Does Division of Labor Increase Productivity? Evidence from Primary Care*

Amanda Dahlstrand[1], Shan Huang[2], Guy Michaels[3], and Nestor Le Nestour[4]

[1] *University of Zurich,* [2] *Stockholm School of Economics; University of Copenhagen,* [3] *London School of Economics,* [4] *Stockholm University*

November 15, 2024

PRELIMINARY - PLEASE DO NOT POST OR CIRCULATE

Abstract

The idea that the division of labor increases productivity is central to economic analyses of countries, industries, social structures, and occupations within organizations. We study the division of labor within an online primary healthcare organization, where an algorithm assigns patient cases between two clinician occupations: nurses and doctors. We compare a knowledge hierarchy, where nurses resolve some cases themselves while referring others to doctors, to a more traditional route where cases are assigned directly to doctors. We analyze about 500,000 cases, using an identification strategy which leverages temporary congestion that increases the odds of cases being assigned to nurses. We find that in the knowledge hierarchy, nurses resolve 70% of cases at the margin, sending the rest to doctors. While this knowledge hierarchy slightly reduces the rates of meaningful diagnosis and prescription, it has no adverse effects on patients' satisfaction, acute care utilization, labor earnings, or mortality. The knowledge hierarchy reduces total costs to the healthcare system by up to 20% without compromising quality at the primary care organization or downstream. We further explore the patterns of comparative advantage of the knowledge hierarchy across case types and the extent to which the allocation of tasks follows comparative advantage.

*We thank Ellen Donath for excellent research assistance. We also thank many colleagues and seminar and conference participants at the HEFUU workshop 2024 in Uppsala, the Joachim Herz Foundation Add-on Fellowship Meeting in Hamburg, the Stockholm School of Economics, the London School of Economics, and the URPP Annual Retreat at the University of Zurich for their insightful comments and feedback. We are grateful to the healthcare company for granting access to their data. We thankfully acknowledge financial support from the ESRC's Centre for Economic Performance and the Swiss National Science Foundation, grant number 100018-228312, and support from the University of Zurich's Research Priority Program "Equality of Opportunity". The research project was approved by the Stockholm Ethics Review Board.

The greatest improvement in the productive powers of labour... seem to have been the effects of the division of labor. — (Smith 1776)

1 Introduction

The division of labor and resulting occupational specialization have long been central to economic studies of productivity (Smith 1776). While occupations have been shaped by history and institutions, technological advancements and decreased communication costs are transforming how work is organized. In modern knowledge-intensive industries, where tasks are complex and communication costs are low, one way to divide labor is through *knowledge hierarchies* (KH), in which less specialized workers perform routine tasks and refer more complex cases to specialized experts (Garicano 2000).[1] But whereas knowledge hierarchies are common today, empirical evidence on their impact on productivity is limited. In this paper, we ask: What are the productivity effects of a knowledge hierarchy compared to sending tasks directly to specialists? And to the extent that the two routes differ in their comparative advantage, does the allocation of tasks between them follow this comparative advantage? We study these questions in the context of healthcare.

Understanding the productivity effects of the knowledge hierarchy is important in healthcare, where costs are rising in aging societies (World Health Organization 2021), and there are acute shortages of specialized medical personnel (Maier and Aiken 2016).[2] Meanwhile, innovations in information technology reshape the costs of coordination and the distribution of expertise in healthcare (Dahlstrand 2024; Dahlstrand et al. 2024; Bronsoler et al. 2022; Rajpurkar et al. 2022). Our study focuses on one lever used to cope with the challenges of modern healthcare: reorganizing the allocation of tasks between healthcare occupations.

To compare the productivity of a knowledge hierarchy to that of sending tasks directly to experts, we need a common set of tasks to be sent via both routes; an empirical strategy to overcome sorting in the allocation of tasks across routes; and a way to measure health outcomes and costs. To this

[1]Knowledge hierarchies arise where higher layers rely on judgment from lower ones to transfer tasks, for example between kitchen workers and chef de cuisines, or in manufacturing between machine operators, supervisors, and engineers (Garicano 2000).

[2]The US, for example, faces a predicted shortage of 86,000 medical doctors by 2036, about one-tenth of the 2021 workforce. See https://www.aamc.org/advocacy-policy/addressing-physician-workforce-shortage (last accessed: 12 Nov 2024).

end, we study a large primary healthcare organization in Sweden, where tasks are patient cases with different combinations of symptoms, demographics, and medical histories, which are handled by doctors and nurses working online. An algorithm allocates some tasks directly to doctors and others to nurses, who may then direct some to doctors. Conditional on case characteristics, the allocation of some tasks between the two routes depends on temporary congestion, which allows us to identify the relative effect of the knowledge hierarchy route. Such congestion builds up and dissipates quickly at the online provider resulting in the allocation to nurses of tasks that would otherwise have been assigned to doctors. We use detailed data on patient health outcomes, labor market outcomes, and costs of healthcare to measure the productivity effect of the knowledge hierarchy and any tradeoffs that it may entail.

We analyze approximately 500,000 patient cases at the margin of being assigned to a nurse. We link each case to comprehensive data from the provider, as well as administrative data on healthcare utilization outside the provider, along with patients' demographic and socioeconomic information. Our rich dataset, along with the empirical strategy exploiting the algorithmic assignment mechanism, allows us to assess the causal effects of the nurse-initiated knowledge hierarchy versus direct-to-doctor assignments on the provider, patients, and the cost burden on the healthcare system. Our design overcomes several common challenges in empirical studies of task organization. First, we use highly granular data on task assignments to assess fine-grained quality and cost metrics within the primary healthcare process. Second, our setting includes a common set of tasks that can be allocated differently, allowing us to compare different work modes. Finally, we make use of an assignment algorithm in order to causally identify the implications of a knowledge hierarchy structure, independently of patient health conditions or preferences which typically drive case assignments in healthcare settings.

Our empirical strategy leverages quasi-exogenous variation from congestion to doctors in two approaches. First, we assume the conditionally random assignment to clinician occupations in a simple linear regression (OLS) model that controls for detailed case characteristics and uses finely-grained date-by-four-hour windows to capture timing. This approach assumes that the algorithm bases assignments only on a limited set of baseline case characteristics, along with quasi-exogenous variation in congestion that otherwise similar cases experience around the same time. However, as the algorithm's exact functional form is unknown, we may be unable to fully exclude residual selective sorting of cases. To address this, our second approach uses an instrumental variables (IV) strategy, where we include the share of prior nurse consultations as a direct measure of congestion to instrument a current case's assignment. Our IV approach relies on the plausible assumption that patient demand in response to staffing is not fully predictable and that cases previously assigned to nurses effectively capture an increase in doctor congestion also relevant to the current case.

Even though the marginal cases we study are handled by both nurses and doctors, these occupations are trained very differently in Sweden. Physicians complete six years of medical school followed by up to ten years of residency and specialization, whereas nurses follow a three-year program, with most not pursuing further specialization. While nurses frequently engage in triage and patient communication, their scope of practice is more limited: For example, nurses in primary care do not have authorization to prescribe medications or refer to specialists. However, they may take on tasks less commonly performed by medical doctors, such as providing socio-emotional support or offering advice on self-care. Yet, in practice, many patient cases may be solved by either nurses, doctors, or both, depending on organizational decisions that are a black box to the researcher, and we do not know whether these decisions represent the most efficient use of available expertise.

Our findings indicate that the large majority of cases in the knowledge hierarchy, about 70%, are directly resolved by nurses without a referral upwards to doctors, while the remaining 30% of the cases are referred to doctors. We also find little evidence that being assigned to the knowledge hierarchy causes patients to seek out doctors independently. We find no evidence of any increase in return visits due to being assigned initially to a nurse. Based on primary care data observable for one region in Sweden, we observe that the rate of external primary care doctor visits (outside the firm we study) increases by at most 2.2 percentage points (from 12%). Moreover, patients are no less satisfied in the nurse-initiated knowledge hierarchy route – if anything, we observe a slight improvement in patient ratings during their care episode at the provider.

Does the knowledge hierarchy restrict access to services requiring specialized doctor skills? Our results show that access is modestly reduced. Cases in the nurse-assigned hierarchy experience are solved with a lower diagnostic specificity. Likewise, patients experience a moderate decline in prescription drug access, with the rate of new prescriptions decreasing by up to 10 percentage points from a 44% baseline. However, our estimates also imply that many patients in nurse-assigned cases, at least 34%, still receive a prescription for their case, suggesting that many referrals up in the knowledge hierarchy may be purely due to occupational licensing restrictions in prescribing.

While the nurse-initiated knowledge hierarchy may limit access to certain doctor-specific services, a reduction in their utilization only suggests lower quality if patient health outcomes worsen. To assess whether the knowledge hierarchy negatively impacts patients, we examine the use of acute care services such as urgent care visits, emergency department visits, and hospitalizations. These outcomes allow us to capture imminent, adverse events during patients' care journeys. We do not observe clear increases in urgent care or emergency department visits due to an initial nurse assignment. We also find little evidence of any increases in hospitalizations, suggesting that, overall, the knowledge hierarchy does not substantially increase short-term acute care needs.

While a patient's care journey allows us to understand patterns of utilization and the occurrence of adverse health outcomes, they would not capture a worsening in health that the patient does not seek care for. To further understand potential adverse outcomes, we thus consider measures that capture adverse medium and long-term impacts on patients' lives, including income reductions and mortality. Our results show that the nurse-initiated knowledge hierarchy, if anything, slightly reduces major income reductions following the initial consultation by 2.1 percentage points (from 22%). Furthermore, we find no significant effect on three-year mortality. These results suggest that the nurse-initiated knowledge hierarchy has no detrimental effects on longer-term economic well-being or mortality.

Our results demonstrate that the knowledge hierarchy in primary care does not adversely affect almost any patient-centered outcome measure we consider. In the final step, we thus turn to its cost implications. We consider the costs of the initial consultation and the downstream healthcare services as imposed on the public payer, who covers the majority of healthcare expenses in Sweden. Relative to a total average cost of 988 SEK per care episode (approximately 90 USD) following direct-to-doctor assignments, we observe marginal cost savings of 12% (OLS) to 20% (IV) in the knowledge hierarchy. After adjusting for the less precisely measured external primary care usage, savings remain at least 7.5% (OLS). These savings primarily result from nurses independently managing 70% of cases and the lower cost of nurse consultations, which cost less than two-thirds of doctor consultations. Overall, the nurse-initiated knowledge hierarchy reduces healthcare costs moderately without negatively impacting patient outcomes – thus suggesting productivity gains.

The concept of specialization and task alignment underpins much of the literature on skill allocation and productivity (Acemoglu and Autor 2011; Jones 2009; Becker and Murphy 1992; Katz and Murphy 1992). Garicano (2000) emphasizes how within organizations, knowledge hierarchies can facilitate efficient task allocation by enabling firms to leverage individual, differentiated skills effectively. Our study empirically tests these principles within a healthcare organization, examining how task alignment among a knowledge hierarchy formed by nurses and doctors with distinct roles impacts patient outcomes and costs.

In providing empirical evidence from within an organization, we also contribute to the large, seminal literature in economics of comparative advantage (Ricardo 1817; Heckscher 1919; Ohlin 1924; Dornbusch et al. 1977; Costinot 2009). The classic concept of comparative advantage has been quintessential for theories on economic transactions, such as for international trade theory. While the underlying micro-foundations typically assume that firms adhere to comparative advantage principles, empirical evidence on its application within organizations remains sparse. Our setup offers a rare opportunity to provide novel insights into how organizations may implement comparative advantages between occupations with differing degrees of specialization in practice. We assess the potential efficiency gains of flexible task allocations rather than strict regulatory scopes at the margin.

A recent body of economic research highlights the productivity implications of improved allocative efficiency among managers (Weidmann et al. 2024), for differentiated skills in education (Biasi et al. 2021), or also in healthcare (Dahlstrand 2024). Yet, evidence specific to task allocation principles between specialized occupations remains limited. Recent studies have begun addressing this gap by evaluating productivity differences between healthcare occupations when scopes of practice become more aligned. Our study is closely related to Chan and Chen (2022), who find that doctors have a comparative and absolute advantage over nurse practitioners (an occupational category in the US which is in between a nurse and a doctor) in the emergency department setting. In contrast, Currie and Zhang (2023) focus on the effectiveness of primary care clinicians in the primary setting, where physicians appear to be somewhat less effective than nurse practitioners in reducing hospitalizations and emergency department visits. Nurses in our setting, as opposed to nurse practitioners in the US, do not compete directly with doctors as their scopes of practice differ. Yet, their tasks may overlap. For example, Almström et al. (2024) shows evidence of task-shifting between nurses and doctors in response to changes in relative reimbursement for large privately owned healthcare organizations in Sweden. Instead of a head-to-head comparison between professions that highlights their overlapping abilities, our study takes a unique perspective by focusing on a knowledge hierarchy where nurses have discretion to solve or refer cases, exploring how nurses can both complement and substitute doctors.

Within the broader health economics literature, we also relate to work on human capital as a crucial factor in healthcare productivity and inefficiency. Previous studies highlight the need for specialization in the provision of healthcare (Arrow 1963; Chandra and Staiger 2007; Doyle et al. 2010; Currie and MacLeod 2017). Crucially, misallocation of specialized resources is found to be the main driver of inefficiencies and rising healthcare costs (Baicker and Chandra 2004; Chandra et al. 2023). The way differentiated, specialized skills are coordinated can thus have major impacts on efficiency in healthcare (Chan 2016; Silver 2021; Dahlstrand 2024; Kelly et al. 2023). We extend the existing discussion by investigating how the knowledge hierarchy structure may reduce inefficiencies through the organization of clinical work between occupations. In our context of primary care, we focus on how task-skill alignment across occupations can shape the clinical path of patients already at their first contact with the healthcare system in a care episode.

Finally, the literature on occupational licensing and scope-of-practice reforms provides a background for changes in historically entrenched task allocations within healthcare (Shapiro 1986; Kleiner 2016; Dillender et al. 2024). Various studies have demonstrated that reductions in licensing restrictions for nurse practitioners have improved access to primary care, mental health, substance abuse treatment, and reduced mortality (Traczynski and Udalova 2018; Alexander and Schnell 2019; Guo et al. 2024; McMichael 2023). We do not study licensing changes or expansions in the scope of practice of nurses; instead, we study a set of tasks that nurses have always had the ability and right to solve but might, in practice, be considered work to be performed by doctors. This approach allows us to investigate the potential for efficiency gains through a structured task allocation in healthcare in a nurse-initiated knowledge hierarchy.

The remainder of the paper is organized as follows. Section 2 introduces the institutional background and our data. Section 3 describes our empirical framework. Section 4 shows and discusses our main results on the productivity effects of the nurse-initiated knowledge hierarchy. Section 5 presents additional evidence on the mechanisms through which the knowledge hierarchy operates. Section 6 concludes.

2 Background and data

To study the effects of division of labor, we examine how patient cases are allocated at an online primary care provider in Sweden. In our setup, patients can either be assigned to doctors, or they can be managed in a knowledge hierarchy structure under a division of labor between nurses and doctors. Our empirical setting provides several advantages. First, the healthcare firm's granular data on case assignments can be linked to regional and national administrative databases that allow us to assess treatment quality and costs. Second, our setting accommodates a common support of tasks that can handled by both doctors and nurses. Finally, we make use of internally set organizational rules in order to causally identify the implications of task allocation to a knowledge hierarchy.

2.1 Institutional setting

In order to study task allocation within a firm, we focus on a large Swedish online healthcare organization that we refer to as *the firm* or *the provider*.[3] This healthcare provider employs two types of clinicians for primary care consultations: *nurses* and *doctors*. Through these consultations, both types of clinicians solve cases of patients seeking healthcare, which we also refer to as *tasks*.

Healthcare in Sweden is primarily publicly funded and has comprehensive and universal coverage. The healthcare system operates on a decentralized model, with regional and local authorities being responsible for organizing healthcare services and reimbursing providers for services delivered. The bulk of healthcare costs is carried by the regions and funded through taxes, whereas patient contributions are low.[4]

Primary care services are typically the initial point of contact between patients and the healthcare system, where clinicians diagnose and treat common ailments and conditions. When needed, patients

[3]Many primary care providers in Sweden are large organizations employing up to hundreds of individual clinicians, comparable to large-scale group medical practices (Almström et al. 2024). Telehealth consultations as offered by the provider we study are an increasingly popular approach to healthcare delivery. In the U.S., half of Medicare and Medicaid patients accessed telehealth services in 2020 (Centers for Medicare & Medicaid Services 2024). In Sweden, where data was first collected in 2021, there were 555 online primary care consultations per 1,000 inhabitants (Sveriges Kommuner och Regioner 2024). The firm we study is one of the leading digital healthcare providers in Sweden, handling a substantial share of all online consultations (we withhold exact market shares to maintain the provider's anonymity).

[4]The national insurance scheme involves small copayments for outpatient visits, including consultations in primary care but also specialty and emergency care visits. The total patient fees for healthcare visits are capped annually, during our sample period at around 1,150 SEK (approximately USD 125). Patient copayments for online primary care ranged between 100 and 200 SEK (about 11 to 22 USD in 2020) for adults, and consultations for children were free.

are referred to specialist or hospital services. Patients select a healthcare center, usually a large team practice with multiple clinicians, as their primary care provider and register with that center.

We study one of these large primary care centers. The provider offers in-person consultations for registered patients, but primarily focuses on delivering primary care online. Consultations at the provider are covered by Sweden's universal public health insurance. The firm can provide us with detailed records of its consultations and clinicians assigned to each patient case.

Patients can book consultations for themselves or their children via a mobile application. To access the service, they log in using an electronic identification system linked to their unique personal identity number (personnummer), which is used across Sweden's healthcare and government systems to identify individuals. This system allows the provider to retrieve key patient information, such as age, gender, and region, from civil records. After completing the identification process, patients select their symptoms from a predefined list and can provide additional information through a questionnaire or free text input. Consultations may be booked immediately as a drop-in or scheduled for a future time. Once the booking is confirmed, patients are prompted to wait while a clinician is assigned to their case.

Whether a patient becomes assigned to a nurse or a doctor is determined by an allocation algorithm. Our analysis focuses on marginal cases where either a nurse or a doctor can be assigned for the initial consultation. Historically, all consultations at the provider were handled by doctors, while nurses were primarily responsible for triaging patients or managing patient tracking (for a chlamydia screening program). However, starting in early 2019, the firm began to explore expanding the role of nurses, allowing them to handle patients with specific symptoms. Although doctors continued to manage the majority of these cases, during periods of temporary congestion among doctors, some patients were allocated to nurse consultations. Both nurses and doctors follow the same treatment protocols during consultations and, in principle, can resolve cases. Unlike most healthcare settings, the allocation of these marginal cases to either a nurse or doctor is not based on the patient's health status or preferences. Instead, we exploit the quasi-exogenous variation in clinician assignment that arises from the firm's congestion management rules.

The case allocation algorithm assigns patients to either a nurse or a doctor based on a score that primarily considers their age, gender, region, and reported symptoms. Additionally, the algorithm adjusts for current congestion levels among doctors, which are influenced by both staff capacity and dynamically predicted patient demand. The exact algorithm has evolved over time and is kept as a trade secret. Physicians are compensated by the hour, whereas most nurses work part-time, hourly, or are self-employed. Because staffing schedules are set in advance, excess demand is common, and during periods of congestion, the system directs patients to nurse consultations to minimize waiting times. When congestion peaks, the provider may also contact off-shift clinicians to handle cases. This process allows congestion to be managed within short time frames of at most a few hours. Our analysis accounts for the firm's dynamic assignment mechanism, patient demographics, and reported symptoms to ensure the comparisons of outcomes at the margin.

When assigned a case, clinicians receive an overview of the patient's demographics, symptoms, and medical history, such as active medications or allergies, reported through previous interactions with the provider. Moreover, when a patient is forwarded to a new clinician, a mandatory text field for notes has to be filled out by the previous clinician. The text field asks the clinician to report the patient's symptoms, the reason for forwarding the case, and highlight other notes from the consultation. When clinicians refer a case, they typically book a consultation in an available time slot with a specified clinician type (such as a doctor), but do not choose the identity of the clinician. During and after a consultation, clinicians may spend additional time managing administrative tasks related to the case.

Importantly, the digital interface for clinicians centers primarily on their personal schedules of booked consultations. On the clinician's homepage, they can see the number of patients currently waiting and the number of clinicians on staff. However, this information does not vary across the provider network and is not directly accessible during individual consultations. This setup contrasts with most in-person healthcare environments, where clinicians generally have greater visibility into patient queues. As a result, we expect clinicians' treatment behaviors in this digital setting to be less influenced by immediate congestion.

We study marginal cases that can be handled by both doctors and nurses. Yet, these occupations are trained very differently. In Sweden, physicians must complete an 11-semester medical degree program and then undergo a full-time supervised internship of at least 18 months (Allmäntjänstgöring, AT) to receive a medical license. To become a specialist doctor, including in general medicine, most physicians then undergo at least 5 more years of full-time residency training (Specialiseringstjänstgöring, ST). Doctors thus spend 7 or more years in education. In contrast, nurses complete a

3-year degree program in order to obtain a nursing license. While nurses with at least one year of professional experience can also pursue a 1.5-year specialty training to work as district nurse (Distriktssköterska), or other specializations such as midwife or anesthesia nurse, the majority of nurses do not seek further specialization.[5] Nurses frequently communicate with patients during medical check-ups and make triaging decisions, but compared to physicians their scope of practice is limited. In particular, nurses in Swedish primary care are not authorized for most medical drug prescriptions or specialist referrals.

Despite the differences, primary care patients in Sweden may be consulted by nurses, doctors, or a combination of both clinician types. If a nurse cannot resolve a case, such as when additional authorizations are required, they can pass the case along to a doctor. At the provider organization we study, the nurse in such a case books a follow-up consultation for the patient with a doctor. We consider cases initially assigned to a nurse – who may either resolve the issue or refer it to a doctor – as being managed within a knowledge hierarchy. This contrasts with cases directly assigned to a doctor. While we will also examine how clinicians at the firm organize follow-up care internally, whenever we refer to a *task assignment*, we specifically mean the initial assignment of the case.

2.2 Sample construction

The definition of our analysis sample starts from the 1.8 million consultations with the primary healthcare provider in the full calendar years 2019 and 2020. Out of these, we only consider initial, unscheduled consultations online. Our time frame encompasses 1 April 2019, when the organization started to employ nurses for patient consultations, until 24 December 2020, such that we can observe subsequent visits to the provider within a week. We exclude patients who are registered with the provider as well as infants (patients strictly below the age of two). These basic restrictions ensure that the algorithmic decision rules described above may assign nurses to a given patient consultation, as different rules govern consultations with patients registered at the provider or infants. We exclude cases for which baseline characteristics, such as patient demographics or length of the initial consultation, are missing. Lastly, we restrict the analysis sample to marginal cases that can be allocated to both, a doctor or a nurse-initiated knowledge hierarchy. This requires us to remove

[5] In our analysis sample, more than 75% of the medical doctors are in or have completed specialty training, whereas around 10% of nurses are specialized.

a set of cases for which the assignment is a priori deterministic and, thus, common support does not hold. We do so by identifying conditions with sufficient consultations held by either clinician type. In particular, we only include cases with symptoms for which we observe at least 5% of initial consultations to be assigned to either nurse or doctor. These restrictions leave us with almost 500,000 marginal patient cases.[6]

We link each case to comprehensive administrative data on patients' previous and subsequent healthcare utilization, demographic, and socio-economic information from the Swedish National Board of Health and Welfare (Socialstyrelsen), Statistics Sweden (SCB), and Region Scania. We obtain information on medical drug prescriptions, as well as utilization and diagnoses in hospital care, urgent care, or specialist care services from Socialstyrelsen for the years 2013–2023. We observe consultations in physical and digital primary care outside of the provider for about 10% of all cases within the region of Scania from 2013-2020.[7] The data from SCB include demographic information, such as age, gender, education level, and immigration status and socioeconomic information, such as employment status and educational attainment, from 2013-2020. In addition, we observe labor market information for patients, including annual sickness benefits paid out by the national social insurance agency and monthly earnings from patients' primary employment. By using patients' national identifiers, we can link data from the primary healthcare organization to external events outside that provider's services.

