We have anonymized codes for the prescribing clinic and the type of care it was prescribed from, such as psychiatric, primary care, or pediatric. We also know the specialization of the prescribing clinician.

Furthermore, we have information on the drugs picked up. For each drug, we have data on its ATC code. We also know the number of pills in the packages prescribed, the number of packages prescribed, and the intended number of daily doses of each package. Finally, we have data on the cost of the picked-up prescription to the patient and the region (the insurer).

A.2 Matching between nurse meetings and doctor consultations

We start with the doctors' visits which are labeled as booked by someone in the firm; these are consultations with doctors who had some preceding (originating) meeting with someone at the firm (online or in person). This originating meeting can be of any type, e.g., nurse meetings, drop-ins, or psychologist visits. To find the originating meeting, we search in a time window of 30 days before the doctor consultation.

Then we utilize two strategies to find this initial meeting. The first strategy is to match the doctor consultation with a preceding meeting with the same symptom, as specified by the patient when seeking care. This label usually follows automatically in a care episode with multiple visits). Sometimes, the symptom is changed in a later visit to "revisit" or "phone triage." The second strategy is developed to deal with these cases. In these cases, we allow the doctor-booked revisit to match the closest preceding meeting with Kry within the 30-day window.

This approach allows for multiple potential matches, and three conflicts may arise. First, a doctor consultation may match with more than one potential originating meeting in the first strategy. To resolve these, we prioritize matches in which the window between the doctor-booked revisit and the originating meeting is as short as possible. Second, two different doctor consultations may match with the same preceding meeting. These conflicts are resolved in favor of earlier doctor-booked revisits over later ones. This rule ensures that matched meetings flow nicely into each other chronologically.

A.3 Defining samples

We document our primary (doctor) sample restrictions in Table A1. We start from the universe of all doctor-booked revisit-type meetings across 2018-2020. We start by requiring that the doctor-booked revisit could be matched to a preceding nurse meeting using the process outlined in section A.2. We also ensured that this nurse meeting was online. This requirement is important because in-person nurse meetings follow different care paths. We also remove patients who sought care for Covid-19 related inquiries, as these can not be treated in person. Furthermore, we must ensure that all patients are at risk of being referred to an in-person consultation by the nurse. To that end, we require that the patient is listed at one of the primary care provider's in-person clinics and that the clinic was opened. The opening was defined as the first in-person consultation held at the clinic. Patient registration is essential because only registered patients can access in-person care. We also remove patients listed at clinics with very few observations leaving only three clinics in Stockholm and one in Lund. Then we removed patients who sought care for chlamydia and breastfeeding and removed infants who were one year old or younger at the time of their consultation. The motivation for all these cases is that we know these patients follow different care paths. So, they may not always be randomly allocated to nurses, who may also be constrained in their decisions to redirect patients. Finally, we ensure that all remaining nurses have at least 20 observations fulfilling these sample restrictions. This requirement ensures we have enough data to create our instrument, the nurses' propensity to refer to an online doctor's consolation.

We also have a secondary sample, a nurse sample, on which we have imposed very similar sample restrictions. The difference is that these nurse meetings do not necessarily culminate in a doctor's consultation but may be entirely solved by the nurse. Therefore, instead of starting from all doctor-booked revisit-type meetings, we start with the universe of nurse meetings. Beyond that, the restrictions are the same. However, note that the 20 minimum observation requirement concerns the number of observations in the primary sample.

The doctor shift sample refers to unregistered and registered patients that had a doctor consultation, both in-person and online. Consultation types exclude nurse meetings, psychologist consultations, prescription renewals or ordering tests. We also removed consultations without any duration time, consultations that take longer than midnight, and all consultations that were on the same day the doctor had mixed work (in-person and online consultations). In the end, the doctor shift sample consists of 1269163 individual consultations, which were then collapsed to the shift level per doctor and day. The most basic approach for a shift is to take the start time of the first consultations and the end time of the last consultations of a doctor within 24 hours in a calendar day. More on the creation of shifts can be found in Section A.6.2. There are 2046 in-person doctor shifts and 76367 online doctor shifts in our doctor shift sample.

A.4 Creating our instrument

After applying our primary sample restrictions, we construct the propensity to online instrument. The instrument is the leave-own-out share of meetings the nurse redirected to an online consultation. That is

$$\hat{\pi}_{ij} \equiv \frac{\sum_{i'=1,\ldots,N_j;i'\neq i} ToOnlineDoctor_{i'j}}{\sum_{i'=1,\ldots,N_j;i'\neq i} ToAnyDoctor_{i'j}} \tag{4}$$

We also use our instrument in the secondary nurse sample for Table A5 (Ref?). We define the instrument differently since not all nurse meetings lead to a doctor's consultation. We define it in two ways depending on whether the meeting was redirected to a doctor. Like our primary sample, the redirected meetings are defined with the leave-own-out propensity. The meetings that are not redirected are given the value of the non-leave-own-out nurse's propensity to redirect to an online consultation among the redirected meetings.

A.5 Matching prescription data

We do not know which prescription is associated with a doctor's consultation in our primary sample with the primary care provider. Therefore we had to estimate which prescriptions pertained to these consultations. To do so, we employed a simple rule. We took the first anonymized prescriber identifier from a prescribing clinic, classified as "primary care" or "other," observed after the doctor's consultation.

A.6 Construction of variables

A.6.1 Summary table of patients

The summary statistics table consists of three columns that explain the mean for three different samples. The first column called "Sample mean" is based on the doctor sample, which consists of 4662 patients that had a nurse meeting and were directed to a doctor (also see restriction table A1). The variables for the doctor sample are based on the year 2019. The "Municipality mean" column takes the same patient observations, but links public data on the municipality level from 2019 to compare patients to their municipality mean. As only 96 of the 290 municipalities in Sweden are represented in the sample, we weight the municipality mean by the frequency of patients in the municipalities. Patients in the doctor sample come primarily from Stockholm and Scania. The "National mean" column is independent of the other columns and shows the mean of all of Sweden based on public data from 2019.

The public data for the municipality and national mean mainly comes from the Swedish government agency Statistics Sweden, also called SCB (2023). All the data was selected for the year 2019. The "Municipality mean" column takes the overall mean over the mean of each municipality. If not otherwise specified, the number of observations in a category have been divided by the total municipality population to obtain the municipality mean. For "Female", population data divided into in one year intervals for the ages 0 to 100+ has been downloaded and summed up for each municipality. The data for "Age" includes the age mean of each municipality for the overall population. This variable already reports the mean age for each municipality and does not need to be divided by the municipality population to obtain the municipality mean. For "University education", data in one year intervals for the ages 16 to 95+ has been downloaded. We took three categories of education levels that indicated education past secondary education: Post-secondary education less than 3 years, post-secondary education 3 years or more and post-graduate. Our variable added the amount of people that were in these three categories together and limited them to 23 years of age or older for each municipality. To obtain the municipality mean for university education, we divided this generalized post-secondary education variable by the municipality population above or equal to 23. The variable "Married" is based on data that has four categories for every age: Unmarried, married, divorced and widowed. The "Municipality mean" column shows the mean for those included in the category married divided by the population above 18 years old as it's only legal to marry in Sweden after turning 18. For "Immigrant background", the data is presented in 5 years intervals from the age of 0 to 95+. Immigrant background is defined as if a person was born outside of Sweden or if both parents were born outside of Sweden. The data for "Income" takes information for everyone above 20 within three categories: Mean income in thousand SEK, median income in thousand SEK and total sum of income in million SEK. We focused on the mean income per municipality, which also includes salary and pension income from other Nordic countries. We did not have to divide by any specific municipality population as the reported income was already the mean income for everyone older than 20 years old. The only variable with another source than SCB is "Big city municipality", which was downloaded from Tillväxtverket (2021). This variable categorizes municipalities into three definitions: Rural municipalities, mixed municipalities and big city municipalities. The categorization was updated in 2021, which is the version we are using. A municipality is a big city municipality if at least 80% of the municipality live in densely populated areas and the municipality also shares a combined area with other municipalities with at least 500000 inhabitants.

The "National mean" column takes the overall mean over all of Sweden and includes the same variable definitions as within the municipality mean. If not otherwise specified, the amount of people in a category have been divided by the total population of Sweden to obtain the national mean. The total population of Sweden in 2019 was 10.32 million inhabitants. For "Female", the total female population of Sweden was divided by the total population of Sweden. The data for "Age" includes the age mean for the overall population. This variable already reports the mean age for Sweden and does not need to be divided by the total population to obtain the national mean. The variable "University education", includes the same three categories for post-secondary education as for the municipality mean and is limited to people above or equal to 23. It as also divided by the total population above 23 to obtain the mean. The same goes for the variable "Married" for the Swedish population above 18 years old. For "Immigrant background", the same definition applies that someone has an immigrant background if this person was born outside of Sweden or if both parents were born outside of Sweden. The variable "Income" is reported as the mean for everyone above or equal to 20. The indicator "Big city municipality" on the national level was created by linking municipality population sizes to each municipality and then collapsing the indicator to its mean with weights on the municipality population.

A.6.2 Creation of doctor shifts for productivity regressions

The shift variables for the doctor shifts are based on the doctor shift sample (see Section A.3) and have been constructed in several ways. The most basic approach for a shift is to take the start time of the first consultations and the end time of the last consultations of a doctor within 24 hours in a calendar day. The time in between the patient consultations varies, but can be in- or excluded from the shift. The time in between the patient consultations varies, but positive times between the consultations have been termed as breaks. Breaks longer than one hour make up 1.63% of the sample. It is important to note that we don't know whether the times between the meetings are actually breaks, waiting times or if the doctor was not working at all. Surprisingly, negative duration times between the consultations also exist. In total, 35.01% of the sample contain negative times in between the consultations. For the online consultations, the negative times between consultations are 34.89% and for in-person consultations 48.61%. The first meeting of a doctor in a day has no duration time in between consultations because there was no previous meeting yet. Based on this approach, there are various possibilities to look at shifts.

