Journal of Development Economics 158 (2022) 102917

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular Article

Overtreatment and benevolent provider moral hazard: Evidence from South African doctors☆

Mylène Lagarde [a,*], Duane Blaauw [b]

[a] Department of Health Policy, London School of Economics and Political Science, UK
[b] Centre for Health Policy, University of the Witwatersrand, South Africa

ARTICLE INFO

JEL classification:
C93
D64
I11
I13
O15

ABSTRACT

Overtreatment is widespread in health, with potentially dire consequences for patients, health systems and public health. It may be fueled by providers when they do not bear the cost of treatment (moral hazard), even they do not profit financially from it (i.e. benevolent providers). We test this hypothesis by creating an exogeneous change in the incentives faced by private doctors in South Africa. We find that provider moral hazard has no effect on overtreatment in volume but fuels overtreatment in cost. By contrast, when they bear the marginal treatment cost, doctors choose cheaper drug. While these results suggest that provider moral hazard contributes to overtreatment in primary care, we consider other plausible channels, such as responses to a perceived demand for high-quality drugs or market segmentation. We discuss the potential scope for supply-side cost-sharing incentives to reduce inefficiency in future health system reforms in South Africa.

1. Introduction

Overtreatment in health care is increasingly recognized as a cause of rising health expenditures in many health systems. For example, the Institute of Medicine suggested that unnecessary care provided in the United States represented US$210 billion each year (Institute of Medicine 2013). Overtreatment, defined as care where volume or cost is higher than appropriate (Emanuel and Fuchs 2008), has received much attention in high-income countries, while concerns in low- and middle income countries (LMICs) have focused on problems of undertreatment – the failure to use effective medical interventions to meet patients' needs (Glasziou et al., 2017). Yet there is growing evidence that overtreatment is also ubiquitous in these settings, with many patients receiving care with little to no benefit for their health (Brownlee et al., 2017). While overtreatment may not seem as damaging as undertreatment, it creates multiple inefficiencies: it translates into considerable waste for health systems with scarce resources; some patients delay needed care because they anticipate high treatment costs partly induced by overtreatment; finally, overconsumption of certain drugs (e.g. antibiotics, anti-malarial treatment) generates negative externalities by fueling antimicrobial resistance, making treatments of deadly infections less effective (WHO 2014).

Theory and evidence suggest that providers play a central role in the provision of unnecessary treatment. Both theoretical models of credence good markets (Dulleck and Kerschbamer 2006) and empirical evidence from a range of healthcare systems (Gruber and Owings 1996; Iizuka 2007; Currie et al., 2014) show how overtreatment occurs when providers both recommend the treatment needed by their patients and sell iit with a profit margin. However, overtreatment is also frequent in healthcare markets when diagnosis and treatment are provided by different parties (for example when doctors write a prescription filled by independent pharmacies), although such separation is supposed to limit

☆ We benefitted from comments and suggestions from: Jishnu Das, Benjamin Daniels, Adrian Brink; from seminar participants at the LSE, and from conference participants at the 2019 Conference of the American Society of Health Economists, the 2019 World Congress on Health Economics, the 2020 AEA-ASSA Conference in San Diego and the 2020 IRDES-Dauphine Workshop on Applied Health Economics and Policy Evaluation. We thank Dr Modi and several other faculty members from the School of Medicine of the University of the Witwatersrand for their contribution to designing the SP methodology and protocols. We are grateful to the Centre for Health Policy of the University of the Witwatersrand for providing logistical support and infrastructure. We also thank Ling Ting for project management support and excellent research assistance. The project would not have been possible without the dedication of all SPs and Braimoh Bello, Rosaline Pineteh and other staff at the Centre for Statistical Analysis and Research in Johannesburg. This study was funded by the UK Economic and Social Research Council through grant ES/P004059/1. The findings, interpretations, and conclusions expressed in this article are those of the authors and do not necessarily represent the views of the ESRC.

* Corresponding author.
E-mail addresses: M.Lagarde@lse.ac.uk (M. Lagarde), Duane.Blaauw@wits.ac.za (D. Blaauw).

https://doi.org/10.1016/j.jdeveco.2022.102917
Received 11 August 2020; Received in revised form 10 May 2022; Accepted 7 June 2022
Available online 10 June 2022

overtreatment in credence good markets (Dulleck and Kerschbamer 2006). Patients' (perceived) demand and providers' inadequate knowledge are reasons often cited to explain this problem. Yet, recent audit studies in LMICs show that high rates of overtreatment occur in the absence of patient request, and often despite providers' knowledge of the appropriate treatment (Mohanan et al., 2015; Daniels et al., 2017; Kovacs et al., 2020; King et al., 2021).

In this paper, we hypothesize that overtreatment can arise from the *absence* of financial incentive to limit unnecessary or unnecessarily expensive drugs. This follows from the seminal model of supply-side cost-sharing showing that providers limit unnecessary treatment when they bear the marginal cost of treatment (Ellis and McGuire 1986, 1993). In our context, we coin this situation 'benevolent provider moral hazard' to emphasize not only that providers do not bear the financial consequences of their treatment recommendations (a classic case of provider moral hazard), but also that they are not driven by profit-making motives ("benevolent provider") – a key distinction from the hospital setting in the Ellis and McGuire model .[1] In addition, following empirical studies of credence good markets which show that overtreatment is higher when customers are insured and do not bear the residual cost of the recommended service (Lundin 2000; Lu 2014; Balafoutas et al., 2017; Balafoutas et al., 2020), we hypothesize that benevolent provider moral hazard may be heightened when patients are insured. This is because providers may care to reduce overtreatment when patients pay for their expenses but not when they are fully insured.[2]

We test these two hypotheses by conducting an experimental audit study in the private healthcare market in South Africa, where we exogenously alter the incentives faced by doctors. As part of the audit study, trained enumerators visited 113 doctors and presented themselves as patients with mild symptoms of a simple medical condition – viral bronchitis. It is an uncomplicated clinical case which requires basic medical knowledge and limited effort to arrive at a correct diagnosis and treatment recommendation. We identify overtreatment in volume (medicines with no medical benefit) by comparing the treatment received by patients to existing clinical guidelines for the case. We measure overtreatment in cost thanks to detailed drug cost data.

Our study takes place in the private market for primary care in South Africa where two types of doctors operate: prescribing and dispensing doctors. Prescribing doctors operate under a classic model of separation between diagnosis and treatment provision: they write a prescription to patients who get it filled at an independent pharmacy. They do not profit from prescribing drugs but they have no private incentive to invest effort to limit unnecessary treatment either, since they do not bear the cost of treatment – a situation of benevolent provider moral hazard. By contrast, dispensing doctors charge a fixed consultation fee for providing both diagnosis and medicines. This unusual institution creates a natural "supply-side cost-sharing" mechanism that creates an incentive for providers to limit unnecessary treatment since the cost of treatment reduces their profit margin, thereby eliminating provider moral hazard (Ellis and McGuire 1993).[3]

To study the effect of benevolent provider moral hazard on overtreatment, our audit study focuses on dispensing doctors and creates an experimental variation in which they are encouraged to behave as prescribing doctors and be free of cost-sharing. Specifically, each doctor receives two simulated patients, one of which is randomly tasked to request a prescription to obtain elsewhere the drugs they would have been dispensed. If they comply with the patient's demand and prescribe the treatment, dispensing doctors no longer bear the financial consequences of their recommendation and have no private incentive to limit overtreatment, in volume or price. In addition, to study the influence of patients' insurance status on providers' treatment decisions, half of doctors were randomized to receive pairs of insured patients, while the other half received patients who were uninsured.

We report four main findings.

First, we find hardly any evidence supporting the hypothesis that benevolent provider moral hazard encourages excess drug supply. When doctors are encouraged to prescribe rather than dispense, and thereby no longer bear the financial cost of treatment, overtreatment in volume does not increase. However, this null effect may be partly due to the near-universal recommendation of unnecessary drugs even as providers bear the financial consequence of treatment. Breaking down unnecessary drugs between those that have potentially harmful consequences for this case (antibiotics) and those that do not (unnecessary 'only'), we find that the probability of recommending 'only' unnecessary drugs increases by 9 percentage points when supply-side cost sharing is lifted but there is no change in the probability of recommending antibiotics. We also find suggestive evidence that the increase in unnecessary 'only' drugs is mostly concentrated in insured patients. Overall, in a context where we observe high rates of overtreatment in the presence of cost-sharing incentives for doctors, these results suggest that provider moral hazard may have a limited role in over-prescription of unnecessary medicines.

By contrast, and this is our second result, we find compelling evidence that benevolent provider moral hazard fuels overtreatment in cost. When providers can choose a treatment without facing its financial cost, the value of medicines recommended to patients increases by 37 percent, leading to an overall increase in the patient's medical expenses by 17 percent. We find no evidence that increased costs are different for insured or uninsured patients, suggesting that doctors do not care to limit expenses of uninsured patients. This result is at odds with situations where providers have been found to limit overtreatment for uninsured patients when they gain financially. However, confidence in our null effects is limited by low power.

Third, we explore the mechanisms behind the difference in treatment costs. We show that it is not driven by a difference in the number of recommended drugs which could be induced by the patient request. Instead, we find strong evidence that doctors prescribe more expensive drugs. We show that this higher cost is not driven by differences in the drug properties (e.g. strength, compound) but by a lower probability to choose generic drugs, drugs from less expensive manufacturers or in cheaper packaging – all of which are deliberate cost-saving strategies used by dispensing doctors to minimize ex-ante the cost of drugs that reduce their profits. While these findings suggest that the absence of incentive to take into account the marginal treatment cost is an important mechanism that fuels overtreatment, we consider other possible channels. First, we rule out that overtreatment is driven by lower provider effort. Next, we acknowledge the possibility that doctors may choose more expensive drugs to respond to a perception that patients

[1] Ellis and McGuire (1986) consider the difference in treatment provision between (1) a fixed patient-based reimbursement where more treatment means a lower net revenue for the hospital and (2) a cost-based reimbursement system, where more treatment means more revenue for the hospital. In our study, we compare a situation similar to (1) – dispensing doctors receive a fixed reimbursement where more expensive drugs mean a lower profit for the consultation – to one where the provision of diagnosis and treatment are separated: when doctors prescribe more (expensive) drugs, it does not mean more revenue for them (hence the absence of profit motives), it means more revenue for pharmacists.

[2] The source of motivation of providers has been qualified in different ways in the literature. While a majority of studies refer to providers' pro-social preferences or altruism to explain why they would reduce their own payoff to the benefit of their clients, others have referred to providers' conscientiousness (Liu 2011) or professionalism (Inderst and Ottaviani 2012) to explain why they may take their customer's utility (including material payout) into account.

[3] In healthcare markets, such bundled payment structure is not uncommon. It is widely used for hospitals, who receive a payment for each patient treated independently of the specific cost of the treatment provided (the payment is a fixed tariff set prospectively at a level reflecting average resource use of similar patients).

want higher-quality drugs, but we argue that provider moral hazard itself increases the likelihood of such response. Finally, we discuss the possibility that overtreatment in cost could be the indirect effect of the segmentation of the private healthcare market. On the one hand, dispensing doctors may strategically choose to signal to patients that opting out of the affordable consultation-and-treatment bundle they offer comes with a higher cost. On the other hand, the choice of more expensive drugs may simply reflect the fact that pharmacies are less likely to stock cheaper drugs, which are less profitable.

Lastly, we explore the consequences of overtreatment in cost in the private primary care market in South Africa. Assuming the patient request in our experiment only influenced treatment choices by inducing doctors to prescribe more, we estimate that the average treatment cost nearly doubles in the absence of financial incentive to limit overtreatment, which currently characterizes the incentives of 60% of doctors working in that market (prescribing doctors). Building on a series of conservative assumptions, policy simulations suggest that this leads to an annual waste of US$367 million and would increase with the expanded contracting of private providers planned as part of the introduction of a new national health insurance scheme. However, the supply-side cost-sharing mechanism embedded in the remuneration structure of the other 40% of doctors (dispensing doctors) provides a natural strategy to simulate the implementation of cost-saving policies. In the most extreme scenario where only dispensing doctors were to be contracted, economic waste due to overtreatment could be reduced by 40%.

Overall our findings show that the separation of diagnosis and treatment in credence goods markets may not necessarily solve the problem of overtreatment, because high levels of inefficiency can still be fueled by benevolent provider moral hazard. We also show that economies of scope combined with a bundled payment, an unusual solution adopted by some providers in the private primary care market in South Africa, eliminates benevolent provider moral hazard and significantly reduces, but does not eliminate, overtreatment.

Our study contributes to several bodies of literature.

First, this paper is related to the empirical literature investigating the drivers of inefficiency in credence good markets, from overprovision of cesarean sections and drugs, to overcharging for computer repairs or taxi rides (Gruber et al., 1999; Iizuka 2012; Balafoutas et al., 2013; Balafoutas et al., 2017; Kerschbamer and Sutter 2017). We add to this literature by studying the drivers of overtreatment in a vertically differentiated market where two types of experts are available: some operate under a model of economies of scope and provide both diagnosis and treatment at a uniform price, while others operate under a separation between the provision of diagnosis and treatment. We show how experts operating in this unusual model of economies of scope internalize the financial consequences of their recommendations and reduce overtreatment in cost. By contrast, our findings suggest that the separation of diagnosis and treatment, seen in theory as a solution to eliminate overtreatment (Dulleck and Kerschbamer 2006), may fail to limit overtreatment in cost due to residual benevolent provider moral hazard, where experts neither gain nor lose financially from overtreating.