2.3 Variable definitions

We examine the differences and trade-offs between two task allocation modes: the knowledge hierarchy, where cases are initially assigned to nurses, and a direct-to-doctor task assignment. Our data enables us to study patients' health outcomes and the subsequent costs incurred in the healthcare system for each case, while accounting for heterogeneity in case characteristics. We outline our outcomes of interest and key control variables below. Appendix E provides additional details on our data.

[6]Appendix Table A1 shows the total number of observations after each sample restriction, as well as separately for cases initially assigned to a nurse or to a doctor.

[7]Scania is the third largest county in Sweden and lies in the aggregated region of South Sweden.

2.3.1 Outcomes

To assess the net efficiency gains from the division of labor in a knowledge hierarchy setup, we consider three distinct perspectives: the firm, the patient, and the public payer who bears the costs of healthcare provision.

Organization of tasks at the firm. First, we examine how cases are managed within the firm to understand its internal organization. The knowledge hierarchy structure enables a flexible coordination of tasks. However, nurses and doctors still operate within distinct scopes of practice, with nurses facing regulatory limitations on their roles and responsibilities. Our first set of outcomes focuses on how tasks at the margin are coordinated within the provider under these regulatory constraints. Specifically, we examine subsequent consultations with the provider, which may occur if patients are dissatisfied with their initial consultation and seek a different clinician or if the original clinician refers the case to another colleague. To differentiate between these scenarios, we focus on the rate of internal referrals for a follow-up consultation with a doctor. For cases initially assigned to nurses, this reflects the knowledge hierarchy concept, where nurses forward cases to doctors. Additionally, we assess the time clinicians spend on each case. These outcome measures seek to provide insight into how tasks are managed within the knowledge hierarchy.

Quality of care for the patient. Second, we evaluate whether the nurse-initiated knowledge hierarchy impacts care quality, focusing on patient-centered outcomes. Our data allows us to track patients' healthcare journeys beyond the provider, enabling us to observe quality signals that may not be captured by the firm.

We begin by studying the use of external primary care services outside the firm, using data from Scania and a subsample of patients registered in the region. This analysis informs us whether patients seek external care due to inconclusive treatment or dissatisfaction with the initial care provided. We also assess patient satisfaction through their ratings to capture whether patients' perceived care quality is affected by an assignment to the knowledge hierarchy.

Next, we evaluate differences in patients' access to services typically provided by doctors, such as diagnoses, drug prescriptions (from both the provider and external sources), and specialist referrals. Introducing an additional step, where nurses make referral decisions, could reduce patient access to doctor-provided services. However, cases that require these services may still be referred up the knowledge hierarchy to a doctor, and a reduction in access may not be inefficient as long as it does not lead to adverse health outcomes.

To assess the impact of the knowledge hierarchy on health outcomes, we examine both immediate and medium- to long-term adverse events. We focus on high-cost care events, such as urgent care visits, emergency department visits, and hospitalizations, which may indicate adverse health outcomes during the care journey. However, utilizing these services may also be beneficial if patients require a higher level of care, or patients may avoid seeking care even while health conditions deteriorate. To gain a more comprehensive understanding, we also study longer-term adverse events, including income reductions and patient mortality, occurring after the healthcare episode at the provider. These unique data allow us to determine whether the mode of task organization affects patients' health outcomes beyond just healthcare utilization.

Costs incurred in healthcare. Lastly, we analyze the effects of the nurse-initiated knowledge hierarchy on the cost of patients' care journeys, both during and after their initial contact with the primary care provider. While examining patient outcomes helps us assess the quality delivered under each work mode, evaluating costs allows us to determine the overall productivity effects of the nurse-initiated knowledge hierarchy compared to the direct-to-doctor route.

In Sweden's publicly funded healthcare system, most healthcare costs are borne by the public payer. We examine the expenses incurred by the public healthcare system, including a patient's consultations with our provider, prescriptions, specialist care, and any urgent or hospital care following the primary care visit. This analysis provides insight into the costs from the perspective of the public payer and allows us to make broader efficiency conclusions about the knowledge hierarchy. Table A16 presents the cost of each healthcare service and the sources of our data. Most cost estimates are derived from public sources and regional announcements on reimbursement rates, except for prescription costs, which we directly obtain from our data for the sample of patients.

Thus, we evaluate multiple outcome sets: task organization within the firm, quality of care, and costs borne by the public system. For most outcomes, except for longer-term measures like income and mortality, we focus on the seven-day period following the initial consultation to clearly link outcomes to that event. Together, these outcomes allow us to draw conclusions on the broader efficiency trade-offs between the knowledge hierarchy and the direct-to-doctor route.

2.3.2 Control variables

While the assignment algorithm takes into account pre-determined case characteristics, such as patients' primary symptoms and demographic information, congestion changes over time. Our main analysis controls for patients' login time, which marks when they submit relevant case information and join the consultation queue, in addition to the pre-determined individual case characteristics. Below is an overview of the control variables included in the analysis.

Login-time. The time at which a patient logs in is crucial for case assignment, as both patient demand and clinician availability fluctuate throughout the day, affecting congestion. To account for this variation, we include fine-grained time controls by segmenting login times into 4-hour blocks starting at midnight of each calendar day.

Symptom categories. Before any clinician contact, patients select a primary symptom from a drop-down list. Appendix Figure A6 presents the number of consultations for each symptom category, separated by initial nurse- and doctor consultations. The largest category, other health inquiries, covers about one-third of the sample and serves as a catch-all category for cases where patients did not choose a symptom from a pre-specified list. The remaining symptom categories in our data include common ailments, such as cold symptoms, infections, and Covid-19.

Demographic information. We control for basic demographics by including, as control variables, patients' gender and indicators for various age categories. The age category variables reflect different life stages: toddlers (2-4 years), children (5-12 years), teens (13-19 years), adults (20-39 years), middle-aged (40-59 years), and seniors (60+ years). We also include indicators for the aggregated regional areas where patients are registered in the year prior to the initial consultation.

Health risk. Health risk is measured by patients' healthcare utilization over the three years prior to the consultation, excluding the 30 days immediately before the consultation to ensure risk is pre-determined. We use separate indicators for inpatient hospitalizations, emergency department (ED) visits, urgent care visits, and specialist visits. Additionally, we account for any comorbidities diagnosed in specialty or hospital care prior to 2019, the start of our analysis period.

Socio-economic background. We account for patients' socio-economic background using data from 2018, the year before the analysis period. Control variables include indicators for income above the sample median (calculated for patients aged 20 or older), employment status (categorized as employed, self-employed, or unemployed), and education level (including basic schooling, secondary education, and further post-secondary education of varying durations). Civil status is captured by indicators for marital status (married, unmarried, or previously married). As employment, education, income, and civil status data are unavailable for certain age groups (such as for underage patients), we include separate indicators for missing data on these dimensions. Finally, we consider migration status by including indicators for first-generation migrants, second-generation migrants, and patients with no foreign background.

Table 1 presents summary statistics for the main analysis sample, along with characteristics of cases initially assigned to nurses versus those assigned directly to doctors. About 63% of the cases are female, and the sample is relatively young, with an average age of approximately 30 years. Cases assigned to nurses differ most notably from those assigned to doctors in terms of symptom composition, reflecting the information used by the assignment algorithm. Nurse-assigned cases have a higher proportion of vaguely defined symptoms, such as "other health inquiries," abdominal pain, fever, or Covid-19. In contrast, cases directly assigned to doctors are more likely to involve cold-related symptoms like cold and flu, sore throat, or sinusitis, as well as specific infections, including eye infections and urinary tract infections. We also observe a higher proportion of women among cases directly assigned to doctors, along with minor imbalances across other characteristics.

Table 1. Characteristics of the analysis sample

	Full sample		Initially to nurse		Direct to doctor		Diff. p-value
Symptom categories							
Abdominal pain	0.030	[0.17]	0.045	[0.21]	0.025	[0.16]	0.00
Cold and flu	0.089	[0.28]	0.024	[0.15]	0.11	[0.31]	0.00
Cold sores	0.025	[0.16]	0.012	[0.11]	0.029	[0.17]	0.00
Constipation	0.0071	[0.084]	0.0062	[0.079]	0.0074	[0.086]	0.00
Covid-19	0.058	[0.23]	0.15	[0.36]	0.031	[0.17]	0.00
Diarrhea or vomiting	0.021	[0.14]	0.024	[0.15]	0.020	[0.14]	0.00
Eye infection	0.075	[0.26]	0.049	[0.22]	0.083	[0.28]	0.00
Fever	0.029	[0.17]	0.034	[0.18]	0.028	[0.16]	0.00
Headache	0.023	[0.15]	0.036	[0.19]	0.019	[0.14]	0.00
Nail problem	0.022	[0.15]	0.010	[0.10]	0.026	[0.16]	0.00
Other health inquiries	0.35	[0.48]	0.45	[0.50]	0.32	[0.47]	0.00
Bites and stings	0.054	[0.23]	0.029	[0.17]	0.062	[0.24]	0.00
Sinusitis	0.032	[0.17]	0.0093	[0.096]	0.038	[0.19]	0.00
Sore throat	0.076	[0.26]	0.057	[0.23]	0.081	[0.27]	0.00
Uncategorized	0.030	[0.17]	0.030	[0.17]	0.030	[0.17]	0.49
Urinary tract infection	0.079	[0.27]	0.033	[0.18]	0.093	[0.29]	0.00
Demographics							
Female	0.63	[0.48]	0.59	[0.49]	0.64	[0.48]	0.00
Patient age	29.2	[16.5]	29.3	[16.6]	29.2	[16.5]	0.51
West Sweden	0.20	[0.40]	0.20	[0.40]	0.20	[0.40]	0.07
Stockholm	0.44	[0.50]	0.45	[0.50]	0.44	[0.50]	0.00
Middle Sweden	0.19	[0.39]	0.18	[0.39]	0.19	[0.39]	0.00
South Sweden	0.095	[0.29]	0.093	[0.29]	0.095	[0.29]	0.02
Norrland	0.034	[0.18]	0.031	[0.17]	0.035	[0.18]	0.00
Småland + the islands	0.042	[0.20]	0.042	[0.20]	0.042	[0.20]	0.31
Health risk							
Any prior hospitalization	0.19	[0.39]	0.18	[0.39]	0.19	[0.39]	0.00
Any prior ED	0.33	[0.47]	0.34	[0.47]	0.33	[0.47]	0.00
Any prior urgent care	0.24	[0.42]	0.23	[0.42]	0.24	[0.43]	0.00
Any prior specialist	0.64	[0.48]	0.63	[0.48]	0.64	[0.48]	0.00
Any comorbidity	0.21	[0.41]	0.20	[0.40]	0.21	[0.41]	0.00
Socio-economic variables							
Income above median	0.32	[0.47]	0.31	[0.46]	0.33	[0.47]	0.00
Any benefits	0.11	[0.32]	0.11	[0.31]	0.11	[0.32]	0.00
Schooling < 9 years	0.049	[0.22]	0.052	[0.22]	0.048	[0.21]	0.00
Middle/High school	0.24	[0.42]	0.23	[0.42]	0.24	[0.42]	0.00
Further educ. < 3 years	0.11	[0.31]	0.10	[0.31]	0.11	[0.31]	0.00
Further educ. <= 3 years	0.19	[0.39]	0.18	[0.39]	0.20	[0.40]	0.00
Education n/a	0.42	[0.49]	0.43	[0.49]	0.41	[0.49]	0.00
Employed	0.57	[0.50]	0.55	[0.50]	0.57	[0.50]	0.00
Self-employed	0.042	[0.20]	0.042	[0.20]	0.042	[0.20]	0.51
Unemployed	0.050	[0.22]	0.052	[0.22]	0.049	[0.22]	0.00
Employment status n/a	0.34	[0.47]	0.36	[0.48]	0.34	[0.47]	0.00
Married	0.23	[0.42]	0.22	[0.41]	0.23	[0.42]	0.00
Unmarried	0.075	[0.26]	0.075	[0.26]	0.075	[0.26]	0.70
Divorced/Widowed	0.42	[0.49]	0.43	[0.49]	0.42	[0.49]	0.15
Civil status n/a	0.27	[0.45]	0.28	[0.45]	0.27	[0.45]	0.00
Not migrated	0.75	[0.43]	0.73	[0.45]	0.76	[0.43]	0.00
Immigrant 1st gen	0.16	[0.36]	0.17	[0.38]	0.15	[0.36]	0.00
Immigrant 2nd gen	0.096	[0.29]	0.100	[0.30]	0.095	[0.29]	0.00
Observations	490,505		111,707		378,798		490,505

Note: This table provides summary statistics for key case characteristics used as control variables in the main analysis sample. Symptom categories reflect the primary symptom reported by patients when requesting a consultation. Demographics include patients' gender and age at the time of consultation, and patients' registered aggregated region as of November 2018, before the start of the analysis period. Health risk variables capture healthcare utilization over the 3 years prior to the consultation, but excluding the 30 days immediately before, and include an indicator for any comorbidity diagnosed in specialist care. Socio-economic variables encompass indicators for above-median income (where the median is 294,700 SEK among patients over the age of 20), benefit receipt, employment status, education level, civil status, and migrant background, all recorded in November 2018. Standard deviations are shown in square brackets. The final column reports the p-value from a t-test comparing the means between patient cases initially assigned to a nurse and those directly assigned to a doctor.

3 Empirical framework

Our analysis examines the trade-offs between different work coordination structures. To causally identify the average effects on quality and cost of the knowledge hierarchy compared to a direct-to-doctor assignment, we must address a key identification challenge: task allocation to occupations is not random, which likely leads to selective sorting of cases and, consequently, selection into different production processes. We tackle this challenge using two quasi-experimental approaches: (1) We exploit variation in congestion, as used by the algorithm, controlling for the timing of cases and detailed characteristics that influence case assignment; and (2) We employ observed assignments just prior to a case as a measure of congestion in a flexible instrumental variable (IV) setup, to account for unobserved residual selection in case assignment. Below, we outline our econometric approach and discuss the validity of our instrument.

3.1 Specification

To quantify the impact of the organizational structure on outcomes of a case i, we use the following empirical specification:

$$Y_i = \delta Nurse_i + X_{i1}\gamma + X_{i2}\beta + \epsilon_i, \tag{1}$$

where Y_i denotes the outcome of interest for case i and $Nurse_i$ denotes the initial assignment to a nurse instead of a doctor. X_{i1} is a vector of indicator variables capturing the baseline characteristics of a case: login time given by date-by-4 hours windows, the patient's main symptom, age and gender, and region. X_{i2} is a vector of additional control variables for case i's characteristics, including the patient's health risk and socioeconomic characteristics, as described in Section 2.3. Finally, ϵ_i is an error term.

Our coefficient of interest is δ, which captures the effect of initial assignment to a nurse, and thus the knowledge hierarchy route, as compared to direct assignment to a doctor. Assignment to a clinician is not random: for instance, nurses may tend to see patients with milder symptoms and less complex conditions. However, in our setting, case assignment is determined algorithmically, based on congestion at the time of case login, along with basic demographic information. This allows us to exploit quasi-experimental variation in case assignment by conditioning on the case's login time and observable characteristics. Based on this reasoning, we estimate δ in Equation 1 by Ordinary Least Squares (OLS). Since cases are not clearly assigned within clusters, we follow Abadie and Cattaneo (2018) and compute robust standard errors.[8]

However, the precise functional form of the assignment algorithm is proprietary to the firm and unknown to us. For example, the provider might incorporate complex interactions between patient age, symptoms, and congestion when assigning cases, and we cannot fully replicate the algorithm's rules. As a result, even with our rich set of control variables, we may not fully account for potential endogeneity in the initial assignment. To address this, we complement our OLS approach with an instrumental variable (IV) strategy that leverages exogenous variation in congestion.

Our IV strategy employs, as an instrument of the current case i's assignment, the *Share of nurse consultations in the past 60 minutes.* The instrument is a leave-own-case-out measure of congestion among cases just prior to the assignment of a given case i. Our first stage thus takes the following form:

$$\begin{aligned} Nurse_i &= \theta PrevNurseShare_i + X_{i1}\zeta + X_{i2}\eta + \nu_i \qquad (2) \\ \text{with } PrevNurseShare_i &= \frac{1}{|N_{t(i)}|} \sum_{k \in N_{t(i)}} Nurse_k. \end{aligned}$$

Here, $PrevNurseShare_i$ corresponds to the share of nurse consultations as a fraction of all initial (nurse or doctor) consultations in a time frame $t(i)$ of 60 minutes prior to the login time of patient case i within our analysis sample. $N_{t(i)}$ defines the set of previous patient cases and strictly excludes i's own consultation. The vectors of time and case characteristics, X_{i1} and X_{i2}, are defined as above, and ν_i is an error term. In the IV specification, we estimate a two-stage least squares (2SLS) model represented by Equations 1 and 2 with robust standard errors.

3.2 Instrument validity

In the IV strategy, δ represents a local average treatment effect (LATE), that is, the average causal effect of being initially assigned to a nurse, as opposed to direct assignment to a doctor, for cases where the initial assignment is affected by congestion measured over prior cases. To interpret δ as the

[8]We also provide robustness checks under alternative assumptions about the structure of the error terms.

LATE, we require that our congestion instrument, the share of nurse consultations among cases in the past 60 minutes, satisfies the four standard IV assumptions: relevance, conditional independence, monotonicity, and exclusion. Below, we provide evidence supporting the validity of our congestion instrument.

Relevance. First, the instrumental variable should have a clear impact on the initial case assignment. We begin by presenting descriptive evidence in Figure 1 that our congestion instrument effectively captures variation in congestion that drives the initial assignment to a nurse. The left subfigure shows that the nurse share in prior cases varies, with most cases experiencing relatively low congestion, but also a considerable amount where nurses handle half or more of the previous cases. The right subfigure demonstrates an almost linear increase in the unconditional probability of an initial nurse assignment as the instrument increases. However, these figures do not account for temporal variation, as both congestion and staffing levels may fluctuate over time, both daily and within a given date. Appendix Figure A7 shows that congestion indeed varies, particularly by time of day and month. To address potential sorting over time correlated with congestion, we include granular fixed effects, using date-by-4-hour windows at the login time.[9]

Appendix Table A2 presents the first-stage regression estimates from Equation 2, when systematically expanding the set of control variables. Once we control for baseline case characteristics used by the algorithm (login date-by-4 hour windows, symptom, and patient demographics), the first-stage coefficient remains large and precisely estimated at 0.38. Moreover, the first-stage estimate remains stable when we include additional controls for patient health risk and socio-economic characteristics. The F-statistic for the first stage exceeds 3,000, confirming that our congestion instrument is both strong and predictive of initial assignment, even after accounting for detailed case characteristics.

Conditional independence. Second, we require that potential outcomes are independent of our congestion instrument, conditional on the control variables. If we had access to the assignment algorithm's exact score for each case based on its pre-determined characteristics, the realized assignment would depend solely on congestion at the patient's login time. Our instrument would then be valid as long as congestion is uncorrelated with potential outcomes under nurse or doctor assignment. In the absence of the exact algorithm, we can condition on the baseline case characteristics

[9] In robustness checks, discussed in Section 4.4, we vary the size of login time windows and show that results remain largely unchanged.

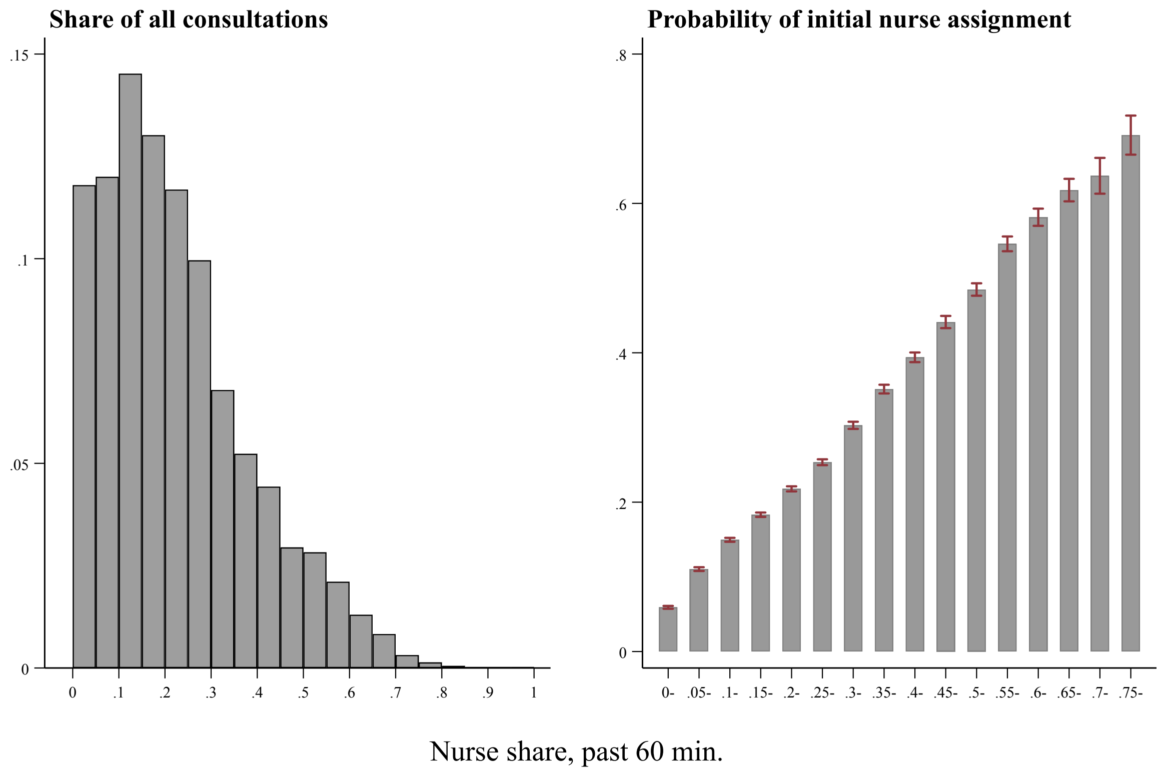


Figure 1. Variation in the congestion instrument

Notes: This figure shows descriptive figures for our congestion instrument, the share of nurse consultations in the past 60 minutes of a given consultation. The left subfigure shows the distribution of the instrument. The right subfigure shows the average probability of an initial assignment to a nurse, our treatment, in categorized values of the instrument. Lines indicate 95% confidence intervals of the mean.

that influence it: login time, symptom categories, and patient demographics. To ensure the validity of our empirical strategy, we assess whether additional case characteristics, such as patients' health risks or socioeconomic backgrounds, which could influence potential outcomes, are balanced relative to our congestion instrument after controlling for these baseline characteristics.

Figure 2 demonstrates that baseline case characteristics alone do not fully eliminate correlations between the initial case assignment and patients' health risks or socioeconomic status. However, once we condition on the baseline controls, the congestion instrument is largely balanced against these additional characteristics, both individually and jointly. Appendix Table A3 shows that congestion may be correlated with case characteristics and thus potential outcomes, but once we account for baseline case characteristics used as algorithm inputs, the instrument is jointly balanced against additional case characteristics.

Monotonicity. Third, if treatment effects can be heterogeneous, we need to assume that the effects of congestion on treatment assignment are unidirectional across individual cases in order to interpret the IV estimate as the average causal effect for compliers. Specifically, our setup assumes that increased congestion does not reduce the likelihood of an initial nurse assignment, implying that cases sent to the nurse-initiated knowledge hierarchy under low congestion are also sent to a nurse under high congestion. This assumption aligns with our institutional context, where direct-to-doctor assignment was the default, and the nurse-initiated knowledge hierarchy route emerged later as a backup option during periods of high demand.