The first shift variable is just the start to the end time of a shift of a doctor with all the (positive) times in between the consultations still included. The second shift variable is the start to the end time of a shift of a doctor with all the (positive) times in between the consultations removed. The third shift variable is the start to the end time of a shift of a doctor, but all removed all (positive) times in between consultations if they were longer than one hour. The fourth shift variable takes a different approach and defines the shift as the sum of patient consultation times. The patient consultation time includes the patient facing time and the administration time for the doctor. Here, there is no decision to make about whether to include (positive) times in between the consultations or not. However, the negative times in between consultations could imply overlapping meetings. With this approach, this could result into a longer shift than actually worked and distort the patient rate.

For all shift variables, prescriptions renewals, test orders have not been taken into account from the start. Further, 8083 consultations without any duration time have been removed. The consultation duration time variable has been winsorized to take the minimum length of 2.58 minutes to avoid unrealistic small consultation duration times. A fraction of 1.02% had duration values below 2.58 minutes. We chose 2.58 minutes because it was the first percentile of the online consultation duration times, which presented a more probable minimum patient stay. Consultations that went past midnight have also been removed. The number of consultations that took longer than midnight is 1173 consultations. After removing those meetings, shifts in which doctors work both in-person and online (4653 consultations) are excluded as well. Consultations that took longer than midnight and also change work types sum up to 9 consultations. In total, 13909 meetings (1.1%) from the original sample without the previously mentioned meeting types were removed before creating the outcome variables for the doctor shift regressions.

The outcome variables were created on the shift level, which means they were estimated per doctor and per day. After collapsing from the meeting to the shift level, most of the outcome variables had no missing values with the exception of the break variables. As previously mentioned, this is because some shifts contain only one consultation and it is not possible to calculate a break without any previous consultation. There are 127 in-person shifts and 306 online shifts with only one consultation. For the control variables, only the gender and age variable for the doctor have both approximately 29% missing values on the shift level. The outcome "Score change over shift" is based on a patient score that is a 1 to 5 measure that patients rate their consultation with. It calculates the average score difference from the first hour of the shift to the last hour to approximate how exhausted a doctor gets during a shift. For in-person shifts, the score difference has 86.69% missing values, while there are 23.21% missing values for shifts online.

A.6.3 Creation of cost table

There are two important costs to consider when looking at digital health - the cost to the health provider and the cost to the patient. In our cost table, we attempt to approximate the costs of these actors to compare online and in-person services. The provider costs come from the same digital health source that we got our other online consultations data from.

The patient costs were calculated manually and are also an generalization of patient costs. The average fees for a primary care visit with a doctor are taken from the regions Stockholm and Scania because the majority of patients in our sample are located in these regions. Patients above the age of 18 and below the age of 85 (73.49% of our sample) have to pay the fee in Stockholm, while patients above 20 years old and below 85 years old (67.67% of our sample) have to pay the fee in Scania. The average percentage of paying patients over these two regions is 70.58%. The fees for online meetings changed for Stockholm in July 2023 to 100 SEK instead of 250 SEK (which is used in the cost table). The paying age was also increased to 20 years instead of 18 years (Kry, 2023). An amount of 95.49% in our drop-in sample didn't have to pay any fee. Paying the fee used to depend on the region in Sweden and our digital health provider was based in a region where no fees applied. This is not the case anymore, as patients have to pay based on their location now and not where the digital health provider gets their medical services from. It's probable that a lot more people have to pay a fee than in our sample. The waiting time in the doctor's office for in-person is an educated guess of 30 minutes based on Ekman (2018) and 15.31 minutes for online.

Transport includes commuting by car, public transport, biking and walking along with their respective probabilities that they are chosen (Rosberg and Enström, 2019). We took time and frequency averages from commuting to work to estimate how likely a patient takes a certain transport to the doctor's office. The commuting costs by car for primary care are based on fuel costs, transport fees, the commuting time and parking fees. The average time to primary care is 11.71 minutes one-way and 23.42 minutes two-way after including frequencies of commuting types. We assume a tempo of 60km/h on average for the fuel cost and a fuel use of 0.5 liter per km. The fuel price was taken from 01 January, 2020 and is 16.03 SEK per liter. The fuel cost is calculated for a 20 minute drive to the doctor and back. We calculate the fuel costs and multiply it with the probability of 57% that they use the car to get to work. Parking fees were taken as 15 SEK/hour on average and weighted by the patient-facing consultation time and multiplied by the probability that patients would take the car. Normally, parking fees differ by municipality and are often active only during the day, not during the night. The parking time is an educated guess: 5 minutes for finding a parking spot before going to the doctor's appointment and 5 minutes after the appointment (incl. the walking distance to the doctor's office). The transport fee is multiplied by the probability that the patient takes public transport to the appointment, which is 24%. We assume that most one-way tickets are valid 90 minutes after the purchase, so that patients are not forced to buy two tickets in total for their in-person doctor appointment. The public transport fee is the average price for a one-way ticket in Sweden in 2021: 30,91 SEK.

The average time to an ED is 15.53 minutes one-way and 31.06 minutes two-way. We assume that commuting to an ED is only done by car; therefore fuel costs are included as well. We assume no parking fees for an ED visit. The median stay time of a patient in an ED is 3.18 hours over all regions in Sweden.

A.6.4 Emergency Department visit categorization

The so-called Billings algorithm for emergency department (ED) visit classification by researchers from the New York University (NYU, 2023). assigns category percentages to four, five or six characters long ICD-10 codes (e.g. A001, C4A72 or E83118). The algorithm database NYU (2023) contains 69823 ICD-10 codes that are classified into nine categories. The categories for the ED algorithm were taken as our variables and are directly from the source. The nine categories consist of ED care needed and unpreventable, ED care needed and preventable, Emergent and primary care treatable, Non-emergent, Alcohol, Drug, Injury, Psych and Unclassified. They are all presented in percentages. The categories are based on almost 6000 New York emergency department records outside of our data, which include data on initial complaint, symptoms, medical history, age, gender, diagnoses and procedures/resources that were used in the emergency department. We added all the categories together in one variable, which confirmed that the categories sum up to 100 percent for each of the almost 70000 ICD-10 codes. However, the registry data we have from out-patient ED visits in Sweden only has three character ICD-10 codes (e.g. M90), which led us to truncate the longer ICD-10 codes from the algorithm into three character codes to match our data. We then collapsed the truncated codes to obtain the category mean for each of the nine categories.

In our out-patient dataset, we created a variable that indicates whether a patient had a first follow-up visit to an emergency department within 30 days. We then matched the truncated ICD-10 codes with our three character long ICD-10 codes, which resulted in a matching rate of 95.32%. After matching the ICD-codes, we linked the out-patient data to our patient main sample. Out of the total 4685 patients in our main sample, 216 patients (5%) had a follow-up visit to an emergency department within 30 days. Overall, 190 patients with ICD-10 codes out of these 216 patients matched with the truncated algorithm code. The other non-matching 26 patients had either Swedish ICD-10 codes (which are not part of the algorithm) or the codes were missing in general. For regressions, the variables have been restricted to the observations that were actually observed for the full 30 days after their first consultation.

A.6.5 Emergency Department distances to municipality centroids

The distance to an emergency department (ED) is an integral part for any patient that needs quick help. There are in total 71 EDs in Sweden (health service, 2023), but we previously excluded specialty clinics, e.g. for eyes from the list of EDs. Sweden has 290 municipalities within 21 regions, therefore, not every municipality has an ED. The distances to an ED in the more rural north of Sweden are bigger than in the south of Sweden. We found an address list of Swedish ED (health service, 2023), which we used to obtain geolocations (latitude/longitude) for each ED with help of the programming language Python. Then, we calculated the shortest (linear) distance of each ED for each of the 290 municipality centroids based on Vincenty's formula. The municipality centroid is the center point of a municipality and represents a generalization of location for every person living in that municipality. They were originally obtained as x-/y-coordinates, but transformed into latitude/longitude as well. The smallest distance is 1.09km for Jönköping, while the largest distance to an ED is 208.74km for Arjeplog. On average, the distance from an ED to a municipality centroid is 31.94km. Weighted by the amount of EDs each municipalities has, the average distance is 34.33km. It is important to note that the closest ED doesn't have to lie within the municipality. It's therefore possible that a municipality centroid could be closer to an ED outside of the municipality even if it has an ED located in the same municipality. We took the log of the calculated distances to make the measurement units more comparable.

We linked the ED distances of each municipality to the rest of our data and only focused on drop-in meetings. We then created a variable that indicated whether a patient had any ED visits within 30 days of the nurse meeting. To obtain the share of ED visits in a municipality, we added up each ED visits in a municipality and divided it by the total of the drop-in sample per municipality. The mean ED visit share for a municipality is around 5.41%. For municipalities without an ED, the average is approximately 5.21%. We again took the log of the ED visit share to ensure better comparability.

We also looked at the ratio of drop-ins per municipality. The ratio is the drop-in meetings divided by the total population per municipality. Two figures show the connection between the distance between municipality centroid and closest ED to the ED share and the drop-in ratio.