This paper also contributes to the literature on the use of financial incentives to encourage efficiency in health care and reduce provider moral hazard. There is an extensive literature on this topic focusing on hospitals' treatment decisions (Moreno-Serra and Wagstaff 2010; Busse 2012), while studies of individual providers tend to focus on the effects of capitation payments compared to fee-for-service arrangements (Lurie et al., 1994; Trottmann et al., 2012).[4] Our study is novel in that it looks at the impact of an unusual supply-side cost-sharing mechanism that can improve the efficiency of treatment decisions made by individual providers in a particular episode of care. It is highly relevant in LMICs where growing evidence points to high levels of inefficiency at the primary care level and suggests a possible solution for policy-makers, already accepted by some providers.

We also contribute to the literature exploring the effect of patient insurance status on experts' advice in credence good markets in general, and health care markets in particular. Several studies in the health economics literature have sought to study the effect of health insurance on doctors' prescribing decisions, with a number exploring whether providers recommend more generic drugs to uninsured patients (Leibowitz et al., 1985; Lundin 2000; Liu et al., 2009; Crea et al., 2019). The evidence from this literature is mixed and plagued by the challenges of observational data to control for endogeneity problems of providers' and patients' decisions, and disentangle completely providers' from patients' (expressed) preferences and choices. Field experiments using audit studies in credence goods markets can overcome these challenges. Their results suggest that clients' insurance coverage increases the likelihood of overtreatment (Lu 2014; Kerschbamer et al., 2016; Balafoutas et al., 2017) Our study is most closely related to the experimental study of Chinese doctors (Lu 2014), where auditors randomly declared to be covered by public health insurance and requested a prescription on behalf of an absent and fictitious family member.[5] Like Lu (2014), our study clearly isolates the decisions made by providers from patients characteristics and demands. Our approach improves the realism of the audit by sending trained standardized patients with real insurance coverage, receiving care for themselves. We also study the effect of patient insurance on overtreatment in the absence of providers' profit motive for overtreatment.

Finally, we add to the growing literature using audit studies to uncover the determinants of otherwise hard-to-observe behaviours. Building on the long tradition of using simulated or standardized patients[6] (SPs) to train and assess medical students, there is a small but growing literature using SPs to explore the quality of medical advice of practicing providers, particularly in low- and middle-income countries (Das, Holla et al. 2012, 2016; Mohanan et al., 2015; Daniels et al., 2017). To explore the response of providers to different stimuli, health economists have started to use SPs to undertake audit studies similar to the ones used, for example, to study discrimination (Bertrand and Duflo 2017). These studies have, for example, sent SPs matched on all observable characteristics except ethnicity (Planas et al., 2015), patient information or requests (Currie, Lin et al. 2011, 2014), and insurance status (Lu 2014).

2. Institutional setting

2.1. The private market for primary care in South Africa

Although public sector primary care is free for all in South Africa, 26 percent of the population (32.2 percent in urban areas) choose to use the fee-charging private sector (NDoH et al., 2019). This is due to the perceived higher quality of care of the private sector, where primary care consultations are done by qualified medical doctors, compared to predominantly nurses in the public sector.[7] Private doctors are used predominantly by the 16.5 percent of the population that is more affluent and covered by private health insurance (Stats SA 2019a, 2019b). However, nearly 30 percent of those regularly consulting private doctors do not have health insurance (NDoH et al., 2019), and instead are "cash patients" paying out-of-pocket for their expenses

[4] In capitation systems, providers receive a lump sum for providing care to a defined population over a defined period.

[5] In their study, patients claim to be covered by health insurance, but they are not, and never have to give proof of health insurance coverage.

[6] SPs are lay individuals trained to describe the clinical symptoms and medical history of a particular clinical case, as a regular patient would do with a doctor.

[7] In large urban centres, a few private clinics and pharmacies also offer consultations with nurses.

(Ataguba and McIntyre 2012, 2018).

The private market for primary care is essentially divided between two types of self-employed doctors: prescribing doctors, who represent 60% of all primary care doctors, and dispensing ones. The practice of prescribing doctors is characterized by the separation of diagnosis and treatment services. Patients pay a fee for the consultation, at the end of which they receive a prescription, which they must then take to a pharmacy to purchase the drugs. Dispensing doctors on the other hand provide both diagnosis and treatment (pharmaceutical drugs).[8] In many health systems, dispensing doctors sell drugs to patients and make a profit on each drug sold, which creates incentives for overtreatment (Iizuka 2007, 2012; Kaiser and Schmid 2016; Goldacre et al., 2019). This is not the case in South Africa. Although in theory dispensing doctors are allowed to charge patients a small regulated dispensing fee for each drug dispensed (see next section), in practice they don't and instead "*provide an all-inclusive service for a flat* fee*, including the provision of basic medicines*" (Gray and Suleman 2015).

The choice to opt for this unusual bundled pricing strategy is the result of the segmentation of the market for primary care between prescribing and dispensing doctors. Prescribing doctors tend to be exclusively located in the most affluent areas (see map in Appendix Fig. 1) and charge consultation rates that are on average 20 percent higher, even though they do not include drugs (Blaauw and Lagarde 2022). These doctors appeal to customers who have a high willingness-to-pay for private primary care and can afford to buy expensive drugs dispensed in pharmacies. By contrast, dispensing doctors are often located in more disadvantaged areas with limited access to pharmacies, and they position themselves as a convenient "one-stop shop" where patients can obtain a comprehensive but basic service at an affordable price. The basic nature of the service they provide is exemplified by their practice – not dissimilar to that of public clinics – to repackage drugs bought in bulk and deliver patients' treatment course in the form of small individual plastic packets (Appendix Fig. 2).

2.2. The market for pharmaceutical drugs

Unlike many other LMICs, the private healthcare market in South Africa is well-regulated and formalized (National Department of Health 2015), and the sale of pharmaceuticals is strictly controlled. Any drug categorized as 'prescription' drug can only be dispensed directly by an authorized healthcare professional or purchased in a pharmacy with a prescription written by an authorized prescriber.[9] Relationships between pharmaceutical companies and doctors are also highly regulated. Sales representatives are allowed to visit doctors to provide information about new products but they cannot give drug samples, and any other gift they make should not exceed R2,000 (approximately USD140).[10] Finally, there is no evidence, anecdotal or otherwise, of any form of kick-back from pharmacies to doctors that could create a financial incentive for doctors to prescribe unnecessary drugs (Inderst and Ottaviani 2012).

The price of drugs is strictly regulated to control pharmaceutical costs. Since the introduction of the single exit pricing (SEP) strategy in 2004, manufacturers can only sell a drug to private dispensers or wholesalers at a uniform price, communicated to all, and rebates linked to volume of sales are strictly prohibited. Dispensing doctors usually procure drugs from wholesalers[11] at a price slightly higher than the SEP, reflecting the small mark-up that wholesalers can add to the SEP to cover their costs. Wholesalers may sometimes offer price reductions to dispensers, for example when drugs come within a few months of their expiry date, but such practice is neither widespread nor systematic, and generally represents minimal reductions.[12] Authorized dispensers (pharmacies or dispensing doctors) are then allowed to sell drugs to individual consumers at a price equal to the SEP plus a mark-up that is strictly regulated and revised yearly. To limit the perverse incentives associated with economies of scope, the mark-ups that dispensing doctors can add to the SEP are capped at a low level, while dispensing fees for pharmacists are a linear function of the drug SEP – see Appendix Fig. 3. Together with the complexity of calculating dispensing fees without investing in a computerized system, these low mark-ups have contributed to the unusual, fixed pricing strategy adopted by dispensing doctors as described above.

This price regulation does not preclude variation in drug prices. In Appendix A2, we provide some evidence of the price variation in the market for some of the most commonn drugs in the study. This evidence suggests that prices vary significantly for common drugs. For example, for Amoxicillin 500 mg, the most expensive drug is worth 15 times the cheapest one. Such price variability means that when they purchase drugs from wholesalers, dispensing doctors are able to choose more or less expensive brands (e.g. from local or Indian manufacturers or European/American ones). Although there is no discount on volume, dispensing doctors can purchase drugs under different packaging, which would come at different unit price. Typically, the unit cost of a drug purchased in bulk (e.g. 500 pills of amoxicillin that will need to be reconditioned for individual patients), would be much lower than that of the same drug pre-packaged in small custom-printed boxes (e.g. a box of 15 pills of Amoxicillin) – see Appendix Fig. 2.

3. Study design

3.1. Clinical case for the audit study

To investigate the treatment choices made by doctors, we conducted an audit study with standardized patients (SPs). SPs are lay individuals recruited from the local community to visit multiple providers and present a clinical case in a systematic and blinded way (Kwan et al., 2019). We recruited twelve female and male enumerators and trained them for 10 days using a multidisciplinary team. The training included rehearsals of a detailed script of the clinical case (including standardized responses to a list of possible questions a doctor could pose), realistic portrayal of the presentation of the symptoms and patient's attitude, construction of a backstory consistent with the socioeconomic status of the individual portrayed (i.e., middle-class person), mock consultations with medical educators as well as unannounced pilot visits to confederate doctors.

SPs were trained to accurately and consistently present the clinical symptoms and history of a viral respiratory infection (acute bronchitis) in a healthy adult in their early 20s – see detailed presentation of the case in Appendix A3. The case was developed in collaboration with several local medical professionals and infectious diseases experts with

[8] To dispense pharmaceutical drugs, a doctor has to obtain a dispensing license from the Department of Health, by completing an online dispensing course for R1,000 (about GDP50). The licence is valid for five years.

[9] Drugs in South Africa fall into eight categories (called schedules), according to their safety, potential for dependency or abuse, and the need for a professional diagnosis. Drugs categorized in schedule 3 and above are 'prescription' drugs. Unlike many other LMICs, South African pharmacists abide strictly by these rules and do not sell prescription drugs directly to customers without a valid prescription.

[10] In addition, any gift should not be for the personal benefit of the doctor, but instead it should benefit his or her medical practice (e.g. water dispenser, stationary items etc.).

[11] All the dispensing doctors we spoke to explained that it was more efficient to procure drugs from wholesalers rather than individual manufacturers.

[12] Because of this well-enforced regulation on pharmaceutical pricing, drug input prices can be taken as being the same everywhere and the comparison of drug costs is not confounded by doctor differences in their ability to influence purchase price levels. The fact that most doctors operate in small practices also limits their ability to exert any market power.

the objective to portray a textbook case of acute viral bronchitis. In their opening statement, SPs described their main complaint ("*I have been coughing for a few days*"), and immediate medical history ("*I had a cold last week, but now it's better*").[13] The persistent cough is potentially consistent with a number of illnesses (e.g. bacterial bronchitis, tuberculosis, pneumonia, asthma), and doctors are trained to rule out alternative diagnoses by questioning and examining the patient (see Appendix Table 14 for a detailed list of possible diagnoses and their recommended treatments). Here, questioning and examination of the patient would uncover that the cough is productive and brings up clear mucus, but other than that the patient does *not* present any symptom consistent with the most likely alternative ailments: the patient has not had any fever; their sputum is not yellow-green (both symptoms would provide a reason to suspect some bacterial infection; their absence rules out bacterial bronchitis), nor does it contain blood (suggestive of tuberculosis); the patient has not experienced any shortness of breath and has a clear chest on examination (ruling out pneumonia); asthma can be ruled out by the absence of wheezing on exhalation (either reported by the patient or checked through auscultation), or broncho-obstruction measured by a peak expiratory flow; and the problem is a once-off episode following a recent cold (which, together with the lack of a history of allergies, rules out allergic concerns). Furthermore, the patient is young and generally healthy with no co-morbidities, which should further alleviate doctors' potential concerns of complications in immune-suppressed or susceptible individuals such as children or the elderly, which often fuel over-prescription of antibiotics.

According to local as well as international evidenced-based clinical guidelines (Woodhead et al., 2011; Smith et al., 2017, Department of Health, 2018), no medication is necessary to treat a patient presenting with such a simple case of viral bronchitis, since it is a self-limiting ailment, especially in a young and healthy subject. Some symptomatic relief treatment can however be prescribed to help the patient and relieve the main symptoms. We come back to the categorization of the different types of drugs that could be given for this case in detail in Section 4.2.

This case was chosen for two main reasons. First, respiratory tract infections (RTIs) remain one of the main reasons for primary care visits, including in South Africa (Brink et al., 2016). During the cold season, such medical problems are commonplace, helping SPs maintain a low profile. Second, since the recommended clinical treatment only includes drugs for symptomatic relief, it is a good case to study overtreatment. As the main symptom (cough) is common to several other conditions (pneumonia, asthma, allergies), it can be tempting for providers to practice defensive medicine and recommend various medicines effective to tackle these alternative conditions. Notably, the clinical literature suggests that bronchitis is a prime candidate for unnecessary prescribing of antibiotics in different settings, including South Africa (Barnett and Linder 2014; Brink et al., 2016).

3.2. Experimental design

The audit study was combined with an experimental approach to create exogenous variations in supply-side cost-sharing and patients' insurance status. To maximise the power of the experiment, we employed a mixed design – see Fig. 1. In the between-subject component of the design we randomly assigned doctors to receive either insured or uninsured patients. In the within-subject component of the design each doctor was visited by a pair of SPs, matched on gender, with visits occurring one week apart. The two SPs visiting a given doctor were randomly assigned to following one of two scenarios, as described below. To avoid confounding each fieldworker individual effect with a particular scenario, fieldworkers played all roles, according to a pre-specified random allocation. Fieldworkers were instructed to check their field journal immediately before each visit to double-check which role they were supposed to play.

3.2.1. Within-subject experiment: pharmacy dispensing encouragement

Through the within-subject component of the design, we explore whether the rationing incentive inherent to the way dispensing doctors operate in South Africa limits overtreatment and inefficient provision of services. In a companion study, we show that the value of drugs given by dispensing doctors is 2.5 times cheaper than those prescribed by prescribing doctors (Blaauw and Lagarde 2022). This suggests that the integration of the diagnosis and dispensing functions under a bundled payment leads to efficiency gains. Yet, due to obvious selection problems, one cannot identify the effect of prescribing by simply comparing the two groups of doctors. It is also not possible to randomize doctors to either dispensing or prescribing, since, by definition, prescribing doctors are not licensed to dispense. Instead, we used an encouragement design to create an exogenous change in the treatment delivery mode of dispensing doctors, allowing us to test if the same doctor chooses different treatment options when they prescribe or dispense.