To verify testable implications of the monotonicity assumption, we first examine the correlation between our congestion instrument and treatment assignment across observable subgroups, which should be weakly positive in each subsamples (Bhuller et al. 2016; Dobbie et al. 2018). Appendix Table A4 presents the first-stage regressions results from Equation 2 for various splits along dimensions of health risk, demographics, and socioeconomic factors. Across these subgroups, we estimate positive and statistically significant first-stage coefficients that are quantitatively similar, indicating no evidence of a violation of monotonicity.

In addition, we can visually inspect the conditional expectation function of the treatment, initial nurse assignment, along the range of the congestion instrument in different subsamples using binscatter estimates (Cattaneo et al. 2024). Appendix Figure A8 shows the conditional expectation functions for subgroups defined by baseline characteristics, including symptom category ("Other

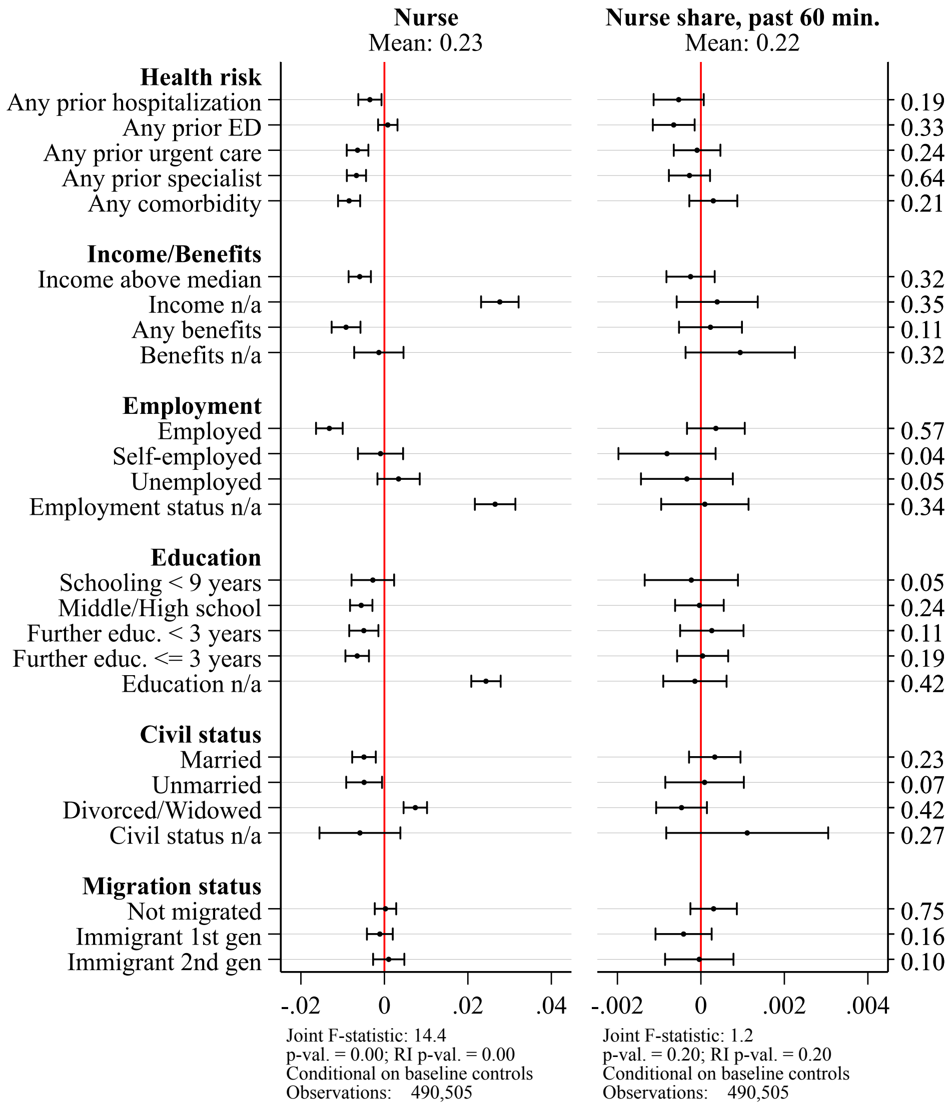


Figure 2. Balance of the treatment and congestion instrument in case characteristics

Notes: This figure presents balance tests for the treatment variable (*Nurse*) in the left and center columns, and for the congestion instrument (*Nurse share, past 60 min.*) in the right column. Each row shows the results of a bivariate regression where the treatment or instrument variable is regressed on a specific case characteristic. All regressions account for baseline case characteristics: login time, main symptom, and patient demographics. Login time indicates date-by-4 hours fixed effects. Symptom categories refer to indicators for the main symptom category reported by the patient when requesting a consultation. Demographics include indicators for patient gender, age categories, and aggregated regions. The horizontal lines represent 95% confidence intervals based on robust standard errors. The F-statistics reflect the joint F-tests for the treatment or instrument variable when regressed on all control variables, conditioned on login time. We report both, conventional as well as randomization inference (RI) p-values from 500 draws, based on Kerwin et al. (2024). The numbers on the far right display the sample means for each case characteristic.

health inquiries" vs. other symptoms), age groups, and regions. In all cases, we observe no downward slope, further supporting the validity of the monotonicity assumption.

Exclusion restriction. Finally, in order to interpret the IV estimates as causal effects of the nurse-initiated knowledge hierarchy, we require that congestion affects case outcomes only through the initial case assignment. While the exclusion restriction assumption is untestable, we can assess potential violations.

The exclusion assumption is violated if clinicians adjust their work to high levels of congestion, which then affect patient outcomes. Although we cannot rule out clinician responses to congestion, we can examine the extent to which the congestion instrument correlates with consultation characteristics. Appendix Table A5 shows that patient waiting times increase (almost by definition) with congestion, but the differences are negligible compared to typical waiting times of patients seeking primary care and unlikely to affect outcomes: in the highest decile of congestion, waiting times are 33 minutes compared to 14 minutes in the lowest decile. Additionally, consultations lasting less than a minute, which may indicate clinicians checking in with patients due to long waiting times, are not more common during high congestion, and drop out rates are, in fact, lower. While clinicians spend slightly less time per case during high congestion, the difference (11.3 minutes vs. 11.5 minutes) is unlikely to affect patient care. Overall, we find little evidence that high congestion meaningfully impacts provider or patient outcomes.

As an additional falsification test, we construct a sample of cases where, by the logic of the case assignment algorithm, congestion does not affect assignment. These cases serve as a "zero first-stage" sample, where congestion should not affect any of the main outcomes (Angrist et al. 2010). We leverage that in a subset of symptom categories, the algorithm automatically assigns cases to a doctor or nurse in a *strictly* deterministic fashion, regardless of case characteristics or congestion. Appendix Figure A9 shows these strictly deterministic symptoms, where fewer than 1% of cases are assigned to a nurse or a doctor. We create the congestion instrument using the share of initial nurse consultations over marginal cases from our analysis sample in the past 60 minutes.

Appendix Figure A10 provides descriptive evidence on the congestion instrument within the zero first-stage sample. The distribution is similar to that in the main analysis (Figure 1), suggesting that we do not capture time periods with markedly different congestion levels. Yet, we observe little systematic relation between the probability of a nurse assignment over the distribution of congestion.

Appendix Table A6 presents first-stage F-statistics of at most 7, compared to over 3,000 in the main analysis. Additionally, the reduced-form regressions show no evidence of the congestion instrument affecting outcomes, and no significant effects from nurse assignments driven by congestion. This confirms that our congestion instrument operates as expected: when congestion does not influence the initial case assignment, it has no effects on outcomes.

4 Main results

We next present our results on the effects of the nurse-initiated knowledge hierarchy for the healthcare provider and patients. We study different outcome dimensions and compare cases initially assigned to a nurse with those assigned directly to doctors. We begin by examining how the healthcare provider organizes tasks at the margin and utilizes its primary resource: clinician time. Our findings show that in 70% of nurse-initiated cases at the margin, a case can be resolved by nurses themselves without referring it up the knowledge hierarchy. We then turn to patients and study their care journey outside the provider following an initial assignment to a nurse versus a doctor. We find that the nurse-initiated knowledge hierarchy impedes patients' access to diagnostic quality and prescriptions. However, when assessing whether these effects translate into worse health outcomes, we find little evidence of reduced care quality in terms of acute care events, income reductions, or mortality. Finally, we compare the marginal cost differences between the nurse-initiated knowledge hierarchy and the direct-to-doctor route. Considering the public payer's contribution to a case's care journey within and outside the organization, we estimate moderate cost savings of between 7.5% and 20%.

We provide both the OLS and IV estimates from the specifications outlined in Section 3.1. The OLS coefficients are informed by our understanding of the inputs into the assignment algorithm. However, since the exact functional form of the algorithm is unknown, we cannot fully rule out selective case sorting. The IV estimates rely on plausibly exogenous variation in congestion and impose weaker structural assumptions on the algorithm but may yield imprecise results.

4.1 Organization of tasks at the firm

We begin by focusing on the primary care provider and how patient consultations are organized within the knowledge hierarchy. Specifically, we examine the impact of an initial assignment to a nurse on internal follow-up consultations and clinician time spent on a case within up to seven days after the initial consultation date, compared to the baseline of a direct-to-doctor assignment. Table 2 presents OLS estimates from Equation 1, as well as IV and reduced-form estimates based on Equation 2, leveraging quasi-exogenous variation in initial case assignments due to congestion. All regressions include the full set of baseline and additional case controls described in Section 3.1. The reduced-form coefficients show that the congestion instrument significantly affects provider outcomes, validating its influence through the probability of initial nurse assignment.

We first analyze follow-up consultations with a doctor, which include referrals from nurses as well as other follow-up appointments that a clinician books, as long as they take place with a doctor. Table 2 indicates that an initial nurse assignment increases the probability of an internal referral to a doctor by 31% (OLS) and 27% (IV), relative to a baseline follow-up rate of 1.4%. This implies that approximately 30% of marginal cases initiated by nurses are referred to doctors, while follow-up consultations with doctors remain rare otherwise. However, these estimates also imply that nurses can solve most of their cases without referring to a doctor: About 70% of the cases on the margin are solved directly by the nurse.

Next, we consider all types of subsequent consultations, which may also capture follow-up or repeat visits sought by patients. As shown in Table 2, baseline rates for subsequent consultations increase to 12% for direct-to-doctor assignments, but the implications of the knowledge hierarchy structure remain consistent. Specifically, the observed increase in subsequent consultations following an initial nurse assignment is not driven by patients whose care was inconclusive and might seek care through repeated drop-ins. The OLS estimate indicates a minor increase in returning cases by 0.84 percentage points against a baseline of 5.5%, whereas the IV estimate is not statistically significant. Thus, we observe that subsequent consultations increase, but this increase is driven by nurse referrals to doctors.

Finally, we examine how the knowledge hierarchy affects the total clinician time spent on a case during the care journey at the provider, accounting for any subsequent consultations. Table

Table 2. Effect of an initial nurse assignment on the organization of tasks at the provider

	Any referral to doctor (7d)			Any subseq. consultation (7d)			Any return drop-in (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	0.31***		0.27***	0.29***		0.22***	0.0084***		-0.014
	(0.0015)		(0.011)	(0.0017)		(0.017)	(0.00091)		(0.011)
Nurse share, past 60 min.		0.10***			0.083***			-0.0055	
		(0.0049)			(0.0068)			(0.0042)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			3,034			3,034			3,034
Baseline mean	0.014	0.014	0.014	0.12	0.12	0.12	0.055	0.055	0.055

	Clinician total time in min. (7d)			Doctor total time in min. (7d)			Nurse total time in min. (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	3.30***		1.53***	-7.98***		-8.93***	11.3***		10.5***
	(0.047)		(0.50)	(0.037)		(0.44)	(0.027)		(0.21)
Nurse share, past 60 min.		0.59***			-3.43***			4.02***	
		(0.19)			(0.18)			(0.11)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			3,034			3,034			3,034
Baseline mean	13.7	13.7	13.7	13.5	13.5	13.5	0.19	0.19	0.19

Note: This table presents regression results for our main specifications on the effects of initial case assignment to a nurse (the *knowledge hierarchy*) rather than directly to a doctor. Each subpanel corresponds to a different outcome as the dependent variable. Patient-provider interactions after the initial consultation are defined as follows: "Any referral to doctor" refers to any internal referrals or revisits with a doctor, "Any subsequent consultations" covers any interaction with the provider after an initial visit, and "Any return drop-in" indicates any unscheduled revisit for the same symptom. "Total time" defines the time that clinicians spend on a case, including consultation time spent with the patient and time on administrative work during the initial and all subsequent consultations. The left column in each subpanel reports the Ordinary Least Squares (OLS) estimate for the treatment variable (*Nurse*) based on Equation 1. The middle column presents the reduced-form regression estimates for the congestion instrument, measured as the share of nurse consultations in the past 60 minutes (*Nurse share, past 60 min.*). The right column provides the Instrumental Variables specification, using *Nurse share, past 60 min.* as an instrument for *Nurse*, estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2. All regressions control for symptom categories, demographics, region, health risk factors, and socioeconomic variables, along with login date-by-4-hour fixed effects. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

2 reveals that total clinician time increases by up to 25% in the knowledge hierarchy: Clinicians spend between 1.53 (IV) and 3.30 (OLS) minutes longer on a case, from a baseline of 13.7 minutes that a directly assigned doctor spends on a case. When splitting by clinician type, we observe that the knowledge hierarchy reduces total doctor time by 7.89 (OLS) to 8.93 (IV) minutes, but nurses effectively overcompensate doctor time by 10.5 (IV) to 11.3 (OLS) minutes.

Appendix Table A8 examines direct consultation time (excluding time spent on administrative work), showing that nurses in the knowledge hierarchy route particularly compensate with additional patient-facing time. Overall, the knowledge hierarchy shifts doctor workload to nurses.

We trace the patient's care journey over the seven calendar days following the consultation date. Appendix Table A7 provides additional results for provider outcomes defined within one day after the consultation date. The coefficient estimates are nearly identical, indicating that most cases are resolved by the next day.

4.2 Effects on the quality of care

We next turn to patient-centered outcomes to analyze how the knowledge hierarchy affects the quality of care provided. We consider a range of outcome measures: patients' use of external primary care; their satisfaction with care at the provider; access to diagnostic information, prescriptions, and specialists; utilization of acute and high-cost secondary care services; and the occurrence of adverse events. By evaluating a diverse set of patient outcomes across their care journey, both within and outside the primary care provider, we aim to provide a comprehensive assessment of any potential quality-of-care differences resulting from the nurse-initiated knowledge hierarchy at a patient's first point of contact with the healthcare system.

Table 3 presents the OLS estimates for our main patient outcomes, alongside the IV and reduced-form estimates based on the congestion instrument, with the full set of case characteristics controlled for as described in Section 3.1. Below, we detail the main results and their implications.

4.2.1 External primary care services

First, we investigate whether an initial assignment to a nurse affects patients' use of external primary care providers (PCPs). A substantial share of nurse-assigned cases at the margin are resolved without a referral to a doctor, yet patients may prefer consulting a doctor and thus seek a primary care

Table 3. Effect of an initial nurse assignment on quality of care

	Any external PCP consultation (7d)			Any external PCP consultation, doctor (7d)			Any external PCP consultation, nurse (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	0.025***		-0.00072	0.022***		-0.028	0.0037		0.024
	(0.0048)		(0.061)	(0.0045)		(0.057)	(0.0027)		(0.034)
Nurse share, past 60 min.		-0.00026			-0.0100			0.0084	
		(0.022)			(0.020)			(0.012)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	42,814	42,814	42,814	42,814	42,814	42,814	42,814	42,814	42,814
First-stage K-P F-statistic			199			199			199
Baseline mean	0.15	0.15	0.15	0.12	0.12	0.12	0.043	0.043	0.043

	Rating: top score (7d)			Rating: physical replacement (7d)			Informative diagnosis (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	0.045***		0.042*	0.015***		0.060***	-0.064***		-0.061***
	(0.0018)		(0.022)	(0.0018)		(0.022)	(0.0016)		(0.018)
Nurse share, past 60 min.		0.016*			0.023***			-0.024***	
		(0.0085)			(0.0085)			(0.0069)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			3,034			3,034			3,034
Baseline mean	0.42	0.42	0.42	0.45	0.45	0.45	0.81	0.81	0.81

	Any new prescription (7d)			Any specialist (7d)			Any urgent care (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	-0.10***		-0.060***	0.0030***		-0.017*	0.0019***		-0.0033
	(0.0017)		(0.021)	(0.00073)		(0.0086)	(0.00051)		(0.0060)
Nurse share, past 60 min.		-0.023***			-0.0064*			-0.0013	
		(0.0080)			(0.0033)			(0.0023)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			3,034			3,034			3,034
Baseline mean	0.44	0.44	0.44	0.036	0.036	0.036	0.017	0.017	0.017

	Any ED (7d)			Any hospitalization (7d)			Any avoid. hospitalization (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	0.0061***		-0.016*	-0.00023		0.0015	0.00026*		0.0014
	(0.00074)		(0.0085)	(0.00039)		(0.0047)	(0.00015)		(0.0018)
Nurse share, past 60 min.		-0.0061*			0.00057			0.00053	
		(0.0033)			(0.0018)			(0.00067)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			3,034			3,034			3,034
Baseline mean	0.030	0.030	0.030	0.0087	0.0087	0.0087	0.0013	0.0013	0.0013

	Income decrease >20 pct. (cal. month after)			Zero income (cal. month after)			Death excl. external causes (3y)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	-0.021***		0.026	-0.0034***		-0.019	-0.000096		-0.0023
	(0.0024)		(0.031)	(0.0013)		(0.016)	(0.00016)		(0.0019)
Nurse share, past 60 min.		0.0091			-0.0069			-0.00087	
		(0.011)			(0.0059)			(0.00072)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	213,425	213,425	213,425	213,425	213,425	213,425	490,505	490,505	490,505
First-stage K-P F-statistic			1,143			1,143			3,034
Baseline mean	0.22	0.22	0.22	0.050	0.050	0.050	0.0016	0.0016	0.0016

Note: This table presents regression results for our main specifications on the effects of initial case assignment to a nurse (the *knowledge hierarchy*) rather than directly to a doctor. Each subpanel corresponds to a different outcome as the dependent variable. Outcomes are defined as follows: "External PCP consultation" refers to any in-person primary care consultation with a doctor or a nurse outside of the provider and is defined for patients registered in Scania; "Informative diagnosis" excludes symptomatic/health status diagnoses (ICD R00-R99/Z00-Z99) and implies that any informative diagnosis was provided within seven days; "Rating" implies that any top rating (top score or survey response that consultation replaced physical care) for a consultation was given within seven days; "New prescription" refers to ATC codes unobserved in the previous 3 months; "Urgent care center" refers to local care centers (Närakut); "ED" refers to emergency departments located at hospitals (Akutmottagning). "Income" refers to gross earnings from the main income source within a calendar year and are reported in each calendar month for employees. Income reductions capture any income drops in the calendar month following the consultation compared to the average in the three months prior, including sick leaves, as the employer replaces 80% of the regular salary in the first 14 days, and employer earnings become zero afterwards as the Swedish Social Insurance Agency then begins to pay out sickness benefits. To study income-related outcomes, we only consider a sample of cases for which patients are reported as employees with income in the three months prior to the consultation exceeding a small threshold of 3,533.33 SEK in every month (in 2010 values), based on the annual threshold of 42,400 SEK used by Saez et al. (2019). "Death excluding external causes" refers 3-year mortality excluding deaths due to accidents, self-inflicted harm, and other causes of death (ICD codes V01-Y89 but excluding X41-X42 and X44-X45, and including U12.9. The left column in each subpanel reports the Ordinary Least Squares (OLS) estimate for the treatment variable (*Nurse*) based on Equation 1. The middle column presents the reduced-form regression estimates for the congestion instrument, measured as the share of nurse consultations in the past 60 minutes (*Nurse share, past 60 min.*). The right column provides the Instrumental Variables specification, using *Nurse share, past 60 min.* as an instrument for *Nurse*, estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2. All regressions control for symptom categories, demographics, region, health risk factors, and socioeconomic variables, along with login date-by-4-hour fixed effects. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

physician outside the provider we study.

We observe external PCPs in the region of Scania, but we lack information outside this region. Thus, we restrict the sample to the roughly 10% of cases from patients registered in Scania in 2018.[10]

Our focus is on substitution to physical primary care, as it is associated with higher costs than online care.[11] Table 3 shows that for baseline cases directly assigned to a doctor, 15% result in a physical consultation with other providers within 7 days. However, the effect of an initial nurse assignment yields conflicting results: the OLS estimate suggests a modest 2.5 percentage-point increase in external PCP consultations, while the IV estimate shows a small and statistically insignificant decrease. Furthermore, external PCP consultations are not significantly correlated with the congestion instrument in reduced-form regressions.

When we disaggregate by external clinician type, the results appear driven by external doctor consultations: 12% of cases at baseline involve an external doctor consultation, with OLS and IV coefficients comparable to those for overall consultations. Appendix Table A8 indicates that including any external PCP contact (such as video or telephone consultations) yields similar results, apart from a higher baseline rate.

In summary, we find inconclusive evidence on the effect of the nurse-initiated knowledge hierarchy on external primary care consultations. Overall, our results suggest at most a modest increase in external PCP consultations under the knowledge hierarchy compared to a direct-to-doctor assignment.

4.2.2 Patient satisfaction

Second, we evaluate patient-reported satisfaction with their care. After each consultation, patients are asked to rate how their case was handled in two ways: first, by providing an overall satisfaction score from one to five stars, and second, by indicating with "yes," "no," or "don't know" whether the consultation replaced the need for in-person care. These ratings allow us to assess whether patients in the nurse-initiated knowledge hierarchy are less satisfied with their care, potentially due to perceived quality differences or an expectation to speak with a doctor rather than a nurse.

We consider any positive rating received over a case's full care journey within seven days of the

[10]As the primary care data from Region Scania also includes children's care centers and maternity services (Barnavårdscentralen, BVC), we may overestimate the baseline odds of actual primary care visits. However, this is unlikely to bias our treatment estimates.

[11]In addition, only 1.2% of overall cases in our sample are followed up by any external online consultation in the 7 days following their consultation at the provider we study.

initial consultation.[12] As shown in Table A8, an initial assignment to a nurse has, at most, a modest positive effect on patient ratings. Patients are 4.2 (IV) to 4.5 (OLS) percentage points more likely to provide a top score during their initial or any follow-up consultation, relative to a baseline of 42% top score ratings. For the question of whether online consultations replace physical care, the effects of an initial nurse assignment are less conclusive. The IV estimates suggest a 6 percentage-point increase in positive responses from a baseline of 45%, while the OLS estimates indicate a smaller increase of 1.5 percentage points. Additional analyses in Appendix Table A8 further show that an initial nurse assignment does not increase negative (below top score) ratings, although there is a small increase in negative responses regarding physical care replacement based on the OLS estimate.

Overall, our results suggest an improvement in patient satisfaction ratings within the knowledge hierarchy, with no definitive impact on patients' perceptions of whether physical care can be replaced.

4.2.3 Diagnosis, prescriptions and specialist visits

Third, we assess whether the nurse-initiated knowledge hierarchy impacts patients' access to services that doctors specialize in: diagnostic information, prescriptions, and referrals to specialists. While nurses are not specialized in diagnosing and not authorized to provide prescriptions or specialist referrals, cases requiring these services may be referred up in the knowledge hierarchy to a doctor.

To measure diagnostic quality, we consider whether any informative ICD diagnosis code was provided in the patient records within seven days following the initial consultation.[13] Table 3 shows that initial assignment to a nurse decreases the likelihood of receiving an informative diagnosis by 6.1 (IV) to 6.4 (OLS) percentage points from a baseline rate of 81%, indicating worse diagnostic quality under the nurse-initiated knowledge hierarchy.

