Quality and Accountability in Healthcare Delivery: Audit-Study Evidence from Primary Care in India

Jishnu Das, Alaka Holla, Aakash Mohpal, and Karthik Muralidharan*

January 16, 2016

Abstract

We present the first direct evidence on the relative quality of public and private healthcare in a low-income setting, using a unique set of audit studies. We sent standardized (fake) patients to rural primary care providers in the Indian state of Madhya Pradesh, and recorded the quality of care provided and prices charged in each interaction. We report three main findings. First, most private providers lacked formal medical training, but they spent more time with patients and completed more essential checklist items than public providers and were equally likely to provide a correct treatment. Second, we compare the performance of qualified public doctors across their public and private practices and find that the *same* doctors exerted higher effort and were more likely to provide a correct treatment in their private practices. Third, in the private sector, we find that prices charged are positively correlated with provider effort and correct treatment, but also with unnecessary treatments. In the public sector, we find no correlation between provider salaries and any measure of quality. We develop a simple theoretical framework to interpret our results and show that in settings with low levels of effort in the public sector, the benefits of higher diagnostic effort in the private sector may outweigh the costs of market incentives to over treat. These differences in provider effort may partly explain the dominant market share of fee-charging private providers even in the presence of a system of free public healthcare.

Keywords: Healthcare quality, healthcare markets, healthcare in low-income and developing countries, audit studies, standardized patients, India

JEL Codes: D40, H10, H42, I11, O15

---

*Das: The World Bank Group, Washington, DC, jdas1@worldbank.org. Holla: The World Bank Group, Washington, DC, aholla@worldbank.org. Mohpal: University of Michigan, Ann Arbor, MI, amohpal@umich.edu. Muralidharan: University of California, San Diego, CA, kamurali@ucsd.edu. Acknowledgements: We thank Brian Chan, Veena Das, Ranendra Kumar Das, Diana Tabak, and Manoj Mohanan for their contribution to designing and implementing the SP methodology, and thank Eun-Young Shim for her contribution to the theoretical appendix. We thank Prashant Bharadwaj, Gordon Dahl, Roger Gordon, Gordon Hanson, Michael Kremer, Paul Niehaus, and Adam Wagstaff for comments. We are grateful to Innovations for Poverty Action, CT, and Pratap Bhanu Mehta and the Center for Policy Research, New Delhi for hosting the project and providing logistical support and infrastructure. We also thank Sreela Dasgupta, L. Ravi and Anvesha Khandelwal for project management support, and Monisha Ashok, Carl Liebersohn, Prerna Mukharya, Suzanne Plant and Anand Shukla for excellent research assistance. The project would not have been possible without the dedication of our SPs and Purshottam, Rajan Singh, Devender, Charu Nanda, Simi Bajaj, Geeta and other staff at Institute of Socioeconomic Research on Democracy and Development (ISERDD) in Delhi. The Social and Rural Research Institute (SRI), New Delhi oversaw data collection of all other aspects other than SPs. This study was funded by the Global Health Program of the Bill and Melinda Gates Foundation through Grant No. 50728 and the Knowledge for Change Program at the World Bank. The findings, interpretations and conclusions expressed in this article are those of the authors and do not necessarily represent the views of the World Bank, its executive directors, or the governments they represent.

1 Introduction

Healthcare is a credence good with substantial information asymmetries between patients and providers. This makes it difficult for patients to determine the quality of care they have received (Dulleck and Kerschbamer, 2006). It is widely believed therefore that unregulated market-based delivery of healthcare is socially undesirable; Arrow (1963), for instance, notes that *"it is the general social consensus, clearly, that the laissez-faire solution for medicine is intolerable."* Further, if optimal care requires the potential denial of services that patients value (such as steroids or antibiotics), market-based healthcare may over-respond to demand, leading to socially inefficient provision (Prendergast, 2003). Partly as a result of these considerations, the default policy approach to delivering healthcare for the poor in most low-income countries is through free or nominally priced medical care in publicly-run facilities staffed by qualified doctors and nurses, who are paid a fixed salary (World Bank, 2003).

However, for primary care services a significant fraction of households in low-income countries choose to visit fee-charging healthcare providers in the private sector; in rural India (the focus of our study), their market share exceeds 70 percent.[1] This is surprising for two reasons. First, private healthcare providers in India face little de facto regulation and most have no formal medical training (Rohde and Viswanathan, 1995; Banerjee, Deaton and Duflo, 2004; CPR, 2011). Second, while the high use of the private sector could, in part, reflect the absence of public options, this cannot be the only explanation. In our data from rural India, the private sector share of primary care visits (constructed from a household census) is 83 percent even in markets with a qualified public doctor offering free care through public clinics, with 60 percent of visits to private providers with no formal qualifications.

The high market share of unqualified private healthcare providers raises a number of questions about the functioning of healthcare markets in low-income settings. First, why would people choose to pay for care from (mostly) unqualified providers when public clinics are staffed with qualified doctors who offer care at a much lower price? Second, how does the quality of care received vary across public and private healthcare providers? Third, what does an unregulated healthcare market reward and how does this compare with the regulated public sector? Specifically, to what extent are prices in the market and wages in the public sector correlated with quality of care? Answers to these questions have been limited by the lack of evidence on the actual quality of care provided in public and private health facilities

[1] The market share of private providers is high in many low-income countries: Data from the DHS show that 50 percent of households seeking pediatric outpatient care in Africa and 70-80 percent in India visit the private sector with little variation over the 20 years that these surveys have been collected (DHS, 2007; Grepin, 2014). The World Health Surveys include adult morbidity and here the numbers vary from 30 percent in Sub-Saharan Africa to between 70 and 80 percent in India (Wagstaff, 2013).

in low-income settings.[2]

This paper uses data from an audit study conducted in rural areas of the Indian state of Madhya Pradesh (MP) to address this gap. Specifically, standardized (fake) patients (SPs) were coached to accurately present symptoms for three different conditions - unstable angina, asthma, and dysentery in a child (who is at home) - to multiple healthcare providers. SPs then made over 1,100 unannounced visits to public and private providers of primary healthcare services and recorded condition-specific metrics of quality of care for each interaction, as well as the price charged.[3] The quality of care metrics include the providers' adherence to a checklist of questions and examinations deemed essential for reaching a correct diagnosis in each case, their likelihood of pronouncing a correct diagnosis, and the appropriateness of the treatments.

We present results from two sets of comparisons. First, we sent SPs to a (nearly) representative sample of public and private health facilities on a walk-in basis, and we use these data to compare the typical patient experience across public and private clinics. However, these differences reflect variation in both provider composition, and differential incentives across public and private clinics. To isolate the effect of practicing in the private sector holding provider characteristics constant, we identified the private practices of qualified public doctors (the majority of whom have one) and sent SPs to present the same medical case to the same set of doctors in both their public and private practices. Our second comparison uses this "dual practice sample" and compares the quality of care across the public and private practices of the same doctors on the same set of cases.

We report three main findings. First, while the majority of private providers in the representative sample have no medical qualifications, they exerted significantly higher effort than public providers and performed no worse on diagnosis and treatment. Private providers spent 1.5 minutes more with patients (62 percent more) and completed 7.4 percentage point more items on a checklist of essential history and examination items (47 percent more) than public providers. They were equally likely to pronounce a correct diagnosis (only 4 percent of public providers do so), to offer a correct treatment (27 percent of public providers do so), and to offer clinically unnecessary treatments (provided by 70 percent of public providers).

[2]Earlier work has highlighted the problem of low doctor effort in the public sector (high absence, low time spent with patients) and low training in the private sector (Banerjee, Deaton and Duflo (2004); Chaudhury et al. (2006); Das and Hammer (2007)). The key evidence gap, however, is the lack of credible estimates of the actual quality of care provided in the public and private sector. For instance, Coarasa, Das and Gummerson (2014) examine 182 cited studies in two systematic reviews of the medical literature and find only one study that adjusts for differences in patients using an audit methodology (as we do here), and no study that adjusts for differences in providers across public and private practices (which we also do here).

[3]Typically used in medical education, SPs are coached to consistently portray a medical case and all of its physical and pyscho-social aspects. When used to evaluate care in hospitals and clinics, they are also trained to accurately recall all aspects of their interactions with the provider. See details in section 3.

These differences do not reflect high patient loads and waiting times in the public sector; neither do they reflect inadequate equipment and facilities. The results hold even after controlling for these factors and after including market fixed effects.

Second, in the dual practice sample the same doctors spent more time with SPs, completed more items on the checklist, and were also more likely to offer a correct treatment in their private practices, relative to their public practices. Notably, we do not find evidence of differential over-treatment under market incentives, with equivalently high rates of unnecessary treatments, use of antibiotics, and total number of medicines in both types of practices. These differences are *conditional* on seeing the doctor and therefore understate the difference in the quality of patient experiences across public and private practices of the same doctor, because the expected number of trips to the clinic to see a qualified doctor is considerably higher in the public practice due to high absence rates.

Third, we find a positive correlation between the fees charged by private providers and measures of quality such as the time spent, the fraction of checklist items completed, and likelihood of providing a correct treatment. However, we also find a positive correlation between prices and the total number of medications given - including unnecessary treatment. In the public clinics, SPs were provided free or nominally priced care. Since there is no variation in prices, we examine the correlation between doctors' compensation and quality of care and find no correlation between salaries (or desirability of posting) in the public sector and any measure of quality of care delivered.

The main limitation of the SP method is that only a few types of cases can be presented. We therefore complement the SP results with direct observations where enumerators recorded observable details of provider-patient interactions from a full day of within-clinic observations of each provider in our sample (as in Das and Hammer (2007)) after SP data collection was completed. We find very similar differences between public and private providers on common measures of quality of care across *all* cases and patient interactions, suggesting that our SP results are likely to be externally valid across a much broader range of cases in this setting.

To help interpret our results, we develop a theoretical framework that models provider-patient interactions in two stages: consultation and treatment. The main insight of the model is that while providers will typically exert more effort in their private practice, the effect on overall patient health is ambiguous. If the default effort level of doctors under low-powered incentives is reasonably high, the marginal gain in diagnostic precision from additional effort in private practice is outweighed by the costs of over-treatment induced by market incentives. On the other hand, if the default effort level is low (as in our setting), the reverse may be true, with better patient health outcomes under market incentives.

Our methodological contribution helps address the fundamental problem of inferring quality in healthcare, where the optimal action is patient and condition specific, and inefficiencies include under-treatment, over-treatment, or both (Pauly, 1980). Specifically, there are four advantages to the use of unannounced SPs relative to existing measures in the literature, which are based on tests of provider knowledge or observation of medical practices.