Each participating doctor was visited in a random order by two SPs, identical in all respects except that one of them was randomly assigned to follow the pharmacy scenario and make a simple request. After the end of the physical examination, the 'pharmacy' patient waited for the doctor to mention the required treatment or fetch drugs from their cabinet and then said: "*If you don't mind, I prefer if you write me a script for the drugs instead*." In other words, they asked for the recommended treatment to be obtained from a pharmacy.[14] SPs were specifically trained on the timing of their request (see Appendix Fig. 4), because we did not want doctors to think that the patient requested any, or more drugs than what the doctor thought was needed. Instead, the intention was to express a preference for drugs to be dispensed from a pharmacy instead of the doctor's cabinet. The doctors we consulted in the design phase of the study thought that this would typically reflect patients' distaste for the appearance of dispensed drugs – typically pills in small plastic bags filled directly by dispensing doctors while pharmacies dispense drugs in folding cartons containing blister packs (see Appendix Fig. 2 for a typical presentation of the repackaged drugs dispensed by doctors).

When they made their request, pharmacy patients suggested that this would allow them to pay a lower consultation fee: "*If you don't mind, I prefer if you write me a script for the drugs instead … and I can pay less for the consultation*." This does not refer to the existence of two tariffs, one with and one without drugs. Instead, it alludes to an ad-hoc practice by some dispensing doctors who, in keeping with the affordable service they want to propose, may offer a reduced consultation fee to patients when they do not dispense any drugs – for example if the patient is referred to the hospital or a specialist doctor, or if doctors do not have the necessary drug in stock.[15] This justification introduced the idea that the patients cared about the financial consequences of their treatment and did not wish to spend more than necessary – an important signal as we did not want doctors to interpret the patient's request as a demand

[13] SPs were trained to say all elements in one opening sentence: "*I have been coughing for a few days. I had a cold about a week ago. The cold is a bit better now, but the cough is not going away*."

[14] This request was similar to the strategy used by Currie et al. (2014). However, in their study, the patient request removed the overtreatment incentives of Chinese doctors who earn a proportion of their income from drug sales. Here, the patient request removes the supply-side cost-sharing incentive.

[15] In a rapid phone survey of 50 dispensing doctors undertaken after this study in Johannesburg, 20% said they gave discounts to patients when they did not dispense drugs. When we created this justification, we worked with confederate doctors to ensure that it would not appear particularly suspicious and unusual. All told us that any dispensing doctor would know about this practice even if they did not offer such discount. They also all mentioned that their patients sometimes request to pay less, as they belong to lower SES groups.

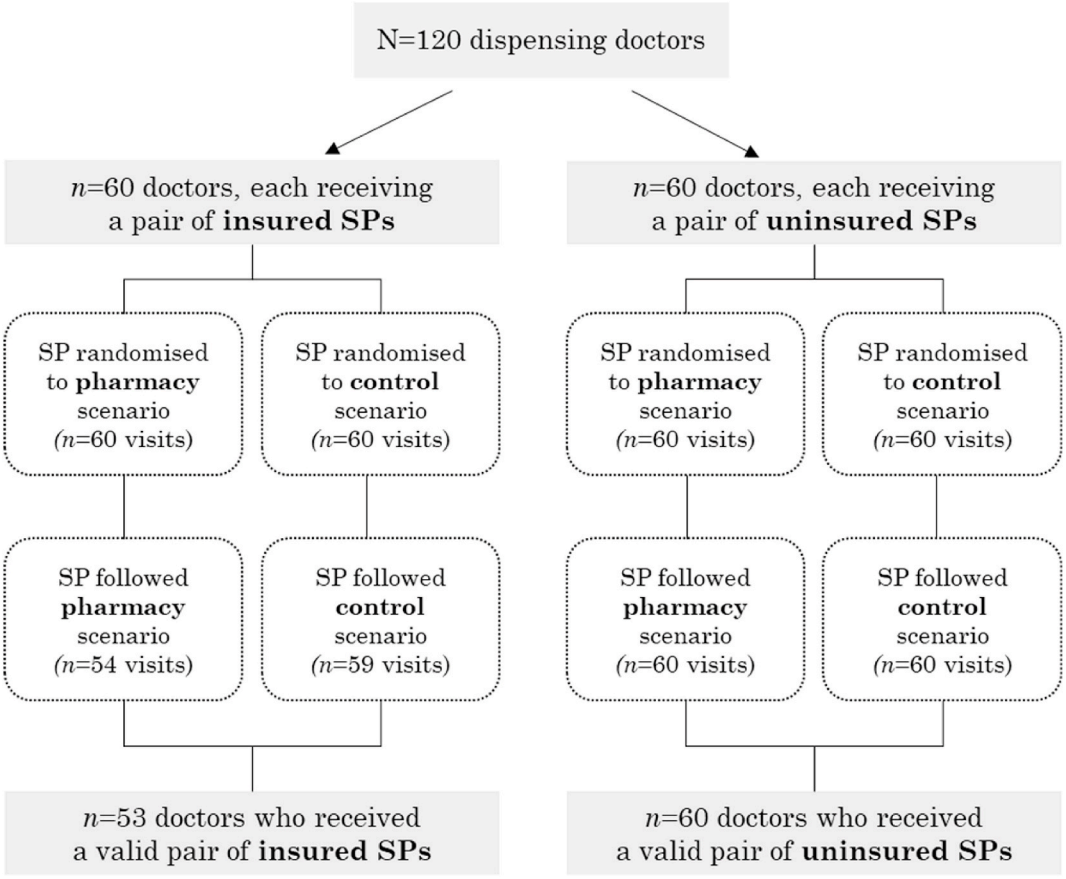

Fig. 1. Experimental design.

for more expensive drugs. If interrogated further about the motives for their request, the SPs were trained to explain that they would get the drugs from a relative who owned a pharmacy and would give them "*a good price for the drugs*" – again conveying the idea that they were price-sensitive.[16]

Each actor could take on both control and pharmacy roles. We designed the within-subject randomization in this way to avoid confounding fieldworker and role effects, in a context where there was a limited number of fieldworkers. The scenario assigned to fieldworkers in each consultation followed a pre-specified random pattern and it was recorded in an individual fieldwork journal. To reiterate the importance of maintaining the same attitude across all providers, before each consultation a field supervisor reminded fieldworkers to review in their journal (1) the whole patient script (opening sentence and responses to providers' questions), and (2) to check whether they should follow the pharmacy scenario and ask for a prescription at the end of the consultation or not.

3.2.2. *Between-subject experiment: patient insurance status*

Half of the doctors were randomized to receive a pair of insured patients while the other half received a pair of uninsured or "cash" patients. We enrolled all SPs into an insurance plan offered by one of the largest private medical insurance in South Africa. This provided real insurance coverage to the SPs, complete with an actual insurance card and electronic file under their own identity.[17] Insurance plans in South Africa are generally differentiated by the type of cover they offer for private hospital care. For primary care specifically, they all offer a system of Medical Savings Account (MSA) or pre-payment system. Upon enrolment in an insurance plan, a beneficiary receives an MSA to which some funds are allocated – the more high-end the insurance plan, the higher the funds allocated. Because it is a pre-payment system, beneficiaries effectively pay back this annual allocation in monthly installments as part of their monthly insurance premium.[18] MSA funds are used to cover primary care expenses such as primary care doctor and specialist consultations, acute medicine, drugs and pathology tests. Once an insured patient has depleted their MSA, they can no longer be reimbursed and have to pay out-of-pocket.

The plan in which we enrolled SPs had R8,316 on its MSA (approximately US$570). Given that SPs consulted many doctors in a limited amount of time, we collaborated with the insurer to ensure that our patients' insurance records would always indicate a full MSA, as if they had not incurred any primary care cost. This was important to ensure that all doctors received a patient with the same available funds, more than sufficient to claim the expenses involved in their episode of care

[16] The doctor could interpret this as the possibility that the pharmacist would probably not charge the dispensing fees (i.e. regulated mark-up) on top of the price of the drugs dispensed.

[17] Unlike past studies varying the insurance status of standardized patients (e.g., Lu 2014), in our setting, the insurance status of the patient is systematically verified by providers. Before a consultation, the doctor's secretary generally asks to see the patient's insurance card and checks with the insurer that they still have enough in their medical savings account to cover the consultation.

[18] Any unused funds are carried over to the next year. If a beneficiary uses more of the funds than the amount they have contributed (e.g. if they leave), they have to pay the difference.

and appear much less sensitive to price than the cash patients.[19]

Finally, our insured SPs always paid the consultation fee themselves and asked for a receipt to file for reimbursement themselves later. This allowed us to rule out two potential effects that could influence providers' behaviours. First, it avoided any potential distorting effect created by a deferred reimbursement of the consultation by the insurance. Second, it ruled out the possibility that doctors' behaviour would be influenced by the hypothetical outcome of the insurance claims adjudication, and concerns that they might not be paid if their recommendation was deemed inappropriate. That being said, the latter concern is mainly theoretical for primary care doctors in South Africa, where insurance companies simply encourage them to follow recommended treatment guidelines, or provide benchmarking feedback comparing their costs and quality to that of their peers. Beyond these strategies, which have largely remained ineffective in controlling costs or improving quality (Ranchod and Dube 2019), insurance companies exert little oversight over primary care doctors,[20] partly for lack of reliable data to judge the adequacy of treatment decisions, and partly because insurers are concerned with more obviously fraudulent behaviours, such as submissions of false claims (Legotlo and Mutezo 2018).

3.3. Minimizing the risk of detection by informed participants

Participating doctors knew they would be visited by fieldworkers posing as real patients. This awareness is uncommon in the audit studies used in economics and social sciences to investigate the discriminatory behaviours of firms (Bertrand and Duflo 2017). Some studies seem to have been exempt from IRB review (Bertrand and Mullainathan 2004; Bauhoff 2012), perhaps on the basis that all of the requirements for the protection of human subjects might not apply to research focusing on the behaviour of firms rather than individuals. Meanwhile, other audit studies have obtained permission from RECs to waive the informed consent of research participants (firms) as long as a number of other criteria were met – from minimizing disruption to normal activities, to guaranteeing research subject confidentiality by reporting results in aggregated form (Zschirnt 2019). The same approach was recently taken by a few medical audit studies using SPs (Mohanan et al., 2015; Planas et al., 2015; Das et al., 2016; Kwan et al., 2018).

We did not try to argue that our project was focusing on the behaviours of firms, rather than individuals, in the private market for healthcare. Hence we cannot know how the RECs would have adjudicated had this definitional question been submitted to them. Instead, we followed the second approach taken by audit studies and requested a waiver of informed consent arguing that the benefits of our research (and its scientific validity) would outweigh the minimal risks posed to providers or other real patients, especially given that provider confidentiality would be maintained. Unfortunately, this request was denied by the REC in South Africa, on the grounds that informed consent was a fundamental individual right protected by the national Bill of Rights that could not be waived. As a result, to mitigate the potential concerns arising from the knowledge of participating doctors, we took a number of precautionary steps which we believe ensured that empirical consequences of doctors' informed consent were kept minimal.

First, we gave doctors minimal information about the SP visits. They were told that we would send "*fieldworkers trained to act as patients*" but they did not know what the patient symptoms or characteristics would be. They also ignored how many patients they would receive, and the timing of the visits was purposefully long and imprecise ("*over the next six months*"). Second, most doctors were recruited well ahead of the patient visits: on average three to four months before the visits.[21] It is unlikely that doctors would have changed their behaviour in a sustained way in anticipation of the SP visits. Empirical evidence of the effect of much more direct observations of clinical consultations provides reassuring evidence that the Hawthorne effect is quite short-lived (Leonard and Masatu 2010).[22] Besides, doctors did not know specifically what outcomes we were interested in, as we told them that the focus of the study was on "*clinical decision-making*". Third, all visits occurred in the middle of the 'cold season' of South Africa (July–August), where doctors see a sharp increase in the number of patient visits, a majority relating to viral respiratory infections. The large volume of patients seen by an average doctor in our sample (26 patients per day) would not only have helped our SPs blend in, but the increased workload would also have reduced doctors' ability to pay attention to individual patients' story details. Fourth, in a context of a busy urban setting like Johannesburg, doctors' clientele is very fluid, and doctors often see new patients. Finally, we worked closely with confederate doctors to ensure that the presentation of the patients and the experimental variations used would not attract any particular attention. As mentioned before, the scripts followed by the patients, especially the one requesting a prescription, were designed and tested with practicing doctors to be credible and avoid drawing particular attention to the SPs.

To verify that SPs succeeded to remain incognito, we phoned all doctors less than two weeks after the last visit. Overall, only four doctors indicated some suspicions, corresponding to five consultations, one of which was ruled out based on the description of the suspected patient's characteristics.[23] We have no reason to believe that this low detection rate (four of the 240 SP visits or 1.66 percent) could be driven by doctors' under-reporting or recall problems. First, this low rate is in line with similar SP studies that have found detection to be below 5% (Kwan et al., 2019). Second, this low detection occurred despite the fact that we asked doctors to keep a record of any patient they suspected to be one of our fieldworkers.[24] Not only did we encourage doctors to record any suspicion (we explained the feedback would help us improve the study), but many were keen to show they could not be easily "duped".[25] This

[19] In general the practice receptionist would check the validity of a patient's insurance before the consultation and communicate to the doctor the patient's insurance status.

[20] For the specific clinical case used in this audit study, the analysis of claims data showed that doctors prescribed antibiotics for 52.9% of patients they diagnosed as having acute bronchitis, even though this was not recommended by national guidelines (Ncube et al., 2017). Insurers paid these claims without comment, and seem to have made little effort to reduce inappropriate antibiotic prescribing more generally (Ncube et al., 2017).