To examine access to prescription medications, we consider whether any new prescriptions, which patients do not regularly purchase for chronic conditions, are filled within seven days after the consultation.[14] Table 3 shows that the initial nurse assignment considerably decreases the odds of

[12]Of initial consultations, 48.22% (45.64%) received any satisfaction score (physical replacement rating); when including subsequent consultations, 50.94% (48.27%) had any score (physical replacement rating).

[13]We consider ICD codes as *informative* if they are non-missing and exclude R00-R99 or Z00-Z99, as these codes only record symptoms and health status updates rather than the diagnosis of a condition. These ICD codes are still commonly used also in physical primary care: Only 68.63% of the diagnoses provided in physical nurse or doctor consultations in Scania are informative.

[14]We exclude any drugs from the Anatomical Therapeutic Chemical level 4 (chemical subgroups), which patients have already purchased at pharmacies within the 6 months prior to the initial consultation. External prescriptions are also included in this analysis.

receiving a prescription by 6 (IV) to 10 (OLS) percentage points from a baseline of 44%. Notably, prescription rates in nurse-initiated cases remain relatively high, at least 34%, suggesting that referrals up in the knowledge hierarchy may primarily stem from occupational licensing restrictions, as nurses cannot prescribe.

For specialist referrals, we observe overall visits to specialists but not direct referrals. Table 3 presents mixed results on the effect of an initial nurse assignment on specialist service utilization within the next seven days: The OLS estimate indicates a small increase of 0.3 percentage points in specialist visits, whereas the IV estimate shows a marginally significant (at the 10% level) decrease of 1.7 percentage points, relative to a baseline of 3.6% for direct-to-doctor assignments.

These results indicate that the nurse-initiated knowledge hierarchy consistently reduces patients' access to high-quality diagnostic information and decreases the likelihood of receiving a prescription, with mixed effects on access to specialist care. However, these findings do not necessarily imply that patients receive less appropriate care.

4.2.4 Acute care

Fourth, we examine the effect of the nurse-initiated knowledge hierarchy on patients' use of acute care services, including high-cost emergency and hospital services. Specifically, we consider visits to urgent care centers – out-of-hours primary care facilities implemented in some regions to alleviate pressure on emergency care – as well as visits to emergency departments (ED), overall hospitalizations, and avoidable hospitalizations. These outcomes serve as direct measures of potential adverse health events that may arise during patients' care journeys.

Table 3 presents conflicting results regarding the effect of an initial nurse assignment on the utilization of higher-cost acute care services. For urgent care services, the OLS estimate indicates a slight increase of 0.19 percentage points from a baseline of 1.7%, whereas the IV estimate shows no effect. For ED visits, the OLS estimate again points to an increase of 0.61 percentage points from a baseline rate of 0.3%, while the IV estimate suggests a marginally significant decrease (at the 10% level) of 1.6 percentage points. Regarding inpatient hospitalizations, we observe no significant effects of the initial nurse assignment in either specification. We also consider avoidable hospitalizations, defined as conditions considered preventable in primary care.[15] Here, the results

[15]Avoidable hospitalizations, also termed hospitalizations for ambulatory care sensitive care conditions, indicate

are again inconclusive, with a marginally significant (at the 10% level) OLS estimate indicating a slight increase due to the initial nurse assignment.

Overall, we do not observe consistent and substantial increases in the uptake of acute care services following the initial consultation for cases that are handled in the knowledge hierarchy instead of being directly managed by doctors.

4.2.5 Adverse events

Lastly, we examine medium- and long-term patient outcomes, specifically income reductions and mortality, which may arise from the insufficient handling of a patient's case. While patterns of utilization and the occurrence of high-cost acute care events provide insights into immediate care issues, some cases may require escalation to a higher level of care and others may not utilize healthcare services despite a deteriorating condition. Therefore, we turn to quality measures that more objectively capture events that adversely affect patients' lives.

We begin by investigating the occurrence of income reductions in the calendar month following the initial consultation, relative to the average income in the three months prior, from the patient's main income source. Our data includes monthly gross earnings from the main employer, which is defined as the largest source of income at the end of each calendar year. To minimize measurement error, we restrict our analysis to cases where patients have reasonably stable employment over the three months prior to the initial consultation.[16] Notably, any observed income reductions also capture sick leave, as employers in Sweden replace only 80% of regular salary for the first 14 days of sick leave, and after 14 days, income from the employer falls to zero as the Swedish Social Insurance Agency begins to pay sickness benefits.

Table 3 shows that, if anything, the nurse-initiated knowledge hierarchy *reduces* the rate of major income reductions. Our OLS estimates suggest a moderate decrease of 2.1 percentage points in the likelihood of any income reduction exceeding 20%, compared to a baseline rate of 22% under the direct-to-doctor assignment. When considering a full reduction to zero income from the main

hospitalizations for diagnoses that are considered preventable under appropriate primary care. We construct avoidable hospitalizations based on a list of ICD diagnostic codes as provided in Page et al. (2007).

[16] Specifically, we consider cases where patients are employed at the end of the year, and their income in the three months preceding the consultation exceeds 3,533.33 SEK each month (in 2010 values). This threshold is based on the annual threshold of 42,400 SEK used by Saez et al. (2019) to ensure we do not capture employees changing jobs or seasonal workers. We focus on employed patients because monthly income from self-employment is calculated as the annual income divided by twelve in our data.

employer, the OLS estimate indicates a small decrease of 0.34 percentage points in the rate of zero income, compared to a baseline rate of 5%. None of the IV estimates are statistically significant.

Next, we examine the effects of the nurse-initiated knowledge hierarchy on patients' three-year mortality following the consultation date. We focus on mortality from health conditions, excluding external causes such as accidents or self-inflicted harm.[17] Table 3 reveals no causal effect of initial nurse assignment on the three-year mortality rate.

In Appendix Table A8, we consider an alternative definition for income reductions. We examine income drops corresponding to the pattern of a full month of sick pay along with sickness benefits received during the calendar year. These additional results reflect the same overall pattern as our main outcomes: If anything, the knowledge hierarchy slightly reduces the rate of income reductions, as suggested by the OLS estimate.

Taken together, our results suggest that the nurse-initiated knowledge hierarchy has no negative impact on medium- or long-term outcomes for cases at the margin, compared to a mode of work in which all cases are handled by doctors.

4.3 Cost analysis

In the last set of main results, we consider the costs associated with service provision throughout patients' journeys in the healthcare system. Given that healthcare in Sweden is publicly funded, the majority of costs are covered through taxpayers' contributions. We take on the perspective of the public payer and examine whether any cost differences, including those from the utilization of healthcare services downstream, arise between the nurse-initiated knowledge hierarchy and direct-to-doctor assignments.

To compute cost differences between the nurse-initiated knowledge hierarchy and direct-to-doctor assignments, we consider the cost categories listed in Table 4. These include the costs of consultations at the provider, along with additional costs from drug prescriptions, specialist visits, and acute care services. As we only observe external primary care for a subsample of approximately 10% of cases from patients in Scania, we account for these costs by extrapolating our previous estimates.

We first provide an estimate of the total costs for a baseline case assigned directly to a doctor at

[17] In particular, we exclude deaths due to accidents, self-inflicted harm, and other external causes of death (ICD codes V01-Y89, excluding X41-X42 and X44-X45, but including U12.9).

our healthcare provider. Table 4 shows that these costs amount to roughly 988 SEK (about 90 USD) when excluding external primary care, or 1,250 SEK (about 115 USD) when including them. These figures are derived from a back-of-the-envelope calculation, where each cost category is weighted by the baseline rate of service usage, based on sample averages under a direct-to-doctor assignment.[18]

Next, Table 5 presents estimates of the cost differential between the nurse-initiated knowledge hierarchy and the direct-to-doctor assignment. We compute total costs for each case based on the categories in Table 4, and then take the natural logarithm. At the provider, we account for the initial consultation and any referral to a doctor, weighted by the costs of an online primary care consultation. Our results indicate moderate cost savings under the knowledge hierarchy. In the full sample (excluding external primary care costs), we estimate savings of 12% in the OLS specification, amounting to about 118.5 SEK from the baseline cost of 988 SEK. When we account for the potential of residual case sorting using our IV strategy, we estimate savings of up to 20%, or 198 SEK.

These cost savings may be reduced when external primary care consultations are included. We consider differences in external primary care use as estimated in Table 3. Savings then remain at about 7.5% in the OLS specification and are unchanged for the IV specification.[19]

In Appendix Table A8, we examine the subset of cases from patients residing in Scania, where we directly observe external primary care use. While our estimates for this subsample are less precise due to the reduced sample size, they support our findings from the full analysis sample: we estimate cost savings under the knowledge hierarchy (compared to direct-to-doctor assignments) of between 6.4% (OLS) and 20% (IV).

As highlighted in previous results, we find little evidence of increased downstream healthcare utilization among patients in the knowledge hierarchy compared to the baseline. However, at the provider level, we observe that nurses are able to resolve up to 70% of marginal cases without referring them up in the knowledge hierarchy. In addition, nurse consultations are substantially cheaper, costing only 55% of the cost of a doctor consultation, as shown in Table 4. Appendix Table A8 further supports this, showing that the cost reductions observed are primarily driven by lower

[18]As nurse consultations at the provider occur in only a small share of direct-to-doctor cases (with nurses spending 0.07 minutes on average), we simplify the cost calculations by excluding nurse consultations in these cases.

[19]Our estimates indicate an increase of at most 2.2 percentage points in the rate of external primary care doctor consultations under the OLS specification. Considering a cost of physical doctor consultations of 2,002 SEK, this would imply additional costs of at most 44.04 SEK under the knowledge hierarchy. The IV estimate is not statistically significant.

costs at our provider, where savings range from 11% (OLS) to 13% (IV).

Overall, our estimates suggest marginal reductions in healthcare costs from the nurse-initiated knowledge hierarchy of up to 20% in the IV and at least 7.5% in the OLS specification. We also find no evidence of decreases in patient satisfaction or an increased occurrence of adverse patient events. Thus, the cost savings from the nurse-initiated knowledge hierarchy indicate marginal efficiency gains from assigning cases initially to a nurse.

Table 4. Baseline costs of the direct-to-doctor assignment

Cost category	Corresponding outcome	Cost (SEK)	Factor	Estimate (SEK)
Online primary care consultation, doctor		500	1	500
Online primary care consultation, nurse[1]		275	0	0
Physical primary care consultation, doctor[2]	Any external PCP consultation, doctor	2002	0.12	240.24
Physical primary care consultation, nurse[2]	Any external PCP consultation, nurse	614	0.043	26.402
Prescription	Any new prescription	260	0.44	114.4
Specialist visit	Any specialist	3594	0.04	143.76
Emergency department visit	Any ED	3911.5	0.03	119.745
Urgent care center visit	Any urgent care	2002	0.02	40.04
Hospitalization	Any hospitalization	7800	0.009	70.2
Total baseline cost in SEK, incl. physical primary care			1254.787	
Total baseline cost in SEK, excl. physical primary care			988.145	

Notes: Costs for the various services are sourced primarily from public reports or regional announcements, with the exception of prescription costs, which are obtained directly for the analysis sample. Additional details are provided in Appendix Table A16.
[1] We assume that a direct-to-doctor assignment involves no nurse consultations at the provider.
[2] Physical primary care is only observed within the region of Scania.

4.4 Sensitivity

We now present several sensitivity tests to assess the robustness of our main results. First, we confirm that there are no pre-existing level or trend differences between the initial case assignments prior to the consultation. Second, we demonstrate that our findings remain consistent when using two alternative measures of congestion as instrumental variables: one based on the number of active clinicians and another based on the number of nurse consultations within cases that are almost always assigned to doctors. Finally, we verify the robustness of our main results across alternative control variables, different time windows for login, and varying assumptions regarding the structure of standard errors.

Table 5. Effect of an initial nurse assignment on healthcare costs

	Log costs, excl. ext. PCP		
	OLS	Red.	IV
Initially to nurse	-0.12***		-0.20***
	(0.0028)		(0.030)
Nurse share, past 60 min.		-0.076***	
		(0.012)	
Baseline characteristics	✓	✓	✓
Additional controls	✓	✓	✓
Observations	490,505	490,505	490,505
First-stage K-P F-statistic			3,034
Baseline mean	6.68	6.68	6.68

Note: This table presents regression results for our main specifications on the effects of initial case assignment to a nurse (the *knowledge hierarchy*) rather than directly to a doctor. Each subpanel corresponds to a different outcome as the dependent variable. Outcomes are defined as follows: Outcomes are defined as follows: "Log costs, excl. ext. PCP" represents the natural logarithm of the estimated care costs. These costs include any downstream expenses from online doctor or nurse primary care consultations at the provider, such as prescriptions, specialist visits, emergency department visits, urgent care visits, and hospitalizations that occur within seven days of the initial consultation, but exclude the costs of external primary care consultations. Estimates for each cost category are provided in Table 4. The left column in each subpanel reports the Ordinary Least Squares (OLS) estimate for the treatment variable (*Nurse*) based on Equation 1. The middle column presents the reduced-form regression estimates for the congestion instrument, measured as the share of nurse consultations in the past 60 minutes (*Nurse share, past 60 min.*). The right column provides the Instrumental Variables specification, using *Nurse share, past 60 min.* as an instrument for *Nurse*, estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2. All regressions control for symptom categories, demographics, region, health risk factors, and socioeconomic variables, along with login date-by-4-hour fixed effects. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

4.4.1 Outcome dynamics

In Figure 3, we present OLS and IV estimates of the effects of the nurse-initiated knowledge hierarchy from separate regressions of patient outcomes in the weeks leading up to and following the initial consultation. Subfigure 3a shows that patients have consultations with the provider in a similar manner prior to the initial consultation. In the week of the consultation, the nurse-initiated knowledge hierarchy significantly increases the likelihood of a follow-up consultation. Similarly, subfigures 3b, 3c, and 3d show no differential pre-trends in the rates of prescribing, emergency department visits, or hospitalizations.[20] Furthermore, any effects of the knowledge hierarchy disappear after the week of the consultation.

These figures thus indicate that health outcomes in the weeks before the initial consultation do not significantly differ between patients initially assigned to a nurse and those directly assigned to a doctor. Moreover, any effects we observe occur within the week of the consultation, supporting our choice of defining a seven-day care episode.

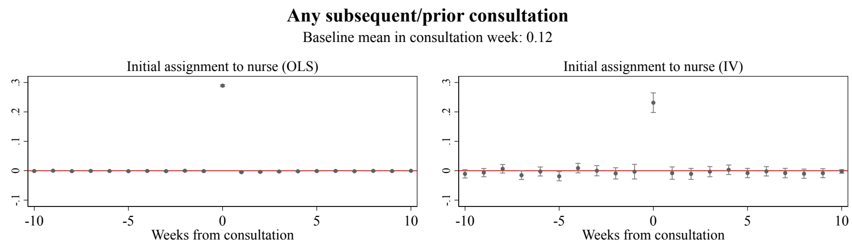
4.4.2 Alternative instrumental variables

To further strengthen our main results regarding the effect of the knowledge hierarchy, we construct two additional congestion measures and estimate our IV specifications using these alternative measures as instrumental variables for the initial case assignment. The results from these alternative instruments are largely consistent with our main findings, supporting the robustness of our conclusions.

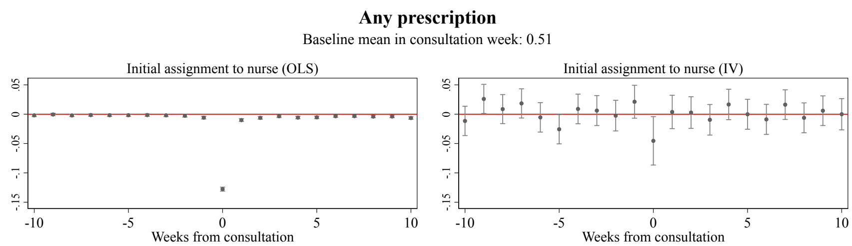
Our first instrument, a *staffing instrument*, is defined as the share of nurses among clinicians who have any consultations for marginal cases in the 60 minutes prior to the current case. We assume that at least one more doctor is staffed than nurses and exclude a case's own clinician from this measure.[21] The staffing instrument is based on the logic that when congestion among doctors increases, more nurses are required to fill in and handle excess demand. Appendix Figure A11 shows the variation in the distribution of this alternative instrument and a strong, non-parametric relationship with an initial nurse assignment. Appendix Table A12 presents IV estimation results for a set of main outcomes, showing that the estimated effects of the knowledge hierarchy are numerically similar to

[20] In the main analysis, we consider prescriptions that have not been filled in the 6 months prior to the initial consultation in order to capture the likelihood of an additional prescription. As we aim to test for general prescribing patterns, including those for chronic medications, we include any prescription in this analysis.

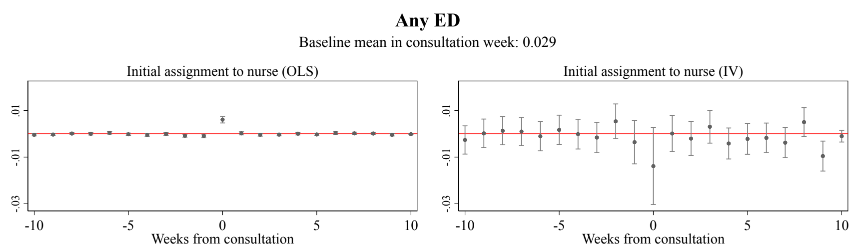
[21] Our data does not allow us to observe clinicians who are staffed but do not take any consultations.



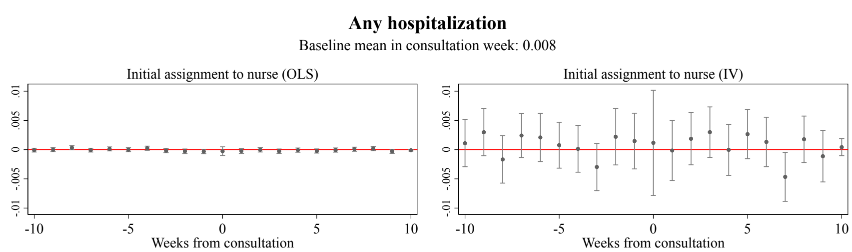
(a) Any consultation at the provider



(b) Any prescription



(c) Any emergency department visit



(d) Any hospitalization

Figure 3. Effect of an initial nurse assignment on lags and leads of patient outcomes

Notes: These figures show the estimated effect of an initial assignment to a nurse (the *knowledge hierarchy*) for outcomes in the 10 weeks prior to or after the initial consultation. "Week 0" marks the week starting with the initial consultation date. Each estimate is based on a separate regression of a lagged or lead outcome on the treatment variable (*Nurse*). Estimates are obtained by Ordinary Least Squares (OLS, left) based on Equation 1, or Two-Stage Least Squares (IV, right) using *Nurse share, past 60 min.* as an instrument for *Nurse* based on Equations 1 and 2. All regressions control for the full set of case characteristics, along with login date-by-4-hour fixed effects. The baseline mean represents the average of an outcome in week 0 for cases directly assigned to a doctor.

those in our main analysis.

Our second instrument, a *doctor shortage instrument*, is the count of nurse-initiated cases in the past 60 minutes that should almost certainly be assigned to doctors. We define symptom categories with an initial consultation share above 95% as cases that are almost always assigned to doctors but still exhibit some variation in their assignment. Given that this variable is right-skewed, we winsorize it at the 99th percentile to reduce noise. Appendix Figure A12 shows that, in over 60% of cases, no doctor-deterministic cases are assigned to a nurse. However, in less than 10% of cases, this occurs 3 or more times, and there is a positive relationship between this instrument and nurse assignment.

Appendix Table A13 shows that, while the first-stage F-statistic is weaker and estimates are less precise with this alternative instrument, the IV results still largely support our main results. Specifically, we find similar effects of the knowledge hierarchy on the rate of referrals to a doctor, a larger reduction in the rate of new prescriptions (18 percentage points instead of 6 percentage points at a baseline of 44%), and a marginally significant but similarly sized reduction in costs. Additionally, there are no adverse effects on measures of patient health or adverse events.

4.4.3 Robustness of the econometric specification

In Appendix Table A9, we assess the robustness of our results to systematic expansions of the set of control variables. The OLS and IV estimates for our main patient outcomes remain nearly unchanged across these specifications.

In Appendix Table A10, we consider alternative sets of fixed effects to control for the login-time of a case. Most of our conclusions remain unchanged when we use year, month, and weekday instead of calendar date or do not consider 4-hour windows. However, we observe a change in the effect of the knowledge hierarchy on income reductions in the IV estimates when we use less granular time windows. Despite this, our OLS estimates continue to show a reduction in the rate of income reductions, regardless of how we specify time-fixed effects.

In Appendix Table A11, we present standard errors computed under alternative assumptions regarding their structure. These do not substantially differ when clustering by login date or by login date within 4-hour windows.

5 Mechanisms

The previous section discussed the causal effects of the nurse-initiated knowledge hierarchy on the provider, patients, and healthcare costs. In this section, we provide further insights into how different tasks are handled within the knowledge hierarchy. First, we descriptively characterize the types of cases that are more likely to be referred up the hierarchy. Specifically, we find that more complex cases – such as those involving older patients or higher health risks – are more often sent to doctors. Additionally, clearly defined symptoms that likely require care outside of a nurse's scope of practice are more likely to be moved up in the hierarchy. Second, we show that these more complex cases are also where the knowledge hierarchy appears to be most effective. Finally, we highlight that cases moved up due to congestion tend to be less clearly defined, suggesting that more cost-efficient cases in the knowledge hierarchy are also more likely to be assigned to it on the margin.

5.1 Characterizing referrals in the knowledge hierarchy

In the previous section, we have shown that about 30% of marginal cases are referred up the knowledge hierarchy. At the same time, the costs imposed on the healthcare system are largely driven by the additional consultations required for nurse-to-doctor referrals. We now aim to better understand what characterizes cases more likely to be referred up the hierarchy, and which cases nurses are able to resolve on their own.

Figure 4 examines the bivariate correlations between referrals up the knowledge hierarchy and various case characteristics. The figure highlights several key differences in referral rates based on symptom type, age, and health risk, allowing us to draw two main conclusions. First, cases involving older patients, those with comorbidities, or those requiring specialist visits or prior hospitalization are more likely to be referred to doctors. This is consistent with the theory behind the knowledge hierarchy, which suggests that more complex cases should be handled by more specialized professionals. Appendix Table A15 supports this notion by showing that doctors spend more time on cases referred to them by nurses compared to cases directly allocated to them.

Second, Figure 4 reveals that certain more clearly defined symptoms – such as urinary tract infections, sore throats, or bites and stings – are more likely to be referred up in the knowledge the hierarchy. In contrast, less clearly defined symptoms, such as uncategorized symptoms, abdominal pain, or health inquiries, are less likely to be forwarded. Additionally, cases involving children are less likely to be referred. These findings suggest a subset of tasks defined by patient age and symptoms clarity may result in larger cost savings in the knowledge hierarchy.