A.6.6 Follow-up ICD-10 codes compared to nurse meeting ICD-10 codes

The initial ICD-10 code given after a nurse meeting does not have to stay the same for the follow-up consultation of a patient if the patient has one. To check how if it's more probable to have the same ICD-10 code after online rather than in in-person consultations, we downloaded public ICD-10-CM data (SOURCE: XXXXX). These ICD-10 codes are 3 characters long and defined not for individual codes, but for broader categories (e.g. D50-D53 stands for "Nutritional anemias"). Follow-ups can either be an emergency department (ED) or a general practice (GP) follow-up within 30 days of the nurse meeting. There are only in-person follow-ups. Some ICD-10 codes in the public data end with a letter (e.g. C7B), but these don't fit the form of the ICD-10 codes in our data and therefore were termed as "Unclassified" (U99). Patients without any ICD-10 code were also given the category of "Unclassified". We linked every ICD-10 code in our sample to those categories conditional on whether they are actually within these broader categories. There are two versions of this approach: (1) We look at the three character letter level (e.g. H09) or (2) We look at the one letter character level (e.g. H). Every ICD-10 in our data was truncated if it didn't fit the level of ICD-10 code we were looking at.

Starting with the three character ICD-10 codes, the most frequent ICD-10 code category in the doctor sample is R50-R69 ("General symptoms and signs") with 10.57% of the sample. This is followed by M50-M54 ("Other dorsopathies") with 8.75% of the sample. Out of the 4662 patients in the doctor sample, 9 patients did not match with the public ICD-10 code categories. In contrast, all of the ICD-10 codes for ED follow-ups match with the public code. There are in total 216 ED follow-ups within 30 days in the doctor sample and the most common ICD-10 code with 11.57% is U99, which we termed as "Unclassified". The second most common ICD-10 group for the ED follow-up is R00-R09 ("Symptoms and signs involving the circulatory and respiratory systems") with 9.26% of the ED follow-ups. For the three character level, 37 ICD-10 codes were the same for the initial nurse meeting and for the ED follow-up within 30 days, which is 17.12%. The most common primary care follow-up ICD-10 code is R50-R69 ("General symptoms and signs") with 9.67%, followed by M50-M54 ("Other dorsopathies") with 7.74% of the sample. Out of 518 GP follow-up, 517 patients matched with the public ICD-10 categories. Overall, 228 from the 517 ICD-10 codes within categories are the same ICD-10 codes as in the nurse meeting, which is 44.1%. For the regression tables, we limited the observations to patients that could be observed for the full 30 days after the initial nurse meeting.

The one letter level ICD-10 code has the same amount of follow-ups in an ED or GP, but the matching rate for the categories is higher. We now have 81 same ICD-10 codes for nurse meetings and ED follow-ups within 30 days instead of 37 for the three character level. This is an increase of 20.37% for ED follow-ups. For GP follow-ups, we have 292 same ICD-10 codes for nurse meetings and follow-ups instead of 228. This is an increase of 12.38% for GP follow-ups. We again limited the observations to patients that could be observed for the full 30 days after the initial nurse meeting.

Data appendix references

Ekman, B. (2018), 'Cost analysis of a digital health care model in Sweden', *Pharmaco Economics Open* **2**, 347–354.

health service, D. (2023), 'List of emergency departments in Sweden from 1177'. Last accessed 08 August 2023.
URL: *www.1177.se/Hitta-vard/?st=c4b57c74-b1f7-4999-b2d1-5ffa0af5a421&nearby=false&s=&g=&lat=&lng=&caretype-list=Akutsjukvrd&location=&caretype=Akutsjukvrd&q=*

Kry (2023), 'Patient fees'. Last accessed 17 July 2023.
URL: *www.kry.se/priser-betalning/*

NYU (2023), 'Commonwealth fund-supported analysis of New York City electronic ED records by the NYU center for health and public service research and the united hospital fund of New York'. Last accessed 07 August 2023.
URL: *www.wagner.nyu.edu/faculty/billings/nyued-background*

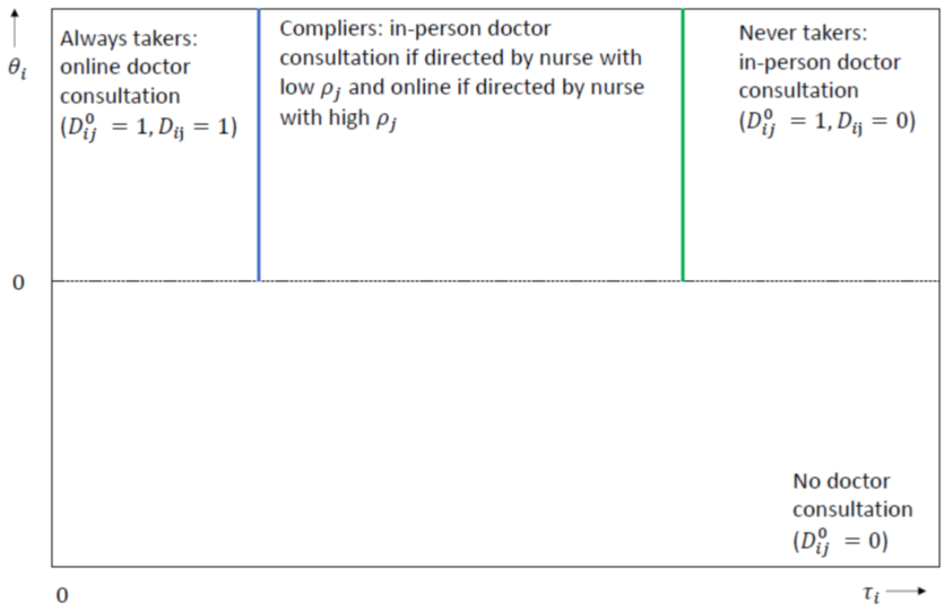
Rosberg, J. and Enström, D. (2019), 'Västsvenskar i rörelse – Samband mellan pendling och tillit', *Göteborgs universitet: SOM-institutet* pp. 113–128.

SCB (2023), 'Statistikmyndigheten SCB (Statistics Sweden) data'. Last accessed 2023.
URL: *www.statistikdatabasen.scb.se/pxweb/sv/ssd/*

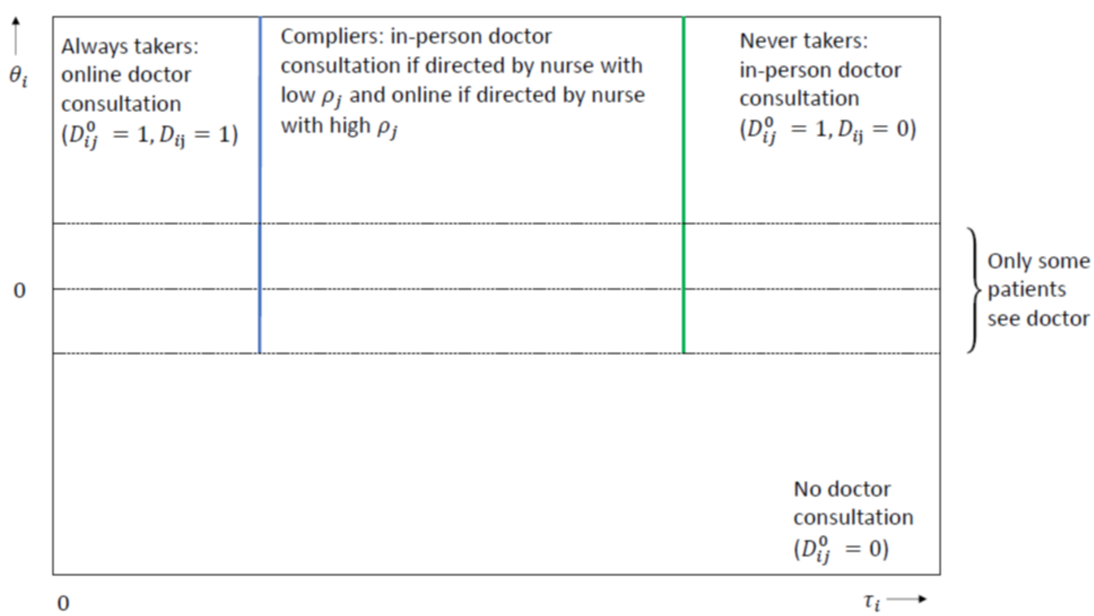
Tillväxtverket (2021), 'Municipality indicator data'. Last accessed 14 August 2023.
URL: *www.tillvaxtverket.se/tillvaxtverket/statistikochanalys/statistikomregionalutveckling/regionalaindelningar/staderochlandsbygder.1844.html*

Figure A1. Sorting of patients in model

(a) Without nurse mistakes



(b) With nurse mistakes



Note: This figure shows the distribution of the instrument $\widehat{\pi}_{ij}$ in the doctor sample.