[21] Recruitment into the study started in March 2018, while the visits occurred in July–August 2018. By the time we contacted them again for a follow-up interview, some doctors had to be reminded that they were part of the study.

[22] Looking at doctors' performance in Tanzania before and after the start of direct observations of consultations by enumerators, Leonard and Masatu (2010) show that although doctors initially improve their performance when they are directly observed, they revert back to a normal behaviour after a few consultations.

[23] In the detection survey, we asked four closed questions about the suspected patients (consultation date, symptoms, age, gender) and one open-ended (reasons for suspicion) that allowed us to cross-check their suspicion with the SP actual visits of the fieldworkers.

[24] When recruited, doctors were told to continue as normal if they suspected a patient to be one of our fieldworkers to protect real patients in case doctors were mistaken. Instead, we suggested that they record the patient's characteristics (e.g. consultation date, symptoms, name, gender, age). This suggestion neither raised concerns related to fear of liability or misconduct (doctors were clearly told to continue any consultation as normal in case they were mistaken), nor did it breach patients' confidentiality – we explained that we would not collect any identifiable information about the suspected patients, simply generic characteristics.

[25] Several recruited doctors, as well as some confederate doctors at the start of the study, were confident they could easily spot 'fake' patients. The views of confederate doctors changed after they received some of our patients and agreed they could realistically pass on as real patients, as they were "looking the part". During the detection survey, several doctors expressed their surprise upon realising they had already received SPs without suspecting anything.

makes it unlikely that doctors responded to some experimenter's demand effect and under-reported their suspicions. Third, even if doctors failed or forgot to record the details of the consultation, suspecting a patient to be 'fake' would have been enough of a memorable – and recent – event that they would remember it, if not the specific patients' details.

Together with the precautionary steps we took, these results provide credible evidence of the validity of our audit study.

4. Data and empirical strategy

4.1. Sample characteristics and balance

We carried out the study in the metropolitan area of Johannesburg, the main urban center in South Africa. To construct a sampling frame of private doctors, we used a commercial national database of practitioners, which includes approximately 80 percent of all registered doctors nationally, and significantly more in urban areas. The database included the contact details of 1012 practicing private primary care general practitioners in Johannesburg, 361 of whom (35.7 percent) were primary care doctors licensed to dispense drugs and eligible to take part in the study.[26] We called all 361 eligible doctors between March and June 2018 to invite them to take part in the study. Of those, 26 percent ($n = 94$) could not be reached despite several attempts[27]; 26.3 percent refused to take part ($n = 95$); 11.6 percent ($n = 42$) requested further information about the study to make their decision but never responded again, and 36 percent agreed to take part ($n = 130$). From this final group, we randomly chose 120 to take part in the study. Panel A in Appendix Table 2 shows the validity of the randomization based on the limited data available from the sampling frame. The only observable difference is that a higher proportion of providers who received insured patients were in the top 20 percent of wealthiest areas.

To preserve the within-doctor comparison between control and pharmacy patients, we dropped the seven pairs of SP visits where at least one enumerator had not complied with their assigned script (see Fig. 1 and Appendix Table 1). This left a final analysis sample of 226 consultations with 113 doctors, of whom 60 received pairs of uninsured patients and 53 received pairs of insured patients.[28] Although we cannot completely rule out that these seven excluded doctors were different from the others, Panel B in Appendix Table 2 shows that the balance between the two groups of doctors was preserved.

Of the 113 providers in the study, 86 (75 percent) agreed to a follow-up face-to-face interview during which we collected detailed information about their characteristics and beliefs. Table 1 presents some basic summary statistics about the 113 providers included in the study, based on the information available from the initial database (column 1) and for the subgroup who took part in the interview data (column 2). The data show that the doctors are mostly male, 51 years old on average, and therefore quite experienced (24 years of practice on average). The doctors work in a wide range of socioeconomic areas, even though 57 percent work in the richest 40 percent of areas.[29] Doctors interviewed reported seeing on average 26 patients in a working day—a plausible number of patients, given that, on average, a consultation for the SPs lasted about 10 min.[30]

Table 1
Provider characteristics.

	(1) Full sample		(2) Interview sample	
	Mean	(SD)	Mean	(SD)
Male	0.73	(0.44)	0.73	(0.45)
Age	51.57	(10.98)	52.48	(10.78)
Practice location, by SE quintile of local area				
Practice is located in Q1 (poorest 20%)	0.11	(0.31)	0.08	(0.28)
Practice is located in a Q2	0.21	(0.41)	0.23	(0.42)
Practice is located in a Q3	0.12	(0.32)	0.12	(0.33)
Practice is located in a Q4	0.41	(0.49)	0.39	(0.49)
Practice is located in a Q5 (richest 20%)	0.16	(0.37)	0.18	(0.39)
Ethnicity				
African			0.38	(0.49)
White			0.12	(0.32)
Asian			0.41	(0.49)
Other			0.09	(0.29)
Experience as a doctor (years)			24.35	(9.12)
No. of patients per day, previous week			26.28	(13.90)
Observations	113		86	

Note: Column 1 presents the basic characteristics of the full sample. Data come from a commercial database (Medpages) we used to construct the sampling frame, which includes basic demographic information and geographical coordinates of the doctor practice. These coordinates were used to assign doctors to the socio-economic quintile of their local area (ward), constructed from the Gauteng City-Region Quality of Life survey. Data in column 2 come from provider interviews undertaken as part of the study with the subset of doctors who agreed to the interview, undertaken in the autumn of 2018, after all SPs had finished their visits.

4.2. Outcomes of interest

Our objective is to study the effects of incentives on overtreatment. Following Emanuel and Fuchs (2008), our definition of overtreatment encompasses two dimensions. First, overtreatment in volume, which relates to the choice of treatment and is defined as a situation where a provider recommends medicines that have no proven clinical benefit for the case. Second, overtreatment in cost, which relates to the cost of treatment, specifically situations where treatment is more expensive than necessary because clinically equivalent alternatives exist that are cheaper.

Choice of treatment. Using the information written on prescriptions or from the drugs directly dispensed, sometimes with the help of pharmacists, we compiled exhaustive information about the drug class and specific compound of all drugs recommended by doctors to SPs. We consulted South African as well as international clinical guidelines and evidence to determine the recommended or appropriate treatment for the case, and from that define overtreatment. There was a clinical consensus that only a symptomatic relief treatment including analgesics or cough suppressants/expectorants would be recognized as beneficial for a young and healthy subject suffering from this self-limiting virus. Any other drug was therefore deemed unnecessary as it would provide no or very limited clinical benefit to the patient. We further split unnecessary drugs into two categories: unnecessary 'only' and unnecessary and harmful drugs. Unnecessary 'only' drugs would provide no clinical benefit for the patient, but their consumption does not cause any harm, either to the patient or others.[31] In this study, they range from probiotics

[26] To be eligible to take part in the study, a doctor had (1) to practice general medicine; (2) to work in a private practice.

[27] Either no one responded, or the receptionist refused to pass the communication to the doctor.

[28] Failure to portray the role assigned occurred more frequently for the insured SPs because they had to remember to indicate their private insurance (something our fieldworkers were not used to), as well as the other elements of their script.

[29] The quintiles were defined based on an analysis of household data from the Gauteng City-Region Observatory.

[30] Assuming some patients take a bit more time, this would therefore represent about 4–5 h of clinical work in a day.

[31] We discussed with medical experts whether the steroids recommended could be considered harmful. The dosage and duration of the recommended courses were such that medical experts ruled out any potential harmful effects.

to bronchodilators, antihistamines or steroids (see Appendix Table 15 for an exhaustive list). By contrast, unnecessary and harmful drugs not only have no proven clinical benefit to the patient, but their consumption also creates a negative externality with damaging effects. There is a consensus in the clinical literature highlighting the lack of clinical benefit of antibiotics in young and healthy patients suffering from bronchitis (Smith et al., 2017). In addition, unnecessary consumption of antibiotics contributes to antimicrobial resistance, which has potentially long-lasting effects on public health. Hence, any antibiotics included in patients' treatment was classified as unnecessary and harmful. Following this classification, we created three binary outcomes denoting whether the patient treatment included (i) any unnecessary drug (overall overtreatment); (ii) any unnecessary 'only' drugs and (iii) any unnecessary and harmful drugs.

Cost of treatment. We look at the financial implications of the treatment recommended to a patient, by considering both its intrinsic monetary value and its cost to the patient. We compute the monetary value of a treatment by multiplying the treatment course recommended by the doctor (e.g. number of tablets) by the regulated price per standard unit (ie, price per tablet) at which the recommended drug is sold by its manufacturer – the Single Exit Price or SEP. When we could clearly identify the name of a brand, we used the SEP of that specific drug. In the absence of specific information about the actual manufacturer, for example when only the name of a molecule was indicated on a dispensed packet or prescription, we followed a conservative approach and chose the cost of the cheapest generic drug available in the market. Finally, in the few instances where we could not identify a drug in the national database of pharmaceutical products, we determined the closest match or used the cheapest retail prices for over the counter or natural products.[32]

Computing the treatment monetary value is best to test for underlying cost differences between the treatment chosen by doctors when dispensing vs. writing a prescription to be filled in pharmacies. However, it is also important to consider the cost of treatment to patients, which incorporates both the value of drugs and their dispensing modes. When medicines are dispensed by a doctor as part of the consultation, the cost of treatment to the patient is simply equal to the consultation fee. By contrast, if the doctor writes a prescription to a patient for (part of) their treatment, the cost of treatment includes the consultation fee plus the expenses incurred at the pharmacy when purchasing the drugs. These pharmacy expenses are obtained by adding the regulated pharmacy mark-up (dispensing fee) to the SEP value of the drug.

Note that pharmacy expenses are estimated based on the value of the drugs written by the doctors on the prescription, not based on filling the actual prescription at pharmacies. It is possible that the latter could be less expensive than the former, because according to the Medicines & Related Substances Act (1965), when a doctor prescribes a brand name drug, the pharmacist should advise the patient about a generic replacement. If the patient agrees to the substitution, the pharmacist is supposed to "*take reasonable steps to inform the doctor that they have substituted for a generic*". It is difficult to know what proportions of pharmacists would comply with this requirement. Studies suggest that this policy is in practice poorly implemented as patients tend to place a greater trust in the specific recommendations made by their doctor rather than the advice of pharmacists (Patel et al., 2010). Nevertheless, our prescription costs are a higher-bound estimate of the actual costs to patients.

[32] In some cases, with the help of doctors and pharmacists, we could identify a close match in the drug database that would typically be considered as a suitable option by a pharmacist. In this case, we used the price of the cheapest match for computing the drug cost. In other cases, especially for nonregulated drugs such as herbal remedies or probiotics, we used the cheapest available product from two big pharmacy chains that list most of their drugs and prices online. In both strategies the prices used are likely to be conservative estimates.

4.3. Empirical strategy

The pharmacy patient's request introduced a partial random manipulation of the delivery mode of the treatment. When receiving the pharmacy patients, dispensing doctors could choose to dispense the treatment themselves, as they would normally, or they could write a prescription that the patient would take to a pharmacy to purchase the drugs. The first stage of our strategy to encourage providers to prescribe and lift supply-side cost-sharing requires that the pharmacy scenario had the intended effect of changing the delivery modes of the patient treatment, and get more dispensing doctors to prescribe. Appendix Fig. 5 shows in detail the large effects of the pharmacy scenario on the delivery mode of drugs. In control consultations, nearly 79 percent of patients obtained drugs exclusively dispensed by the doctor while this proportion falls to 40 percent when patients followed the pharmacy script. Appendix Table 3 confirms that the pharmacy scenario increases the likelihood that the doctor prescribed at least one drug by nearly 41pp. Overall, there is clear evidence that the encouragement strategy worked: we created an exogeneous variation in the cost-sharing incentive faced by doctors.

We use this exogenous variation to identify the effect of the patient request, which is also the intent-to-treat (ITT) effect of prescribing, or in our setting, of lifting doctors' cost-sharing payment structure.[33] We estimate the following fixed-effects model to account for the hierarchical nature of the data, since each doctor saw two patients:

$$Y_{ij} = \beta_0 + \beta_1 Pharmacy_{ij} + \gamma_j + order_i + fw_i + \varepsilon_{ij}$$

where Y_{ij} is the outcome for patient i seen by doctor j, and $Pharmacy_{ij}$ takes the value 1 if patient i told doctor j that they would prefer to get a prescription to obtain their treatment from a pharmacy, and 0 otherwise. In addition, we control for provider (γ_j), visit order ($order_i$), and fieldworker (fw_i) fixed effects. The coefficient β_1 captures the ITT effect of removing doctors' cost-sharing payment structure. We hypothesize that overtreatment is lower when the provider bears the cost of treatment, hence it should increase when providers are encouraged into a situation of benevolent moral hazard.

Next, we explore the effects of the patient request and insurance status by estimating the following specification:

$$Y_{ij} = \beta_0 + \beta_1 Pharmacy_{ij} + \beta_2 Insured_i + \beta_3 Pharmacy_{ij} \times Insured_i + \gamma_j + order_i + fw_i + \varepsilon_{ij}$$

where $Insured_i$ takes the value 1 if patient i was insured (and 0 otherwise). Coefficient β_2 captures the impact of the patient's insurance status on doctors' treatment choices in their default incentive environment (i.e. facing a cost-sharing incentive). The coefficient on the interaction term β_3 captures the effect of the patient request when it is made by insured patients. We would expect that overtreatment increases even more when neither the provider, nor the patient bears the cost of treatment.

5. Results

We first consider the effects of our treatment variations on overtreatment in volume, by looking at treatment choices, before looking at the effects on overtreatment in cost.