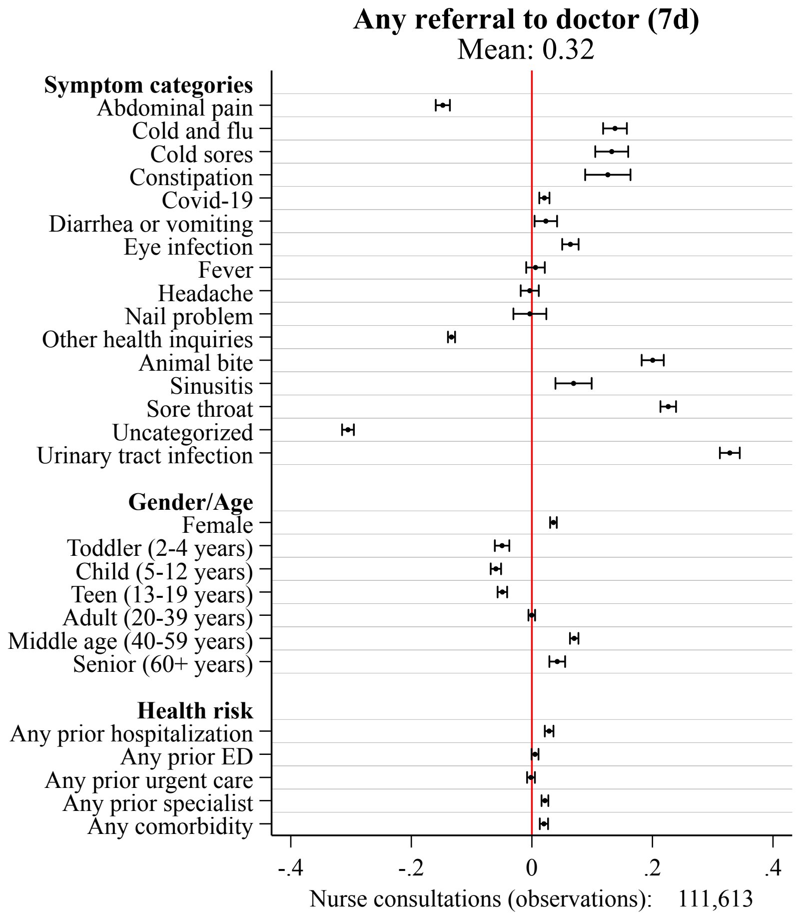


Figure 4. Correlates of referrals from nurses to doctors

Notes: This figure presents, for the subset of cases initially assigned to a nurse, correlates for referrals to doctors (*Referral to doctor*) and thus up in the knowledge hierarchy. Each row shows the results of a bivariate regression where *Referral to doctor* is regressed on a specific case characteristic. All regressions account for the login date-by-4 hours fixed effects. Symptom categories refer to indicators for the main symptom category reported by the patient when requesting a consultation. The horizontal lines represent 95% confidence intervals based on robust standard errors. Mean refers to the average rate at which cases are forwarded from a nurse to a doctor in the sample of nurse-initiated cases.

5.2 Task heterogeneity

Building on our finding that cases with unclear symptoms or involving children are less likely to be referred up the knowledge hierarchy, we now explore whether these lower referral rates reflect a comparative advantage of the knowledge hierarchy in these case types. Specifically, we ask: Are these the cases where the knowledge hierarchy is most effective in reducing healthcare costs?

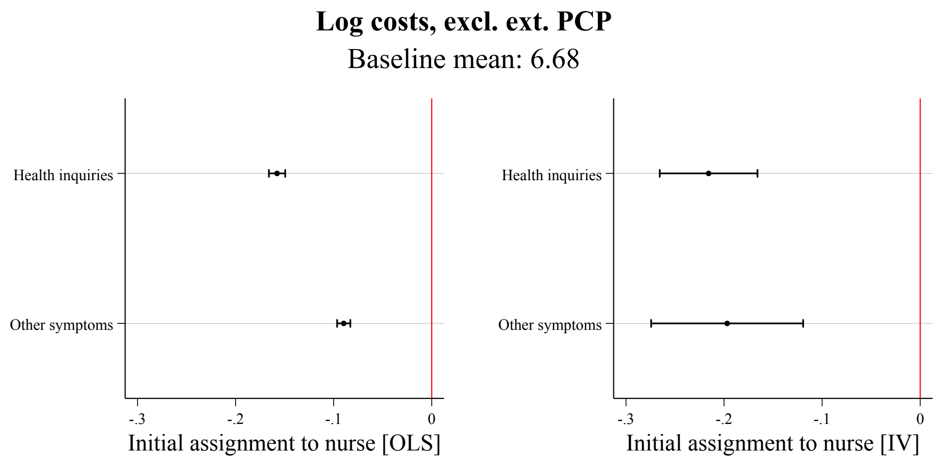
To this end, we estimate the following specification:

$$Y_i = \sum_g \delta_g (I_{g(i)=g} \times Nurse_i) + \kappa_g + X_{1i}\tilde{\beta}_1 + X_{2i}\tilde{\beta}_2 + \tilde{\epsilon}_i, \tag{3}$$

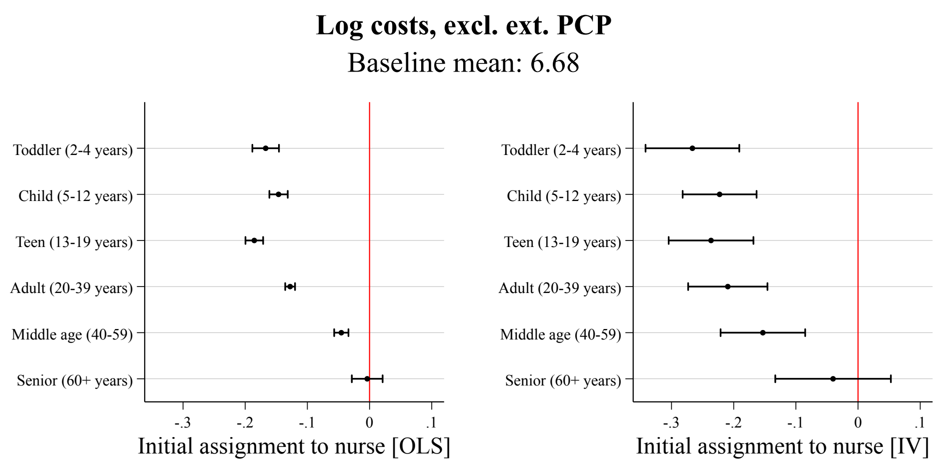
where g represents a subgroup, $I_{g(i)=g}$ indicates that case i belongs to subgroup g, and the other variables are defined as in 1. In the IV approach, we instrument $(I_{g(i)=g} \times Nurse_i)$ by $(I_{g(i)=g} \times PrevNurseShare_i)$, as described in 2. Estimates of δ_g provide insight into the subgroup-specific effects of the knowledge hierarchy on costs.

Figure 5 presents the estimates of δ_g, comparing subgroups g based on symptom categories, specifically "Other health inquiries" versus other symptoms, and comparing across age groups. The figure supports the idea that cost savings from the knowledge hierarchy are most pronounced for cases with unclear symptoms or for younger patients. In subfigure 5a, our OLS estimates show that the knowledge hierarchy reduces costs for "other health inquiries" by nearly twice as much as for other symptoms (16% vs. 9%), although the IV estimates suggest smaller cost differences. Subfigure 5b reveals an age gradient in cost reductions: Among older patients, the knowledge hierarchy shows no effect on costs, while the strongest cost reductions are observed for toddlers, children, and teens (up to 18% in the OLS and 26% in the IV approach). Appendix Figure A13 further shows that these cost savings are driven by lower rates of referrals up the knowledge hierarchy.

One potential explanation for the observed heterogeneity across symptoms and age is that some cases may involve greater uncertainty about the patient's condition. For example, children are often brought in through their parents and may not be able to clearly articulate their symptoms, limiting the information available. In these cases, where the severity of the condition is uncertain, nurses may be capable of providing appropriate care and prevent unnecessary doctor visits, which results in cost savings.



(a) Heterogeneity by main symptom category



(b) Heterogeneity by age group

Figure 5. Heterogeneity across cases in total costs, excluding external primary care

Note: These figures show the heterogeneous effects of an initial assignment to a nurse (the *knowledge hierarchy*) on costs, broken down by subgroups defined either by whether the main reported symptom is "Other health inquiries" (subfigure 5a) or by age category (subfigure 5b). The outcome "Log costs, excl. ext. PCP" represents the natural logarithm of the estimated care costs. These costs include any downstream expenses from online doctor or nurse primary care consultations at the provider, such as prescriptions, specialist visits, emergency department visits, urgent care visits, and hospitalizations that occur within seven days of the initial consultation, but exclude the costs of external primary care consultations. Each subfigure presents estimates from a regression of the outcome on all interactions of the initial nurse assignment (*Nurse*) with each subgroup, while controlling for subgroup fixed effects, the full set of case characteristics, and login date-by-4-hour fixed effects, following Equation 3. The estimates are obtained by Ordinary Least Squares (OLS, left) or Two-Stage Least Squares (IV, right), where interactions of *Nurse share, past 60 min.* with each subgroup serving as instruments for the interactions between *Nurse* and each subgroup. Each row represents the estimated subgroup-specific effect of an initial nurse assignment, with horizontal lines representing 95% confidence intervals based on robust standard errors.

5.3 Complier characteristics

Finally, we seek to better understand which cases are moved to the nurse-initiated knowledge hierarchy under increasing congestion. Our IV estimates, as presented in Section 4, represent a local average treatment effect (LATE), that is, the average causal effect of the nurse-initiated knowledge hierarchy on the subgroup of compliers whose probability of initial assignment to a nurse is influenced by congestion. We now characterize this group of complier cases to investigate whether the allocation of tasks corresponds to the heterogeneous effects of the knowledge hierarchy on costs.

Appendix Table A14 compares the mean characteristics of the complier group with those of the overall analysis sample, following (Frandsen et al. 2023) and (Abadie 2003). For complier cases, the reported symptom is substantially more likely to be classified as "Other health inquiries" (a catch-all category) and somewhat more likely to be Covid-19. In contrast, symptoms with very low rates in the complier group, such as urinary tract infections, sore throats, or eye infections, all which may require a prescription, tend to be more specific. Complier cases are also less likely to be female, slightly younger on average, and show small differences in other risk or socio-economic characteristics compared to the full sample. However, these differences are less pronounced than those given by the symptom category. Overall, these results suggest that cases assigned to the knowledge hierarchy are less precisely defined.

6 Conclusion

In knowledge-intensive sectors where expertise is scarce, and costs are escalating, optimizing task allocation across differentiated skills is becoming increasingly critical. We study the effects of the division of labor when tasks are organized in a knowledge hierarchy. We focus on an industry particularly affected by high costs and scarce resources: healthcare.

Our study offers evidence that a nurse-initiated knowledge hierarchy in primary care reduces costs imposed on the public healthcare system without compromising care quality. We find that this approach achieves cost savings of 7.5 to 20%, driven by nurses' ability to resolve 70% of cases on their own, with little evidence of lower quality. The knowledge hierarchy maintains high levels of patient satisfaction and does not adversely affect health outcomes, as measured by acute care needs, labor market outcomes, or patient mortality. We find that the cost of healthcare is particularly reduced for cases with undefined symptoms or involving younger patients when allocated to the nurse-initiated knowledge hierarchy.

Our findings contribute to the understanding of how the division of labor can enhance productivity by aligning tasks to occupational competencies in healthcare, within the organization, and downstream. Beyond healthcare, our results have implications for similar knowledge-intensive sectors with highly specialized occupations, suggesting that the organization of tasks in a knowledge hierarchy rather than under rigid occupational norms may improve efficiency when differentiated tasks require varying levels of expertise.

References

Abadie, Alberto, "Semiparametric instrumental variable estimation of treatment response models," *Journal of Econometrics*, April 2003, *113* (2), 231–263.

_ **and Matias D. Cattaneo**, "Econometric Methods for Program Evaluation," *Annual Review of Economics*, August 2018, *10* (1), 465–503.

Acemoglu, Daron and David Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 1043–1171.

Alexander, Diane and Molly Schnell, "Just what the nurse practitioner ordered: Independent prescriptive authority and population mental health," *Journal of Health Economics*, July 2019, *66*, 145–162.

Almström, Axel Hellbom, Lina Maria Ellegård, Klara Stromberg, and Andreea Enache, "Mixed Payment and Mixed Objectives: Insights from the Ownership Structure in Swedish Primary Care," 2024.

Angrist, Joshua, Victor Lavy, and Analia Schlosser, "Multiple Experiments for the Causal Link between the Quantity and Quality of Children," *Journal of Labor Economics*, October 2010, *28* (4), 773–824. Publisher: The University of Chicago Press.

Arrow, Kenneth J., "Uncertainty and the Welfare Economics of Medical Care," *The American Economic Review*, 1963, *53* (5), 941–973.

Baicker, Katherine and Amitabh Chandra, "The Productivity of Physician Specialization: Evidence from the Medicare Program," *American Economic Review*, April 2004, *94* (2), 357–361.

Becker, Gary S. and Kevin M. Murphy, "The Division of Labor, Coordination Costs, and Knowledge," *The Quarterly Journal of Economics*, 1992, *107* (4), 1137–1160. Publisher: Oxford University Press.

Bhuller, Manudeep, Gordon B Dahl, Katrine V Løken, and Magne Mogstad, "Incarceration, Recidivism, and Employment," *Journal of Political Economy*, 2016.

Biasi, Barbara, Chao Fu, and John Stromme, "Equilibrium in the Market for Public School Teachers: District Wage Strategies and Teacher Comparative Advantage," *NBER Working Paper*, 2021.

Bronsoler, Ari, Joseph Doyle, and John Van Reenen, "The Impact of Health Information and Communication Technology on Clinical Quality, Productivity, and Workers," *Annual Review of Economics*, August 2022, *14* (Volume 14, 2022), 23–46. Publisher: Annual Reviews.

Cattaneo, Matias D., Richard K. Crump, Max H. Farrell, and Yingjie Feng, "On Binscatter," *American Economic Review*, May 2024, *114* (5), 1488–1514.

Centers for Medicare & Medicaid Services, "Medicare Telehealth Trends," 2024. Accessed: 2024-10-28.

Chan, David and Yiqun Chen, "The Productivity of Professions: Evidence from the Emergency Department," October 2022.

Chan, David C., "Teamwork and Moral Hazard: Evidence from the Emergency Department," *Journal of Political Economy*, June 2016, *124* (3), 734–770.

Chandra, Amitabh and Douglas O. Staiger, "Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks," *Journal of Political Economy*, February 2007, *115* (1), 103–140. Publisher: The University of Chicago Press.

_ , Carrie Colla, and Jonathan Skinner, "Productivity Variation and Input Misallocation: Evidence from Hospitals," Technical Report w31569, National Bureau of Economic Research, Cambridge, MA August 2023.

Costinot, Arnaud, "An Elementary Theory of Comparative Advantage," *Econometrica*, 2009, *77* (4), 1165–1192.

Currie, Janet and Jonathan Zhang, "Doing More with Less: Predicting Primary Care Provider Effectiveness," *The Review of Economics and Statistics*, February 2023, pp. 1–45.

_ and W Bentley MacLeod, "Diagnosing Expertise: Human Capital, Decision Making, and Performance among Physicians," *Journal of labor economics*, 2017, *35* (1), 1–43.

Dahlstrand, Amanda, "Defying Distance? The provision of services in the digital age," 2024.

_ , Nestor Le Nestour, and Guy Michaels, "Online versus In-Person Services: Effects on Patients and Providers," 2024.

Dillender, Marcus, Anthony T. Lo Sasso, Brian J. Phelan, and Michael Richards, "Occupational Licensing and the Healthcare Labor Market," *Journal of Human Resources*, February 2024, pp. 0722–12450R2.

Dobbie, Will, Jacob Goldin, and Crystal S. Yang, "The Effects of Pre-Trial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges," *American Economic Review*, February 2018, *108* (2), 201–240.

Dornbusch, R., S. Fischer, and P. A. Samuelson, "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods," *The American Economic Review*, 1977, *67* (5), 823–839.

Doyle, Joseph J., Steven M. Ewer, and Todd H. Wagner, "Returns to physician human capital: Evidence from patients randomized to physician teams," *Journal of Health Economics*, December 2010, *29* (6), 866–882.

Frandsen, Brigham, Lars Lefgren, and Emily Leslie, "Judging Judge Fixed Effects," *American Economic Review*, January 2023, *113* (1), 253–277.

Garicano, Luis, "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 2000, *108* (5), 874–904. Publisher: The University of Chicago Press.

Guo, Jiapei, Angela E. Kilby, and Mindy S. Marks, "The impact of scope-of-practice restrictions on access to medical care," *Journal of Health Economics*, March 2024, *94*, 102844.

Heckscher, Eli, "The Effect of Foreign Trade on the Distribution of Income," *Ekonomisk Tidskrift*, 1919, *21*, 497–512.

Jones, Benjamin F., "The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?," *The Review of Economic Studies*, January 2009, *76* (1), 283–317.

Katz, Lawrence F. and Kevin M. Murphy, "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, February 1992, *107* (1), 35–78.

Kelly, Elaine, Carol Propper, and Ben Zaranko, "Team Composition and the Returns to Human Capital: Evidence from Nursing Teams," *Centre for Economic Policy Research*, 2023.

Kerwin, Jason, Nada Rostom, and Olivier Sterck, "Striking the Right Balance: Why Standard Balance Tests Over-Reject the Null, and How to Fix it," *SSRN Electronic Journal*, 2024.

Kleiner, Morris M., "Battling over Jobs: Occupational Licensing in Health Care," *American Economic Review*, May 2016, *106* (5), 165–170.

Maier, Claudia B. and Linda H. Aiken, "Task shifting from physicians to nurses in primary care in 39 countries: a cross-country comparative study," *European Journal of Public Health*, December 2016, *26* (6), 927–934.

McMichael, Benjamin J., "Supply-side health policy: The impact of scope-of-practice laws on mortality," *Journal of Public Economics*, June 2023, *222*, 104901.

Ohlin, Bertil, "Handelns teori (The Theory of Trade)." PhD dissertation, Stockholm University, Stockholm, Sweden 1924.

Page, Anthea, Sarah Ambrose, John Glover, and Diana Hetzel, *Atlas of Avoidable Hospitalisations in Australia: Ambulatory Care-Sensitive Conditions*, Adelaide: PHIDU, University of Adelaide, 2007.

Rajpurkar, Pranav, Emma Chen, Oishi Banerjee, and Eric J. Topol, "AI in health and medicine," *Nature Medicine*, January 2022, *28* (1), 31–38. Publisher: Nature Publishing Group.

Ricardo, David, "The Theory of Comparative Advantage," in Pierro with Collaboration of M.H. Dobb Sraffa, ed., *Principles of Political Economy and Taxation*, Vol. 1, Cambridge, London: Cambridge University Press, 1817.

Saez, Emmanuel, Benjamin Schoefer, and David Seim, "Payroll Taxes, Firm Behavior, and Rent Sharing: Evidence from a Young Workers' Tax Cut in Sweden," *American Economic Review*, May 2019, *109* (5), 1717–1763.

Shapiro, Carl, "Investment, Moral Hazard, and Occupational Licensing," *The Review of Economic Studies*, 1986, *53* (5), 843–862. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

Silver, David, "Haste or Waste? Peer Pressure and Productivity in the Emergency Department," *The Review of Economic Studies*, May 2021, *88* (3), 1385–1417.

Smith, Adam, *An Inquiry Into the Nature and Causes of the Wealth of Nations* 1776.

Sveriges Kommuner och Regioner, "Ekonomi- och verksamhetsstatistik inom hälso- och sjukvården," 2024. Accessed on October 28, 2024.

Traczynski, Jeffrey and Victoria Udalova, "Nurse practitioner independence, health care utilization, and health outcomes," *Journal of Health Economics*, March 2018, *58*, 90–109.

Weidmann, Ben, Joseph Vecci, Farah Said, David J. Deming, and Sonia R. Bhalotra, "How Do You Find a Good Manager?," July 2024.

World Health Organization, "Global expenditure on health: Public spending on the rise?," December 2021.

Appendix

Does Division of Labor Increase Productivity? Evidence from Primary Care

A Additional sample descriptives

Table A1. Restrictions on the main analysis sample

	Doctor	**Nurse**	**Total**
All observed consultations	86.6%	13.4%	1,814,706
+ Restrict to unscheduled online consultations	80.5%	19.5%	1,229,383
+ Restrict to analysis time window	78.8%	21.2%	1,118,738
+ Exclude patients registered with the provider	79.4%	20.6%	1,047,474
+ Exclude infants	78.7%	21.3%	988,042
+ Exclude follow-ups within two hours	78.8%	21.2%	963,734
+ Exclude consultations with missing characteristics	78.8%	21.2%	959,244
+ Exclude rare symptom categories	79.0%	21.0%	953,359
+ Restrict to marginal symptom categories	77.2%	22.8%	490,507
+ Exclude singleton observations in date-shift cells	77.2%	22.8%	490,505

Note: This table presents the number of observations after applying our sample restrictions. The columns show the share of consultations with a doctor or a nurse, as well as the total sample size. Each row introduces an additional sample restriction: We restrict the sample to unscheduled online consultations with doctors or nurses from 1 April 2019 to 24 December 2020, during which nurse consultations were active and follow-up consultations within 7 days can be observed. We exclude patients registered with the provider as well as infant patients (age below one year), as different internal protocols apply to them. Consultations are excluded if they involve the same patient within two hours, or if data is missing for login time, patient age, gender, region, or migration status. Symptom categories with fewer than 1,000 observations are excluded, and we focus on marginal categories handled by both doctors and nurses, where nurses manage between 5% and 95% of cases. From the remaining sample, we drop login-time cells with singleton observations in our baseline specification. After imposing all restrictions, we are left with 490,505 observations of initial consultations, which we refer to as the main analysis sample.

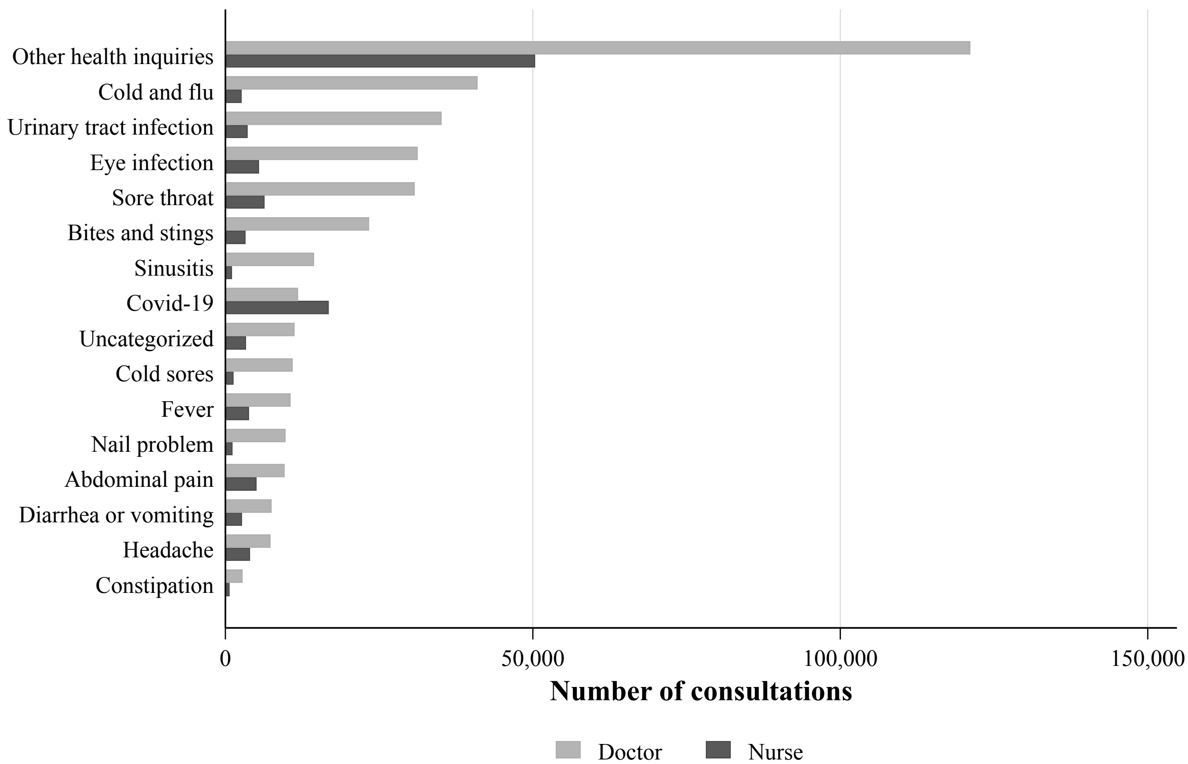


Figure A6. Symptom categories in the analysis sample

Notes: This figure shows the number of initial consultations across symptom categories as reported by patients in our analysis sample, separately for consultations assigned to a doctor or a nurse.