Figure A2. Distribution of instrument

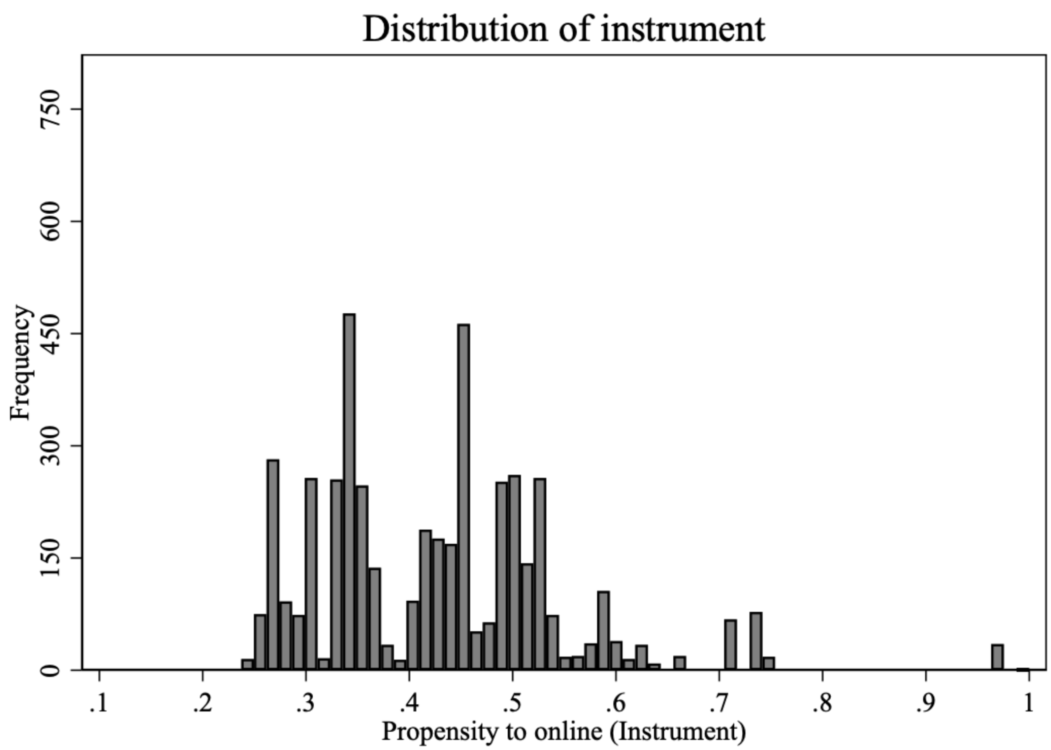


Table A1. Imposition of restrictions on primary sample

	In-person Count/(Percent)	Online Count/(Percent)	Total Count/(Percent)
All revisits with doctors	12017	93553	105570
	(11.4)	(88.6)	
+ Preceded by a nurse meeting	4605	51030	55635
	(8.3)	(91.7)	
+ Nurse meeting was online	4460	50987	55447
	(8.0)	(92.0)	
+ Patient was listed at clinic	2991	4154	7145
	(41.9)	(58.1)	
+ Patient did not seek care for Covid-19	2960	3636	6596
	(44.9)	(55.1)	
+ Clinic was open	2960	2389	5349
	(55.3)	(44.7)	
+ Removed clinics with very few observations	2953	2307	5260
	(56.1)	(43.9)	
+ Removed chlamydia and breastfeeding	2936	2294	5230
	(56.1)	(43.9)	
+ Removed infants	2922	2246	5168
	(56.5)	(43.5)	
+ Nurse has at least 20 observations	2668	1994	4662
	(57.2)	(42.8)	

Note: This table shows the number of observations after imposing restrictions on our primary sample. The three columns are divided into in-person and online meetings, which sum up to the meeting total. Each row in the table adds another restriction to the original sample. The observation count is shown for each row, while the percentages are in parentheses. The percentages of the last third row always add up to 100%. Prior to the first row, there is an implicit restriction that the revisit with the doctor was able to be accurately matched with a previous referring meeting. This is the case for 95% of the meetings, and for 99% of meetings preceded by a meeting with a nurse. The nurse meetings were predominantly online, but 145 nurse meetings were in-person. The next requirement for patients is to be listed at a clinic from our data provider. Such patients can go to an in-person consultation unlike unlisted patients. Patients also should not have gone to a consultation based on Covid-19 symptoms. Another condition for the sample is that the clinic was open at the time of the nurse meeting. We additionally removed centers with very few observations. Symptoms that were defined as chlamydia or breastfeeding were also removed. We decided to only keep patients above or equal to 2 years old as smaller infants might get treated differently by doctors. After imposing all restrictions, the table ends with 4662 observations, which we call the "doctor sample".

Table A2. Variable descriptions

Variable	Description
Total consultation time, minutes:	Provided by primary care provider data
Patient-facing consultation time, minutes:	Provided by primary care provider data
Administrative consultation time, minutes:	Total consultation time − Patient-facing consultation time
Days between meetings:	Number of calendar days between Nurse meeting and doctor consultation
doctor set an informative diagnosis:	ICD-10 code not in "R" (Symptoms (cough, rash, etc.)) or "Z" (Health status factors) categories.
Patient received a prescription:	Provided by primary care provider data
patient collected prescription within 30 days:	Patient is observed in the prescription data picking up a prescription we tied to the primary care meeting (see. section ?? for details)
doctor makes a specialist referral:	The internal meeting outcome of the primary care provider specifies there was a specialist referral. This is only applicable for Stockholm-based clinics
Patient score, 1-5:	The patient's score on a 1-5 scale of the meeting asked in a voluntary post-consultation survey. The best score is 5, while the worst score is 1.
Avoidable hospitalizations within 30 days:	We observe the patient in an in-patient hospitalization (they are in a ward for observation or treatment) within 30 days of the doctor's consultation where the ICD-10 code set at the hospital is from a list of ICD-10 codes known (**source??**) to have been preventable in primary care if caught earlier.
Any hospitalization within 30 days:	We observe the patient in an in-patient hospitalization (they are in a ward for observation or treatment) within 30 days of the doctor's consultation.
Any Emergency Department (A&E) visit within 30 days:	We observe the patient in an out-patient emergency department visit within 30 days of the doctor's consultation.
New visit to primary care provider within 30 days:	We observe the patient in a second doctor consultation within the primary care provider company within 30 days of the doctor's consultation. We restrict to actual doctor consultations, so we do not include meetings with psychologists, or prescription renewals or ordered tests.
Replacement for in-person answer:	The patients' response to a post-online consultation survey asking the patients whether the online consultation replaced an in-person consultation. Positive answers to the question are coded as 1, "Don't know" responses as 0.5, and negative responses as 0.
Doctor books a revisit within 30 days:	We matched the doctor's consultation to a doctor-booked revisit meeting within 30 days, which means that the second consultation is likely a follow-up from the first. See section **??** for details on the matching.

Table A2. Variable descriptions (continued)

Variable	Description
Shift (start to end w/o long breaks), hours:	Doctor online or in-person shift in hours. Start of shift is the first consultation and end of shift is the end time of the last consultation. Breaks longer than one hour have been removed. Consultations that took longer than midnight have also been removed, which restricts the shift to 24 hours in a calendar day. See section **??** for details on the definitions of shifts.
Shift (sum of duration times), hours:	Doctor online or in-person shift in hours. The duration time of a consultation includes patient-facing time and administration time. We then summed them up for this variable, after winsorising them. Consultations that took longer than midnight have been removed, which restricts the shift to 24 hours in a calendar day. See section **??** for details on the definitions of shifts.
Consultations per shift (shift based on start to end):	Total number of consultations per doctor online or in-person shift.
Consultations per shift (shift based on sum of duration times):	Total number of consultations per doctor online or in-person shift.
Consultations per hour (shift based on start to end):	Consultations per shift (shift based on start to end) / Length of shift in hours (shift based on start to end)
Consultations per hour (sum of duration times):	Consultations per shift (sum of duration times) / Length of shift in hours (shift based on sum of duration times)
Break (shift based on start to end):	The total positive time in between the start time of a consultation and the end time of the previous consultation. Breaks can only be calculated when the shift is defined from start to end time.
Longest break (shift based on start to end):	The longest positive time in between the start time of a consultation and the end time of the previous consultation in a shift.
Administration time:	Total consultation time − Patient-facing consultation time of the doctor. The same for all shift variables.
Total shift time from start to end, hours:	Total time from start of first consultation until the end of the last consultation. No breaks have been removed.
Score change over shift, -4 to 4:	The patient score difference of the first hour average to the last hour average of a shift. The difference is negative when the score has worsened over the course of a shift. The patient score is a 1 to 5 measure that patients rate their consultation with.

Table A2. Variable descriptions (continued)

Variable	Description
(In FE) Nurse 4h time blocks:	This is a factor variable categorizing the time of day of the nurse meetings for our primary sample (doctor sample) into one of six 4h blocks.
(In FE) Nurse day of the week:	This is a factor variable for the weekday during which the nurse meetings for our primary sample (doctor sample) were held.
(In FE) User listing center:	This is a factor variable for the primary care clinic at which the patient is registered. The data is from the primary care provider.
(In FE) Nurse year X Nurse month:	This is the interaction of factor variables for the year and the month during which the nurse meetings for our primary sample (doctor sample) were held.
(In Demographics) age and age^2:	The age and squared age of patients. We use the 2018 age, where we have set patients not yet born to age = 0.
(In Demographics) Foreign born:	Patient is a first-generation immigrant, that is, the patient was born outside of Sweden.
(In Demographics) Second generation immigrant:	Patient is a second-generation immigrant. The patient was born in Sweden, with both parents born outside Sweden.
(In Demographics) Born outside EU15 and Scandinavia:	The patient was born outside the EU15 countries and Scandinavia.
(In Demographics) Born outside EU15 and Scandinavia:	The patient was born outside the EU15 countries and Scandinavia.
(In Demographics) Married and divorced:	Two variables on the patients' marriage status in 2018. Children have a separate category for each variable.
(In Demographics) Employment status:	An indicator for whether the patient was in employment during 2018. This only applies if the patient is between 16 and 74; the remaining patients receive their own category.
Any comorbidity:	An indicator for whether the patient has been observed with any comorbidities prior to the sample period 2018-2020. The comorbidities are defined from the Elixhauser comorbidity index (**ref?**) using all in-patient and out-patient care data from 2013 to 2017.
Nurse-set ICD group:	This is the letter level category of the ICD-10 code set by the nurse who redirected the patient to the doctor consultation in our primary sample (doctor sample).