5.1. Choice of treatment

The results on treatment choices are presented in Table 2. In Column 1, we consider effects on the likelihood to recommend unnecessary

[33] This is the ITT effect because providers may still choose to dispense drugs despite the request for a prescription made by patients.

Table 2
Effects on doctors' treatment choices.

	Overtreatment			Undertreatment
	(1)	(2)	(3)	(4)
	All unnecessary drugs	Unnecessary only drugs	Unnecessary & harmful drugs	Appropriate drugs
Panel A. Effects of patient's request				
Pharmacy patient	0.001	0.091*	0.001	0.028
	(0.015)	(0.047)	(0.045)	(0.038)
Mean for control patients	0.991	0.805	0.708	0.805
Observations	226	226	226	226
Panel B. Effects of patient's request and patient insurance				
Pharmacy patient	−0.000	0.035	0.031	0.019
	(0.022)	(0.062)	(0.060)	(0.051)
Insured patient	0.020	−0.057	0.001	−0.112
	(0.023)	(0.070)	(0.088)	(0.073)
Pharmacy X insured patient	−0.014	0.104	−0.065	0.012
	(0.031)	(0.090)	(0.086)	(0.074)
Mean for control 'cash' patients	0.983	0.833	0.717	0.850
Observations	226	226	226	226

Notes: Data come from the SP debriefing questionnaire. Estimates are from fixed-effects linear probability models looking at the effect of the pharmacy scenario and the insurance status randomly assigned to standardized patients. Results in columns 1–3 consider changes in overtreatment in volume. In column 1, the dependent variable is a dummy variable that takes the value 1 if the patient treatment included any drug deemed unnecessary or potentially harmful. In column 2, the dependent variable is a dummy variable that takes the value 1 if the patient treatment included any drug deemed unnecessary (this category includes primarily steroids, antihistamines, bronchodilators – see details in Appendix Table 15). In column 3, the dependent variable is a dummy variable that is equal to 1 if the treatment included unnecessary and potentially harmful drugs (in this clinical case, antibiotics) and 0 otherwise. In column 4, the dependent variable is a dummy variable that takes the value 1 if the patient received any drugs deemed appropriate, i.e. to alleviate symptoms (analgesics or cough suppressant). A negative coefficient associated with the patient request or insurance status would be a reduction in appropriate treatment (under-treatment). All regressions include provider, fieldworker and visit order fixed effects, as well as a constant term. Observations are at the SP-provider interaction level. Standard errors are in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

drugs; in Columns 2 and 3 this is split respectively between unnecessary only drugs, and unnecessary and harmful drugs (antibiotics).

We find hardly any evidence that overtreatment in volume increases when providers no longer have to bear treatment costs (Panel A). There is no evidence that the patient request changes the overall probability of recommending any unnecessary treatment (p = 0.954) or the probability of recommending an antibiotic increased ($p = 0.983$). There is only some borderline significant evidence suggesting an increase in the probability of receiving some unnecessary 'only' drugs, by about 9.0 percentage points ($p = 0.055$), an increase by 11 percent compared to control patients. This evidence is reassuring as it suggests that, in this context, doctors' treatment choices were not fundamentally altered by the change in financial incentives. However, the result could also be driven by the high rates of overtreatment observed, even in the presence of supply-side cost-sharing. Virtually all (99 percent) control patients' treatments include at least one unnecessary drug and nearly 71 percent include some antibiotics.

Given that insured patients are less price-sensitive, we could expect the rate of overtreatment to be greater for them than for cost-sensitive cash patients. The results shown in Panel B of Table 2 do not lend much support to this idea. Overall, in the default situation where doctors face a cost-sharing incentive, there is no evidence that patients' insurance status influences the treatment choices made by doctors, across all outcomes considered. However, the fact that the increase in the probability that a patient receives some unnecessary drug (Column 2) is no longer statistically significant in this specification ($p = 0.572$) suggests that this effect is more concentrated in insured patients, although the experiment may not be powered to detect such effect ($p = 0.247$).

Overall, this first set of results shows hardly any evidence supporting the notion that doctors recommend unnecessary drugs due to benevolent provider moral hazard. The lack of evidence may be partly driven by the fact that rates of overtreatment are already very high in the clinical context studied, even in the presence of cost-sharing incentives for doctors.

5.2. Cost of treatment

We now consider the effects of provider moral hazard and patient insurance on treatment costs. We first plot the cumulative distribution functions (CDF) of the value of drugs recommended by doctors to control and pharmacy patients (Top Fig. 2) and the CDF of the total cost of the episode of care to patients (Bottom Fig. 2). Fig. 2 shows that the distribution of treatment value for pharmacy patients first-order stochastically dominates that of the control patients (Kolmogorov-Smirnov [K–S] test, $p = 0.008$). Hence, when doctors no longer have to bear the cost of treatment, they recommend more expensive treatment than when they face a cost-sharing incentive. Given that pharmacy patients are more likely to have to purchase drugs from a pharmacy where they incur the pharmacy dispensing fees, the distribution of the total cost of the episode of care is significantly higher for pharmacy patients than for control patients (K–S test, $p = 0.002$).

In Panel A of Table 3 we present regression results estimating the ITT effect of provider financial indifference on the total value of drugs recommended by the doctor (Column 1) and the total cost of the episode of care (Column 2) as well as the consultation fee paid by the patient (Column 3). The latter is included to better interpret any change in the cost of the episode of care, given that doctors who prescribed may have granted a fee reduction to patients.

The results confirm that when dispensing doctors can be released from the cost-sharing they usually face, they choose more expensive treatments, which translates into higher costs to the patient. Estimates in Column 1 show that the treatment recommended to patients who requested a prescription was worth R38 more on average than that of control patients ($p = 0.005$), corresponding to a 37 percent increase. A disaggregated analysis by category of drugs (Appendix Table 4) shows that the value of medicines chosen by doctors increased by a similar amount across all three drug categories (unnecessary, harmful, and appropriate), meaning that the overall increase in driven by the higher value of unnecessary and harmful medicines. Moreover, the cost of the episode of care for patients who requested a prescription is R85 higher ($p < 0.001$) than those who did not – corresponding to a 17 percent difference. This large increase is driven by the cost of prescribed

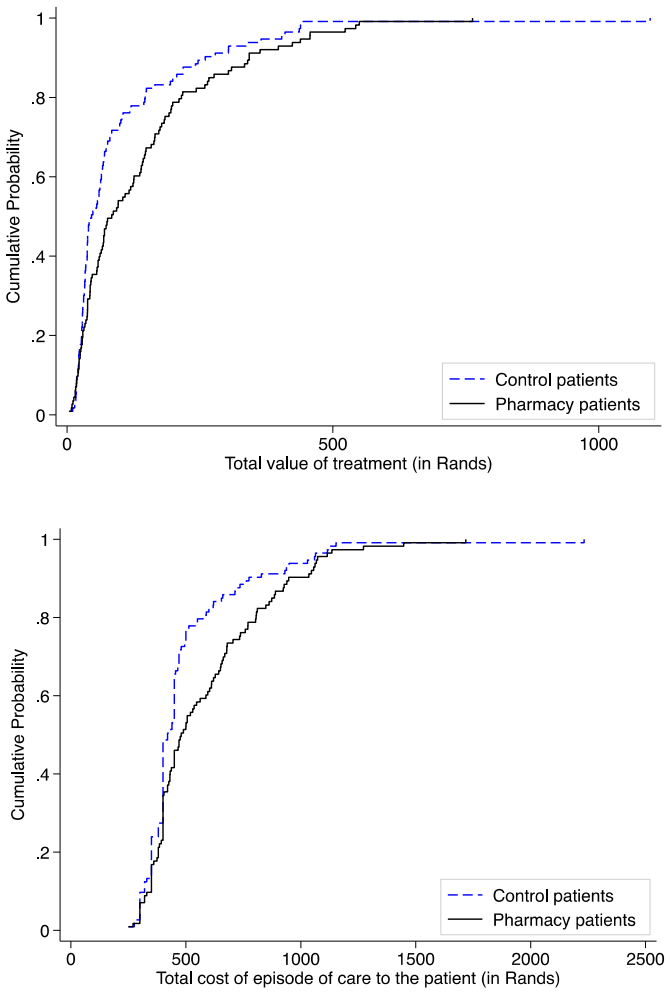

Fig. 2. Distribution of the treatment value (top) and cost to patient (bottom) of recommended treatment.
Note: this graph shows the cumulative distribution functions (CDF) of the value of treatment recommended to control and pharmacy patients (top) and the total cost of care for the two types of patients (control patients in dashed line, pharmacy patients in straight line). The total value of treatment is calculated by multiplying each patient's recommended treatment by its regulated single exit price. The total cost of care includes the consultation fee and the cost of purchasing drugs from a pharmacy if any was prescribed by the doctor.

medicines which have to be bought in pharmacies at a price that includes pharmacy dispensing fees, added to the value of medicines. Results from Column 3 show that this increase in cost is not compensated by the small and insignificant reduction in fees.[34]

Together, these results suggest that when providers are indifferent to the cost of treatment, they recommend more expensive medicines, leading to significantly higher healthcare expenditures for patients.

Is this increase more concentrated in insured patients? Looking at the effects of patients' insurance status in Panel B of Table 3, providers do not seem to treat differently insured and uninsured patients. There is no statistically significant evidence ($p = 0.867$) that the value of drugs chosen is different for insured patients compared to uninsured ones

[34] The estimated reduction was approximately worth R10 or 2.5 percent of the average consultation fees of doctors receiving control patients. This finding is consistent with some complementary evidence that we gathered showing that only a minority of providers (about 20 percent) would consider a reduction in fee in case they did not dispense any drug, and that the reduction would generally be small (between R20-50).

Table 3
Effects on treatment value and cost of care.

	(1) Total value of drugs	(2) Total cost of episode of care	(3) Consultation fee
Panel A. Effect of patient's request			
Pharmacy patient	37.689***	84.623***	−9.935
	(13.172)	(20.735)	(6.815)
Mean for control patients	101.35	500.53	411.60
Observations	226	226	226
Panel B. Effect of patient's request and patient insurance			
Pharmacy patient	32.927*	84.381***	−13.973
	(17.679)	(27.873)	(9.283)
Insured patient	−4.566	27.956	27.823*
	(27.308)	(50.677)	(16.330)
Pharmacy X insured patient	7.107	−1.709	7.514
	(25.415)	(39.939)	(13.310)
Mean for control 'cash' patients	103.2	485.8	396.9
Observations	226	226	226

Notes: Data come from the SP debriefing questionnaire. Estimates are from fixed-effects linear probability models, looking at the effect of the pharmacy scenario and the insurance status randomly assigned to standardized patients. In column 1, the dependent variable is the total value of the treatment recommended by the doctor, measured by multiplying, for all drugs, the quantity of drugs recommended for the treatment duration by the standard exit price. In column 2, the dependent variable is the total cost of the episode of care, which includes the consultation fee and the cost of any drugs to be purchased in a pharmacy, inclusive of the pharmacy dispensing fee. In column 3, the dependent variable is the consultation fee paid by patients. All regressions include provider, field-worker and visit order fixed effects, as well as a constant term. Observations are at the SP-provider interaction level. Standard errors are in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

(Column 1). In addition, although the coefficient associated with the interaction term is positive, which suggests a higher increase in treatment value for insured patients, this is not statistically significant ($p = 0.780$). It follows that the increase in the cost of care resulting from the provider's isolation from cost is the same regardless of the patient's insurance status (Column 2).

6. Mechanisms

6.1. *Responding to a demand for (more) drugs?*

A concern is whether doctors interpreted the request of pharmacy patients as a demand for *any* drug, or *more* drugs than the doctor was planning on recommending. To avoid this, recall that we trained patients to time their request after doctors had clearly formed an opinion about the treatment of the patient (see section 3.2.1). SPs were also coached to put the emphasis of the request on the delivery mode.[35] Still, if doctors interpreted the request of the patient as a demand for some or more drugs, we should see a change in (1) the likelihood of having at least one drug recommended as part of the treatment or (2) the number of drugs included in the patient's treatment. This is not the case. First, regardless of the script followed by the patients, all treatments included at least one drug. Second, as illustrated in Appendix Fig. 6, there is no difference in the distribution of the quantity of drugs received by patients in the control and pharmacy treatments (K–S: $p = 0.991$). Appendix Table 5 formally confirms this result, for all drugs together and each drug category separately.

[35] In practice, they often said: "*for the treatment … could you write me a prescription instead?*"

6.2. Decomposing the gap in treatment cost

Looking at the average drug price by delivery mode for the six most popular drug categories (Fig. 3), prescribed medicines appear consistently more expensive than dispensed drugs.[36] However, price is not the only difference between dispensed and prescribed treatments. Compared to dispensed drugs, prescribed drugs are seven times more likely to be branded and from manufacturers from high-income countries. There are also some differences in the type of drugs chosen in terms of active ingredient (e.g. less paracetamol and more expectorants amongst prescribed drugs) or strength of the active ingredient – see Appendix Table 6 for a comparison of the characteristics of drugs dispensed and prescribed. To determine how much these differences drive the gap in costs, we undertake a Blinder-Oaxaca (BO) decomposition (Blinder 1973; Oaxaca 1973). This approach allows us to determine how much of the price gap is driven by (1) differences in the characteristics of the two groups of drugs (characteristics), (2) differences in the relationship between price and drug characteristics (coefficients), or (3) some interaction between the first two differences – details of the approach are provided in Appendix A4.

The results of the decomposition, presented in Table 4, Panel A, show that 51.2% of the cost difference between prescribed and dispensed drugs comes from differences in the drug characteristics, and 37% from differences in the relationship between price and drug characteristics. Fig. 4 shows the detailed results of the decomposition, with the contribution of each drug characteristic to the cost gap in characteristics (left) and coefficients (right). The fact that prescribed drugs are more likely to be branded (as opposed to generic) and from high-income country manufacturers explains respectively 28% and 43% of the price gap in characteristics. The remainder of the gap comes from the greater likelihood of receiving certain drug compounds (e.g. paracetamol, NSAIDs, expectorants) and a higher compound strength as part of a prescribed treatment compared to a dispensed one. The decomposition of the cost gap in 'coefficients' also shows that prescribed drugs from high-income country manufacturers and from South Africa tend to be more expensive than dispensed ones.