First, the use of SPs ensures a common set of patient and illness characteristics, which limits concerns about differential patient sorting across clinics on the basis of personal or illness characteristics, as might be the case when observing real patient-provider interactions. Second, the SP method allows us to objectively score the quality of care using condition-specific metrics (checklist completion, diagnosis, and treatment) because we *know* the actual illness being presented and the optimal care associated with the case. In the case of observations with real patients, we would observe only the presenting symptoms and would have to speculate about the true underlying illness.[4] Third, we are able to observe prices charged for completed transactions, which allows us to study the extent to which the unregulated market rewards quality and which improves upon audit studies in other settings that obtain price quotes but do not complete the purchase.[5] Finally, Hawthorne effects are not a concern in the SP context because providers do not know that they are being observed.

Substantively, the advances in measurement above combined with our ability to observe the *same* doctor across public and private practices allow us to provide the first direct comparison of the quality of care across public and private sectors.[6] We also provide the first evidence on how market prices for healthcare behave in an unregulated setting and show that there is a positive correlation between price and checklist completion (and correct treatment), but also between price and unnecessary treatments. This suggests that while unregulated market prices do reflect some information on the quality of care, patients cannot evaluate whether they are being over-treated and charged for unnecessary treatments.

These findings are consistent with the broader empirical literature on credence goods that has demonstrated over-provision of services to the detriment of customer welfare in settings ranging from caesarian sections to car repairs and cab rides for tourists (Gruber and

[4]Medical vignettes, which measure provider knowledge, also allow for standardization of case-mix and knowledge of the actual illness underlying the presented symptoms, but do not measure actual provider practice, which has been shown to differ markedly from provider knowledge in multiple contexts (Rethans et al., 1991; Leonard and Masatu, 2005; Das and Hammer, 2007).

[5]For instance, first price offers can be very different from the price of the completed transaction if the distribution of willingness to pay is different across populations. See for instance, Ayres and Siegelman (1995) and Goldberg (1996) for an example of how the lack of completed sales data can lead to misleading conclusions in audit studies of car sales. In our case, the "sale" is always completed as the SP leaves only after the provider has completed the interaction and the price has been paid.

[6]Our approach parallels a literature that tests for moral hazard in developing-country labor markets by comparing worker effort and output under different contractual arrangements (Shaban, 1987; Foster and Rosenzweig, 1994), and extends it to a credence good setting where output is harder to measure for both customers and researchers, and where there is substantial direct provision of the good by the public sector.

Owings, 1996; Schneider, 2012; Balafoutas et al., 2013). However, inefficiencies in market provision do not imply that public provision will do better, and a key contribution of our paper is the ability to compare public and private provision of a canonical credence good such as healthcare.

Combined with the theoretical framework, our results suggest that in settings of poor governance and administrative accountability in the delivery of primary healthcare services through the public sector (Banerjee, Deaton and Duflo, 2004; Banerjee, Duflo and Glennerster, 2008), market-based provision of healthcare may present a legitimate alternative in spite of its many theoretical (and empirical) weaknesses. Further, while public healthcare is free to the consumer, it is not free to the taxpayer. We calculate the per-patient cost in the public sector and conservatively estimate it to be four times higher than the fees charged by private providers in our sample. Thus, the unregulated private market for healthcare, which is mainly staffed by unqualified providers, appears to deliver higher provider effort and comparable quality of care, at a much lower cost per patient. Our results have direct implications for global policy debates on the organization and delivery of healthcare services in low-income countries with low state capacity to deliver effective oversight over public healthcare systems. We discuss these along with caveats in the conclusion.

The rest of this paper is organized as follows. Section 2 describes healthcare provision in rural India and Madhya Pradesh; section 3 describes the standardized patient (SP) methodology, sampling, data, and measures of healthcare quality; section 4 presents results on quality of care; section 5 covers pricing and cost-effectiveness; section 6 discusses robustness to alternative explanations; section 7 presents a theoretical framework to interpret our results; and section 8 concludes with a discussion of policy implications and caveats.

2 Context

2.1 Healthcare in Rural India

Healthcare in India is delivered by both public and private clinics and hospitals. In the public sector, patients can obtain primary care on a walk-in basis in facilities differentiated by their level of specialization ranging from district hospitals and Community Health Centers (CHCs) to Public Health Centers (PHCs) and sub-centers.[7] PHCs, CHCs, and hospitals are supposed to be staffed with trained doctors, who are expected to make diagnoses and either treat or refer patients as appropriate (although in practice, doctor positions are often vacant). Sub-centers are supposed to be staffed with qualified nurses with doctors visiting on a fixed

[7]Official guidelines stipulate that there should be a sub-center for every 5,000 people, a primary healthcare center for every 25,000 people, and a community health center for every 100,000 people.

rotation. Most doctors hold a Bachelor of Medicine and Bachelor of Surgery (MBBS) degree, the rough equivalent of an MD in the US, and receive a fixed salary from the government, with no variable compensation based on either patient load or quality of care.[8]

Consultations in public clinics are provided on a walk-in basis during opening hours (appointments are rarely used), and are free or nominally priced. Patients are also supposed to receive free medication, if available. Although a federally-funded insurance program for inpatient hospital care was introduced in 2007, the tax-funded public system of care was the only source of (implicit) public insurance for primary care.

In theory, public facilities are accountable to administrative norms and procedures (documented in the *Civil Service Codes* for each state). In practice, administrative accountability of public health-care providers is weak. Nationwide, doctor absences in public clinics averaged 43 percent on any given day in 2003 and 40 percent in 2010 (Muralidharan et al., 2011; CPR, 2011). These absences do not occur on predictable days or hours (Banerjee, Deaton and Duflo, 2004) and they are not easy to address at a system-level (Banerjee, Duflo and Glennerster, 2008; Hanna and Dhaliwal, 2015). When asked about adherence to administrative rules, more than 80 percent of public sector doctors agree that the rules and norms are frequently flouted and that appropriate 'payments' can allow providers to circumvent disciplinary proceedings, even for grave negligence (La Forgia and Nagpal, 2014).

While official policy documents of the Government mainly focus on improving the public system of primary healthcare (Planning Commission of India, 2013), data from household surveys consistently show that the fee-charging private sector accounts for over 70 percent of primary care visits (DHS, 2007; Selvaraj and Karan, 2009; CPR, 2011). Barriers to entry for private healthcare providers are low. Provider qualifications range from MBBS degrees to no medical training at all, and clinics can range from well-equipped structures to small one-room shops, the provider's residence, or the patients' home for providers that make home visits. Providers operate on a fee-for-service basis, and prices often include the cost of medicines. While providers operating without a medical license are not legal and face the threat of being shut down, they have come to be the dominant source of care in these markets (as the data below will show).

[8]India also recognizes medical degrees from alternative schools of medicine including Ayurveda, Homeopathy, and Unani. However, providers with these qualifications are only licensed to prescribe medication in line with their training and are not licensed to prescribe allopathic medicine. They also are not typically posted in the frontline healthcare system of PHCs, CHCs, and district hospitals that prescribe allopathic medicine.

2.2 Sampling of Healthcare Markets, and Summary Statistics

We carried out the SP study in the Indian state of Madhya Pradesh (MP), one of India's poorer states, with a GDP/capita of ∼\$600/year (or ∼\$1500/year in PPP terms) in 2010-11 (the period of the study). We first drew a representative sample of 100 villages across 5 districts, stratified by geographic regions and an index of health outcomes. We then conducted a household *census* in these villages, where respondents named all providers from whom they sought primary care in the previous thirty days and their locations (including providers practicing outside the village). We then surveyed all providers in all of these locations, regardless of whether or not the providers themselves had been mentioned in the sample villages, thereby obtaining a census of all providers in the healthcare market that catered to sampled villages (see Figure A.1).

Table 1 (columns 1-3) presents summary statistics based on the provider census (Panel A) and the household census (Panel B) in these markets; columns 4-6 compare villages sampled for the SP study to the representative villages. The table highlights three key features of health markets in rural India. First, villages are served by a large number of providers once the health market is correctly accounted for by including locations that are nearby but outside village boundaries. There are 11 primary care providers per market and 46 percent of households reported visiting a primary care provider in the 30 days prior to the survey.

Second, the majority of providers are private (7 out of 11 or 64 percent), and they account for 89 percent of household visits; excluding paramedical public health workers (typically responsible for preventive, maternity and child care) increases the fraction further to 93 percent. The share of visits to private providers (with or without qualifications) is 88 percent when there is a public provider in the market, and is 83 percent even when there is a public MBBS doctor in the same market.

Third, 46 percent of all providers and 70 percent of all private providers (5.4 per village) have no formal medical training, yet they account for 77 percent of household visits. There is less than one MBBS doctor per market, and one is rarely available within the village. The distribution of MBBS providers is uneven. Only 30 percent of all villages have recourse to an MBBS provider (public or private) in their market, and only 5 percent have one within village boundaries. Private unqualified providers remain the dominant providers of care in most settings, accounting for 74 percent of all visits even when there is a public provider in the same market, and 60 percent even when there is a public MBBS doctor in the same market.[9] MBBS doctors account for only 4 percent of all patient interactions (Panel B).

[9]Note that even public facilities have many unqualified providers. While these are typically support staff (who are only supposed to assist a qualified doctor), we find that it is very common for these staff to act as the main healthcare providers in public clinics and prescribe medication (given high doctor absence rates).

3 Measuring Healthcare Quality Using Standardized Patients

3.1 The Standardized Patient (SP) Methodology

Used routinely in the training and evaluation of medical students in high-income countries, including the United States, SPs are highly-trained 'fake patients' who present symptoms of an illness to a physician like any other normal patient. Details of the interactions when SPs are unknown or unannounced to the providers beforehand can be used to evaluate the quality of care received by a typical patient (Rethans et al., 1991). SPs are coached to present their initial symptoms and answer any questions that the physician may ask as part of history taking, in a manner consistent with the underlying condition. We followed the same method (adapted to local conditions) and sent unannounced SPs to healthcare providers in our sample during the course of a normal working day.

A total of 15 SPs were recruited from the districts where the study was conducted. Using a team that included a professional SP trainer, two medical doctors, and a medical anthropologist familiar with local forms of presenting symptoms and illnesses, SPs were coached to accurately and consistently present one of three cases - unstable angina in a 45 year-old male, asthma in a 25 year-old female or male, and dysentery in a child who was at home presented by the father of the child (see Das et al. (2012) and Appendix B for details on SP protocols).[10] SPs visited sampled providers, who did not know they were receiving standardized patients and therefore should have treated them as new patients.[11] After the interaction, SPs were debriefed within an hour with a structured questionnaire that documented the questions and examinations that the provider completed or recommended, the treatments provided, and any diagnoses offered. The SPs retained any medicines dispensed in the clinic and paid all fees charged by providers at the end of the interaction.