Table A2. Variable descriptions (continued)

Variable	Description
(In Time FE) Shift year X Shift month:	Interaction of factor variables for the year and the month during which the shift took place. The range goes from January 2019 to December 2020.
(In Time FE) Shift day of the week:	Interaction of factor variable for the day of the week during which the doctor shift took place (Monday to Sunday).
(In Doctor FE) Doctor ID:	The identification number of the doctor of the shift.
(In Demographics) Doctor female:	The gender of the doctor.
(In Demographics) Doctor age and age^2:	The age and age squared of the doctor.
(In Demographics) Doctor age X Shift online:	Interaction of factor variables for the age of the doctor and whether the shift was online.
(In Specifications) Non-EU language:	Indicates whether the doctor speaks any language outside of EU-15 (e.g. Russian or Turkish). EU-15 refers to the time back when the EU had only 15 members.
(In Specifications) Doctor specialty:	Indicates if the doctor is specialized in any of the 21 categories outside of primary care.
(In Specifications) Seniority doctor:	Indicates whether the doctor has any specializations further than a medical degree or is in training for a specialization.
(In Specifications) COVID-19 period:	Indicates in which COVID-19 time period the shift took place. The base category is pre-pandemic period.

Table A2. Variable descriptions (continued)

Variable	Description
ED care needed and unpreventable, percent:	ED care was required and ambulatory care treatment could not have prevented the condition.
ED care needed and preventable, percent:	ED care was required based on the complaint or procedures performed, but the emergent nature of the condition was potentially preventable if timely and effective ambulatory care had been received during the episode of illness.
Emergent and primary care treatable, percent:	Treatment was required within 12 hours, but care could have been provided effectively and safely in a primary care setting. The complaint did not require continuous observation, and no procedures were performed or resources used that are not available in a primary care setting.
Non-emergent, percent:	The initial condition, presenting symptoms, vital signs, medical history, and age indicated that immediate medical care was not required within 12 hours.
Alcohol, percent:	The condition involves primary diagnosis of alcohol abuse.
Drug, percent:	The condition involves primary diagnosis of drug or substance abuse.
Injury, percent:	The condition involves primary diagnosis of an injury.
Psych, percent:	The condition involves primary diagnosis of mental health problems.
Unclassified, percent:	The condition was not classified in any of the previously mentioned categories.
Sum of categories:	Adds up all nine previously mentioned categories. Sums up to 100 percent in total for each ICD-10 code.
First ED (A&E) visit within 30 days:	Patient has an out-patient ED visit within 30 days of the doctor's consultation.

Table A2. Variable descriptions (continued)

Variable	Description
Municipality:	Observation unit. Contains the 290 municipalities of Sweden that lie within 21 regions.
ED latitude:	The latitude of an emergency department.
ED longitude:	The longitude of an emergency department.
Centroid latitude:	The latitude of the center point of a Swedish municipality.
Centroid longitude:	The longitude of the center point of a Swedish municipality.
Shortest distance, km:	The distance of a municipality centroid to its closest ED.
Log of shortest distance:	The log of the distance of a municipality centroid to its closest ED.
ED visit share:	The share of ED visits in the sample for each municipality.
Log of ED visit share:	The log of the share of ED visits in the sample for each municipality.
Ratio:	Drop-in meetings divided by the total population per municipality.

Table A3. Key variable summary statistics

	Mean	SD	Min	Max	Observations
Consultation was online	0.43	0.49	0	1	4662
Instrument ($\widehat{\pi}_{ij}$)	0.43	0.12	0.24	1	4662
Propensity to redirect	0.54	0.092	0.28	0.85	4660
Nurse "mistake" share	0.11	0.048	0	0.33	4660
Patients with a comorbidity	0.19	0.40	0	1	4662
Has some university education	0.58	0.49	0	1	3396
Income > sample median	0.50	0.50	0	1	4112
Immigrant (non-EU15/Scandi.)	0.24	0.43	0	1	4661
Other physical health issue	0.30	0.46	0	1	4662
Period of low covid spread	0.49	0.50	0	1	4662
Patient female	0.49	0.50	0	1	4662
Age	33.0	13.7	0	85	4662

Note: This table presents summary statistics of some key variables used in our analysis. The propensity to redirect is the share of meetings the nurse redirected to a doctor. The nurse "mistake" share is the share of the nurses patients (within our sample) which were observed in A&E care within 10 days of the meeting. The median income in our sample is 27,416 USD (Jan. 2023 rate). "Other physical health issue" is a label suggesting the patient should see an in-person doctor, given by the algorithm that takes patient symptoms as inputs. suggested should see an in-person clinician. The low Covid-19 spread consultations are before 11th of March 2020 and between the 6th of July 2020 until the 24th of October 2020.

Table A4. First Stage Results

LHS variable: Online consultation				
Nurse propensity for online referrals	0.77	0.70	0.69	0.67
	(0.055)	(0.058)	(0.059)	(0.058)
	[0.053]	[0.057]	[0.058]	[0.056]
Fixed Effects		✓	✓	✓
Demographics			✓	✓
Has a comorbidity			✓	✓
Nurse-set ICD group				✓
Observations	4662	4662	4529	4516
F-stat	197	145	138	133
Clustered F-stat	211	153	141	141
Baseline mean	0.43	0.43	0.43	0.43

Note: This table is based on the doctor sample. The F-statistic is the Kleibergen-Paap rk Wald F statistic. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses and standard errors clustered by nurse are in brackets.

Table A5. Instrument independence

A. Balance of instrument on patient characteristics in nurses' sample

Demographics	✓	✓	✓	✓	✓	✓
Has a comorbidity			✓	✓	✓	✓
Nurse-set ICD group					✓	✓
Fixed Effects		✓		✓		✓
Joint test p-value	0.53	0.72	0.58	0.77	0.44	0.88
Observations	8639	8639	8639	8639	8609	8609

B: Balance of propensity to direct online on nurse propensity to redirect

Propensity to redirect	0.18	0.078
	(0.22)	(0.18)
Weighted by num. meetings:	Redirected	Total
Observations	62	62
Baseline mean	0.43	0.43

C: Balance of instrument on patient characteristics in doctors' sample

Demographics	✓	✓	✓	✓	✓	✓
Has a comorbidity			✓	✓	✓	✓
Nurse-set ICD group					✓	✓
Fixed Effects		✓		✓		✓
Joint test p-value	0.37	0.32	0.45	0.39	0.16	0.25
Observations	4529	4529	4529	4529	4516	4516

Note: This table is based on the doctor sample and shows tests on instrument independence. Panel A and Panel B show instrument balance tests on patient and nurse meeting characteristics. The p-values for the joint tests always control for the fixed effects when these are included. In Panel B, we collapse the sample to the doctor level. The estimates in Panel B show the correlations between the nurses' propensity to direct patients to online doctor consultations with the propensity to redirect to any doctor (in-person or online). We present two different weighting schemes on the estimates: (1) The total number of meetings observed by the nurse, and (2) The total number of patients a nurse has redirected to a doctor. Both schemes are conditional on our sample restrictions. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Table A6. Average exclusion table

A. Nurse meetings are short (and shorter than doctor consultations)

	Mean	Quartiles Q_{25}	Q_{50}	Q_{75}	Count
Nurse patient-facing time		2.5	4	6.1	4266
Doctor patient-facing time		4.2	12.6	30.1	4266

B. Nurse mistake share uncorrelated with instrument

Nurse 'mistake' share	-0.34	-0.31
	(0.33)	(0.31)
Weighted by num. meetings:	Redirected	Total
Observations	62	62
Baseline mean	0.43	0.43

Note: This table is based on the nurse sample with a total of 62 nurses and shows tests of the average exclusion assumption. In Panel A, we show the large difference in meeting duration, defined in minutes. In Panel B, we collapse the sample to the doctor level. The estimates in Panel B show the correlations between the nurses' propensity to direct patients to online doctor's consultations with the share of the nurses' patients who appeared in any hospital care (both A&E/ED and hospitalizations) during the ten days following the meeting ("Nurse 'mistake' share"). We define a nurse mistake as an instance where the nurse did not refer a patient to a doctor consultation, but the patient appeared in any emergency care within 10 days (both A&E/ED and hospitalizations). We present two different weighting schemes in Panel B: (1) The total number of meetings observed by the nurse, and (2) The total number of patients a nurse has redirected to a doctor. Both schemes are conditional on our sample restrictions. Robust standard errors are in parentheses.