B Tests on the instrument validity

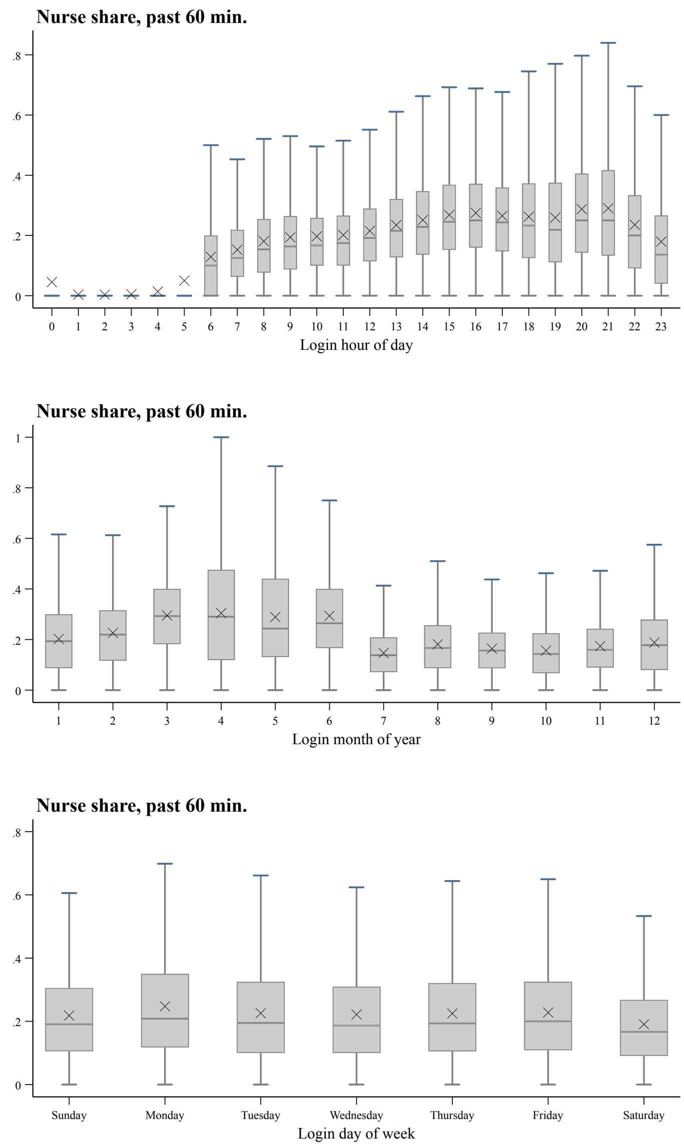


Figure A7. Distribution of the congestion instrument across different time intervals

Notes: This figure presents the distribution of the congestion instrument, the share of nurse consultations in the past 60 minutes, across three dimensions: hour of the day (top panel), month of the year (middle panel), and day of the week (bottom panel). The box plots represent the interquartile range for each time period. The median is shown as a horizontal line inside the box, and the mean is indicated by a cross. The whiskers extend to the upper and lower adjacent values.

Table A2. First stage

	Initial assignment to nurse					
	(1)	(2)	(3)	(4)	(5)	(6)
Nurse share, past 60 min.	0.84	0.38	0.38	0.38	0.38	0.38
	(0.0038)	(0.0073)	(0.0070)	(0.0070)	(0.0070)	(0.0070)
Login time		✓	✓	✓	✓	✓
Symptom categories			✓	✓	✓	✓
Demographics				✓	✓	✓
Health risk					✓	✓
Socio-economic variables						✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505
R-squared	0.10	0.12	0.19	0.19	0.19	0.19
K-P F-statistic	49,277	2,722	3,026	3,028	3,030	3,034

Notes: This table reports the first stage from regressing the treatment (*Initial assignment to nurse*) on the congestion instrument, the share of nurse consultations in the past 60 minutes (*Nurse share, past 60 min.*). Login time indicates login date-by-4-hours fixed effects. Symptom categories refer to indicators for the main symptom category reported by the patient when requesting a consultation. Demographics include indicators for patient gender, age categories, and aggregated regions. Health risk includes indicators for any prior hospitalization, ED visit, urgent care center visit, and specialist visit in the 3 years prior to the consultation, but excluding the 30 days immediately before, as well as an indicator for any comorbidity. Socio-economic variables include indicators for above-median income, benefit receipt, employment type, education level, civil status, and migrant background. Robust standard errors are in parentheses.

Table A3. Conditional balance of the congestion instrument

	Nurse share, past 60 min.			
	(1)	(2)	(3)	(4)
Login time	✓	✓	✓	✓
Symptom categories		✓	✓	✓
Demographics			✓	✓
Health risk				✓
Socio-economic variables				
Observations	490,505	490,505	490,505	490,505
R-squared	0.7	0.7	0.7	0.7
Joint K-P F-statistic	2.93	2.50	1.24	0.94
Joint F-test p-val.	0.00	0.00	0.20	0.52
Joint F-test RI p-val.	0.00	0.00	0.20	0.50

Notes: This table reports balance tests of the congestion instrument, the share of nurse consultations in the past 60 minutes (*Nurse share, past 60 min.*), conditional on varying sets of case characteristics. Control characteristics that the congestion instrument is conditioned on are marked by checkmarks. Missing checkmarks indicate case characteristics that the congestion instrument is balanced against. Login time indicates date-by-4 hours fixed effects. Symptom categories refer to indicators for the main symptom category reported by the patient when requesting a consultation. Demographics include indicators for patient gender, age categories, and aggregated regions. Additional case characteristics include patient health risk and socio-economic characteristics. Health risk includes indicators for any prior hospitalization, ED visit, urgent care center visit, and specialist visit in the 3 years prior to the consultation, but excluding the 30 days immediately before, as well as an indicator for any comorbidity. Socio-economic variables include indicators for above-median income, benefit receipt, employment type, education level, civil status, and migrant background. The test statistics are based on robust standard errors. Based on Kerwin et al. (2024), randomization inference p-values from 500 repetitions are provided alongside conventional ones.

Table A4. Monotonicity subgroup test

	Prior hospitalization		Prior ED visit		Prior specialist visit	
	Any	None	Any	None	Any	None
Nurse share, past 60 min.	0.39***	0.38***	0.41***	0.37***	0.38***	0.39***
	(0.016)	(0.0078)	(0.012)	(0.0086)	(0.0087)	(0.012)
Baseline characteristics	✓	✓	✓	✓	✓	✓
Observations	92,562	397,778	164,242	326,205	312,315	178,108
K-P F-statistic	579	2,442	1,163	1,862	1,938	1,086
	Prior urgent care		**Comorbidity**		**Gender**	
	Any	None	Any	None	Female	Male
Nurse share, past 60 min.	0.40***	0.38***	0.40***	0.38***	0.34***	0.46***
	(0.015)	(0.0080)	(0.016)	(0.0079)	(0.0087)	(0.012)
Baseline characteristics	✓	✓	✓	✓	✓	✓
Observations	115,792	374,587	102,329	388,024	307,543	182,881
K-P F-statistic	755	2,266	672	2,328	1,562	1,505
	Income		**Migration background**		**Further education**	
	$>$ median	$\leq$ median	Any	None	Any	None
Nurse share, past 60 min.	0.38***	0.39***	0.43***	0.37***	0.38***	0.39***
	(0.012)	(0.0085)	(0.014)	(0.0082)	(0.0083)	(0.013)
Baseline characteristics	✓	✓	✓	✓	✓	✓
Observations	159,220	331,156	123,212	367,218	351,374	139,023
K-P F-statistic	938	2,094	1,001	2,023	2,112	895

Note: This table reports the results of Equation 2, which represents the first stage of the IV regression, across various subsamples, including the baseline case characteristics as control variables. The baseline case characteristics includ the symptom categories, demographics, region, along with login date-by-4 hour fixed effects. *Further educ.* refers to any post-secondary education reported for patients aged 25 or older, while income is reported for patients over the age of 20. The subsamples split by further education and median income exclude patients with missing or undefined information. Descriptions of all variables are provided in Appendix Table A17. The baseline mean refers to the mean of the variable *Initial assignment to nurse.* The First-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

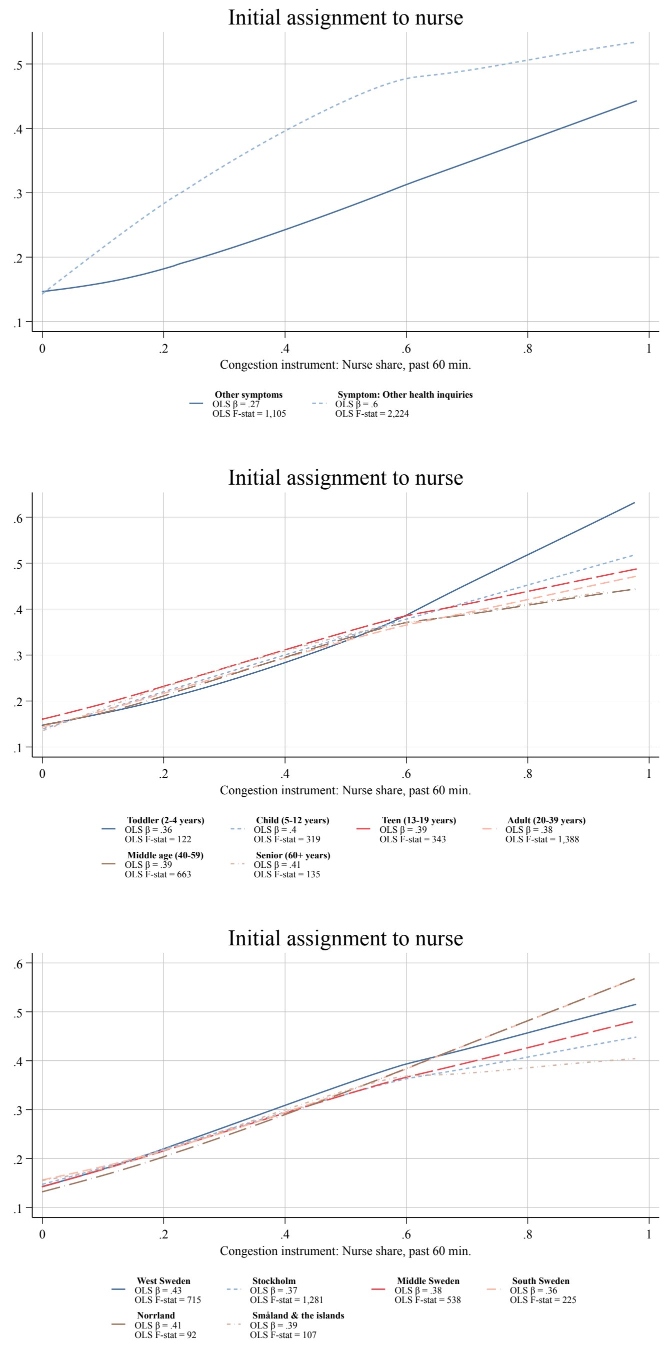


Figure A8. Instrument monotonicity test

Notes: We show non-parametric estimates within subsamples of the conditional expectation function (CEF) for the treatment, initial nurse assignment, given our congestion instrument, *Nurse share, past 60 min.*. The CEF is estimated within each subsample as a smoothed line over a binscatter. The estimates are conditional on our baseline set of controls (login time, symptom, and demographic controls), and bins are selected to minimize the integrated mean square error (Cattaneo et al. 2024). The OLS β and F-statistics are estimated with linear regressions in each subsample conditional on our baseline set of controls.

Table A5. Differences in consultation characteristics under low and high congestion

	High congestion		Low congestion		T-test	
	Mean	SD	Mean	SD	Diff.	p-val.
Patient waiting time in min.	33.1	46.0	14.1	23.9	19.0	0.00
Consultation time below one minute	0.046	0.21	0.044	0.20	0.0020	0.16
Drop out	0.013	0.11	0.037	0.19	-0.024	0.00
Clinician total time in min.	11.3	6.74	11.5	6.87	-0.23	0.00
Observations	49233		49436			

Notes: This table presents summary statistics for various consultation characteristics during low and high congestion. For low (high) congestion, we consider the lower (upper) decile of the congestion instrument, *Nurse share, past 60 min.*. The last two columns perform t-tests on the difference in means between consultation characteristics at low and high congestion.

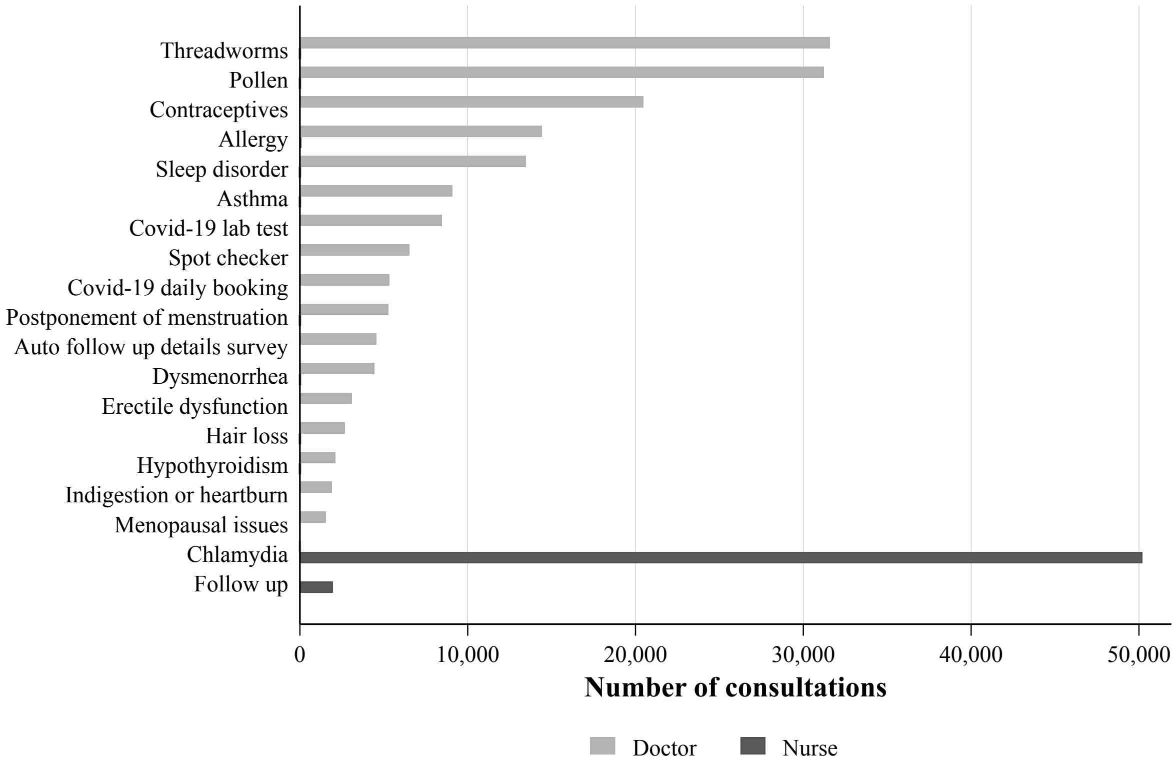


Figure A9. Strictly deterministic symptom categories in the sample with no compliers

Notes: This figure shows the number of initial consultations across strictly deterministically assigned symptom categories as reported by patients, separately for consultations exclusively assigned to a doctor or a nurse. In these symptom categories, fewer than 1% of cases are assigned to either clinician type and congestion should not influence the case assignment.

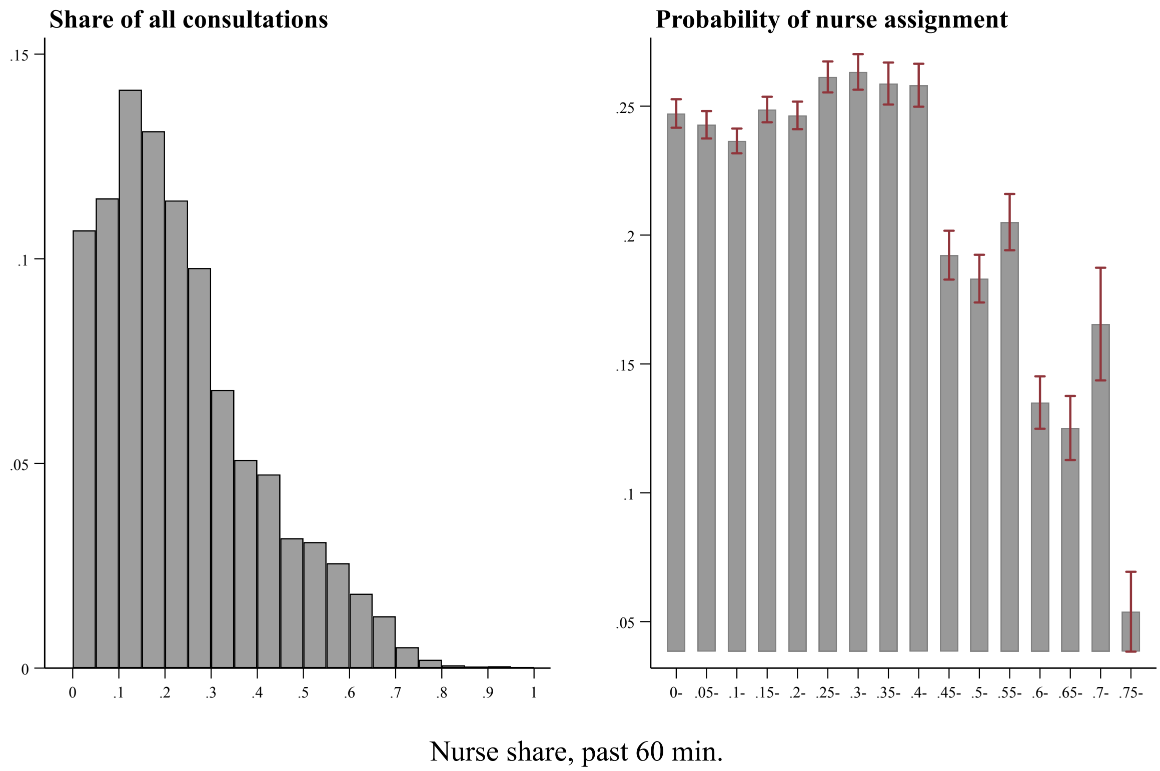


Figure A10. Variation in the congestion instrument in a sample with strictly deterministic clinician assignment

Notes: This figure shows descriptive figures for the congestion instrument, *Nurse share, past 60 min.*, as constructed in our analysis sample based on Equation 2. However, the descriptive figures are shown for a sample of cases where symptoms are strictly deterministically assigned to either a nurse or a doctor, with fewer than 1% of cases assigned to either clinician type. In these cases, congestion should not influence case assignment. The left subfigure shows the distribution of the instrument. The right subfigure shows the average probability of an initial assignment to a nurse, our treatment, plotted against categorized values of the instrument. Lines indicate 95% confidence intervals of the mean.

Table A6. Instrumental variable regressions in a sample with strictly deterministic clinician assignment

	Any referral to doctor (7d)			Any external PCP consultation (7d)			Rating: top score (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initial assignment to nurse	0.47***		-2.79	-0.15***		83.0	-0.024		4.69
	(0.048)		(2.20)	(0.053)		(245.8)	(0.047)		(8.23)
Nurse share, past 60 min.		-0.0045			0.031			0.0075	
		(0.0029)			(0.021)			(0.013)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	218,293	218,293	218,293	19,591	19,591	19,591	218,293	218,293	218,293
First-stage K-P F-statistic			7			.12			7
Baseline mean	0.012		0.012	0.059		0.059	0.48		0.48
	Informative diagnosis (7d)			**Any new prescription (7d)**			**Any ED (7d)**		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initial assignment to nurse	-0.059		-11.5	-0.15***		0.47	-0.0075		-1.09
	(0.037)		(7.27)	(0.046)		(7.39)	(0.0097)		(1.80)
Nurse share, past 60 min.		-0.018*			0.00075			-0.0017	
		(0.0094)			(0.012)			(0.0028)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	218,293	218,293	218,293	218,293	218,293	218,293	218,293	218,293	218,293
First-stage K-P F-statistic			7			7			7
Baseline mean	0.81		0.81	0.54		0.54	0.011		0.011
	Any hospitalization (7d)			**Income decrease >20 pct. (cal. month after)**			**Log total costs**		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initial assignment to nurse	-0.0034***		0.22	0.068		-23.4	-0.16***		5.16
	(0.00094)		(0.90)	(0.18)		(25.7)	(0.057)		(8.54)
Nurse share, past 60 min.		0.00035			-0.021			0.0083	
		(0.0014)			(0.016)			(0.013)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	218,293	218,293	218,293	92,283	92,283	92,283	218,293	218,293	218,293
First-stage K-P F-statistic			7			1.8			7
Baseline mean	0.0031		0.0031	0.19		0.19	6.63		6.63

Note: This table presents regression results for a sample of cases where symptoms are strictly deterministically assigned to either a nurse or a doctor, with fewer than 1% of cases assigned to either clinician type. In these cases, congestion should not influence case assignment. Each subpanel corresponds to a different outcome as the dependent variable. The left column in each subpanel reports the Ordinary Least Squares (OLS) estimate for the treatment variable (*Nurse*) based on Equation 1. The middle column presents the reduced-form regression estimates for the congestion instrument, measured as the share of nurse consultations in the past 60 minutes (*Nurse share, past 60 min.*). The right column provides the Instrumental Variables (IV) specification, using *Nurse share, past 60 min.* as an instrument for *Nurse*, estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2, but applied the sample of strictly deterministically assigned cases. All regressions control for symptom categories, demographics, health risk factors, and socioeconomic variables, along with along with login date-by-4-hour shifts fixed effects. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

C Robustness of the main results

Table A7. Effect of an initial nurse assignment on the organization of tasks at the provider (within +1 day)

	Any referral to doctor (1d)			Any subseq. consultation (1d)			Any return drop-in (1d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	0.32***		0.28***	0.32***		0.26***	0.010***		-0.0046
	(0.0015)		(0.011)	(0.0016)		(0.014)	(0.00069)		(0.0081)
Nurse share, past 60 min.		0.11***			0.100***			-0.0018	
		(0.0047)			(0.0059)			(0.0031)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			3,034			3,034			3,034
Baseline mean	0.0059	0.0059	0.0059	0.053	0.053	0.053	0.028	0.028	0.028
	Clinician total time in min. (1d)			**Doctor total time in min. (1d)**			**Nurse total time in min. (1d)**		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	3.28***		1.67***	-7.83***		-8.75***	11.1***		10.4***
	(0.039)		(0.40)	(0.031)		(0.35)	(0.024)		(0.18)
Nurse share, past 60 min.		0.64***			-3.36***			4.00***	
		(0.16)			(0.15)			(0.10)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			3,034			3,034			3,034
Baseline mean	12.5	12.5	12.5	12.4	12.4	12.4	0.065	0.065	0.065

Note: This table presents regression results for our main specifications on the effects of initial case assignment to a nurse (the *knowledge hierarchy*) rather than directly to a doctor. Each subpanel corresponds to a different outcome as the dependent variable. Patient-provider interactions after the initial consultation are defined as follows: "Any referral to doctor" refers to any internal referrals or revisits with a doctor, "Any subsequent consultations" covers any interaction with the provider after an initial visit, and "Any return drop-in" indicates any unscheduled revisit for the same symptom. "Total time" defines the time that clinicians spend on a case, including consultation time spent with the patient and time on administrative work during the initial and all subsequent consultations. The left column in each subpanel reports the Ordinary Least Squares (OLS) estimate for the treatment variable (*Nurse*) based on Equation 1. The middle column presents the reduced-form regression estimates for the congestion instrument, measured as the share of nurse consultations in the past 60 minutes (*Nurse share, past 60 min.*). The right column provides the Instrumental Variables specification, using *Nurse share, past 60 min.* as an instrument for *Nurse*, estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2. All regressions control for symptom categories, demographics, region, health risk factors, and socioeconomic variables, along with login date-by-4-hour fixed effects. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A8. Effect of an initial nurse assignment on additional outcomes