The SPs depicted uncomplicated textbook presentations of the cases, and a panel of doctors who advised the project concurred that appropriate history taking and examinations should lead providers towards the correct diagnosis and treatment. Cases were specifically chosen so that the opening statement by the SPs would be consistent with multiple underlying illnesses, but further questioning should have led to an unambiguous (correct) diagnosis. This allows us to measure provider quality through adherence to an essential checklist of

[10] Das et al. (2012) discusses the SP methodology in further detail and presents summary statistics on overall quality of care in this setting. The current paper focuses on the economics of unregulated healthcare markets and we do not replicate the analysis in Das et al. (2012). See Appendix B for further details on how the SP method was implemented, including further discussion on the choice of cases and their relevance. Details on case presentations and instruments are posted on `www.healthandeducationinindia.org`

[11] The research ethics board of Innovations for Poverty Action approved this design following a successful pilot in Delhi, where the detection rate of SPs was extremely low even among a set of doctors who were informed that they would receive an SP at some point in the next month.

questions and examinations that would allow them to accurately make a diagnosis and provide a correct treatment. We also chose these cases since they represented conditions with high or growing incidence in India and other middle- and low-income countries, and they minimized risk to SPs that could arise from unsafe invasive examinations, such as a blood test with an unsterilized needle.

In these cases the role of suitable medical advice was important because real patients would be unlikely to be able to categorize the symptoms as "life threatening" or "potentially non-harmful" and triage themselves into clinics or hospitals. For instance, the SP with unstable angina complains of chest pain which, even in countries with advanced health systems, is often mistaken by patients as arising from heartburn, exertion or muscle strain.[12] Similarly, wheezing and shortness of breath in asthma may arise from short-term allergies to environmental contaminants. Finally, for any child with diarrhea, a key contribution of a healthcare provider is to assess whether the symptoms reflect a bacterial or viral infection (and thus whether the patient requires antibiotics) and the degree of dehydration - each of which may be difficult for parents to assess.

3.2 Healthcare Provider Sampling and Summary Statistics

Our study first uses the census of healthcare providers described earlier to construct a near representative sample of public and private healthcare providers in three of the five sampled districts in rural MP. While our SPs were recruited from the districts in our sample, they were never residents of the villages where they presented themselves to health providers. Since providers in rural areas might know their patients, the SPs had to justify their presence in the area by mentioning, for example, work-related travel or visits to relatives. For such excuses to be plausible, our final sample dropped villages that could not be accessed by paved roads and comprised a total of 46 villages across three districts. While these sampled villages have more providers on average than the entire representative set of villages, there is no difference in the composition of providers across the frame and sample (Table 1).

Since SPs visited clinics to obtain primary care, we excluded community health workers, midwives, and providers that only made home visits. We then sampled all public clinics (some large ones were sampled twice), and a maximum of six private providers in each market for a total of 235 clinics, and SPs completed interactions with 224 providers.[13]

Data from this 'representative sample' allow us to compare care provided across typical

[12] The REACT study in the United States found that many chest pain patients delayed calling 911 because they confused their symptoms with heartburn (Faxon and Lenfant, 2001).

[13] In one case, a sampled village was near a market with over a hundred different healthcare providers. In this one case, we sampled over 20 private providers. See Appendix A for further details on sampling.

public and private clinics in rural MP (all estimates are re-weighted by the inverse of the sampling probabilities to provide population representative averages). However, this comparison would reflect a combination of any compositional differences among providers across public and private clinics, as well as the effect of practicing in the private sector.

To isolate the role of private sector practice, we identified the universe of public MBBS doctors posted to PHCs and CHCs from all five study districts, even if these clinics were not located in the village-based sampling scheme. We then identified the private practices of these doctors (we found a private practice for 61 percent). We sampled and successfully administered SP visits to 116 public MBBS doctors. Our 'dual sample' consists of the 91 doctors in this MBBS sample who also have a private practice, and for 70 of these, SPs presented cases in both their public and private practices. The 'dual sample' enables a comparison of the quality of care provided by the same doctor on the same case across his public and private practices. SP completion rates in the dual sample were higher in the private (92 percent) compared to public practices (78 percent), due to higher doctor absence rates in their public practice, leading to non-completion despite multiple attempts. We show that all our results are robust to adjustments for differential non-completion rates (see section 5.6 and Appendix D.1).

Note that in the representative sample, the unit of analysis is the *clinic* and the SP experience is recorded based on whoever they saw in the clinic. In the dual sample, the unit of analysis is the *doctor* and the SP made repeat visits to see the sampled doctor if needed (especially in the public practice). Appendix A and Tables A.1 and A.2 provide further details on the sampling and construction of the representative and dual samples.

Table 2 (columns 1-3) provides summary statistics for the representative sample of providers. The providers are mostly middle-aged men and just under 60 percent have completed 12 or more years of education (Table 2, Panel A). Their practices have been open for 13-15 years, and private and public providers self-report an average of 16 and 28 patients per day, respectively. Most practices (82 percent of private and 100 percent of public) dispense medicines in the clinic itself and are equipped with the infrastructure and medical devices required for routine examinations, such as stethoscopes and blood pressure cuffs. In the representative sample, public providers are more likely to have an MBBS degrees (26 percent vs. 8 percent). Private providers charged an average of Rs.51 per interaction. Consistent with nominally priced public care, our SPs paid Rs.3.7 on average in public clinics.

Strikingly, 70 percent of private providers in the representative sample report no formal or unverifiable medical qualifications. However, most of them do have non-credentialed medical training. Table A.3 presents details of medical training in the representative sample, and we see that 86 percent of unqualified providers report having received additional training, with the average duration being 32 months. Similarly, 75 percent of providers with unrecognized qualifications report non-credentialed medical training averaging 37 months. The most common form of training is from being an assistant in another doctor's practice. Field interviews suggest that these providers also receive informal continuing medical education from pharmaceutical sale representatives. Thus, while they have no formal qualifications and are not legally licensed to practice, these providers do have considerably greater medical knowledge than a lay person and command considerable credibility in their communities (as suggested by their high market share).

Column 4 presents summary statistics on the universe of public MBBS doctors, while columns 5-7 present these for the 88 public MBBS doctors in the dual sample and test if they are comparable. Overall, doctors with and without dual practices are similar on observable characteristics, but the former have a longer tenure at their current location. There is no significant difference in the equipment reported across these practices (Columns 8-10), although the overall number of patients seen is higher in the public practice and the fees charged are higher in the private practice.

We randomly assigned three SPs to each sampled clinic in the representative sample, one presenting each of the three cases. For the dual sample, we sent SPs presenting the asthma and dysentery cases to both practices of the same provider.[14] Since the rarity of unstable angina could have raised suspicions if providers saw two travelers presenting the same case (even though visits were typically separated by a few weeks), we randomized the providers into two groups - one that received an unstable angina patient in his/her private practice and another that received the case in the public clinic. We show that the randomization was valid in Table A.4.

3.3 Measuring Quality of Care

We use three measures of quality of care. Our first metric is the extent to which the provider adhered to a checklist of questions and examinations required for making a differential diagnosis on each of the presented cases. For instance, these questions and exams would allow a doctor to distinguish between heartburn (that has gastrointestinal origins) and a heart attack, or between viral diarrhea and dysentery. These items represent a parsimonious subset

[14]Since we had 15 SPs and 3 cases, we made sure that the same case was presented by different SPs in the public and private practices. To ensure that our standardized patients saw the sampled provider when (s)he visited the public clinic and not a substitute, we first interviewed all providers in their private practices or residences without revealing that we knew they also worked in the public sector, and we obtained either their photograph or a detailed description of their physical appearance. SPs portrayed a dummy case (e.g. headache) if the doctor was absent when they visited the public clinic, and we sent in other SPs on our subsequent attempts. As we discuss later, it took significantly more trips to complete an SP case in the public practice relative to the private one, due to the high rates of provider absence in the public practice.

of the Indian government's own guidelines, and the list we use was developed by a panel of Indian and American doctors (the items are described for each case in Table A.5).[15] While the most transparent measure of checklist adherence is the percentage of checklist items completed, we also compute an index score using Item Response Theory (IRT), which gives more weight to items that discriminate better among providers. Developed in the context of educational testing, IRT allows us to create a composite measure of provider quality based on questions asked across all three cases, with lower weights on checklist items that are less essential and higher weights on more essential questions that do a better job of discriminating between low and high quality providers (see Das and Hammer (2005) for details). We report both measures in our analysis.

Second, we examine diagnoses - whether one was provided and whether it was correct. We only classify a diagnosis as correct if the provider specified the actual ailment that the SP presented or a functional equivalent. Table A.5 - Panel B presents the diagnoses that were considered correct for each case, and also provides a sense of the wide range of incorrect diagnoses that were seen in practice.

Third, we evaluate the quality of treatment provided. SPs noted all treatment instructions received and retained all prescriptions and medication dispensed in the clinic. These were then classified as correct, palliative, or unnecessary/harmful, based on inputs from our panel of doctors, pharmacists, and a pharmaceutical company (see Appendix B.4 for details; Table A.5 - Panel C lists specific treatments in each category). Since providers can dispense or prescribe multiple medicines, we classify each medicine as correct, palliative, or unnecessary/harmful and thus allow the total treatment protocol to be classified into multiple categories at the same time.

Correct treatment refers to a treatment that is clinically indicated for the specific case and that would relieve/mitigate the underlying condition. Palliative treatments are those that may provide symptomatic relief, or treatments where the providers correctly identified which system was being affected, but which on their own would not cure the patient of the condition that was being presented - for example, allergy medicine for the asthma patient. Treatments classified as unnecessary/harmful were neither correct nor palliative. We group these two potentially distinct categories together because it was difficult to achieve consensus among doctors on what should be considered harmful. Some, for example, would consider antibiotics for the unstable angina patient unnecessary. Others took a longer view with

[15] The Indian government's National Rural Health Mission (NRHM) has developed triage, management, and treatment protocols for unstable angina, asthma, and dysentery in public clinics, suggesting clear guidelines for patients presenting with any of these conditions. The checklist we use is more parsimonious. If we had used the more extensive checklist and asked the SPs to recall adherence to more items, it is likely that checklist adherence would be lower than the numbers that we document.

antibiotic resistance in mind and considered it as ultimately harmful. However, none of the treatments we observed were directly contra-indicated, and hence most of these represent unnecessary treatments as opposed to directly harmful ones.[16]

However, even after classifying all medicines as correct, palliative, and unnecessary/harmful, there are two challenges in coding the "correctness" of a treatment. The first is: How should we interpret a referral when incentives are very different? In some cases, this may be a good thing (if, for example, the provider refers a heart attack patient to a hospital). In other cases, a "referral" may simply reflect a provider who deflected the case without directing the patient usefully.[17] Since we did not send the SPs to the place that was referred, there is no obvious way of coding the quality of referrals. We therefore try to be conservative in our main analysis and do not treat referrals as correct treatments. When we repeat the analysis treating referrals as correct in the angina case, our results are unchanged (results below).