Table A7. Average Monotonicity

	Patient female	Patient male	Age > median	Age ≤ median
Propensity for online	0.73	0.67	0.67	0.72
	(0.085)	(0.081)	(0.083)	(0.082)
Fixed Effects	✓	✓	✓	✓
Observations	2298	2364	2246	2416
First-stage K-P F-statistic	74	68	67	77
Baseline mean	0.45	0.40	0.42	0.43
	Uni. education	No Uni. education	Income > median	Income ≤ median
Propensity for online	0.70	0.69	0.80	0.59
	(0.087)	(0.083)	(0.085)	(0.091)
Fixed Effects	✓	✓	✓	✓
Observations	2126	2301	2055	2057
First-stage K-P F-statistic	65	68	88	41
Baseline mean	0.39	0.47	0.40	0.43
	Immigrant (non-EU15/Scandi.)	All other	Any comorbidity	No comorbidities
Propensity for online	0.76	0.69	0.82	0.67
	(0.12)	(0.068)	(0.14)	(0.065)
Fixed Effects	✓	✓	✓	✓
Observations	1140	3521	908	3754
First-stage K-P F-statistic	43	103	37	109
Baseline mean	0.41	0.43	0.46	0.42
	"Other physical health issue"	Not "Other physical health issue"	Low covid periods	All other
Propensity for online	0.80	0.65	0.67	0.72
	(0.10)	(0.071)	(0.10)	(0.071)
Fixed Effects	✓	✓	✓	✓
Observations	1410	3252	2299	2363
First-stage K-P F-statistic	61	85	43	103
Baseline mean	0.42	0.43	0.40	0.45

Note: This table shows the first stage of our IV in different sub-samples. For Panel A, the median age in our sample is 33, while Panel D also includes children. For Panel F, we use the median income in our sample which is 27,416 USD (Jan. 2023 rate). The district nurses in Panel B are specialized nurses who have undergone additional training and can carry some additional responsibilities. "Other physical health issue" is a label suggesting the patient should see an in-person doctor, given by the algorithm that takes patient symptoms as inputs. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Standard errors are in parentheses and clustered by the 62 nurses in our sample.

Table A8. In-consultation probability that patient scores the consultation

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Consultation was online	0.21	0.21	0.22	0.21	0.16	0.20	0.23	0.23
	(0.014)	(0.014)	(0.014)	(0.015)	(0.073)	(0.084)	(0.086)	(0.089)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4662	4662	4529	4516	4662	4662	4529	4516
First-stage K-P F-statistic					197	145	138	133
Baseline mean	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2

Note: This regression table is based on the doctor sample and shows the estimated probability that the patient answers the post-meeting survey to score the consultation based on the patient's satisfaction level. Each panel documents OLS and IV estimates with four different sets of controls each. The instrument of the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. For a description of the control variables we use, please see main text. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Table A9. Patient outcomes in 30 days after nurse meeting

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Avoidable hospitalizations within 30 days								
Consultation was online	-0.000093	-0.00014	-0.00029	-0.00039	0.0021	0.0019	0.0017	0.0022
	(0.0011)	(0.0010)	(0.0011)	(0.0011)	(0.0035)	(0.0052)	(0.0057)	(0.0053)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4048	4048	3939	3926	4048	4048	3939	3926
First-stage K-P F-statistic					146	104	97	92
Baseline mean	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
B: Any hospitalization within 30 days								
Consultation was online	0.0026	0.0024	0.0022	0.0020	0.033	0.034	0.037	0.039
	(0.0031)	(0.0030)	(0.0031)	(0.0033)	(0.018)	(0.021)	(0.023)	(0.023)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4048	4048	3939	3926	4048	4048	3939	3926
First-stage K-P F-statistic					146	104	97	92
Baseline mean	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008
C: Any Emergency Department (A&E) visit within 30 days								
Consultation was online	0.013	0.0096	0.010	0.012	0.12	0.11	0.12	0.12
	(0.0070)	(0.0071)	(0.0073)	(0.0078)	(0.044)	(0.053)	(0.056)	(0.058)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4048	4048	3939	3926	4048	4048	3939	3926
First-stage K-P F-statistic					146	104	97	92
Baseline mean	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044

Note: This regression table is based on the doctor sample. Each panel documents OLS and IV estimates with four different sets of controls each. The instrument of the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. For a description of the control variables we use, please see main text. In Panel A and Panel B, we show estimates of hospitalizations (i.e., when a patient was registered as an in-patient in a hospital). In Panel A, we restrict these hospitalizations to avoidable hospitalizations where the patient received a diagnosis that could have been handled by a primary care physician if caught earlier. In Panel C, we show estimates of emergency visits, which are out-patient hospital visits, i.e., Emergency Department (A&E). There are 37 hospitalizations, 5 avoidable hospitalizations, and 199 Emergency Department (A&E) visits in our sample. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Table A10. Patient outcomes in primary care 30 days after doctor consultation

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Doctor books a revisit within 30 days								
Consultation was online	0.10	0.10	0.10	0.11	0.17	0.22	0.23	0.24
	(0.012)	(0.012)	(0.013)	(0.013)	(0.070)	(0.082)	(0.085)	(0.088)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
B: Doctor books an in-person revisit within 30 days								
Consultation was online	0.098	0.100	0.10	0.11	0.20	0.25	0.25	0.26
	(0.010)	(0.011)	(0.011)	(0.012)	(0.064)	(0.075)	(0.077)	(0.079)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
C: Doctor books an online revisit within 30 days								
Consultation was online	0.0039	0.0025	0.0016	-0.00097	-0.032	-0.023	-0.018	-0.018
	(0.0065)	(0.0066)	(0.0068)	(0.0069)	(0.034)	(0.041)	(0.043)	(0.045)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
D: Patient initiated follow-up visit								
Consultation was online	0.025	0.031	0.034	0.040	0.027	0.076	0.10	0.11
	(0.014)	(0.015)	(0.015)	(0.016)	(0.080)	(0.094)	(0.098)	(0.10)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27

Note: This regression table is based on the doctor sample. Each panel documents OLS and IV estimates with four different sets of controls each. The instrument of the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. For a description of the control variables we use, please see main text. In Panel A, Panel B, and Panel C, we show estimates on whether the doctor booked a second meeting for the patient within 30 days. In Panel B and Panel C, we also condition on the consultation format. Panel D shows estimates on whether the patient contacted the primary care provider to book another meeting within 30 days. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Table A11. Medium-run results

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Avoidable hospitalizations after at least 30 days								
Consultation was online	-0.00027	-0.00030	-0.00027	-0.00020	-0.0027	-0.0021	-0.0024	-0.0017
	(0.00085)	(0.00087)	(0.00090)	(0.0010)	(0.0041)	(0.0053)	(0.0056)	(0.0056)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Mean days observed	111	111	111	111	111	111	111	111
B: Any hospitalization after at least 30 days								
Consultation was online	0.0022	-0.00028	0.000013	-0.0019	0.029	-0.0053	-0.0051	-0.0086
	(0.0041)	(0.0042)	(0.0043)	(0.0043)	(0.028)	(0.035)	(0.037)	(0.038)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.016
Mean days observed	111	111	111	111	111	111	111	111
C: Any Emergency Department (A&E) visit after at least 30 days								
Consultation was online	0.013	0.0069	0.0097	0.0083	0.13	0.053	0.082	0.085
	(0.0074)	(0.0074)	(0.0076)	(0.0079)	(0.046)	(0.053)	(0.055)	(0.058)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.048	0.048	0.046	0.046	0.048	0.048	0.046	0.046
Mean days observed	111	111	111	111	111	111	111	111
D: New visit to primary care provider after at least 30 days								
Consultation was online	-0.021	-0.035	-0.031	-0.029	0.22	0.0093	0.044	0.045
	(0.016)	(0.015)	(0.015)	(0.016)	(0.091)	(0.096)	(0.099)	(0.10)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
Mean days observed	111	111	111	111	111	111	111	111

Note: This regression table is based on the doctor sample. Each panel documents OLS and IV estimates with four different sets of controls each. The instrument of the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. For a description of the control variables we use, please see main text. In Panel A and Panel B, we show estimates of hospitalizations (i.e., when a patient was registered as an in-patient in a hospital). In Panel A, we restrict these hospitalizations to avoidable hospitalizations where the patient received a diagnosis that could have been handled by a primary care physician if caught earlier. In Panel C, we show estimates of emergency visits, which are out-patient hospital visits, i.e., Emergency Department (A&E). There are 66 hospitalizations in our sample, 3 avoidable hospitalizations, and 216 emergency visits. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

3 Auxiliary appendix

Table A12. NYU ICD-10 Emergency department (A&E) algorithm results

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Non-emergent								
Consultation was online	0.0026	0.0021	0.0023	0.0018	-0.0023	-0.0068	-0.0066	-0.0060
	(0.0017)	(0.0017)	(0.0018)	(0.0018)	(0.0076)	(0.0099)	(0.010)	(0.011)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.004	0.004	0.005	0.005	0.004	0.004	0.005	0.005
B: Emergent and PC treatable								
Consultation was online	0.0016	0.0013	0.0016	0.0021	0.023	0.029	0.030	0.032
	(0.0018)	(0.0019)	(0.0020)	(0.0021)	(0.013)	(0.017)	(0.017)	(0.018)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
C: ED needed, unpreventable								
Consultation was online	0.00031	-0.000027	-0.000027	0.00053	-0.00083	-0.0052	-0.0054	-0.0061
	(0.0012)	(0.0012)	(0.0012)	(0.0013)	(0.0064)	(0.0095)	(0.0098)	(0.010)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
D: ED needed, preventable								
Consultation was online	0.0011	0.00092	0.00082	0.00054	-0.0023	-0.0041	-0.0040	-0.0039
	(0.00086)	(0.00081)	(0.00085)	(0.00088)	(0.0041)	(0.0073)	(0.0076)	(0.0077)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: This regression table is based on the doctor sample shows results based on the NYU algorithm. Each panel documents OLS and IV estimates with four different sets of controls each. The instrument of the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. For a description of the control variables we use, please see main text. The outcome variables have the following descriptions. Non-emergent: The patient's initial complaint, presenting symptoms, vital signs, medical history, and age indicated that immediate medical care was not required within 12 hours. Emergent and Primary Care (PC) treatable: Based on information in the record, treatment was required within 12 hours, but care could have been provided effectively and safely in a primary care setting. ED needed, unpreventable: Emergency department care was required and ambulatory care treatment could not have prevented the condition (e.g. trauma). ED needed, preventable: Emergency department care was required based on the complaint or procedures performed/resources used, but the emergent nature of the condition was potentially avoidable if timely and effective ambulatory care had been received during the episode of illness (e.g. the flare-ups of asthma). More on the NYU variables can be found in the data appendix. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Table A13. NYU ICD-10 Emergency department (A&E) algorithm (continued)