	Doctor consultation time in min. (1d)			Nurse consultation time in min. (1d)			Any external PCP contact (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	-3.08***		-3.22***	4.45***		4.54***	0.025***		-0.024
	(0.014)		(0.17)	(0.010)		(0.081)	(0.0057)		(0.073)
Nurse share, past 60 min.		-1.24***			1.75***			-0.0084	
		(0.069)			(0.044)			(0.026)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	42,814	42,814	42,814
First-stage K-P F-statistic			3,034			3,034			199
Baseline mean	4.70	4.70	4.70	0.026	0.026	0.026	0.23	0.23	0.23

	Rating: below top score (7d)			Rating: no physical replacement (7d)			Any prescription (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	0.00098		0.020	0.0080***		-0.0048	-0.13***		-0.054**
	(0.0011)		(0.013)	(0.00081)		(0.0092)	(0.0017)		(0.021)
Nurse share, past 60 min.		0.0076			-0.0018			-0.021**	
		(0.0051)			(0.0035)			(0.0082)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			3,034			3,034			3,034
Baseline mean	0.096	0.096	0.096	0.039	0.039	0.039	0.52	0.52	0.52

	Sick pay 1m (cal. month after)			Log cost, incl. ext. PCP			Log costs of provider visits		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	-0.0043***		-0.00021	-0.064***		-0.20	-0.11***		-0.13***
	(0.00066)		(0.0085)	(0.011)		(0.13)	(0.0016)		(0.014)
Nurse share, past 60 min.		-0.000076			-0.070			-0.048***	
		(0.0030)			(0.045)			(0.0054)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	213,425	213,425	213,425	42,814	42,814	42,814	490,505	490,505	490,505
First-stage K-P F-statistic			1,143			199			3,034
Baseline mean	0.011	0.011	0.011	6.84	6.84	6.84	6.37	6.37	6.37

Note: This table presents regression results for our main specifications on the effects of initial case assignment to a nurse (the *knowledge hierarchy*) rather than directly to a doctor. Each subpanel corresponds to a different outcome as the dependent variable. Outcomes are defined as follows: "Consultation time" for the time clinicians spend directly with the patient during the initial and all subsequent consultatinos; "Admin time" for time spend on administrative work without patient contact; "Rating" implies that any suboptimal rating (below top score or survey response that consultation could not replace physical care) for a consultation was given within seven days; "External PCP contact" refers to any contact (personal, mail, telephone contact, online or in-person consultation) with a doctor or a nurse outside of the provider and is defined for patients registered in Scania; "Sick pay 1m (post-1m)" refers to the receipt of sickness benefits in a given calendar year, as well as a drop in income in the calendar month following the consultation such that a full month of sick pay is captured: Income drops to below 37% of the average income in the three months prior, which corresponds to an employer-paid replacement rate of 80% for 14/30 days and no income from employment for 16/30 days, followed by a rebound to more than 37% of the pre-consultation income. "Log costs, incl. ext. PCP" includes the costs of primary care consultations at the provider, prescriptions, specialist visits, emergency department visits, urgent care visits, hospitalizations, and external physical primary care within seven days of the initial consultation, observed for a subsample of patients registered in Scania, for whom information on external primary care is available. "Log costs of provider visits" includes costs associated with online consultations at the provider we study. The left column in each subpanel reports the Ordinary Least Squares (OLS) estimate for the treatment variable (*Nurse*) based on Equation 1. The middle column presents the reduced-form regression estimates for the congestion instrument, measured as the share of nurse consultations in the past 60 minutes (*Nurse share, past 60 min.*). The right column provides the Instrumental Variables specification, using *Nurse share, past 60 min.* as an instrument for *Nurse*, estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2. All regressions control for symptom categories, demographics, region, health risk factors, and socioeconomic variables, along with login date-by-4-hour fixed effects. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A9. Alternative set of controls

	Any referral to doctor (7d)		Any ext. PCP cons. (7d)		Any new prescription (7d)		Any ED (7d)		Any hospitalization (7d)		Informative diagnosis (7d)		Rating: top score (7d)		Income drop >20% (1 mo. after)		Log total costs	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Panel A: Login time																		
Initially to nurse	0.30***	0.27***	0.023***	-0.0075	-0.18***	-0.064***	0.016***	-0.024***	0.0025***	-0.00065	-0.14***	-0.042**	0.028***	0.047**	0.0067***	0.026	-0.13***	-0.21***
	(0.0014)	(0.012)	(0.0046)	(0.063)	(0.0016)	(0.022)	(0.00071)	(0.0088)	(0.00037)	(0.0048)	(0.0016)	(0.019)	(0.0018)	(0.023)	(0.0023)	(0.032)	(0.0027)	(0.030)
Panel B: Login time + Symptoms																		
Initially to nurse	0.31***	0.27***	0.023***	-0.0061	-0.11***	-0.066***	0.0060***	-0.020**	-0.00032	0.00088	-0.065***	-0.059***	0.044***	0.045**	-0.020***	0.027	-0.12***	-0.20***
	(0.0015)	(0.011)	(0.0048)	(0.061)	(0.0017)	(0.021)	(0.00075)	(0.0086)	(0.00039)	(0.0047)	(0.0016)	(0.018)	(0.0019)	(0.022)	(0.0024)	(0.031)	(0.0027)	(0.029)
Panel C (Baseline controls): Login time + Symptoms + Demographics																		
Initially to nurse	0.31***	0.27***	0.024***	-0.0027	-0.10***	-0.059***	0.0060***	-0.017**	-0.00029	0.0012	-0.065***	-0.060***	0.045***	0.044**	-0.020***	0.028	-0.12***	-0.19***
	(0.0015)	(0.011)	(0.0048)	(0.061)	(0.0017)	(0.021)	(0.00074)	(0.0086)	(0.00039)	(0.0047)	(0.0016)	(0.018)	(0.0019)	(0.022)	(0.0024)	(0.031)	(0.0027)	(0.029)
Panel D: Login time + Symptoms + Demographics + Health risk																		
Initially to nurse	0.31***	0.27***	0.024***	0.00097	-0.10***	-0.059***	0.0061***	-0.016*	-0.00023	0.0014	-0.065***	-0.060***	0.045***	0.043**	-0.020***	0.027	-0.12***	-0.19***
	(0.0015)	(0.011)	(0.0048)	(0.061)	(0.0017)	(0.021)	(0.00074)	(0.0086)	(0.00039)	(0.0047)	(0.0016)	(0.018)	(0.0019)	(0.022)	(0.0024)	(0.031)	(0.0027)	(0.029)
Panel E (Full controls): Login time + Symptoms + Demographics + Health risk + Socio-economic variables																		
Initially to nurse	0.31***	0.27***	0.025***	-0.00072	-0.10***	-0.060***	0.0061***	-0.016*	-0.00023	0.0015	-0.064***	-0.061***	0.045***	0.042*	-0.021***	0.026	-0.12***	-0.19***
	(0.0015)	(0.011)	(0.0048)	(0.061)	(0.0017)	(0.021)	(0.00074)	(0.0085)	(0.00039)	(0.0047)	(0.0016)	(0.018)	(0.0018)	(0.022)	(0.0024)	(0.031)	(0.0027)	(0.029)
Observations	490,505	490,505	42,814	42,814	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	213,425	213,425	490,505	490,505
Baseline mean	0.01	0.01	0.15	0.15	0.44	0.44	0.03	0.03	0.01	0.01	0.81	0.81	0.42	0.42	0.22	0.22	6.67	6.67

Notes: This table presents regression results for different sets of control variables. We show the effects of initial case assignment to a nurse rather than directly to a doctor, with each column corresponding to a different outcome as the dependent variable. Each table panel then increases the number of control variables included in the specification, successively including symptom categories, demographics, health risk factors, and socioeconomic variables. All regressions control for login time based on date-by-4 hour fixed effects. Each cell shows estimates for *Nurse* from the Ordinary Least Squares (OLS) or the Instrumental Variables (IV) specification using *Nurse share, past 60 min.* as an instrument and estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A10. Alternative time fixed effects

	Any referral to doctor (7d)		Any ext. PCP cons. (7d)		Rating: top score (7d)		Informative diagnosis (7d)		Any new prescription (7d)		Any ED (7d)		Any hospitalization (7d)		Income drop >20% (1 mo. after)		Log total costs	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Panel A: Year, month, and weekday fixed effects																		
Initially to nurse	0.31***	0.29***	0.026***	0.032*	0.046***	0.042***	-0.065***	-0.077***	-0.10***	-0.12***	0.0053***	-0.00075	-0.00033	-0.00030	-0.015***	0.064***	-0.12***	-0.15***
	(0.0014)	(0.0036)	(0.0045)	(0.018)	(0.0018)	(0.0065)	(0.0015)	(0.0053)	(0.0016)	(0.0061)	(0.00072)	(0.0025)	(0.00038)	(0.0014)	(0.0023)	(0.0089)	(0.0027)	(0.0089)
Panel B: Year, month, weekday, and 4-hour fixed effects																		
Initially to nurse	0.31***	0.30***	0.025***	-0.0034	0.047***	0.056***	-0.063***	-0.055***	-0.10***	-0.086***	0.0053***	-0.0068***	-0.00033	-0.00076	-0.015***	0.069***	-0.12***	-0.15***
	(0.0014)	(0.0038)	(0.0045)	(0.019)	(0.0018)	(0.0070)	(0.0015)	(0.0056)	(0.0016)	(0.0065)	(0.00072)	(0.0026)	(0.00038)	(0.0014)	(0.0023)	(0.0094)	(0.0027)	(0.0094)
Panel C: Date fixed effects																		
Initially to nurse	0.31***	0.27***	0.023***	-0.0074	0.046***	0.055***	-0.063***	-0.041***	-0.10***	-0.059***	0.0057***	-0.0069*	-0.00026	0.00050	-0.021***	-0.0035	-0.12***	-0.16***
	(0.0015)	(0.0053)	(0.0046)	(0.028)	(0.0018)	(0.010)	(0.0016)	(0.0082)	(0.0016)	(0.0095)	(0.00073)	(0.0039)	(0.00038)	(0.0021)	(0.0024)	(0.014)	(0.0028)	(0.014)
Panel E (Baseline): Date-by-4 hour fixed effects																		
Initially to nurse	0.31***	0.27***	0.025***	-0.00072	0.045***	0.042*	-0.064***	-0.061***	-0.10***	-0.060***	0.0061***	-0.016*	-0.00023	0.0015	-0.021***	0.026	-0.12***	-0.20***
	(0.0015)	(0.011)	(0.0048)	(0.061)	(0.0018)	(0.022)	(0.0016)	(0.018)	(0.0017)	(0.021)	(0.00074)	(0.0085)	(0.00039)	(0.0047)	(0.0024)	(0.031)	(0.0028)	(0.030)
Observations	490,505	490,505	42,814	42,814	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	213,425	213,425	490,505	490,505
Baseline mean	0.01	0.01	0.15	0.15	0.42	0.42	0.81	0.81	0.44	0.44	0.03	0.03	0.01	0.01	0.22	0.22	6.68	6.68

Notes: This table presents regression results for different fixed effects to control for login time. We show the effects of initial case assignment to a nurse rather than directly to a doctor, with each column corresponding to a different outcome as the dependent variable. Each table panel then progressively narrows the time cells, with our baseline specification presented by Panel D which includes date-by-4 hour fixed effects. All regressions control for the full set of case characteristics beside the time fixed effects. Each cell shows estimates for *Nurse* from the Ordinary Least Squares (OLS) or the Instrumental Variables (IV) specification using *Nurse share, past 60 min.* as an instrument and estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A11. Alternative standard error specifications

	Any referral to doctor (7d)		Any ext. PCP cons. (7d)		Any new prescription (7d)		Any ED (7d)		Any hospitalization (7d)		Informative diagnosis (7d)		Rating: top score (7d)		Income drop >20% (1 mo. after)		Log total costs	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Panel A: No adjustments																		
Initially to nurse	0.31***	0.27***	0.025***	-0.00072	-0.10***	-0.060***	0.0061***	-0.016**	-0.00023	0.0015	-0.064***	-0.061***	0.045***	0.042*	-0.021***	0.026	-0.12***	-0.19***
	(0.00092)	(0.011)	(0.0047)	(0.062)	(0.0017)	(0.021)	(0.00067)	(0.0081)	(0.00036)	(0.0043)	(0.0015)	(0.018)	(0.0018)	(0.022)	(0.0024)	(0.031)	(0.0023)	(0.028)
(Baseline) Panel B: Robust standard errors																		
Initially to nurse	0.31***	0.27***	0.025***	-0.00072	-0.10***	-0.060***	0.0061***	-0.016*	-0.00023	0.0015	-0.064***	-0.061***	0.045***	0.042*	-0.021***	0.026	-0.12***	-0.19***
	(0.0015)	(0.011)	(0.0048)	(0.061)	(0.0017)	(0.021)	(0.00074)	(0.0085)	(0.00039)	(0.0047)	(0.0016)	(0.018)	(0.0018)	(0.022)	(0.0024)	(0.031)	(0.0027)	(0.029)
Panel C: Clustering by login date																		
Initially to nurse	0.31***	0.27***	0.025***	-0.00072	-0.10***	-0.060***	0.0061***	-0.016*	-0.00023	0.0015	-0.064***	-0.061***	0.045***	0.042*	-0.021***	0.026	-0.12***	-0.19***
	(0.0032)	(0.013)	(0.0050)	(0.059)	(0.0023)	(0.021)	(0.00082)	(0.0087)	(0.00039)	(0.0048)	(0.0028)	(0.020)	(0.0021)	(0.023)	(0.0026)	(0.032)	(0.0038)	(0.030)
Panel D: Clustering by login date-by-4 hour windows																		
Initially to nurse	0.31***	0.27***	0.025***	-0.00072	-0.10***	-0.060***	0.0061***	-0.016*	-0.00023	0.0015	-0.064***	-0.061***	0.045***	0.042*	-0.021***	0.026	-0.12***	-0.19***
	(0.0023)	(0.013)	(0.0048)	(0.061)	(0.0019)	(0.021)	(0.00076)	(0.0086)	(0.00039)	(0.0048)	(0.0023)	(0.019)	(0.0020)	(0.022)	(0.0025)	(0.031)	(0.0031)	(0.029)
Full controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	490,505	490,505	42,814	42,814	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	213,425	213,425	490,505	490,505
Baseline mean	0.01	0.01	0.15	0.15	0.44	0.44	0.03	0.03	0.01	0.01	0.81	0.81	0.42	0.42	0.22	0.22	6.67	6.67

Notes: This table presents regression results for alternative standard errors estimators. We show the effects of initial case assignment to a nurse rather than directly to a doctor, with each column corresponding to a different outcome as the dependent variable. Each table panel uses a different specification to estimate standard errors. Each cell shows estimates for *Nurse* from the Ordinary Least Squares (OLS) or the Instrumental Variables (IV) specification using *Nurse share, past 60 min.* as an instrument and estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2. All regressions control for symptom categories, demographics, health risk factors, and socioeconomic variables, along with login time fixed effects based on date and 4-hour shifts. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

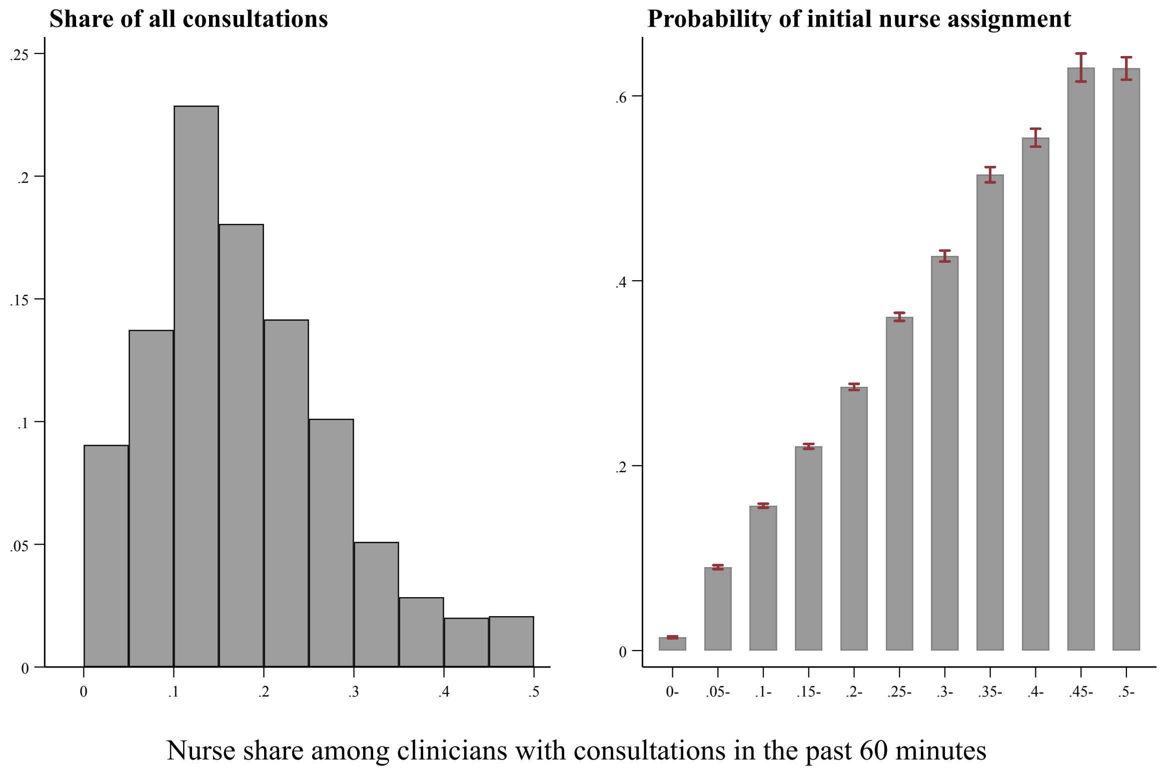


Figure A11. Variation in the staffing instrument

Notes: This figure shows descriptive figures for an alternative instrument based on pre-determined staffing. This staffing instrument is constructed as the share of nurses among active clinicians in the past 60 minutes of a given consultation, where we assume that no more nurses are staffed compared to doctors and exclude an own case's clinician. The left subfigure shows the distribution of the instrument. The right subfigure shows the average probability of an initial assignment to a nurse, our treatment, in categorized values of the instrument. Lines indicate 95% confidence intervals of the mean.

Table A12. Main outcomes: Alternative staffing instrument

	Any referral to doctor (7d)			Any external PCP consultation, doctor (7d)			Any external PCP consultation, nurse (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	0.31***		0.30***	0.022***		0.0086	0.0037		0.017
	(0.0015)		(0.0060)	(0.0045)		(0.027)	(0.0027)		(0.016)
Nurse staff share, past 60 min.		0.31***			0.0092			0.019	
		(0.0070)			(0.029)			(0.018)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	42,814	42,814	42,814	42,814	42,814	42,814
First-stage K-P F-statistic			11,964			1,006			1,006
Baseline mean	0.014	0.014	0.014	0.12	0.12	0.12	0.043	0.043	0.043

	Rating: top score (7d)			Rating: physical replacement (7d)			Informative diagnosis (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	0.045***		0.046***	0.015***		0.031***	-0.064***		-0.070***
	(0.0018)		(0.012)	(0.0018)		(0.012)	(0.0016)		(0.0095)
Nurse staff share, past 60 min.		0.049***			0.033***			-0.073***	
		(0.012)			(0.012)			(0.0099)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			11,964			11,964			11,964
Baseline mean	0.42	0.42	0.42	0.45	0.45	0.45	0.81	0.81	0.81

	Any new prescription (7d)			Any specialist (7d)			Any ED (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	-0.10***		-0.083***	0.0030***		0.00029	0.0061***		-0.0045
	(0.0017)		(0.011)	(0.00073)		(0.0046)	(0.00074)		(0.0046)
Nurse staff share, past 60 min.		-0.087***			0.00031			-0.0047	
		(0.012)			(0.0049)			(0.0048)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			11,964			11,964			11,964
Baseline mean	0.44	0.44	0.44	0.036	0.036	0.036	0.030	0.030	0.030

	Any hospitalization (7d)			Any urgent care (7d)			Income decrease >20 pct. (cal. month after)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	-0.00023		-0.0010	0.0019***		0.0011	-0.021***		-0.0066
	(0.00039)		(0.0025)	(0.00051)		(0.0032)	(0.0024)		(0.016)
Nurse staff share, past 60 min.		-0.0011			0.0011			-0.0067	
		(0.0026)			(0.0034)			(0.016)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	213,425	213,425	213,425
First-stage K-P F-statistic			11,964			11,964			5,024
Baseline mean	0.0087	0.0087	0.0087	0.017	0.017	0.017	0.22	0.22	0.22

	Zero income (cal. month after)			Death excl. external causes (3y)			Log costs, excl. ext. PCP		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	-0.0034***		-0.014*	-0.000096		-0.0010	-0.12***		-0.13***
	(0.0013)		(0.0085)	(0.00016)		(0.00095)	(0.0028)		(0.016)
Nurse staff share, past 60 min.		-0.014*			-0.0011			-0.14***	
		(0.0086)			(0.00100)			(0.017)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	213,425	213,425	213,425	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			5,024			11,964			11,964
Baseline mean	0.050	0.050	0.050	0.0016	0.0016	0.0016	6.68	6.68	6.68

Note: This table presents results for the effects of initial case assignment to a nurse (the *knowledge hierarchy*) rather than directly to a doctor with an alternative instrumental variable: The share of nurses among active clinicians in the past 60 minutes (*Nurse staff share, past 60 min.*). Each subpanel corresponds to a different outcome as the dependent variable. Outcomes are defined as in Tables 2, 3, and 5. The left column in each subpanel reports the Ordinary Least Squares (OLS) estimate for the treatment variable (*Nurse*) based on Equation 1. The middle column presents the reduced-form regression estimates for the alternative instrument. The right column provides the Instrumental Variables specification, using *Nurse staff share, past 60 min.* as an instrument for *Nurse*, estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2, replaced by the alternative instrument. All regressions control for symptom categories, demographics, region, health risk factors, and socioeconomic variables, along with login date-by-4-hour fixed effects. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

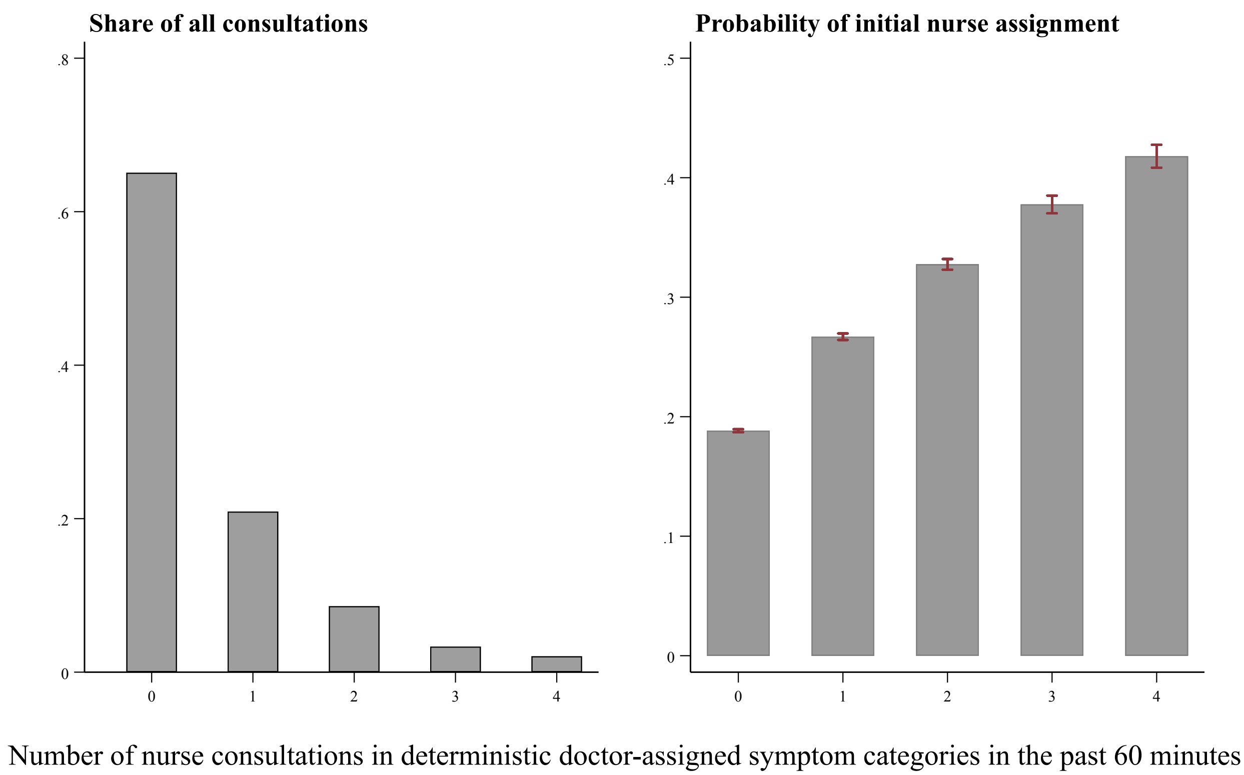


Figure A12. Variation in the doctor shortage instrument

Notes: This figure shows descriptive figures for an alternative instrument based on doctor shortages relative to patient demand. This doctor shortage instrument is constructed in two steps. First, it considers the number of cases in the past 60 minutes of a given consultation assigned to nurses, even though their symptoms are otherwise deterministically assigned to doctors (share of direct-to-doctor assignment above 95%). Then, this number is winsorized at the 99th percentile. The left subfigure shows the distribution of the instrument. The right subfigure shows the average probability of an initial assignment to a nurse, our treatment, in categorized values of the instrument. Lines indicate 95% confidence intervals of the mean.