A second challenge arises from the proxy nature of the dysentery case. Many providers did not provide a treatment because the child was not presented and instead asked to see the child. We therefore report results for 'checklist completion' using all three cases, but drop the dysentery case for 'diagnosis' and 'treatment' because the patient (the sick child) was not actually presented for this case. All results are robust to dropping the case completely.

4 Results - Quality of Care across Public and Private Providers

4.1 Estimation Framework

Our main interest is in estimating differences in the quality of care that patients received from providers in the public and private sectors. In the representative sample, we estimate:

$$q_{(i(scp)m)} = \beta_0 + \beta_1 \text{Private}_{ip} + \beta_2 X_p + \delta_s + \delta_c + \delta_m + \epsilon_{i(scp)m} \quad (1)$$

where we regress each measure of quality q (checklist completion, diagnosis, and treatment) in interaction i between a standardized patient s presenting case c and a provider p in market m on an indicator for the sector (Private), with β_1 being the coefficient of interest. Since we pool cases and SPs and there may be systematic differences across them, all our

[16] If the overall quality of care were higher, we could have designed the SP case with a patient who is allergic to certain kinds of antibiotics or who is on regular medication for another illness. In this case, many treatments would have been harmful and the case would have required the doctor to watch out for drug interactions. Given the low-level of overall quality of care, designing such an SP case would not have been very useful at discriminating quality because SPs were never asked about existing allergies or whether they were currently taking any medication.

[17] Field notes suggest that this often happened in public clinics where the doctor was absent. The available provider did not ask questions or conduct any examinations, and told the SP to go elsewhere. By necessity, this is coded as a "referral" in our data, although the patient received no information from the interaction.

specifications include SP and case fixed effects (δ_s and δ_c). We report three sets of estimates for each quality measure. First, we include only SP and case fixed effects; then we add market fixed effects so that comparisons reflect relative performance in the same market (note that not all markets had both types of providers); finally, we add controls for provider and practice characteristics X_p, to adjust for observable differences across providers including demographics, reported qualifications, and number of patients waiting during the visit.

While β_1 provides a useful estimate of the differences in quality across public and private providers in a representative sample of providers, it is a composite estimate that includes differences in unobservable provider characteristics, as well as the effect of practicing in the private sector. To isolate the impact of private sector practice, we re-estimate equation 1 in the dual sample that only includes data from the cases where we sent the SPs to the public and private practices of the same MBBS doctor. We report three sets of estimates here as well. First, we include only SP and case fixed effects;[18] then we add district fixed effects (since the dual practice sample was drawn from the universe of public MBBS doctors practicing in each district rather than the universe of providers practicing in sampled villages, as was the case for the representative sample); finally, we include controls for observable differences across the public and private practices of the doctors.

4.2 Completion of Essential Checklist of History Taking and Examinations

Columns 1-3 in Table 3 present results from estimating equation 1 in the representative sample. Our outcome variable is 'provider effort', measured by consultation length and checklist completion. While the results are similar across the three specifications, we focus our discussion on the estimates in Panel B, because they compare relative performance within the same market (without controlling for provider characteristics), which is the relevant choice set for patients. The base level of effort among representative public providers was low. The average public provider spent 2.4 minutes with the SP in a typical interaction and completed 16 percent of checklist items. Private providers spent 1.5 minutes more per patient and completed 7.4 percentage points more items on the checklist (62 percent and 47 percent more than the public providers respectively). When evaluated on the IRT scaled score, private providers scored 0.61 standard deviations higher. Figure A.2 shows that time spent with the patient is strongly correlated with the number of checklist items completed,

[18]Note that we do not include provider fixed effects since the angina case was not presented in both the public and private practices of the same doctor and will drop out if we do so. Since the case was randomly allocated across the public and private practices of the doctor and assignment was balanced on measures of quality of other cases (see Table A.4), our estimates will be an unbiased estimate of the average quality difference across the public and private practices of public MBBS doctors. We also estimate equation 1 with provider fixed effects and the results are unchanged (but driven by variation in the asthma case).

which points to the credibility of the SP presenting the case, as more time spent with the patient led to greater checklist completion.

Columns 4-6 repeat the analysis in the dual sample, with similar results. Public MBBS doctors appear to be more productive than the typical public provider in the representative sample (many of whom are unqualified) because they complete a slightly higher fraction of checklist items (18 percent) in 35 percent less time (0.8 minutes less). However, this additional productivity is not used to complete more checklist items in the public practice, but rather to reduce the time spent with patients (1.56 minutes versus 2.4 minutes in the representative sample). In their private practices, the same doctors doubled consultation length, completed 60 percent more checklist items, and scored 0.76 standard deviations higher on the IRT-scaled measure of quality. It is worth comparing these differences with those obtained in interventions that are regarded as highly successful. For instance, Gertler and Vermeersch (2013) look at checklist completion as a result of the introduction of performance pay in Rwanda. They find that performance pay increased checklist completion by 0.13 standard deviations; we find that the difference in checklist completion across public and private practices of the same doctor is over five times larger.

These differences are seen clearly in Figures 1-3. Figure 1 plots the cumulative distribution functions (CDF) of the IRT-score (based on checklist completion) of public and private providers in the representative sample, Figure 2 does so for the dual sample, and Figure 3 pools all four samples together (Figures A.3 - A.5 plot the corresponding distributions). The distribution of checklist completion for private providers first-order stochastically dominates that of the public providers (Figure 1) and the corresponding distribution for the private practices of public providers also first-order stochastically dominates that of their public practices (Figure 2). Finally checklist completion is higher for public MBBS doctors than a representative public provider (as would be expected given that the former are more qualified), but it is lower for the public MBBS doctors even relative to a representative sample of private providers (most of whom are unqualified, Figure 3).

Focusing on individual checklist items (Table A.6) shows that private providers in both samples are significantly more likely to perform several items on the checklist on all three cases and are no less likely to perform any of the items (except for one in asthma). In addition to β_1, Table 3 (columns 1-3) also shows that there is no statistically significant correlation between the possession of any formal medical qualification and checklist completion, suggesting that formal qualifications may be a poor predictor of provider effort.

4.3 Diagnosis

Results for diagnosis (Table 4) follow the same format as Table 3 but the dependent variables of interest are whether any diagnosis was given and whether a correct diagnosis was given (both conditional and unconditional on uttering a diagnosis). In the representative sample, 26 percent of public providers offer a diagnosis, of whom only 15 percent offer a correct one. The unconditional probability of a correct diagnosis was only 4 percent.

Private providers in the representative sample are more likely to offer a diagnosis but are not more likely to offer a correct one. The probability of offering a correct diagnosis is higher in the dual practice sample (15 percent vs. 4 percent), which is not surprising since these providers are all trained MBBS doctors. Even among these doctors, however, there is no difference in the rate of correct diagnosis between their public and private practices. Overall, the summary statistics, our price regressions (seen later), and our field work suggest that pronouncing a correct diagnosis (or even just a diagnosis) is not seen by providers (and the market) as being essential in this setting. Of course, *pronouncing* a correct diagnosis is neither necessary nor sufficient for providing a correct or palliative treatment.[19]

4.4 Treatment

Table 5 reports on several outcomes related to the treatment offered, coded as discussed in section 3.3. The probability of receiving at least one correct treatment from a representative public provider was 21 percent. However, they offered non-indicated treatments at much higher rates, with a 53 percent probability of providing a palliative treatment and a 74 percent probability of providing an unnecessary treatment. Since the majority of providers provide unnecessary treatments, the probability of receiving only a correct treatment and nothing more is 2.6 percent. We can also examine two potential proxies for over-treatment - the rate of antibiotic prescriptions and the total number of medicines provided. Antibiotics were prescribed or dispensed in 26 percent of interactions (though they were not indicated for the asthma and angina cases), and an average of 2 medicines per interaction were dispensed.

In the representative sample, we do not find a significant difference between public and private providers on the probability of providing a correct, palliative, or unnecessary treatment; however, point estimates suggest that private providers have a higher probability of providing both correct and unnecessary treatments. Private providers in the representative sample also provide significantly more medicines (over 3 medicines on average, which is 50 percent greater than the public clinics).

[19]Since the providers are usually much more educated than the typical patient, field interviews suggest that they often feel no need to explain themselves to the patients. Thus, providers may have an implicit diagnosis in their minds before they treat, but appear to feel no need to pronounce a diagnosis.

In the dual practice sample, treatments provided in the private practice strictly dominate those provided in the public practice of the same doctor. The rate of correct treatment is 42 percent higher (16 percentage points on a base of 37 percent), the rate of providing a clinically non-indicated palliative treatment is 20 percent lower (12.7 percentage points on a base of 64 percent). The rate of antibiotic provision is 28 percent lower (13.9 percentage points on a base of 49 percent) in the private relative to the public practice of the same doctor.

As discussed in section 3.3, the results reported here are based on treatments that were dispensed at the clinic as well as those that were prescribed. In public clinics, medicines were typically dispensed within the premises, and provided free.[20] In the private clinics, we observed both dispensing and prescribing behavior. While the fees charged included the medicines provided at the clinic, patients would have had to pay separately for prescribed medications. Since we do not observe the typical rate of patient adherence to prescription protocols, our results should be interpreted as referring to the quality of medical advice provided as opposed to the quality of realized health outcomes.

4.5 Knowledge and Effort of Public and Private Providers

There is a strong correlation between higher provider effort and the probability of giving a correct treatment (Figure 4). Nevertheless, the results in Tables 3 and 5 suggest that the higher effort exerted by private providers in the representative sample does not translate into better treatment outcomes. A natural explanation is that the representative private provider has a lower level of medical knowledge but compensates with higher effort, yielding comparable overall levels of treatment accuracy (in line with our theoretical framework below). To examine this possibility further, we use the 'discrimination' parameter of each checklist item (as estimated by the IRT-model; see Table A.6), to classify individual items into terciles of low, medium, and high discrimination items.[21] Here, higher discrimination items are those that are more effective at distinguishing provider quality.