Overall, these results provide strong evidence that prescribed treatments are more expensive because of a choice of more expensive drugs, not due to clinical differences in treatment choices, such as differences in drug compound, strength or drug class.

6.3. The benefits of supply-side cost-sharing

One way of exploring further what drives the difference in treatment costs when doctors dispense and when they prescribe is to explore further the impact of the specific steps taken by doctors to limit the marginal costs they purchase drugs. The doctors we interviewed reported using three strategies to minimize *ex-ante* how much they spend on drugs. First, they only purchase the more common and basic drugs. Second, they procure almost exclusively generic drugs, often from the cheaper manufacturers - which may incur time and effort if they need to contact a range of wholesalers. Third, whenever possible, they buy drugs in bulk, since those come at a much cheaper unit cost - doctors later repackage the drugs into individual plastic packets corresponding to a patient's recommended treatment.

The decomposition analysis already showed the impact of the choice of generic drugs. To explore how much the last practice contributes to differences in treatment costs, we undertake the BO decomposition analysis in the sub-sample of solid oral drug products (i.e. tablets and pills), which are the most amenable to being repackaged and represent 69% of all drugs in our sample. The results, presented in Table 4 Panel B, are striking. Differences in drug characteristics now explain nearly 76% of the difference in treatment cost, and 42% of this gap comes from the (repackaged) drugs bought in bulk by doctors when they dispense (see detailed results in Appendix Table 18). As before, the greater proportion of generics amongst dispensed drugs explains another 25% of the cost gap, while the greater use of high-income countries manufacturers when prescribing explains another 17%.

7. Alternative explanations

While our results suggest that the absence of incentive for providers to consider the marginal treatment cost is an important mechanism that fuels overtreatment in primary care, other possible channels could drive the difference observed. In this section we consider the following alternative explanations: (1) a difference in provider effort; (2) a perceived demand for high-quality drugs; and (3) the indirect effects of market segmentation.

7.1. Provider effort during the consultation

A potential concern is whether the differences in drugs recommended to the patient who requested a prescription and the one who didn't are driven by differences in the level of effort exerted by providers during the consultation. For example, the choice of more expensive drugs for the pharmacy patients' treatment could be the result of a less thorough examination or could be a way to compensate patients for a shorter consultation – assuming that doctors believe patients' utility increases in drug price, an assumption we explore further below. Because the experimental variation creating the change in incentive structure (patient request) occurred at the end of the consultation, systematic differences in doctor behaviour during the patient interview or examination are unlikely. Yet, this experimental variation was created by the visits of two different SPs, one following the pharmacy scenario and another one following the control script. Even if we minimized the within-SP pair variation by matching fieldworkers on gender, age and ethnic group, and standardizing the case presentation, one cannot rule out that providers exerted less effort when they received the pharmacy patients. To test this hypothesis, we use the fact that standardized patients were trained to record detailed aspects of the consultation – see Appendix Table 16 for more details about the post-consultation questionnaire. Analysis of these data rules out that there were differences in provider effort, measured by the number of questions asked and examinations done, or duration of the consultation (Appendix Table 7). In other words, overtreatment was not fueled by lower provider effort.

7.2. A perceived demand for higher quality drugs?

There is anecdotal evidence suggesting that South African patients may view generic or cheap drugs as a signal of low quality (Patel, Gauld et al. 2010, 2012; Mpanza et al., 2019). In qualitative interviews, dispensing doctors indicated that patients may frown at the drugs packaged in small plastic packets, which look like the free treatment obtained in public clinics (Appendix Fig. 2).[37] Although the same doctors denied that they would have interpreted the patient request as a concern for the quality of the drugs they dispense, or a willingness-to-pay for what could be perceived (wrongly) as higher-quality drugs, we cannot rule out that other doctors interpreted the patient's request like that. Whilst there is no evidence that counterfeit or sub-standard drugs are a problem in South Africa (Patel et al.,

[36] We did not include the other three drug categories (vitamins, probiotics, nasal sprays) as there were too few observations in one of the two sub-groups (dispensed or prescribed), which led to less meaningful comparisons, although differences followed the same pattern.

[37] Such packaging pales in comparison to the that of drugs dispensed by retail pharmacies (blister packs in small custom-printed boxes), where the name and brand of the drug appear clearly, and where a notice provides reassuring and extensive information on the drug.

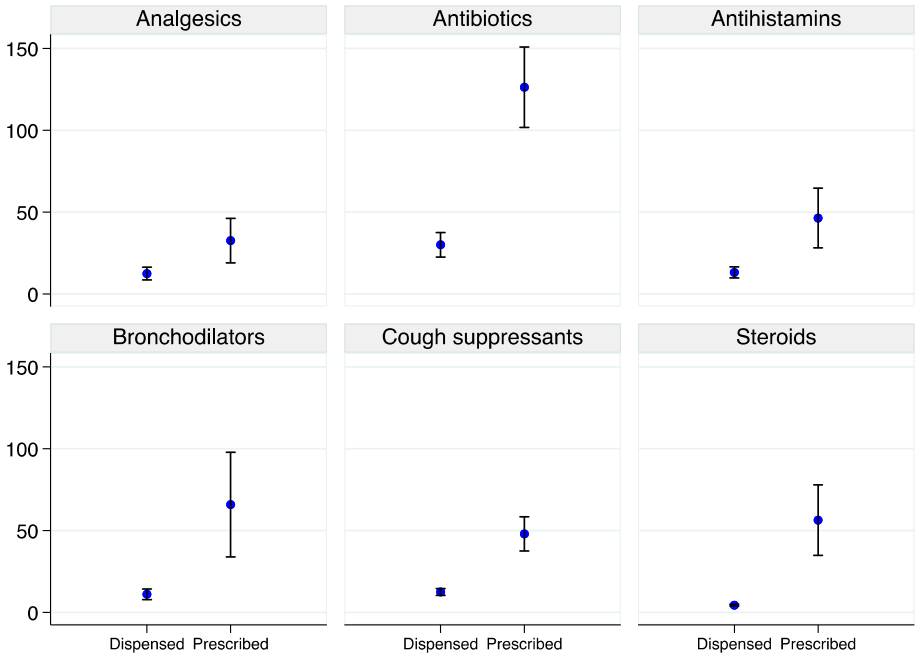

Fig. 3. Average value of the six most popular drug categories, by delivery mode.
Note: this graph shows, for the six categories of drugs most frequently recommended to patients, the average value of drugs (and associated 95% confidence intervals) depending on whether the drugs were directly dispensed by doctors, or whether the doctor wrote a prescription for the drugs. The drug value is calculated based on the standard exit price.

Table 4
Blinder-Oaxaca decomposition of the difference in treatment costs.

	(1)	(2)	(3)	(4)
	Overall gap	Gap in characteristics	Gap in coefficients	Interaction
Panel A: All drugs				
Mean gap in (log) price (prescribed – dispensed)	1.402*** (0.098)	0.718*** (0.098)	0.518*** (0.098)	0.166 (0.111)
Proportion of overall gap	–	51.2%	36.9%	11.8%
Observations	766	766	766	766
Panel B: Tablets and pills only				
Mean gap in (log) price (prescribed – dispensed)	1.503*** (0.132)	1.137*** (0.129)	0.247** (0.122)	0.118 (0.130)
Proportion of overall gap	–	75.6%	16.4%	7.8%
Observations	527	527	527	527

Notes: The table show the overall results of a three-way Blinder-Oaxaca decomposition of the value of drugs received by patients, depending on whether the treatment was dispensed or prescribed. In Panel A the analysis focuses on six categories of drugs most frequently recommended to patients (analgesics, antibiotics, antihistamines, bronchodilators, cough suppressants and steroids). In Panel B, the sample is restricted to drugs provided in the form of tablets or pills, which can be purchased in bulk by dispensing doctors. In column 1, the estimate shows the difference in the (log) cost of drugs prescribed and drugs dispensed. In column 2, the estimate shows the difference in the (log) cost of drugs prescribed and drugs dispensed which is driven by difference in drug characteristics. Those include: whether the drug is generic or not; the geographical origin of the manufacturer (high-income country, South Africa, or other low- and middle-income country); details of the drug compound; the relative strength of the main active ingredient; whether the drug requires a prescription to be purchased. In column 3, the estimate shows the difference in the (log) cost of drugs prescribed and drugs dispensed which is driven by differences, in the two groups, in the relationship between price and drug characteristics. In column 4, the estimate shows the interaction effect between the differences due to characteristics and coefficients. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

2012), educating or reassuring patients that such concerns are unfounded would require time and effort (Patel et al., 2010). If doctors have nothing to gain from exerting such effort, they are more likely to prescribe the drugs that they believe will meet patients' expectations. In other words, if the patient request was interpreted as a demand for higher-quality drug, part of the induced effect would be driven by provider moral hazard.

7.3. The indirect effects of market segmentation

Another possible interpretation of the decision made by dispensing doctors to prescribe more expensive drugs is that it was either a strategic or passive response to the market structure in which they operate. Recall that the private market for primary care services is segmented between dispensing doctors, who offer an affordable consultation-and-treatment bundle, and prescribing doctors, who serve a more affluent clientele paying higher consultation rates and purchasing their treatment separately in retail pharmacies. In this context, dispensing doctors could strategically recommend more expensive treatment to signal to patients that being "*elite patient*" who purchases drugs from a pharmacy means paying for an expensive treatment and facing higher medical expenses (Patel et al., 2012). Such behaviour would aim to discourage patients to choose the upmarket segment in the future, and ensure that they continue to choose the affordable offer provided by dispensing doctors.

Alternatively, doctors may simply adapt to the fact that pharmacies are likely to carry more expensive drugs. Although identifying the type of drugs stocked by community pharmacies was beyond the scope of this study, anecdotal evidence supports the notion that "*pharmacies rarely stock the lowest-priced medicines*" (Mpanza et al., 2019). In a small survey of pharmacies we conducted, very few said that they would buy and repackage drugs in bulk,[38] a clear difference with the procurement decisions of dispensing doctors. Hence, when doctors prescribe drugs to be

[38] We called 20 pharmacies randomly chosen in the less affluent areas of the city, to maximise the likelihood that pharmacies might stock cheaper drugs to serve their clientele. The pharmacists confirmed that, unless there was a specific reason (patient demand or lack of availability of alternative drugs), they would prefer to procure pre-packaged drugs and avoid the hassle to have to repackage drugs bought in bulk.

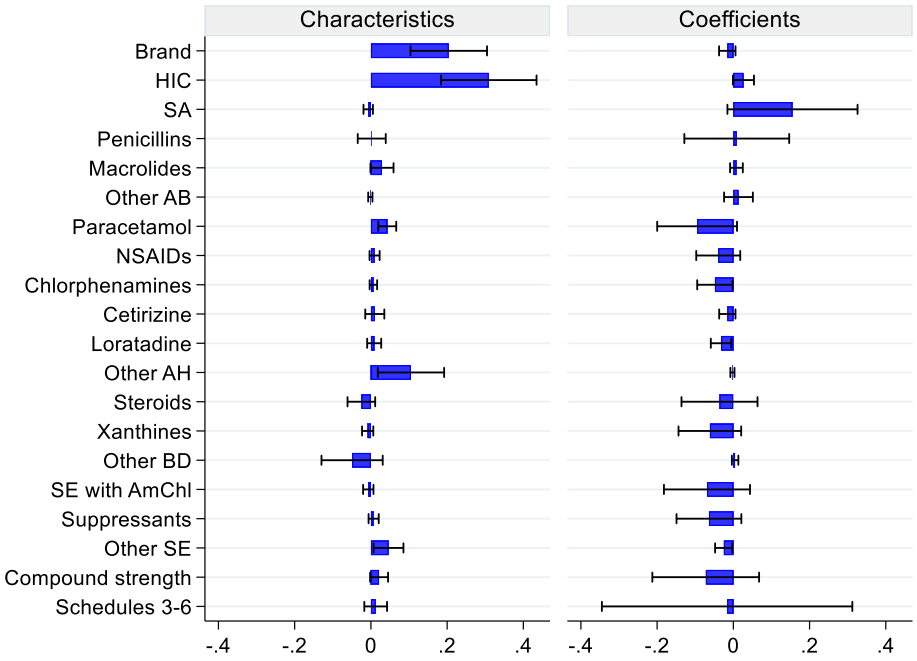

Fig. 4. Decomposition of the price difference between prescribed and dispensed drugs.
Note: the figure reports the coefficients (and 95% confidence intervals) of the Blinder-Oaxaca decomposition of the difference in (log) price between prescribed and dispensed drugs for the six categories of drugs most frequently recommended to patients (analgesics, antibiotics, antihistamines, bronchodilators, cough suppressants and steroids). The full results of the regression analysis can be found in Appendix Table 17. Abbreviations: HIC=High-Income country manufacturer; SA=South African manufacturer; AB = antibiotics; NSAID= Non-steroidal anti-inflammatory drugs; AH = antihistamines; BD = bronchodilators; SE=Suppressants and Expectorants; AmChl = Ammonium Chloride.

purchased by their patients in pharmacies, they may simply choose the drugs that are more likely to be in stock, for example to ensure that patients can access the necessary treatment without delay. Whilst this adaptation of doctors' treatment choices to pharmacies' behaviours is distinct from provider moral hazard, it reflects the incentives prevailing for those delivering treatment (pharmacies) in that higher-end of the market: on the supply-side, pharmacies have an incentive to sell more expensive drugs as their profit is increasing in drug price[39]; on the demand-side, those purchasing prescription drugs in pharmacies are not very price sensitive given that they are often partly or fully insured. Hence, overtreatment in cost when doctors prescribe is indirectly fueled by incentives born out of the separation of diagnosis and treatment provision – which requires a profit for diagnosis sellers *as well as* treatment sellers – in contrast to the model of economies of scale associated with cost-sharing prevailing in the lower-end of the market, when doctors dispense.