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Psych								
Consultation was online	0.0019	0.0021	0.0020	0.00037	0.0021	-0.00085	-0.000044	-0.00092
	(0.0013)	(0.0013)	(0.0013)	(0.0011)	(0.0034)	(0.0044)	(0.0050)	(0.0053)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
B: Injury								
Consultation was online	-0.0012	-0.0017	-0.0018	-0.0013	0.030	0.033	0.036	0.037
	(0.0016)	(0.0016)	(0.0017)	(0.0018)	(0.014)	(0.016)	(0.017)	(0.018)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
C: Drug								
Consultation was online	-0.000062	-0.000071	-0.000079	-0.000071	0.00028	0.00033	0.00036	0.00038
	(0.000062)	(0.000071)	(0.000078)	(0.000071)	(0.00028)	(0.00033)	(0.00036)	(0.00038)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
D: Alcohol								
Consultation was online	-0.000062	-0.000071	-0.000079	-0.000071	0.00028	0.00033	0.00036	0.00038
	(0.000062)	(0.000071)	(0.000078)	(0.000071)	(0.00028)	(0.00033)	(0.00036)	(0.00038)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
E: Unclassified								
Consultation was online	0.010	0.0093	0.0092	0.012	0.071	0.068	0.074	0.077
	(0.0046)	(0.0045)	(0.0047)	(0.0052)	(0.029)	(0.035)	(0.037)	(0.038)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	4002	4002	3893	3880	4002	4002	3893	3880
First-stage K-P F-statistic					148	102	95	90
Baseline mean	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020

Note: This regression table is based on the doctor sample shows results based on the NYU algorithm. Each panel documents OLS and IV estimates with four different sets of controls each. The instrument of the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. For a description of the control variables we use, please see main text. More on the NYU variables can be found in the data appendix and in Table A11. The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Table A14. Meeting ICD codes vs Follow-up ICD codes (Three character level)

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Same ICD code as ED follow-up within 30 days								
Consultation was online	0.101	0.0968	0.113	0.131	-0.194	-0.149	0.0540	0.0802
	(0.0527)	(0.0582)	(0.0620)	(0.0688)	(0.284)	(0.301)	(0.343)	(0.334)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	189	189	187	186	189	189	187	186
First-stage K-P F-statistic					7	6	4	5
Baseline mean	0.108	0.108	0.109	0.109	0.108	0.108	0.109	0.109
B: Same ICD code as GP follow-up within 30 days								
Consultation was online	-0.00733	-0.0165	0.00145	-0.0150	0.125	0.157	0.130	0.257
	(0.0496)	(0.0530)	(0.0549)	(0.0562)	(0.255)	(0.268)	(0.293)	(0.288)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	436	436	420	418	436	436	420	418
First-stage K-P F-statistic					25	21	16	15
Baseline mean	0.436	0.436	0.437	0.437	0.436	0.436	0.437	0.437

Note: This regression table is based on the doctor sample. Each panel documents OLS and IV estimates with four different sets of controls each. The instrument of the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. For a description of the control variables we use, please see main text. Panel A shows whether an Emergency Department (ED) follow-up visit within 30 days had the same ICD-10 code as the previous first meeting. Panel B shows the same for General Practice (GP). The ICD-10 code's precision level is on the three character level (e.g. H09). The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Table A15. Meeting ICD codes vs Follow-up ICD codes (Letter level)

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Same ICD code as ED follow-up within 30 days								
Consultation was online	0.0733	0.0513	0.0576	0.0547	-0.0829	-0.0285	0.215	0.168
	(0.0699)	(0.0776)	(0.0802)	(0.0882)	(0.426)	(0.402)	(0.494)	(0.474)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	189	189	187	186	189	189	187	186
First-stage K-P F-statistic					7	6	4	5
Baseline mean	0.323	0.323	0.326	0.326	0.323	0.323	0.326	0.326
B: Same ICD code as GP follow-up within 30 days								
Consultation was online	0.0626	0.0572	0.0682	0.0621	0.441	0.471	0.421	0.488
	(0.0497)	(0.0533)	(0.0545)	(0.0565)	(0.243)	(0.273)	(0.308)	(0.310)
Fixed Effects		✓	✓	✓		✓	✓	✓
Demographics			✓	✓			✓	✓
Any comorbidity			✓	✓			✓	✓
Nurse-set ICD group				✓				✓
Observations	437	437	421	419	437	437	421	419
First-stage K-P F-statistic					25	20	15	14
Baseline mean	0.516	0.516	0.520	0.520	0.516	0.516	0.520	0.520

Note: This regression table is based on the doctor sample. Each panel documents OLS and IV estimates with four different sets of controls each. The instrument of the IV specifications is the propensity to online $\widehat{\pi}_{ij}$. For a description of the control variables we use, please see main text. Panel A shows whether an Emergency Department (ED) follow-up visit within 30 days had the same ICD-10 code as the previous first meeting. Panel B shows the same for General Practice (GP). The ICD-10 code's precision level is on the letter level (e.g. H). The baseline mean is the mean of the dependent variable for in-person doctor consultations. Robust standard errors are in parentheses.

Table A16. Summary statistics for doctor shifts

	In-person	Online
Hours worked (incl. all breaks)	5.03	5.19
Hours worked (excl. all breaks)	2.82	3.17
Consultations per hour (incl. all breaks)	1.81	3.68
Consultations per hour (excl. all breaks)	2.88	5.22
Consultations per shift	5.83	16.5
N	2046	76367

Note: This table reports mean summary statistics for doctor shifts in our sample, comparing in-person meetings to online meetings per doctor and shift. The observation unit is shift per doctor. The time unit is in hours. A shift starts with the first consultation and ends with the end time of the last consultation within 24 hours in a calendar day. The doctor sample was collapsed to the doctor level per day to obtain the shift level as the observational unit. More on the creation of doctor shifts can be found in the data appendix. Meetings with consultation times below 2.58 minutes were winsorized to the minimum of 2.58 minutes for both in-person and online. Meetings without consultation times have been removed. Meetings later than midnight and shifts with mixed working types (in-person and online) were not taken into account. The "In-person" column reports unweighted means of the sample observations for only in-person meetings. The "Online" column takes unweighted means from the same sample for only online meetings. "Consultations per shift" is the number of consultations with patients per shift. The table is not internally consistent: Multiplying the mean of hours worked with the mean of consultations per hour does not precisely lead to consultations per shift.

Table A17. Patient score change over doctor shift

	OLS							Specifications			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
A: Any shift variable											
Shift was online	0.094	0.090	0.010	0.077	0.073	0.072	-1.34	0.092	0.091	0.090	0.090
	(0.065)	(0.065)	(0.079)	(0.29)	(0.29)	(0.29)	(1.36)	(0.065)	(0.065)	(0.065)	(0.065)
Doctor female					0.013	0.012					
					(0.0078)	(0.0078)					
Doctor age						-0.0024	-0.035				
						(0.0024)	(0.030)				
Doctor age squared						0.000023					
						(0.000024)					
Online*Age							0.035				
							(0.030)				
Time fixed effects		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Doctor fixed effects			✓								
Non-EU language								✓			
Doctor specialty									✓		
Seniority doctor										✓	
COVID-19 period											✓
Observations	57914	57914	57914	39631	39631	39631	39631	57914	57914	57914	57914
Baseline mean	-0.11	-0.11	-0.11	-0.10	-0.10	-0.10	-0.10	-0.11	-0.11	-0.11	-0.11

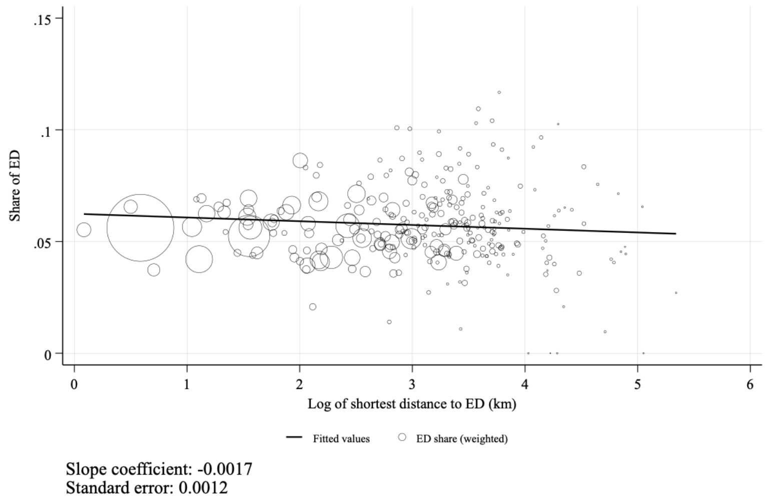
Note: This table shows results based on doctor shifts in the doctor shift sample. The doctor shift sample consists of registered and non-registered patients collapsed to doctor and day, excluding meeting types such as prescription renewals. The outcome variable is the difference between the patient score average of the first hour in the shift compared to the last hour in the shift. If the score difference is negative, the score has worsened over the course of a doctor shift. This outcome variable is based on a patient score that ranges from 1 to 5 and measures how patients rate their consultation. The patient score change is the same for all shift variables. See more on the creation of doctor shifts in the data appendix. Time fixed effects include Year*Month fixed effects and day of the week fixed effects. The base category for the former is January 2019 and for the latter, it is Sunday. Doctor fixed effects controls for the doctor of the shift and is only included in column 3. It includes age and gender and is dropped for the OLS regressions for age and gender. Doctor female controls for the gender of the doctor. Doctor age controls for the age of the doctor. This includes main effects and interactions terms for the age. The sample starting column 4 to column 7 is restricted to observations that are non-missing for age and gender of the doctor. Non-EU language indicates whether the doctor speaks any language outside of EU-15. The base category for the doctor speciality is primary care compared to other fields such as surgeons or neurologists. The base category for seniority doctor is a doctor that has only a medical degree compared to a specialist or a specialist in training. The base category for COVID-19 period is the pre-pandemic period. More on the control variables can be found in the data appendix. The baseline mean is the mean of the dependent variable for in-person doctor shifts. Robust standard errors are in parentheses.