Table A.7 reports the same specifications as in Table 3 but compares public and private providers on checklist completion for different levels of item discrimination. All providers are less likely to complete high discrimination items on the checklist (consistent with low overall quality of care). In the representative sample, private providers complete 11 percentage points more of the low-discrimination checklist items but are no more likely to complete

[20] Medicines were provided free in 92 percent of interactions in the representative sample and 97 percent in the dual sample. Thus, for the most part, care in the public sector was free as it is meant to be.

[21] The classification of items into terciles of difficulty is done within each case, but the results are robust to classifying the items jointly across all cases as well. The terciles for each item are indicated in Table A.6.

high-discrimination items. However, doctors in the dual sample are significantly more likely to complete both low and high-discrimination items in their private practice. These results suggest that while private providers do exert more effort, their lower knowledge leads to this effort being directed towards questions that are easy to ask and interpret, and may limit the marginal productivity of their effort. The results also highlight the importance of using the dual sample for holding provider knowledge and unobservable characteristics constant, and isolating the effect of market incentives on quality of care provided.

4.6 Robustness of checklist and treatment results

Our main results pool data across cases to maximize power. For completeness, we also show the results from Tables 3-5 by case (Table A.8). The superior performance of private providers on consultation length and checklist completion is seen in each of the three cases and in both the representative and the dual samples. Consistent with the overall results, private providers in the representative sample do not do better on diagnosis or treatment in any of the individual cases. In the dual sample, MBBS doctors were 14 percentage points more likely to correctly diagnose and 29 percentage points more likely to correctly treat the unstable angina (heart attack) case in their private practice relative to their public practices. In the asthma case, they are 13 percentage points more likely to offer a correct treatment (but this is not statistically significant given the smaller case-specific sample size).

We confirm that the results in Table 5 are robust to alternative definitions of correct treatment. Table A.9 shows the specific treatments offered by case, including referral frequency. Table A.10 shows that the results in Table 5 are robust to treating all referrals as a correct treatment. As discussed earlier, we include the dysentery case for the analysis of checklist completion but exclude it from the analysis of correct diagnosis and treatment because of the large (and differential) fraction of cases where the provider did not provide these and instead asked to see the child (see Table A.9). Since checklist completion may also be censored in such cases, we also present the checklist completion results without the dysentery case and the results of Table 3 continue to hold (Table A.11). We also show the core results with controls for clinic-level infrastructure and facilities (Table A.12), and all the results continue to hold, suggesting that the results are not being driven by differences in facilities and infrastructure across public and private clinics. The final concern is that of differential completion rates of cases across public and private practices in the dual sample. We discuss this issue in detail in Appendix D.1 and show that our estimates are likely to be a lower bound of the public private differences (Table A.13 and A.14).

5 Results - Pricing and Cost Effectiveness

5.1 Correlates of Prices Charged among Private Providers

Private providers in this setting do not typically have fixed consulting fees that patients see or pay before entering the clinic. Rather, patients walk in to the clinic and describe their symptoms. Pricing takes place at the end of the transaction with the provider asking the patient to pay a certain amount for the entire transaction including medicines dispensed. Our SPs followed this same protocol and paid the prices charged at the end of each interaction with a provider. We now examine the correlates of prices charged for completed transactions to understand what the market rewards in this setting.

Table 6 presents correlations between prices charged by private providers and our various metrics of healthcare quality in the representative sample, dual sample, and pooled sample. The odd columns present binary correlations, while the even columns present multiple regressions. The market rewards several measures of quality of care including time spent, checklist completion rates, and provision of a correct treatment (Table 6, Columns 1, 3 and 5). On the other hand, there is no price premium for pronouncing a correct diagnosis and a price penalty for referrals; whether this penalty is optimal (without a penalty, every provider should just refer the patient) or reduces provider incentives to refer patients adequately is unclear. Finally, there is a price premium for dispensing medicines, but not for prescribing them.[22] The price charged is increasing in the total number of medicines dispensed, which may provide incentives for the provision of excessive medication.

Most of these patterns are repeated in the multiple regressions (Table 6, Columns 2, 4 and 6). Note, however, that correct treatment is no longer rewarded in the multiple regressions. This is likely due to the high correlation between the provision of a correct treatment and the checklist completion rate (Figure 4) and between correct treatment and the use of medicines. Thus the market appears to reward observable measures of quality such as time spent, checklist completion, and dispensing medicines (which are correlated with the provision of correct treatment), but patients do not appear to be able to discern whether they received the correct treatment conditioning on these observable measures.

The correlates of pricing observed in Table 6 point to both strengths and weaknesses of market-based incentives for healthcare provision. On one hand, there appear to be positive incentives for the provision of better quality care (including more effort and providing the correct treatment). On the other hand, the results are consistent with evidence from other

[22] Note that we cannot rule out the possibility that pharmacists provide doctors with a commission for prescriptions that they fill out, which would increase the incentives to over-treat as shown in China by (Currie, Lin and Meng, 2014).

settings, which show that markets for credence goods with asymmetric information between providers and customers often reward over-provision to the detriment of customer welfare. Overall, the results suggest that the market rewards providers who "do more", which is correlated with doing more "good" things as well as more "unnecessary" things.[23]

In sharp contrast to the market for private healthcare, the public sector rewards qualifications and age (experience), but there is no correlation between provider wages and any of our measures of quality including the time spent, checklist completion, or correct treatment (Table 7). Since public employees receive non-pecuniary rewards for better performance through more desirable job postings, we also present correlations between the desirability of a posting and measures of quality and again find that the only significant correlate of a better posting is age - suggesting that the public sector does not reward the quality of care provided by doctors with either more pay or with more desirable job postings.[24]

5.2 Comparative Cost Effectiveness

While healthcare in the public sector is free or nominally priced to the user, it is not cost-free to the tax payer. Table A.16 presents estimates of the cost per patient in the public sector, and calculates that the cost per patient interaction is around Rs.240. This is a conservative calculation because it uses only the wage cost in the public sector and does not include any cost of infrastructure, facilities, equipment, medicines or administration. By contrast, the fees charged are the only source of revenue for private providers and hence will cover all operating costs. Thus, even though private providers charge higher consultation rates than public providers (as seen in Table 2), the per-consultation fee of Rs.51 charged by private providers is less than a fourth of the cost of a patient interaction in the public sector.[25]

[23] Note that the results are robust to excluding observations where we were not able to identify the medicines provided and classify them as correct or not (see Table A.15).

[24] These results are similar to those found in publicly-provided education in India and Pakistan, where teacher salaries increase with qualifications and seniority, but are not correlated with their effectiveness at raising test scores (Muralidharan, 2013; Das and Bau, 2014). Note also that our results add to a very limited evidence base (outside education) on the correlation between pay and productivity in the public sector, since worker-level productivity is typically not observed (see Muralidharan (2016) for a review of the evidence)

[25] Note that we assume that there is a comparable case mix for primary-health visits across public and private facilities, as is standard in comparative cost effectiveness analysis of this sort. This is also consistent with our data from observing real patients (see section 7 below) where we observe considerable overlap in the symptoms presented across public and private clinics.

6 Robustness

6.1 Real Patients

The use of SPs to measure quality of healthcare presents several advantages over the method of clinical observations. However, SPs are limited in the number and types of cases that can be presented. Further, we may worry that the SPs present “off equilibrium” situations in the market that do not extend to its general functioning. We therefore supplemented our data collection after completing the SP modules by conducting day-long clinical observations to code actual provider-patient interactions. We conducted these observations in both the representative and dual samples and in the latter observed a provider in both his/her private and public practices. While we cannot code the actual quality of care from these observations (since we do not observe underlying illnesses), we record several observable characteristics of each patient interaction based on over 1000 interactions in both samples.

Table 8 reports results from estimating Eq. 1 with data from real patient interactions. Private providers spend more time with patients, ask more questions, and are more likely to conduct a physical exam. They also give out more medicines on average. Results from the dual sample are also remarkably similar to those in Tables 3-5, with private providers still exhibiting higher effort but not providing more medicines. Thus, while our SPs present only three specific cases, our results from observing real interactions between patients and providers across the entire set of cases seen in a typical day are very similar to those from the SPs, suggesting that our SP-based results may be valid for a wider range of cases.

6.2 Statistical Discrimination

Another issue in interpreting our dual-sample results is the possibility that doctors expect to see different patients and cases across their public and private practices, and that the differences we observe do not reflect market incentives as much as statistical discrimination.

We address this concern in three ways. First, we note that the cases are both standard and ubiquitous in our setting, and it is therefore unlikely to be "off the equilibrium" path for a provider to see a patient with these symptoms in either public or private clinics. Second, the cases were chosen such that the optimal diagnosis effort and treatment protocol for an initial consultation for these symptoms should not vary by the affluence level of the patient or their ability to afford follow up treatments. Third, we conducted detailed exit interviews with a sample of patients from each clinic that we conducted physician observations in. While patients visiting private clinics are wealthier and have more education (in the dual sample), we find that there aren’t many differences on average in case characteristics across public and private clinics (see Table A.17). In other words, for the majority of observable symptoms and patient characteristics, it is not the case that patients go exclusively to a public or private clinic, suggesting that our results are unlikely to be explained by statistical discrimination (see Appendix D.2 for a more detailed discussion).

6.3 Strategic Diversion of Effort in the Dual Sample

A further issue in interpreting our dual-sample results is the possibility that doctors with private practices may deliberately under-provide effort in their free public practices to shift demand to their fee-for-service private practices (see Jayachandran (2014) for a similar example from education). While we cannot fully rule out this possibility, there is suggestive evidence against this. We compare public providers with and without a private practice and find that providers with a private practice are not any more likely to refer away an SP (Table A.18). Providers with a dual practice do provide less effort in their public practices relative to those without a private practice, but the lack of any evidence of differences in referral rates suggest that these differences may reflect selection rather than strategic behavior, with more publicly conscientious doctors less likely to have a private practice.

The relevant policy question is whether doctors will start exerting more effort in their public practice if the option of private practice did not exist. But it is worth noting that private practice by public MBBS doctors was illegal in MP during the time of our study and that over 60 percent of providers still had a private practice, consistent with the idea that this is a low accountability environment.

6.4 Alternative Comparisons in the Representative Sample

Finally, our representative sample analysis compares the average public and private provider in a market, but it is not clear if the average is the correct metric for quality since patients can choose the best provider in the market. We therefore present an alternative comparison between the best public and best private provider in *each* market in Table A.19 and find that our results are very similar to those in Tables 3-5.