Table A13. Main outcomes: Alternative doctor shortage instrument

	Any referral to doctor (7d)			Any external PCP consultation, doctor (7d)			Any external PCP consultation, nurse (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	0.31***		0.25***	0.022***		-0.10	0.0037		0.016
	(0.0015)		(0.040)	(0.0045)		(0.16)	(0.0027)		(0.095)
Doctor shortage, past 60 min.		0.0038***			-0.0016			0.00025	
		(0.00068)			(0.0025)			(0.0015)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	42,814	42,814	42,814	42,814	42,814	42,814
First-stage K-P F-statistic			294			20			20
Baseline mean	0.014	0.014	0.014	0.12	0.12	0.12	0.043	0.043	0.043
	Rating: top score (7d)			Rating: physical replacement (7d)			Informative diagnosis (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	0.045***		0.058	0.015***		0.035	-0.064***		-0.050
	(0.0018)		(0.069)	(0.0018)		(0.069)	(0.0016)		(0.056)
Doctor shortage, past 60 min.		0.00089			0.00053			-0.00077	
		(0.0011)			(0.0011)			(0.00087)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			294			294			294
Baseline mean	0.42	0.42	0.42	0.45	0.45	0.45	0.81	0.81	0.81
	Any new prescription (7d)			Any specialist (7d)			Any ED (7d)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	-0.10***		-0.18***	0.0030***		-0.021	0.0061***		-0.018
	(0.0017)		(0.065)	(0.00073)		(0.026)	(0.00074)		(0.026)
Doctor shortage, past 60 min.		-0.0027***			-0.00033			-0.00028	
		(0.00100)			(0.00040)			(0.00040)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			294			294			294
Baseline mean	0.44	0.44	0.44	0.036	0.036	0.036	0.030	0.030	0.030
	Any hospitalization (7d)			Any urgent care (7d)			Income decrease >20 pct. (cal. month after)		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	-0.00023		0.0052	0.0019***		-0.013	-0.021***		-0.13
	(0.00039)		(0.014)	(0.00051)		(0.018)	(0.0024)		(0.097)
Doctor shortage, past 60 min.		0.000079			-0.00021			-0.0018	
		(0.00022)			(0.00027)			(0.0014)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	490,505	490,505	490,505	490,505	490,505	490,505	213,425	213,425	213,425
First-stage K-P F-statistic			294			294			113
Baseline mean	0.0087	0.0087	0.0087	0.017	0.017	0.017	0.22	0.22	0.22
	Zero income (cal. month after)			Death excl. external causes (3y)			Log costs, excl. ext. PCP		
	OLS	Red.	IV	OLS	Red.	IV	OLS	Red.	IV
Initially to nurse	-0.0034***		0.0026	-0.000096		-0.0011	-0.12***		-0.17*
	(0.0013)		(0.048)	(0.00016)		(0.0062)	(0.0028)		(0.093)
Doctor shortage, past 60 min.		0.000036			-0.000017			-0.0026*	
		(0.00068)			(0.000096)			(0.0014)	
Baseline characteristics	✓	✓	✓	✓	✓	✓	✓	✓	✓
Additional controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	213,425	213,425	213,425	490,505	490,505	490,505	490,505	490,505	490,505
First-stage K-P F-statistic			113			294			294
Baseline mean	0.050	0.050	0.050	0.0016	0.0016	0.0016	6.68	6.68	6.68

Note: This table presents results for the effects of initial case assignment to a nurse (the *knowledge hierarchy*) rather than directly to a doctor with an alternative instrumental variable: The number of nurse consultations in almost deterministically doctor-assigned symptom categories in the past 60 minutes, winsorized at the 99 percentile (*Doctor shortage, past 60 min.*). Each subpanel corresponds to a different outcome as the dependent variable. Outcomes are defined as in Tables 2, 3, and 5. The left column in each subpanel reports the Ordinary Least Squares (OLS) estimate for the treatment variable (*Nurse*) based on Equation 1. The middle column presents the reduced-form regression estimates for the alternative instrument. The right column provides the Instrumental Variables specification, using *Doctor shortage, past 60 min.* as an instrument for *Nurse*, estimated via Two-Stage Least Squares (2SLS), following Equations 1 and 2, replaced by the alternative instrument. All regressions control for symptom categories, demographics, region, health risk factors, and socioeconomic variables, along with login date-by-4-hour fixed effects. The baseline mean represents the average of the dependent variable for cases directly assigned to a doctor. The first-stage K-P F-statistic refers to the Kleibergen-Paap F-statistic. Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A14. Mean characteristics of compliers

	Sample mean	Complier mean
Symptom categories		
Abdominal pain	0.030	0.046
Cold and flu	0.089	0.032
Cold sores	0.025	0.011
Constipation	0.0071	0.0084
Covid-19	0.058	0.13
Diarrhea or vomiting	0.021	0.015
Eye infection	0.075	-0.0039
Fever	0.029	0.042
Headache	0.023	0.037
Nail problem	0.022	0.027
Other health inquiries	0.35	0.57
Bites and stings	0.054	0.051
Sinusitis	0.032	0.019
Sore throat	0.076	-0.0025
Uncategorized	0.030	0.018
Urinary tract infection	0.079	-0.00059
Demographics		
Female	0.63	0.56
Patient age	29.2	28.6
West Sweden	0.20	0.22
Stockholm	0.44	0.44
Middle Sweden	0.19	0.18
South Sweden	0.095	0.090
Norrland	0.034	0.031
Småland + the islands	0.042	0.038
Health risk		
Any prior hospitalization	0.19	0.18
Any prior ED	0.33	0.34
Any prior urgent care	0.24	0.24
Any prior specialist	0.64	0.64
Any comorbidity	0.21	0.22
Socio-economic variables		
Income above median	0.32	0.28
Any benefits	0.11	0.11
Schooling < 9 years	0.049	0.052
Middle/High school	0.24	0.22
Further educ. < 3 years	0.11	0.10
Further educ. <= 3 years	0.19	0.17
Education n/a	0.42	0.45
Employed	0.57	0.51
Self-employed	0.042	0.038
Unemployed	0.050	0.058
Employment status n/a	0.34	0.40
Married	0.23	0.20
Unmarried	0.075	0.095
Divorced/Widowed	0.42	0.38
Civil status n/a	0.27	0.32
Not migrated	0.75	0.70
Immigrant 1st gen	0.16	0.19
Immigrant 2nd gen	0.096	0.11

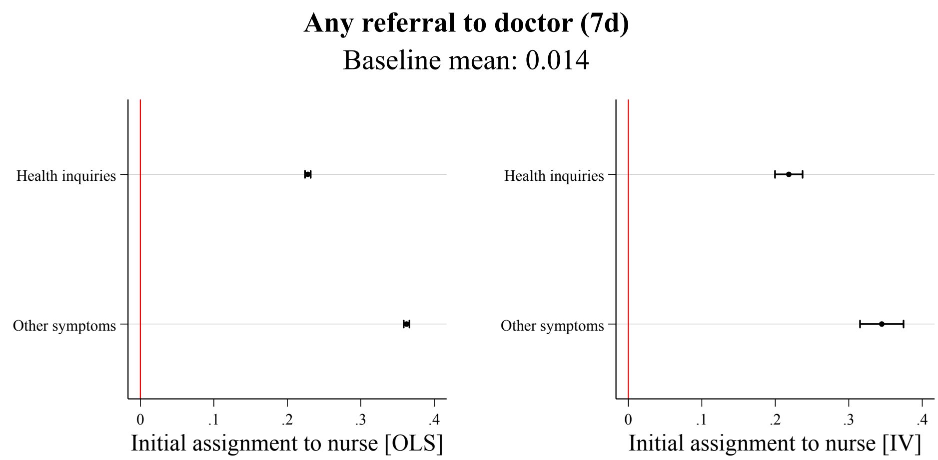
Note: This table reports average characteristics for the overall analysis sample and the population of compliers. The first column reprints the overall sample mean. The second column shows the estimated average of a case characteristic x_i in the complier population. We follow the procedure described in Frandsen et al. (2023) based on Abadie (2003) and report the mean characteristic among compliers as the estimate of $\frac{E[\omega_i X_i]}{E[\omega_i]}$, where ω_i is the weight given to case i by the IV. We estimate complier characteristics by Two-Stage Least Squares (2SLS) regressions of the interaction between a predetermined case characteristic and treatment *Initial assignment to nurse* ($x_i \times Nurse_i$) on to the treatment ($Nurse_i$), instrumented by our congestion instrument *Nurse share, past 60 min.* and the baseline time controls (login date-by-4-hours fixed effects).

D Additional results

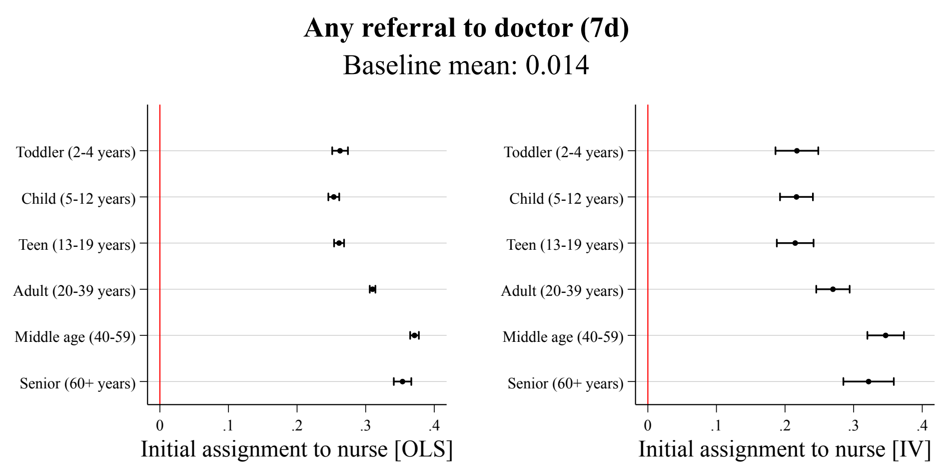
Table A15. Time spent on patients seen by a doctor

	Doctor time spent in min. (7d)		
	Total time	Consultation time	Admin time
Case referred by nurse	1.34***	0.53***	0.81***
	(0.044)	(0.026)	(0.036)
Baseline controls	✓	✓	✓
Additional controls	✓	✓	✓
Observations	414352	414352	414352
Baseline mean	11.7	4.44	7.30

Note: This table presents additional results on cases that are seen by a doctor, which include either cases directly assigned to a doctor or cases referred to a doctor by a nurse. We consider the closest consultation with a doctor within +7 days for cases that are referred from a nurse. The dependent variables measure the amount of time a doctor spends on a case, defined as: "Total time" for overall case duration, "consultation time" for time spent with the patient, and "admin time" for administrative work without patient contact. Doctor time is regressed on an indicator for whether a case was referred by a nurse, and estimated by Ordinary Least Squares (OLS). Robust standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01.



(a) Heterogeneity by main symptom category



(b) Heterogeneity by age group

Figure A13. Heterogeneity across cases in the propensity to be referred to a doctor

Note: These figures show the heterogeneous effects of an initial assignment to a nurse (the *knowledge hierarchy*) on the propensity to refer to a doctor, broken down by subgroups defined either by whether the main reported symptom is "Other health inquiries" (subfigure A13a) or by age category (subfigure A13b). The outcome "Any referral to doctor" refers to any internal referrals or revisits with a doctor. Each subfigure presents estimates from a regression of the outcome on all interactions of the initial nurse assignment (*Nurse*) with each subgroup, while controlling for subgroup fixed effects, the full set of case characteristics, and login date-by-4-hour fixed effects, following Equation 3. The estimates are obtained by Ordinary Least Squares (OLS, left) or Two-Stage Least Squares (IV, right), where interactions of *Nurse share, past 60 min.* with each subgroup serving as instruments for the interactions between *Nurse* and each subgroup. Each row represents the estimated subgroup-specific effect of an initial nurse assignment, with horizontal lines representing 95% confidence intervals based on robust standard errors.

E Data Appendix

The primary data is supplied by a large healthcare firm in Sweden. Their data contains comprehensive records of all of the online primary care consultations within the firm for the years 2019 and 2020. For each consultation, we observe, among other information, the date, duration, and purpose of the consultation, the physician type, the contact method, and primary diagnoses recorded in the form of the ICD-10 codes, the standardized system of disease classification. In addition, we observe patients' demographic characteristics, such as age and gender. For clinicians, demographic information is limited. Each consultation record includes patients' and clinicians' unique personal identifiers, which have been pseudonymized for us.

The personal identifiers enable us to link patients' consultations to additional individual-level information from Statistics Sweden (SCB). SCB has supplied us with basic information about patients' immigration background. In addition, we obtain information about patients from three key datasets from SCB: the Population Statistics Register (RTB), the Longitudinal Integrated Database for Health Insurance and Labour Market Studies (LISA), and the Education Register (UREG). RTB covers patients' civil status and municipality of residence from 2013 to 2018. LISA provides socio-economic and demographic data for patients, including income, rehabilitation benefits, employment status, paid sickness, and income-related benefits, for the years 2013 to 2020, with additional income and employment data available for 2019 to 2022. Yearly data at LISA is typically updated in November of each year. Finally, UREG offers data on patients' educational backgrounds for the years 2013 to 2020.

We also observe patients' healthcare utilization by using the unique identifiers to link their consultations to data from the National Board of Health and Welfare (Socialstyrelsen, NBHW). Data from NBHW contains prescriptions, mortality, outpatient, and inpatient healthcare data. The prescription data, covering 2013 to 2023, includes medications collected by patients at pharmacies but excludes medical drugs administered within healthcare facilities. Mortality data, available for 2019 to 2023, includes the date and cause of death. Outpatient and inpatient data, spanning 2013 to 2023, include visit and admission dates, ICD-10 diagnostic codes, and, for out-patient data, additional information on contact types.

The data from Region Scania contains all primary care visits (not only within the firm we study but also for other providers) in Scania, a region in southern Sweden where 13% of the country's population resides. These data cover the period from 2019 to 2020, and can be linked to our other data sources through the unique patient identifiers. Patient information in the Scania data includes all primary care visits with the date of the contact, clinician type, contact type, healthcare clinic, and main diagnoses in the form of the ICD-10 code.

Our cost estimates are obtained from various resources, primarily based on public reports. We present an overview of the source material in Appendix Table A16.

Finally, Appendix Table A17 presents an overview of the key case characteristics we use in our main analysis.

Table A16. Costs of healthcare services covered by Sweden's public healthcare system

Service	Cost in SEK	Year	Source	Link
Online primary care consultation				
Doctor	500	2019	Reimbursement rate for digital care services in primary care.	vardanalys.se[1] Last access: 5 Nov 2024
Nurse	275	2020	Recommendation on remuneration for digital healthcare services to healthcare providers (SKR).	skr.se[2] Last access: 5 Nov 2024
Physical primary care consultation				
Doctor	2002	2019/2020	Average cost of a physical doctor's visit, national estimates, $(1,838 + 2,166) \times 0.5 = 2,002$.	vardanalys.se[1] Last access: 5 Nov 2024
Nurse	614	2019	Average cost of a physical primary care nurse visit, estimates from Region Östergötland.	vardanalys.se[1] Last access: 5 Nov 2024
Other healthcare				
Prescription	260	2019/2020	Average cost to the region over all prescriptions in 2019 or 2020 among patients in our analysis sample.	Own calculation. Data from Socialstyrelsen.
Specialist visit	3594	2019	Unweighted average cost of a doctor's visit across various specialties in Northern Sweden (Norra Sjukvårdsregionen): medicine, pulmonology, infectious diseases, dermatology, urology, orthopedics, ophthalmology, otolaryngology, pediatrics, and gynecology. We consider specialties for acute and routine specialty services for an average one-time cost and exclude, for example, oncology and psychiatry which likely involve ongoing therapies.	norrasjukvardsregionforbundet.se[3] Last access: 5 Nov 2024
Emergency department visit	3991.5	2019/2020	Average costs in Southern Sweden, $(3,963 + 4,020) * 0.5 = 3,991.5$.	sodrasjukvardsregionen.se[4] Last access: 5 Nov 2024
Urgent care center visit	2002	2019/2020 information from 2024	Using the information provided—"Without an EU card, you must pay the full cost of the care yourself. [...] An appointment with a doctor at a vårdcentralen (healthcare centre) or a visit to the närakuten (urgent care centre) costs SEK 2,093."—we base our analysis on 2019/2020 costs, assuming that urgent care visits have the same cost as regular primary care visits.	1177.se[5] Last access: 5 Nov 2024
Hospitalization	7800	2020	Stated average cost of a hospitalization.	kristianstadsbladet.se[6] Last access: 5 Nov 2024

[1] Full source: https://www.vardanalys.se/rapporter/besok-via-natet/
[2] Full source: https://skr.se/download/18.32563d7d1784aa279ece294c/1618741364556/11-2020-WEBB-Rek-om-ersattning-for-digitala-vardtjanster.pdf
[3] Full source: https://www.norrasjukvardsregionforbundet.se/halso-och-sjukvard/avtal-och-priser/arkiv/
[4] Full source: https://sodrasjukvardsregionen.se/verksamhet/avtal-priser/regionala-priser-och-ersattningar-foregaende-ar/
[5] Full source: https://www.1177.se/en/Stockholm/other-languages/other-languages/soka-vard/hitta-ratt-vard-nara-dig—andra-sprak—stockholms-lan/
[6] Full source: https://mosaik.kristianstadsbladet.se/nyheter/varden-kostar-mycket-mer-an-du-betalar-for/

Table A17. Description of key case characteristics

Variable	Description	Data source
Treatment and instrument		
Initial assignment to a nurse; Nurse	Treatment: Indicator variable that is one if the initial consultation of a case is assigned to a nurse, and zero if the case is directly assigned to a doctor.	Provider
Nurse share, past 60 min	Congestion instrument: The share of initial nurse consultations among cases in the past 60 minutes of a given case	Provider
Login time		
Login date-by-4 hours	Fixed effects for the login time of a case when a consultation is requested. Login time is constructed as calendar date-by-4 hour window: 0 am - 4 am; 4 am - 8 am; 8 am - 12 pm; 12 pm - 4 pm; 4 pm - 8 pm; 8 pm - 12 am of a given date, such as April 1, 2020, 12 am - 4 am.	Provider
Symptoms		
Symptom category	A set of indicator variables for the main symptom category of the case, as one of 16 symptoms included in our analysis sample. The full list of symptoms is provided in A6.	Provider
Demographics		
Gender	Indicator variable that is one if the patient is female.	Provider
Age group	A set of indicator variables for the patient's age group, as one of the following categories: Infant (0-1 year); Toddler (2-4 years); Child (5-12 years); Teen (13-19 years); Adult (20-39 years); Middle age (40-59 years); Senior (60+ years).	Provider
Region	A set of indicator variables for the patient's registered region as reported in November of the year prior to the initial consultation. We consider six regional areas based on Sweden's national areas (riksområden): West Sweden, Stockholm, Middle Sweden (combining East Middle Sweden and West Middle Sweden), South Sweden, Norrland (combining Middle Norrland and Upper Norrland), and Småland and the islands.	SCB
Health risk		
Any prior hospitalization	Indicator for any inpatient hospitalization within the 3 years prior to the initial consultation, excluding visits in the 30 days immediately preceding it.	NBHW
Any prior ED	Indicator for any emergency department (akutmottagning) visit within the 3 years prior to the initial consultation, excluding visits in the 30 days immediately preceding it.	NBHW
Any prior urgent care	Indicator for any urgent care center (närakuter) visit within the 3 years prior to the initial consultation, excluding visits in the 30 days immediately preceding it.	NBHW
Any prior specialist	Indicator for any specialist visit within the 3 years prior to the initial consultation, excluding visits in the 30 days immediately preceding it.	NBHW
Any comorbidity	Indicator variable for an Elixhauser comorbidity index of 1 or higher, as diagnosed within specialty or hospital care in 2018 or earlier.	NBHW
Socio-economic variables		
Income above median	Indicator variable for the patient's annual income reported in Nov 2018 being above the median (non-missing) income in-sample: Income above 294,700 SEK (approximately 28,000 USD). Income is reported for patients $\geq$ age 20 in Nov 2018.	SCB
Income missing	Indicator variable for information on income being unavailable (either missing or unreported).	SCB
Any benefits	Indicator variable for the receipt of any social security benefits in 2018. Benefits are reported for patients $\geq$ age 20 in Nov 2018.	SCB
Benefits missing	Indicator variable for information on benefits being unavailable (either missing or unreported).	SCB
Employment status	Indicator variable for the main employment type reported in Nov 2018, as one of the following categories: employed; self-employed; unemployed. Employment status is reported for patients aged 20 to 67.	SCB
Employment status missing	Indicator variable for information on employment status being unavailable (either missing or unreported).	SCB
Education level	Indicator variable for the highest educational degree reported in Nov 2018, as one of the following categories: schooling below 9 years; middle/high school; secondary education $<$ 3 years (further educ. $<$ 3 years); secondary education $\geq$ 3 years (further educ. $\geq$ 3 years). Educational degrees are reported for patients $\geq$ age 25.	SCB
Education level missing	Indicator variable for information on education being unavailable (either missing or unreported).	SCB
Civil status	Indicator variable for the civil status reported in Nov 2018, as one of the following categories: married; divorced or widowed; unmarried. Civil status is reported for patients $>$ age 18.	SCB
Civil status missing	Indicator variable for information on civil status being unavailable (either missing or unreported).	SCB
Migration status	A set of indicator variables for the migration status of the patient, as one of the following categories: Not migrated; 1st generation immigrant (born outside of Sweden); 2nd generation immigrant (born in Sweden to two foreign-born parents). Migration information is available for the entire sample.	SCB

Notes: This table presents the variable definition and main data source for key case characteristics.