8. Policy simulations

To explore the financial implications of our findings for the South African health system, we estimate the waste generated by overtreatment in cost under the current situation, and under alternative situations varying the relative share of dispensing and prescribing doctors. All calculations and assumptions are detailed in Appendix A5 and summarized below.

To undertake the simulations, we first obtain an estimate for the average value of treatment for a patient seen by prescribing and dispensing doctors. For dispensing doctors, we use the average value of the treatment provided to control patients by dispensing doctors in our sample (R101.35 or about US$7). For prescribing doctors, we use the results of the ITT analysis presented in section 5.2. This analysis, comparing pharmacy and control patients, provides an estimate of the causal effect of the patient request, or from the doctor's viewpoint, being able to prescribe and lift the cost-sharing incentive faced when dispensing. Yet not every doctor decided to prescribe to the pharmacy patient. Therefore, to retrieve the causal effect of prescribing (rather than the causal effect of the patient request), we use an instrumental-variable approach with the pharmacy request as an instrument for prescribing. This analysis uses the patient request's random assignment to isolate the effect of prescribing. Intuitively, since the patient request increased the chance of obtaining the treatment through a prescription by about 41 percentage points (Appendix Table 19), and assuming the patient request affected outcomes only by changing the delivery mode of treatment (see Appendix A5 for a more detailed discussion of this exclusion restriction assumption), the effect of prescribing is simply nearly 2.5 times (i.e., 1 divided by 0.41) as high as the ITT effect. Hence we assume that the average value of treatment for a patient seeing a prescribing doctor would be about R194,[40] nearly twice as high as the value of treatment of dispensing doctors.

Next, we estimate that the cheapest available option for the recommended treatment by experts, which includes only palliative drugs (paracetamol and cough suppressant), is worth R17.86. Using this as a benchmark, the average economic waste for a consultation for a viral bronchitis is respectively R83 (≈US$5.8) with a dispensing doctor and R176 (≈US$12.2) with a prescribing doctor. For the purpose of illustration, we assume that the same level of overtreatment estimated for bronchitis would be observed in any primary care consultation. Assuming that individuals make 2.5 primary care visits per year (OECD 2017), this amounts to an annual waste of about R209 (≈US$14.4) and R440 (≈US$30.5), for patients of a dispensing and prescribing doctor respectively.[41]

[39] This is an indirect consequence of the separation of advice and treatment provision in that market segment: as pharmacies' revenues almost entirely depend the regulated mark-up added to drug prices (SEP), they are allowed to charge dispensing fees that are much higher than those of dispensing doctors.

[40] This is the average of dispensing doctors (R101.35) plus the increase due to the prescribing effect (R37.689 × (1/0.41)).

[41] Even though we extrapolate from one particular condition, we believe that the economic waste observed for bronchitis may in fact be a lower bound estimate, for three reasons. First, all unnecessary drugs recommended for bronchitis are both common, with multiple generics, and inexpensive. This limits the scope for expensive overtreatment, which might not be the case for other conditions where unnecessary drugs will be more expensive. Second, the clinical case is simple enough that no doctor recommended any unnecessary investigation or tests, eliminating a common source of overtreatment. Third, survey data suggest that private sector users, concentrated in urban areas, have on average much more than 2.5 visits per year (McIntyre 2010).

Next, we extrapolate this waste at the national level, using the latest 2019 official population estimates (Stats SA 2019a,b), considering two scenarios of expansion of access to the private sector under a future national health insurance, and two alternative compositions of the medical workforce.[42] Detailed results are presented in Appendix Table 22. Under the current level of use of the private sector for primary care and the 40%–60% split between dispensing and prescribing doctors, we estimate that about R5.3billion (≈US$367 million) or 2.20% of the national public sector health budget or 1.27% of total national health expenditure is wasted annually in the private primary care market alone.

As the government of South Africa is considering the introduction of a single-payer system that would expand access to private care by contracting private providers (Republic of South Africa 2019), it is important to consider the financial implications of such waste under alternative scenarios.[43] If 50 percent of individuals sought primary care from private doctors, R10.2billion (≈US$796million), or 4.22% of the national budget, would be wasted annually. This figure goes up to R13.2 billion (≈US$1035 million) if 65% of the population was covered by the scheme. Our simulation also shows that contracting with a workforce composed of 80% (100%) of dispensing doctors would reduce this waste by 27% (40%). This provides further evidence of the importance of designing appropriate contractual arrangements with private providers, that not only avoid perverse incentives to encourage overtreat but also embed incentives to limit overtreatment.

9. Robustness checks

In this section, we test the robustness of the results to alternative specifications, explore issues of generalizability of our findings and the null effects of insurance.

9.1. Generalizability of results

A concern for generalizability of our results is whether the doctors who responded to our encouragement strategy have specific characteristics. For example, doctors more motivated by profit could be more inclined to agree to the patient's request, since the consultation profit is higher when prescribing. If these doctors are also more likely to adopt cost-saving strategies to maximise their profit, the effect size observed would overestimate the average difference between dispensing and prescribing. By contrast, if doctors who respond to our encouragement are more likely to recommend expensive treatments to control patients, we would underestimate the true effect. To the first concern, we explore the correlates of the decision to prescribe to pharmacy patients, including a range of socio-demographic characteristics, as well as two measures of altruism.[44] We find no evidence that doctors who prescribed were different from those who dispensed to pharmacy patients (Appendix Table 8). Next, we test whether doctors who prescribed to the pharmacy patients were more likely to recommend more expensive treatments to control patients. We find no evidence that this is the case (Appendix Table 9). Although it is not possible to completely rule out that some unobserved traits may play a role in the decision to respond to the patient's request, these findings provide reassuring evidence that it was probably not the case.

9.2. Alternative specifications

We test the robustness of our findings to alternative econometric specifications. Data with repeated measures for the same subjects can be analyzed in different ways to account for the correlation structure of the multiple observations (consultation) for the same subject (doctor). Alternative approaches to the fixed-effect model include an OLS regression clustering errors at the physician level; a marginal model estimated through the generalized estimating equations (GEE) method, and a mixed effects model where individual physicians are treated as random rather than fixed effects. We use these three alternative approaches to check the robustness of our main results. The results shown in Appendix Tables 10–11 are nearly identical in direction and size to the main effects presented in Tables 2 and 3

9.3. No effect of insurance?

If the observed effects in an experiment are small relative to the "noise" in the data, there is a risk of falsely concluding that there are no treatment effects, while in fact there might be limited power to detect non-trivial effects. In our study, this is especially a concern for the null effect of insurance status on treatment costs, where some of the coefficients are large, suggesting that there was an effect but that it could not be detected because of a combination of large variation in outcomes and small sample size. To explore the power of our experiment to detect changes in the main outcomes and help calibrate some of our null findings, we follow Andrews (1989) and calculate inverse power functions in Appendix A5.

The results suggest that one should be cautious in concluding that insurance had limited effect on treatment costs. Specifically, the evidence obtained is only strong enough to rule out alternative effects as large as a *doubling* of the value of drugs and an increase by 37% of the total cost of the episode of care. Hence the evidence is not able to distinguish the null hypothesis of zero insurance effect from alternative hypotheses of quite large, economically meaningful insurance effects.

10. Discussion and conclusion

It is often assumed that overtreatment is driven by patients when they do not face the full cost of care (moral hazard) or by providers when they benefit financially from over-treating (supplier-induced demand). Yet, our findings show that overtreatment can be fueled by providers not because they profit from such choices, but because they do not bear the financial risks associated with them. In the context of the clinical case studied, this situation of benevolent provider moral hazard does not lead to excess drug supply but it increases inefficiency through excessive treatment costs. These findings occur in the context of a segmented market for private primary care, where doctors operating under a classic separation of diagnosis and treatment, while others run an unusual model combining the provision of diagnosis and treatment at a fixed price. While benevolent moral hazard occurs under the separation of diagnosis and treatment, we show that economies of scale associated with cost-sharing encourage providers to internalize the cost of treatment and lead to large efficiency gains.

The external validity of the results is a limitation of the study, since it focuses on a specific segment of providers (dispensing doctors), who agreed to take part in study, and saw a unique patient case (a viral respiratory infection). The focus on dispensing doctors was necessary to carry out the study, since it was the specificity of these doctors' remuneration that created the opportunity to test its effects on treatment choices. Although these doctors represent about 40 percent of all primary care providers, they have distinct characteristics. According to

[42] Throughout this simulation, we assume that the contractual arrangements and oversight mechanisms under which doctors operate now remain the same with an expansion of a national health insurance. This is unlikely. For example, new contracts might create a requirement for generic prescribing.

[43] We ignore savings from unnecessary treatment no longer given in the public sector because of the shift from patients from public to private providers. However, the value of unnecessary treatment in the public sector is negligible compared to that of the private sector because public providers only deliver generic drugs.

[44] Our measures of altruism include (1) effort exerted in the consultation and (2) whether the doctor accepts patients with a low-cost insurance, which requires agreeing to the lowest reimbursement rates in the market.

data collected as part of a related study, dispensing doctors are older, located in less affluent areas, accept poorer patients, are more altruistic, and charge lower rates. These observable characteristics suggest that these dispensing doctors are likely to be less sensitive to profit concerns. Hence the effect of a cost-sharing incentive in this group probably underestimates the impact it could have for more profit-oriented prescribing doctors.

To mitigate concerns of self-selection into the study given that we had to obtain doctors consent, we provided minimal information at enrolment, only mentioning that the study would look at the determinants of "clinical decision making" in primary care, with no mention of the outcomes of interest. Still, it is possible that those who agreed to participate were more competent and altruistic doctors – a hypothesis we cannot rule out for lack of data.

Regarding concerns about generalizing from a single clinical case, it is useful to remember that RTIs are the most common reason for primary care consultations in South Africa, and in many other LMICs. Therefore, they represent a large volume of all primary care consultations. There is also little reason to believe that doctors would make inefficient treatment decisions for these illnesses but not for others. If anything, the treatment for RTIs is simple and does not require any costly investigation or long courses of expensive drugs, hence the scope for overtreatment in cost is limited. This suggests that the estimates built on this case are conservative, and that the financial implications of provider moral hazard in primary care are likely much larger than those calculated.

At a time when the government of South Africa is considering how to contract private providers to expand access to primary care services to the population, this study provides both concerning and encouraging results. As mentioned, evidence of widespread and costly overtreatment of patients is concerning for the financial sustainability of a national health insurance reform. Worryingly, the rates of unnecessary drugs recommended are extremely high not only when doctors have no private incentive to limit overtreatment, but also when they face a cost-sharing incentive to reduce unnecessary treatment. Blaming insufficient knowledge for overtreatment may be tempting but it probably does not explain everything. For example, the high rates of antibiotics recommended do not seem explained simply by misdiagnosis (84 percent of doctors identified a virus as the cause of the clinical case) or inadequate beliefs (only 22 percent believed the patient would probably recover more quickly with antibiotics). Instead, there is suggestive evidence that competitive pressures combined with wrong beliefs about the patients' preferences (57 percent of doctors thought the patient described in a hypothetical vignette of the case would probably not come back if they did not receive an antibiotic) and treatment norms (61 percent said other doctors would give the patient some antibiotics)[45] might driving the high rates of overtreatment observed in the data. Overall, our findings point to the need to study further the role and interaction of market and behavioural drivers of overtreatment in health.

On a more encouraging note, our results suggest that bundled payments for consultation and drugs that shift some of the financial risk of treatment decisions onto primary care providers could limit waste.[46] A potential pitfall of these cost-sharing arrangements is that, in the absence of monitoring, they rely on professionalism or providers' altruistic concerns for patient welfare not to skimp on quality and underprovide efficient treatment. In the case of the self-limiting viral infection studied here, it was not a real concern, and we found no evidence that doctors reduced palliative treatment. Besides, the existing cost-sharing arrangement for dispensing doctors in South Africa limits this risk by allowing some flexibility for doctors, who can release themselves from the financial constraint if necessary. However, our results would benefit from being replicated with other clinical cases where the risk of quality skimping is greater.

Our study suggests that plans for expanding coverage should include discussions to introduce incentives to reduce the large levels of inefficiency observed. Although this paper focused on private providers, a companion study found comparable levels of inefficiency in treatment decisions in the public sector (Blaauw and Lagarde 2022), where providers are equally isolated from the financial consequences of their decisions. While much attention has been focused on increasing access to more services to a greater number in low- and middle-income countries, this study cautions against the financial implications of such reforms when providers have no incentive to limit cost and overtreatment. At the same time, it shows how supply-side cost-sharing measures can reduce waste, although not through a reduction of overtreatment in volume. Future studies could draw on more complex clinical cases to explore the potential trade-offs between under- and overtreatment created by cost-sharing measures in primary care, and the reasons why financial incentives are not more able to reduce the rate of unnecessary drugs.

Credit author statement

Mylene Lagarde: Conceptualization, Methodology, Data curation, Formal analysis, Visualization, Writing – original draft preparation, Writing - Reviewing and Editing, Project administration, Funding acquisition. **Duane Blaauw:** Methodology, Data curation, Writing-Reviewing and Editing, Project administration, Funding acquisition.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jdeveco.2022.102917.