7 Theoretical Framework

The main contribution of this paper is in establishing key facts about the functioning of healthcare markets for primary care in settings of low-income and low state capacity for administering high-quality public health systems. In Appendix C, we present a simple theoretical framework to help interpret the facts that we document. We model provider-patient interactions as comprising of two stages: consultation, and treatment; and characterize the optimal effort and treatment choices that a provider is likely to make with and without market incentives, and the effects of their choices on patient health outcomes.

We model the consultation stage as one of Bayesian learning for the provider. Patients present their initial symptoms to the provider, based on which he forms a prior distribution regarding the true ailment. Higher effort in the consultation stage yields a more precise posterior distribution of beliefs regarding the true ailment (provider effort and knowledge are complements). In the treatment stage, the choice of treatment is determined by a combination of the physician's desire to cure the patient (which is facilitated with a more precise diagnosis), and market incentives for over-treatment. Consistent with our empirical results on correlates of prices charged, we assume that providers in the private sector receive compensation for (observable) effort, as well as a piece-rate for medicines dispensed.

The main insight from the model is that market incentives will typically lead to higher diagnostic effort on the part of the provider (which is observable to patients and rewarded with higher prices) but that the impact on health outcomes is ambiguous. In settings where the default effort level of doctors under low-powered incentives is reasonably high (as may be true in many high-income country settings), the costs of over-treatment under market incentives will likely exceed the benefits of higher diagnostic effort (which has diminishing returns). However, in settings where the default effort level is very low (as is true for public-sector healthcare providers in our setting), the benefits of higher diagnostic effort in the private sector may outweigh the costs of over-treatment under market incentives.

The framework may help shed light on why the quote from Arrow (1963) regarding the undesirability of market-based provision of healthcare may not fully apply to our setting, though it may be highly relevant to developed country contexts.[26] More generally, it highlights the need for caution in extrapolating insights obtained in high-income settings to developing country settings where there is much lower state capacity to implement policies as designed (Muralidharan, Niehaus and Sukhtankar, 2014).

8 Discussion and Conclusion

We present the first set of results on the quality of public and privately provided primary healthcare in a low-income country that features a de facto unregulated private sector, using an audit methodology that accounts both for differences in provider and patient composition. The audit results based on three tracer conditions are similar to those using observations

[26]For instance, the US healthcare literature has paid considerable attention to the problem of over-treatment induced by fee-for-service compensation of providers (see Clemens and Gottlieb (2014) for an illustration). But the default level of provider effort in the US is reasonably high even without market incentives (through a combination of higher-quality medical training and accreditation, peer monitoring of practice standards, and a functioning liability regime for malpractice), making over-treatment the more salient concern. Over-treatment may also be a more first-order concern when patients are insured and do not face the marginal financial costs of over-treatment, which is not true in our setting.

of real patients in the same clinics. Extrapolating beyond the context of our study to critical illnesses or tertiary care requires caution; nevertheless, it is worth highlighting that primary care is an essential first step in identifying the need for tertiary care and directing patients accordingly. In particular, our angina case represents a situation where self-triage was difficult and where the patient would not know if he needed to go the hospital.

Our data suggest that patients in our setting have few good options for healthcare - public or private. Private sector providers exert higher effort but their effectiveness is ultimately limited by low level of medical knowledge, as the majority does not have formal medical qualifications. Public sector clinics, though theoretically staffed by qualified providers, are characterized by lower provider effort. Posts are vacant and doctors frequently absent, so that even in a public sector clinic, the patient often sees a provider without formal training. Lower effort compared to the private sector offsets the benefits of more qualified providers in the public sector, and ultimately there is little difference in correct treatment or the overuse of incorrect medicines across a representative sample of public and private providers. Further, our preferred estimates suggest that the public healthcare system in India spends at least four times more per patient interaction but does not deliver better outcomes than the private sector.[27]

Comparing the same provider in the public and private sector allows us to isolate the comparative effectiveness of market-based accountability in the private sector to administrative accountability in the public sector. The former performs better on all counts. Adherence to checklists and correct treatment rates are higher in the provider's private clinic with no differences in the extent of unnecessary treatments.

These results are consistent with the hedonic earnings-effort relationship in the private sector, which is absent in the public sector. Providers in the private sector earn more when they complete more of the medically necessary checklist and when they provide a correct treatment, showing that the market rewards certain key aspects of high quality. However, the market also rewards unnecessary treatments and patients frequently receive and pay for treatments that they do not need, a finding that mirrors consistent concerns regarding over-treatment in the literature on credence goods.

Despite market incentives for over-treatment, one surprising result is that the rate of provision of unnecessary medication is equally high in the public clinics. Our theoretical

[27]These results mirror recent experimental evidence on primary education. Muralidharan and Sundararaman (2015) find that private schools in rural India deliver equal or superior learning outcomes than public schools, even though public schools spend three times more per student. Private school teachers are less qualified than public teachers, but exert much higher levels of effort. Thus, private providers in both primary health and education appear to make up for lower qualifications with higher effort, yielding outcomes no worse than those provided by the public sector - which have much higher costs per student/patient.

framework provides a possible explanation for this result by showing that unnecessary treatments are not only driven by market incentives, but can also arise from low diagnostic effort. In our setting of low default effort in the public sector, the increase in diagnostic precision enabled by higher effort in the private sector may offset the incentives for over-treatment under market incentives, yielding no net difference in the provision of unnecessary treatment. Overall, our results suggest that in settings with low state-capacity for high-quality public service delivery, the effort advantage of the private sector may outweigh the credence good costs of privately-provided healthcare.

Indian and global health policy debates have been hampered by a lack of empirical evidence on the quality of clinical interactions in the public and private sectors. Under the status quo, considerable attention has been focused on improving access and spending for publicly-provided healthcare (Planning Commission of India, 2013). Our results suggest that enthusiasm for the public sector as the primary source of primary care services in resource poor settings has to be tempered by the extent to which administrative accountability is enforced in the system and that poor incentives for effort may be a binding constraint to quality in the public system of healthcare delivery.

On the other hand, the marginal returns to better training and credentialing may be higher for private healthcare providers who have stronger incentives for exerting effort. Current policy thinking often points in the opposite direction, with a focus on hiring, training, and capacity building in the public sector on one hand (without much attention to their incentives for effort), and considerable resistance to training and providing legitimacy to unqualified private providers on the other (Reddy et al., 2011; Shiva Kumar et al., 2011; Planning Commission of India, 2013).

This viewpoint is often justified by assuming that patients - particularly those who are poor and illiterate - make poor decisions regarding their health care. While certainly possible, a more nuanced understanding of patient behavior in low-income settings requires better empirical evidence on the actual quality of care obtained from different types of healthcare providers. Our paper presents some of the first evidence on this question, and expanding this methodology to other conditions and settings will allow for a richer understanding of the functioning of healthcare systems in settings with low resources and poor administrative capacity.

References

Arrow, Kenneth J. 1963. "Uncertainty and The Welfare Economics of Medical Care." *American Economic Review*, 63(5): 941–973.

Ayres, Ian, and Peter Siegelman. 1995. "Race and Gender Discrimination in Bargaining for a New Car." *American Economic Review*, 85(3): 304–21.

Baker, George. 1992. "Incentives Contracts and Performance Measurement." *Journal of Political Economy*, 100: 598–614.

Balafoutas, Loukas, Adrian Beck, Rudolf Kerschbamer, and Matthias Sutter. 2013. "What Drives Taxi Drivers? A Field Experiment on Fraud in Market for Credence Goods." *Review of Economic Studies*, 80: 876–891.

Banerjee, Abhijit, Angus Deaton, and Esther Duflo. 2004. "Wealth, Health, and Health Services in Rural Rajasthan." *American Economic Review, Papers and Proceedings*, 94(2): 326–300.

Banerjee, Abhijit V, Esther Duflo, and Rachel Glennerster. 2008. "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Public Health Care System." *Journal of the European Economic Association*, 6(2-3): 487–500.

Black, R E, S Cousens, H L Johnson, and et al. 2010. "Global, regional and national causes of child mortality in 2008: a systematic analysis." *The Lancet*, 375(9730): 1969–87.

Chaudhury, Nazmul, Jeffery Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives*, 20(1): 91–116.

Clemens, Joshua, and Jeffrey Gottlieb. 2014. "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review*, 104(4): 1320–49.

Coarasa, Jorge, Jishnu Das, and Elizabeth Gummerson. 2014. "Evaluating the Evidence on Public and Private Sector Quality of Care in Low and Middle Income Countries."

CPR. 2011. "Mapping Medical Providers in Rural India: Four Key Trends." Center For Policy Research.

Currie, Janet, Wanchuan Lin, and Juanjuan Meng. 2014. "Addressing Antibiotic Abuse in China: An Experimental Audit Study." *Journal of Development Economics*, 110: 39–51.

Das, Jishnu, Ada Kwan, Ben Daniels, Srinath Satyanarayana, Ramnath Subbaraman, Sofi Bergkvist, Ranendra K Das, Veena Das, and Madhukar Pai. 2015*a*. "First use of the standardized patient methodology to assess quality of Tuberculosis care." *Lancet Infectious Diseases*, forthcoming.

Das, Jishnu, Alaka Holla, Michael Kremer, Aakash Mohpal, and Karthik Muralidharan. 2015*b*. "Quality and Accountability in Healthcare Delivery: Audit-Study Evidence from Primary Care in India." *NBER Working Paper 21405.*

Das, Jishnu, Alaka Holla, Veena Das, Manoj Mohanan, Diana Tabak, and Brian Chan. 2012. "The Quality of Medical Care in Clinics: Evidence from a Standardized Patients Study in a Low-Income Settting." *Health Affairs*, 31(12): 2274–2784.

Das, Jishnu, and Jeffery Hammer. 2005. "Which Doctor? Combining Vignettes and Item Response Theory to Measure Doctor Quality." *Journal of Development Economics*, 78(3): 348–383.

Das, Jishnu, and Jeffery Hammer. 2007. "Money for Nothing: The Dire Straits of Medical Practice in Delhi, India." *Journal of Development Economics*, 83(1): 1–36.

Das, Jishnu, and Natalie Bau. 2014. “The Misallocation of Pay and Productivity in the Public Sector: Evidence from the Labor Market for Teachers.” *Working Paper.*

DHS. 2007. “National Family Health Survey (NFHS-3) 2005-2006: Key Findings.”

Dulleck, Uwe, and Rudolf Kerschbamer. 2006. “On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods.” *Journal of Economic Literature*, 44: 5–42.

Faxon, David, and Claude Lenfant. 2001. “Timing is everything - motivating patients to call 911 on onset of acute myocardial infarction.” *Circulation*, 104(11): 1210.