Table A18. Consultations per hour - Alternative specifications

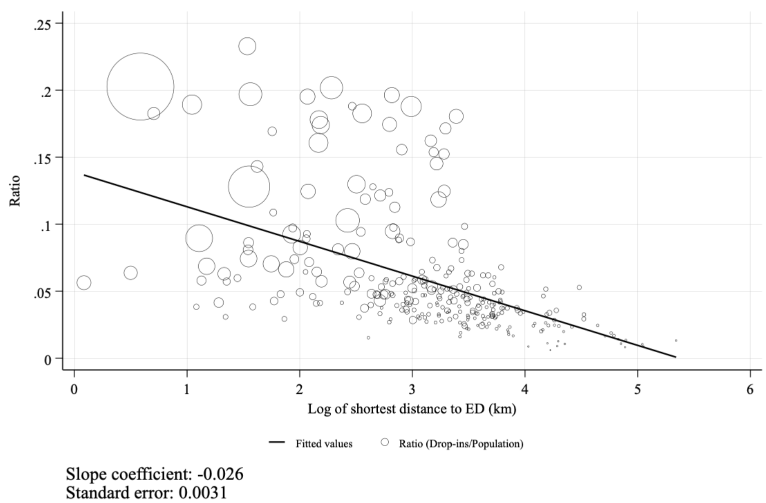
	OLS							Specifications			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
A: Shift goes from start to end with all breaks											
Shift was online	1.88	1.94	1.68	2.43	2.45	2.43	2.60	1.95	1.90	1.92	1.94
	(0.078)	(0.078)	(0.052)	(0.084)	(0.084)	(0.084)	(0.32)	(0.078)	(0.079)	(0.079)	(0.078)
Doctor female					-0.044	-0.047					
					(0.011)	(0.011)					
Doctor age						-0.011	0.00041				
						(0.0039)	(0.0082)				
Doctor age squared						0.000073					
						(0.000040)					
Online*Age							-0.0043				
							(0.0083)				
Time fixed effects		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Doctor fixed effects			✓								
Non-EU language								✓			
Doctor specialty									✓		
Seniority doctor										✓	
COVID-19 period											✓
Observations	78413	78413	78413	51956	51956	51956	51956	78413	78413	78413	78413
Baseline mean	1.81	1.81	1.81	1.40	1.40	1.40	1.40	1.81	1.81	1.81	1.81
B: Shift goes from start to end w/o any breaks											
Shift was online	2.34	2.47	2.00	2.99	3.01	2.94	3.22	2.47	2.47	2.46	2.47
	(0.084)	(0.084)	(0.051)	(0.13)	(0.13)	(0.13)	(0.60)	(0.084)	(0.085)	(0.085)	(0.084)
Doctor female					-0.058	-0.070					
					(0.013)	(0.013)					
Doctor age						-0.059	-0.013				
						(0.0046)	(0.013)				
Doctor age squared						0.00042					
						(0.000047)					
Online*Age							-0.0067				
							(0.013)				
Time fixed effects		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Doctor fixed effects			✓								
Non-EU language								✓			
Doctor specialty									✓		
Seniority doctor										✓	
COVID-19 period											✓
Observations	78413	78413	78413	51956	51956	51956	51956	78413	78413	78413	78413
Baseline mean	2.88	2.88	2.88	2.43	2.43	2.43	2.43	2.88	2.88	2.88	2.88

Note: This regression table shows results based on the doctor shift sample. The outcome variable is consultation per hours for different shift variables. A shift starts with the first consultation and ends with the end time of the last consultation within 24 hours in a calendar day. Breaks are positive times in between the consultations. More on the creation of doctor shifts can be found in the data appendix. Time fixed effects include Year*Month fixed effects and day of the week fixed effects. The base category for the former is January 2019 and for the latter, it is Sunday. Doctor fixed effects controls for the doctor of the shift and is only included in column 3. It includes age and gender and is dropped from the OLS regressions for age and gender. Doctor female controls for the gender of the doctor. Doctor age controls for the age of the doctor. This includes main effects and interactions terms for the age. The sample starting column 4 to column 7 is restricted to observations that are non-missing for age and gender of the doctor. Non-EU language indicates whether the doctor speaks any language outside of EU-15. The base category for the doctor speciality is primary care compared to other fields such as surgeons or neurologists. The base category for seniority doctor is a doctor that has only a medical degree compared to a specialist or a specialist in training. The base category for COVID-19 period is the pre-pandemic period. More on the control variables can be found in the data appendix. The baseline mean is the mean of the dependent variable for in-person doctor shifts. Robust standard errors are in parentheses.

Figure A3. Share of ED visits and shortest distance to an ED per municipality



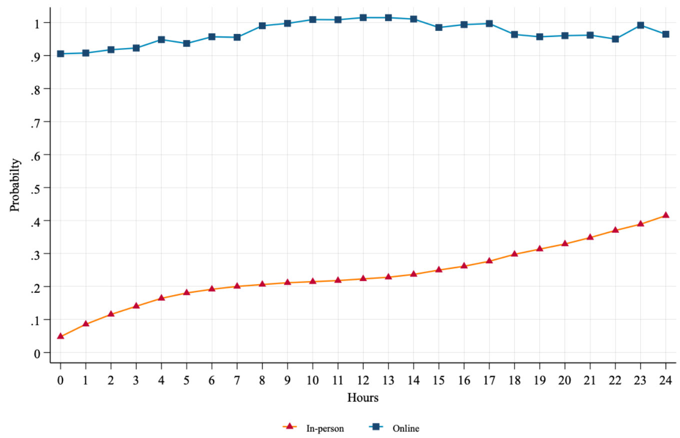
(a) Share of ED visits and shortest distance to an ED per municipality



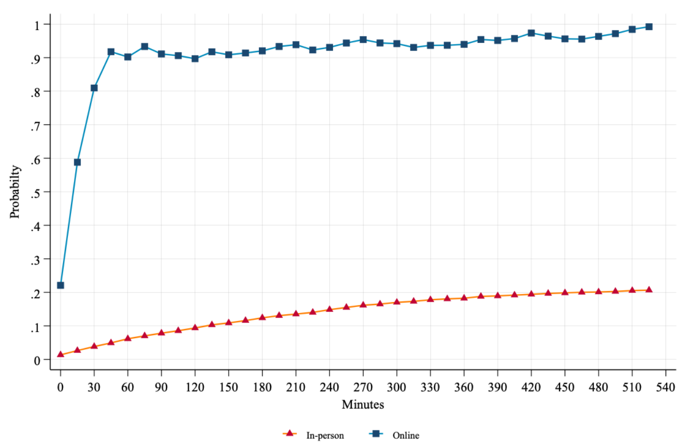
(b) Ratio of drop-ins and shortest distance to an ED per municipality

Note: This figure is based on the doctor sample. It shows the link between the log of the distance of a municipality centroid to its closest ED and (a) the share of ED visits in the sample for each municipality or (b) the ratio of drop-ins per municipality. The scatter plots are weighted by municipality populations.

Figure A4. Time from nurse meeting to doctor consultation



(a) Time from nurse meeting to doctor consultation, hours



(b) Time from nurse meeting to doctor consultation, minutes

Note: This figure is based on the doctor sample. It shows the probability of time between nurse meeting to doctor consultation (a) within 24 hours in 1 hour intervals and (b) within 9 hours in 15 minute intervals. The probabilities are taken from the IV regression with all controls and can be above 1 for online due to over-saturation of the model. At 0, the graph shows the probabilities if the patient has (a) less than 1 hour or (b) less than 15 minutes between nurse meetings and doctor consultation. The probabilities for in-person meetings are marked with triangles, while online meetings are characterized with squares.

Table A19. Shortest distance from municipality centroid to ED regressions

	(1) ED share (Unweighted)	(2) ED share (Weighted)	(3) Ratio (Unweighted)	(4) Ratio (Weighted)
Log distance	-0.0017	-0.00016	-0.026	-0.033
	(0.0012)	(0.00066)	(0.0031)	(0.0091)
Constant	0.062	0.056	0.14	0.19
	(0.0035)	(0.0016)	(0.011)	(0.024)
Observations	290	290	290	290
R-squared	0.007	0.000	0.267	0.282

Note: This table is based on the sample of (un-)registered online drop-ins collapsed to the municipality level in Sweden. Log distance is the log of the distance of a municipality centroid to its closest Emergency department (ED). The ED share is the share of ED follow-ups after the drop-ins within 30 days for each municipality. Ratio denotes the drop-ins per population for each municipality. The weights in Column (2) and Column (4) are the population size per municipality. Robust standard errors are in parentheses.