References

Andrews, D.W.K., 1989. Power in Econometric Applications. Econometrica 57 (5), 1059–1090. https://doi.org/10.2307/1913623.
Ataguba, J.E., McIntyre, D., 2012. Modelling the affordability and distributional implications of future health care financing options in South Africa. Health Pol. Plann. 27 (Suppl. 1_1), i101–i112.
Ataguba, J.E., McIntyre, D., 2018. The incidence of health financing in South Africa: findings from a recent data set. Health Econ. Pol. Law 13 (1), 68–91.
Balafoutas, L., Beck, A., Kerschbamer, R., Sutter, M., 2013. What drives taxi drivers? A field experiment on fraud in a market for credence goods. Rev. Econ. Stud. 80 (3), 876–891.
Balafoutas, L., Kerschbamer, R., Sutter, M., 2017. Second-degree moral hazard in A real-world credence goods market. Econ. J. 127 (599), 1–18.
Balafoutas, L., Fornwagner, H., Kerschbamer, R., Sutter, M., Tverdostup, M., 2020. Diagnostic Uncertainty and Insurance Coverage in Credence Goods Markets. IZA DP No. 13848.
Barnett, M.L., Linder, J.A., 2014. Antibiotic prescribing for adults with acute bronchitis in the United States, 1996–2010. JAMA 311 (19), 2020.
Bauhoff, S., 2012. Do health plans risk-select? An audit study on Germany's Social Health Insurance. J. Publ. Econ. 96 (9–10), 750–759.
Bertrand, M., Duflo, E., 2017. Field experiments on discrimination. Handbook of field experiments. E. Duflo and A. Banerjee.
Bertrand, M., Mullainathan, S., 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. Am. Econ. Rev. 94 (4), 991–1013.
Blaauw, D., Lagarde, M., 2022. Better, but at what Cost? an Audit Study of the Performance of Private and Public Healthcare Providers in South Africa. London School of Economics.
Blinder, A., 1973. Wage discrimination: reduced form and structural estimates. J. Hum. Resour. 8 (4), 436–455.
Brink, A.J., Van Wyk, J., Moodley, V.M., Corcoran, C., Ekermans, P., Nutt, L., Boyles, T., Perovic, O., Feldman, C., Richards, G., Mendelson, M., 2016. The role of appropriate

[45] In the companion study to this paper, we observed evidence of treatment patterns emerging between public and private sector providers: while we observe high rates of unnecessary prescribing of steroids by private sector doctors, SPs visiting public sector clinics never received steroids.

[46] Alternatively, and for doctors who do not dispense drugs, an equivalent solution would be to regulate doctors' choice of drugs by restricting the list of drugs to be reimbursed by a national health insurance fund.

diagnostic testing in acute respiratory tract infections: an antibiotic stewardship strategy to minimise diagnostic uncertainty in primary care. S. Afr. Med. J. 106 (6), 30–37.
Brownlee, S., Chalkidou, K., Doust, J., Elshaug, A.G., Glasziou, P., Heath, I., Nagpal, S., Saini, V., Srivastava, D., Chalmers, K., Korenstein, D., 2017. Evidence for overuse of medical services around the world. Lancet 390 (10090), 156–168.
Busse, R., 2012. Do diagnosis-related groups explain variations in hospital costs and length of stay? – analyses from the Eurodrg project for 10 episodes of care across 10 European countries. Health Econ. 21 (S2), 1–5.
Crea, G., Galizzi, M.M., Linnosmaa, I., Miraldo, M., 2019. Physician altruism and moral hazard: (no) Evidence from Finnish national prescriptions data. J. Health Econ. 65, 153–169.
Currie, J., Lin, W., Zhang, W., 2011. Patient knowledge and antibiotic abuse: evidence from an audit study in China. J. Health Econ. 30 (5), 933–949.
Currie, J., Lin, W., Meng, J., 2014. Addressing antibiotic abuse in China: an experimental audit study. J. Dev. Econ. 110, 39–51.
Daniels, B., Dolinger, A., Bedoya, G., Rogo, K., Goicoechea, A., Coarasa, J., Wafula, F., Mwaura, N., Kimeu, R., Das, J., 2017. Use of standardised patients to assess quality of healthcare in Nairobi, Kenya: a pilot, cross-sectional study with international comparisons. BMJ Glob. Health 2 (2).
Das, J., Holla, A., Das, V., Mohanan, M., Tabak, D., Chan, B., 2012. In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps. Health Aff. 31 (12), 2774–2784.
Das, J., Holla, A., Mohpal, A., Muralidharan, K., 2016. Quality and accountability in health care delivery: audit-study evidence from primary care in India. Am. Econ. Rev. 106 (12), 3765–3799.
Department of Health, 2018. Standard Treatment Guidelines and Essential Medicines List for South Africa. Primary Health Care Level, 2018 edition. Department of Health, Pretoria.
Dulleck, U., Kerschbamer, R., 2006. On doctors, mechanics, and computer specialists: the economics of credence goods. J. Econ. Lit. 44 (1), 5–42.
Ellis, R.P., McGuire, T.G., 1986. Provider behavior under prospective reimbursement. Cost sharing and supply. J. Health Econ. 5 (2), 129–151.
Ellis, R.P., McGuire, T.G., 1993. Supply-side and demand-side cost sharing in health care. J. Econ. Perspect. 7 (4), 135–151.
Emanuel, E.J., Fuchs, V.R., 2008. The perfect storm of overutilization. JAMA 299 (23), 2789–2791.
Glasziou, P., Straus, S., Brownlee, S., Trevena, L., Dans, L., Guyatt, G., Elshaug, A.G., Janett, R., Saini, V., 2017. Evidence for underuse of effective medical services around the world. Lancet 390 (10090), 169–177.
Goldacre, B., Reynolds, C., Powell-Smith, A., Walker, A.J., Yates, T.A., Croker, R., Smeeth, L., 2019. Do doctors in dispensing practices with a financial conflict of interest prescribe more expensive drugs? A cross-sectional analysis of English primary care prescribing data. BMJ Open 9 (2), 026886.
Gray, A., Suleman, F., 2015. Pharmaceutical Pricing in South Africa. Pharmaceutical Prices in the 21st Century. Springer, pp. 251–265.
Gruber, J., Owings, M., 1996. Physician financial incentives and cesarean section delivery. Rand J. Econ. 27 (1), 99–123.
Gruber, J., Kim, J., Mayzlin, D., 1999. Physician fees and procedure intensity: the case of Cesarean delivery. J. Health Econ. 18, 473–490.
Iizuka, T., 2007. Experts' agency problems: evidence from the prescription drug market in Japan. Rand J. Econ. 38 (3), 844–862.
Iizuka, T., 2012. Physician agency and adoption of generic pharmaceuticals. Am. Econ. Rev. 102 (6), 2826–2858.
Inderst, R., Ottaviani, M., 2012. Competition through commissions and kickbacks. Am. Econ. Rev. 102 (2), 780–809.
Kaiser, B., Schmid, C., 2016. Does physician dispensing increase drug expenditures? Empirical evidence from Switzerland. Health Econ. 25 (1), 71–90.
Kerschbamer, R., Sutter, M., 2017. The economics of credence goods - a survey of recent lab and field experiments. CESifo Econ. Stud. 63 (1), 1–23.
Kerschbamer, R., Neururer, D., Sutter, M., 2016. Insurance coverage of customers induces dishonesty of sellers in markets for credence goods. Proc. Natl. Acad. Sci. USA 113 (27), 7454–7458.
King, J.J.C., Powell-Jackson, T., Makungu, C., Hargreaves, J., Goodman, C., 2021. How much healthcare is wasted? A cross-sectional study of outpatient overprovision in private-for-profit and faith-based health facilities in Tanzania. Health Policy Plan 36 (5), 695–706.
Kovacs, R.J., Lagarde, M., Cairns, J., 2020. Overconfident health workers provide lower quality healthcare. J. Econ. Psychol. 76, 102213.
Kwan, A., Daniels, B., Saria, V., Satyanarayana, S., Subbaraman, R., McDowell, A., Bergkvist, S., Das, R.K., Das, V., Das, J., Pai, M., 2018. Variations in the quality of tuberculosis care in urban India: a cross-sectional, standardized patient study in two cities. PLoS Med. 15 (9), e1002653.
Kwan, A., Daniels, B., Bergkvist, S., Das, V., Pai, M., Das, J., 2019. Use of standardised patients for healthcare quality research in low- and middle-income countries. BMJ Glob. Health 4 (5), e001669.
Legotlo, T.G., Mutezo, A., 2018. Understanding the types of fraud in claims to South African medical schemes. S. Afr. Med. J. 108 (4), 299–303.
Leibowitz, A., Manning, W.G., Newhouse, J.P., 1985. The demand for prescription drugs as a function of cost-sharing. Soc. Sci. Med. 21 (10), 1063–1069.
Leonard, K.L., Masatu, M.C., 2010. Using the Hawthorne effect to examine the gap between a doctor's best possible practice and actual performance. J. Dev. Econ. 93 (2), 226–234.
Liu, T., 2011. Credence goods markets with conscientious and selfish experts. Int. Econ. Rev. 52 (1), 227–244.
Liu, Y.-M., Yang, Y.-H.K., Hsieh, C.-R., 2009. Financial incentives and physicians' prescription decisions on the choice between brand-name and generic drugs: evidence from Taiwan. J. Health Econ. 28 (2), 341–349.
Lu, F., 2014. Insurance coverage and agency problems in doctor prescriptions: evidence from a field experiment in China. J. Dev. Econ. 106, 156–167.
Lundin, D., 2000. Moral hazard in physician prescription behavior. J. Health Econ. 19 (5), 639–662.
Lurie, N., Christianson, J., Finch, M., Moscovice, I., 1994. The effects of capitation on health and functional status of the Medicaid elderly. A randomized trial. Ann. Intern. Med. 120 (6), 506–511.
McIntyre, D, 2010. Private sector involvement in funding and providing health services in South Africa: Implications for equity and access to health care. https://www.equinetafrica.org/sites/default/files/uploads/documents/DIS84privfin%20mcintyre.pdf. (Accessed 13 June 2022). EQUINET Discussion Paper 84.
Mohanan, M., Vera-Hernandez, M., Das, V., Giardili, S., Goldhaber-Fiebert, J.D., Rabin, T.L., Raj, S.S., Schwartz, J.I., Seth, A., 2015. The know-do gap in quality of health care for childhood diarrhea and pneumonia in rural India. JAMA Pediatr. 169 (4), 349–357.
Moreno-Serra, R., Wagstaff, A., 2010. System-wide impacts of hospital payment reforms: evidence from central and Eastern Europe and central Asia. J. Health Econ. 29 (4), 585–602.
Mpanza, N.M., Bradley, H., Laing, R., 2019. Reasons why insured consumers co-pay for medicines at retail pharmacies in Pretoria, South Africa. Afr. J. Prim. Health Care Fam. Med. 11 (1).
National Department of Health, 2015. White Paper on National Health Insurance. Towards Universal Health Coverage. National Department of Health, Pretoria.
Ncube, N.B., Solanki, G.C., Kredo, T., Lalloo, R., 2017. Antibiotic prescription patterns of South African general medical practitioners for treatment of acute bronchitis. S. Afr. Med. J. 107 (2), 119–122.
NDoH, Stats, S.A., SAMRC and, I.C., 2019. South Africa Demographic and Health Survey 2016. South Africa and Rockville, Maryland, USA, NDoH, Stats SA, SAMRC, and ICF, Pretoria.
Oaxaca, R., 1973. Male-female wage differentials in urban labor markets. Int. Econ. Rev. 14 (3), 693–709.
OECD, 2017. Health at a Glance 2017: OECD Indicators. OECD Publishing, Paris. https://doi.org/10.1787/health_glance-2017-en.
Institute of Medicine, 2013. Best Care at Lower Cost: the Path to Continuously Learning Health Care in America. The National Academies Press, Washington, DC.
Patel, A., Gauld, R., Norris, P., Rades, T., 2010. This body does not want free medicines": South African consumer perceptions of drug quality. Health Policy Plan 25 (1), 61–69.
Patel, A., Gauld, R., Norris, P., Rades, T., 2012. Quality of generic medicines in South Africa: perceptions versus Reality – a qualitative study. BMC Health Serv. Res. 12 (1), 1–8.
Planas, M.-E., García, P.J., Bustelo, M., Carcamo, C.P., Martinez, S., Nopo, H., Rodriguez, J., Merino, M.-F., Morrison, A., 2015. Effects of ethnic attributes on the quality of family planning services in Lima, Peru: a randomized crossover trial. PLoS One 10 (2), e0115274.
Ranchod, S., Dube, N., 2019. When Good Work Fails: Why Provider Profiling in South Africa Doesn't Effect the Change it Should. Actuarial Society of South Africa's 2019 Convention, Sandton.
Republic of South Africa, 2019. National Health Insurance Bill.
Smith, S.M., Fahey, T., Smucny, J., Becker, L.A., 2017. Antibiotics for Acute Bronchitis. Cochrane Database of Systematic Reviews(6).
Stats, S.A., 2019a. General Household Survey 2018. Satistics South Africa, Pretoria.
Stats, S.A., 2019b. Mid-Year Population Estimates 2019. Satistics South Africa, Pretoria.
Trottmann, M., Zweifel, P., Beck, K., 2012. Supply-side and demand-side cost sharing in deregulated social health insurance: which is more effective? J. Health Econ. 31 (1), 231–242.
WHO, 2014. Antimicrobial Resistance Global Report on Surveillance. Geneva, Switzerlnad.
Woodhead, M., Blasi, F., Ewig, S., Garau, J., Huchon, G., Ieven, M., Ortqvist, A., Schaberg, T., Torres, A., van der Heijden, G., Read, R., Verheij, T.J.M., , J. T. o. t. E. R. Society, Microbiology, E.S.f.C., Diseases, I., 2011. Guidelines for the management of adult lower respiratory tract infections–full version. Clin. Microbiol. Infect. 17 (Suppl. l), 6.
Zschirnt, E., 2019. Research ethics in correspondence testing: an update. Res. Ethics 15 (2), 1–21.