Foster, Andrew D., and Mark R Rosenzweig. 1994. “A Test for Moral Hazard in Labor Market: Contractual Agreements, Reputation, and Competition.” *Review of Economics and Statistics*, 76(2): 213–227.

Gertler, Paul J, and Christel Vermeersch. 2013. “Using Performance Incentives to Improve Medical Care Productivity and Health Outcomes.” *NBER Working Paper No. 19046.*

Goldberg, Penelopi K. 1996. “Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey.” *Journal of Political Economy*, 104(3): 622–54.

Grepin, Karen. 2014. “The role of the private sector in delivering maternal and child health services in low-income and middle-income countries: an observational, longitudinal analysis.” *The Lancet*, 387(S7).

Gruber, Jonathan, and Maria Owings. 1996. “Physician Financial Incentives and Cesarean Section Delivery.” *RAND Journal of Economics*, 27(1): 99–123.

Hanna, Rema, and Iqbal Dhaliwal. 2015. “Making a Deal with the Devil: Experiental Evidence on Bureaucratic Reform in India.” *Working Paper.*

Holmstrom, Bengt, and Paul Milgrom. 1991. “Multi-Task Principal-Agent Problems: Incentive Contracts, Asset Ownership and Job Design.” *Journal of Law, Economics and Organization*, 7(Special Issue): 24–52.

Jayachandran, Seema. 2014. “Incentives to Teach Badly: After-School Tutoring in Developing Countries.” *Journal of Development Economics*, 27(1): 99–123.

Jindal, S K, D Gupta, A N Aggarwal, and R Agarwal. 2005. “Guidelines for management of asthma at primary and secondary levels of health care in India.” *Indian Journal of Chest Disease and Allied Science*, 47(4): 308–343.

La Forgia, Gerrard, and Somil Nagpal. 2014. *Government Sponsored Health Insurance in India: Are You Covered?* The World Bank, Washington, DC.

Leonard, Kenneth L, and Melkiory C Masatu. 2005. “The Use of Direct Clinical Observation and Vignettes for Health Services Quality Evaluation in Developing Countries.” *Social Science and Medicine*, 61(9): 1944–1951.

Muralidharan, Karthik. 2013. “Priorities for Primary Education Policy in India’s 12th Five Year Plan.” *India Policy Forum 2012-13*, 9: 1–46.

Muralidharan, Karthik. 2016. “A New Approach to Public Sector Hiring in India for Improved Delivery.” *India Policy Forum.*

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. “Teacher Performance Pay: Experimental Evidence from India.” *Journal of Political Economy*, 119(1): 39–77.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2015. "The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India." *Quarterly Journal of Economics*, 130(3): 1011–66.

Muralidharan, Karthik, Nazmul Chudhury, Jeffrey Hammer, Michael Kremer, and F Halsey Rogers. 2011. "Is there a Doctor in the House? Absent Medical Providers in India." *Working Paper.*

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2014. "Building State Capacity: Evidence From Biometric Smartcards in India." *NBER Working Paper No. 19999.*

Norris, Pauline. 2002. "Reasons why mystery shopping is a useful and justifiable research method." *The Pharmaceutical Journal*, 272: 746–747.

Patel, V, S Chatterji, D Chisholm, and et al. 2011. "Chronic diseases and injuries in India." *The Lancet*, 377(9763): 413–428.

Pauly, Mark. 1980. *Doctors and Their Workshops: Economic Models of Physician Behaviors.* National Bureau of Economic Research Monograph, Boston, MA.

Planning Commission of India. 2013. "Health, Nutrition and Family Welfare."

Prendergast, Canice. 2003. "The Limits of Bureaucratic Inefficiency." *Journal of Political Economy*, 111(5): 929–958.

Reddy, S K, Vikram Patel, Prabhat Jha, Vinod K Paul, A K Shiva Kumar, and Lalit Dandona. 2011. "Towards Achievement of Universal Health Care in India by 2020: A Call to Action." *The Lancet*, 377(9767): 760–768.

Rethans, Jan-Joost, Ferd Sturmans, Riet Drop, Cees van der Vlueten, and Pie Hobus. 1991. "Does Competence of General Practitioners Predict Their Performance." *British Medical Journal*, 303(6814): 1377–1380.

Rohde, Jon E, and Hema Viswanathan. 1995. *The Rural Private Practitioner.* Oxford University Press, New Delhi.

Rosenzweig, Mark R. 1995. "Why are there Returns to Schooling?" *American Economic Review, Papers and Proceedings*, 85(2): 153–158.

Schneider, Henry S. 2012. "Agency Problems And Reputation in Expert Services: Evidence from Auto Repair." *The Journal of Industrial Economics*, 60(3): 406–433.

Selvaraj, S, and Anup K Karan. 2009. "Deepening Health Insecurity in India: Evidence from National Sample Surveys since 1980s." *Economic and Political Weekly*, 44(18): 55–60.

Shaban, Radwan Ali. 1987. "Testing between Competing Models of Sharecropping." *Journal of Political Economy*, 95(5): 983–20.

Shiva Kumar, A K, Lincoln C Chen, Mita Choudhury, and et al. 2011. "Financing Health Care for All: Challenges and Opportunities." *The Lancet*, 377(9766): 668–679.

Wagstaff, Adam. 2013. "What exactly is the public-private mix in health care?" *Lets Talk Development (blog), Dec 2.*

World Bank. 2003. *World Development Report 2004.* The World Bank, Washington, DC.

Figures and Tables

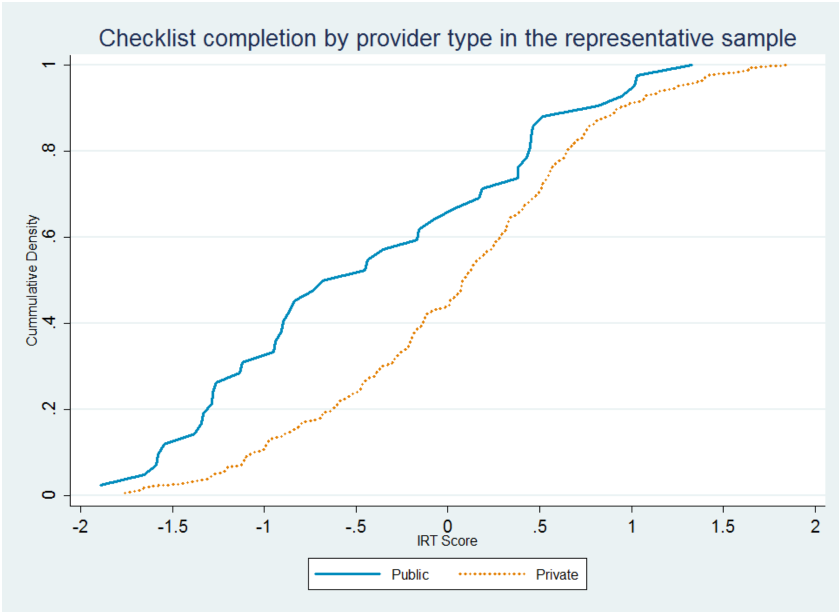


Figure 1

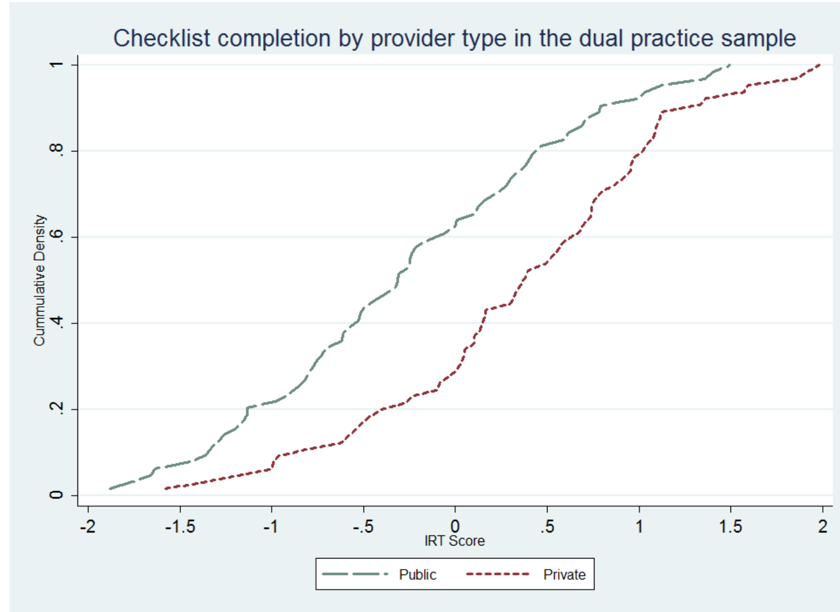


Figure 2

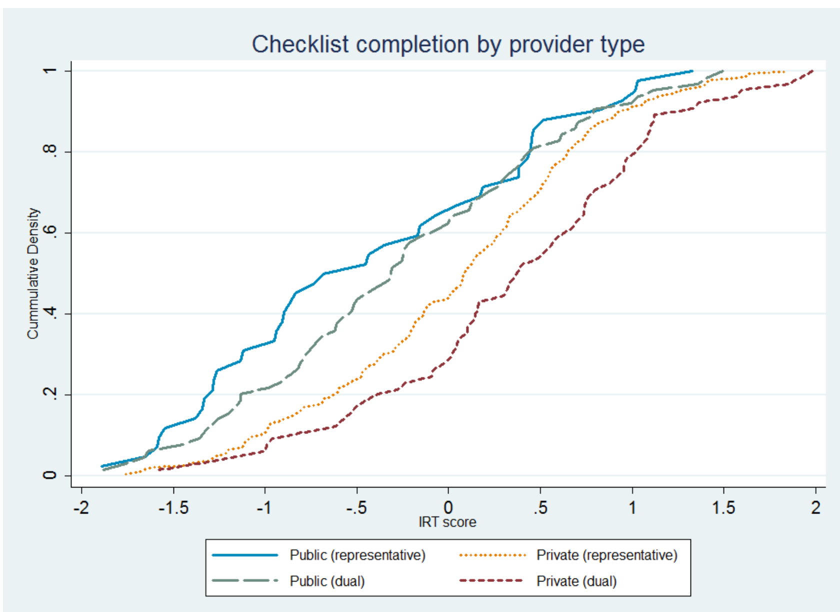


Figure 3

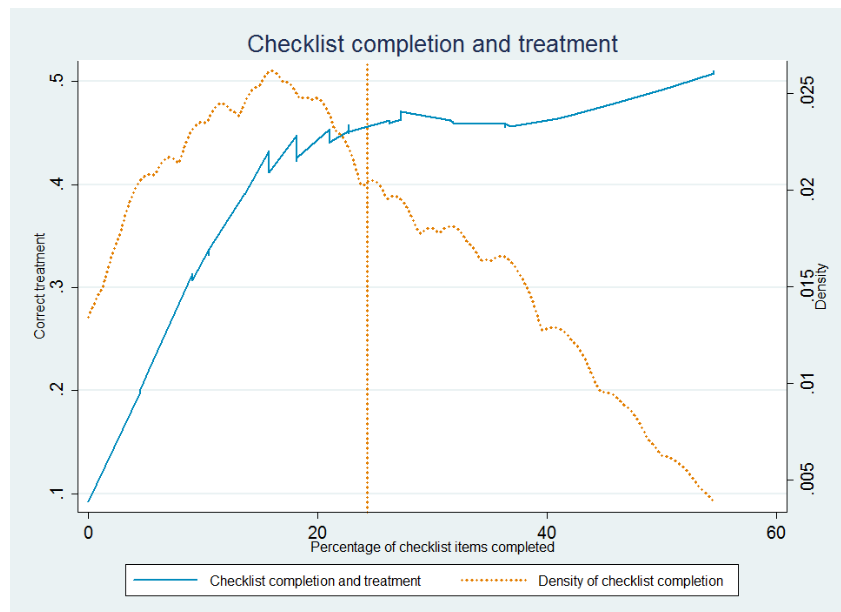


Figure 4

Table 1: Health market attributes

	(1)	(2)	(3)	(4)	(5)	(6)
	Madhya Pradesh (5 districts, 100 markets)			**SP Sample Villages** (3 districts, 46 markets)		
	All	Inside village	Outside village	All	Inside village	Outside village
Panel A: Composition of markets based on census of providers						
Total	11.68	3.97	7.71	16.02	4.65	11.37
	(12.06)	(4.49)	(12.17)	(15.81)	(5.41)	(16.42)
Public MBBS	0.45	0.05	0.40	0.50	0.02	0.48
	(0.97)	(0.22)	(0.93)	(1.11)	(0.15)	(1.11)
Public alternative qualification	0.22	0.07	0.15	0.24	0.07	0.17
	(0.48)	(0.29)	(0.39)	(0.52)	(0.33)	(0.44)
Public paramedical	1.58	1.13	0.45	1.98	1.30	0.67
	(1.90)	(1.46)	(1.33)	(2.12)	(1.49)	(1.59)
Public unqualified	1.71	0.68	1.03	2.07	0.67	1.39
	(1.75)	(1.04)	(1.54)	(2.05)	(1.12)	(1.94)
Total public	3.96	1.93	2.03	4.78	2.07	2.72
	(3.20)	(2.28)	(2.63)	(3.53)	(2.45)	(3.17)
Private MBBS	0.40	0.00	0.40	0.59	0.00	0.59
	(1.57)	(0.00)	(1.57)	(2.15)	(0.00)	(2.15)
Private alternative qualification	1.92	0.23	1.69	2.67	0.33	2.35
	(3.65)	(0.66)	(3.65)	(4.86)	(0.90)	(4.89)
Private unqualified	5.40	1.81	3.59	7.98	2.26	5.72
	(6.01)	(2.23)	(6.14)	(7.88)	(2.74)	(8.32)
Total private	7.72	2.04	5.68	11.24	2.59	8.65
	(10.54)	(2.69)	(10.81)	(14.31)	(3.38)	(14.87)
Panel B: Composition of demand from census of households in sampled villages						
Fraction of households that visited a	0.46			0.58		
provider in last 30 days	(0.50)			(0.49)		
Fraction provider visits inside/outside village		0.66	0.34		0.69	0.31
		(0.47)	(0.47)		(0.46)	(0.46)
Distance traveled to visited provider (km)	1.61	0.40	3.83	1.37	0.38	3.51
	(2.14)	(0.65)	(2.14)	(2.37)	(1.16)	(2.84)
Fraction of visits to MBBS doctor	0.04	0.01	0.09	0.02	0.00	0.06
	(0.19)	(0.09)	(0.29)	(0.13)	(0.00)	(0.23)
Fraction of visits to private sector	0.89	0.92	0.85	0.96	0.97	0.93
	(0.31)	(0.28)	(0.36)	(0.21)	(0.18)	(0.26)
Fraction of visits to private sector	0.88	0.89	0.83	0.95	0.96	0.91
(conditional on public availability)	(0.33)	(0.31)	(0.38)	(0.22)	(0.20)	(0.28)
Fraction of visits to private sector	0.83	0.84	0.79	0.93	0.98	0.90
(conditional on public MBBS availability)	(0.37)	(0.36)	(0.41)	(0.25)	(0.15)	(0.30)
Fraction of visits to unqualified providers	0.77	0.87	0.55	0.82	0.89	0.64
	(0.42)	(0.34)	(0.50)	(0.39)	(0.31)	(0.48)
Fraction of visits to unqualified providers	0.74	0.82	0.54	0.81	0.86	0.64
(conditional on public availability)	(0.44)	(0.38)	(0.50)	(0.39)	(0.35)	(0.48)
Fraction of visits to unqualified providers	0.60	0.77	0.38	0.66	0.81	0.39
(conditional on public MBBS availability)	(0.49)	(0.42)	(0.48)	(0.47)	(0.39)	(0.49)
Panel C: Sample Characteristics from household census of provider choice						
Number of villages	100			46		
Average village population	1,149			1,199		
Average number of households per village	233			239		
Number of reported provider visits	19,331			12,122		
Average number of visits per household per	0.83			1.10		

Notes: Standard deviations in parentheses. The number of providers available to a village was determined by a provider census, which surveyed all providers in all locations mentioned by households in 100 sample villages, when asked where they seek care for primary care services, regardless of whether or not the particular provider was mentioned by households. Unqualified providers report no medical training. All others have training that ranges from a correspondence course to a medical degree. "Outside villages" are typically adjacent villages or villages connected by a major road. The 30-day visit rate was calculated from visits to providers reported by households in a complete census of households in the 100 sample villages. The type of provider they visited was determined by matching reported providers to providers surveyed in the provider census.

Table 2: Characteristics of providers and practices where SPs were administered

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Representative sample (3 districts)			**Representative sample of Public MBBS** (5 districts)				**Dual practice sample** (5 districts)		
	Public	Private	p-value of (1)-(2)	All public	Non-dual public	Dual public	p-value of (5)-(6)	Public	Private	p-value of (8)-(9)
Panel A: Provider characteristics										
Age of provider	46.92	43.51	0.10	44.52	44.74	44.43	0.89			
Is male	0.86	0.96	0.02	0.87	0.96	0.84	0.10			
More than 12 years of basic education	0.58	0.52	0.48	0.64	0.52	0.69	0.09			
Has MBBS degree	0.25	0.07	0.00	1.00	1.00	1.00				
Has alternative medical degree	0.11	0.21	0.18	0.00	0.00	0.00				
Has no formal medical qualification	0.61	0.68	0.42	0.00	0.00	0.00				
Has non-credentialed medical training	0.63	0.78	0.05	0.23	0.22	0.23	0.96			
Number of practices	1.14	1.07	0.21	1.83	1.16	2.13	0.00			
Tenure in years at current location	15.22	13.70	0.42	6.15	5.11	6.56	0.28			
Panel B: Clinic characteristics										
Dispense medicine	1.00	0.81	0.00							
Consultation fee (Rs.)	3.65	51.24	0.00	3.75	3.15	3.92	0.00	3.92	57.93	0.00
Number of patients per day (self reported in census)	28.06	15.74	0.00	31.85	31.30	35.00	0.74	35.00	17.59	0.07
Number of patients per day (from physician observations)	5.72	5.75	0.98	16.04	13.72	16.86	0.31	16.86	5.63	0.00
Electricity	0.94	0.95	0.93	1.00	1.00	1.00		1.00	1.00	
Stethoscope	0.97	0.94	0.47	1.00	1.00	1.00		1.00	1.00	
Blood pressure cuff	0.83	0.75	0.34	1.00	1.00	1.00		1.00	1.00	
Thermometer	0.94	0.92	0.64	0.97	0.94	0.98	0.20	0.98	0.97	0.63
Weighing Scale	0.86	0.52	0.00	0.94	0.94	0.94	0.96	0.94	0.82	0.04
Handwash facility	0.89	0.81	0.30	0.84	0.84	0.85	0.93	0.85	0.81	0.56
Number of providers	36	188		103	31	72		72	84	

Notes: Standard deviations are in parentheses. Unit of observation is a provider. The dual practice sample consists of providers who received a standardized patient in both their public and private practices. Provider mapping and complete provider census yielded information about whether or not a provider operates more than practice. The representative sample did not employ the intense reconnaisance to find both the public and private practices of the same provider, and thus the proportion of dual practice providers can be considered self-reported. In the dual practice sample, however, the existence of additional medical practices was verified by repeated observation. Alternative qualifications are as follows: BAMS, BIMS, BUMS, BHMS/DHMS, DHB, BEHMS, BEMS, B.Sc. Nursing/M.Sc. Nursing, B.Pharma/M.Pharma. In the public sector of the representative sample, there are 3 providers with BAMS and 1 with B.Pharma/M.Pharma. In the private sector, there are 21 with BAMS, 9 with BHMS/DHMS, 3 each with BIMS and DHB, 2 with B.Pharma/M.Pharma and 1 with BUMS. No medical training includes providers with unverifiable degrees and providers who self-reported no formal training. In the public sector of the representative sample, there are 22 with no formal qualifications and 5 who reported other degree. In the private sector, there are 128 with no formal qualfication and 56 who reported other unverifiable degrees. Means for consultation fee were calculated from direct observations of clinical interactions. All other variables derive from a survey administered during the census of providers.

Table 3: Effort in the public and private sectors

	(1)	(2)	(3)	(4)	(5)	(6)
	Representative sample			Dual practice sample		
	Time Spent (mins)	Percentage of checklist items	IRT score	Time Spent (mins)	Percentage of checklist items	IRT score
Panel A: SP and case fixed effects						
Is a private provider	1.222***	6.758***	0.551**	1.507***	8.977***	0.755***
	(0.250)	(2.488)	(0.212)	(0.298)	(1.935)	(0.207)
R-squared	0.305	0.160		0.241	0.220	
Number of observations	662	662	233	331	331	138
Mean of public	2.388	15.287		1.561	17.720	
Mean of private	3.703	22.302		2.983	28.308	
Mean of sample	3.603	21.764		2.274	23.030	
Panel B: SP, case and market/district fixed effects						
Is a private provider	1.486***	7.352***	0.668**	1.514***	8.977***	0.759***
	(0.333)	(2.705)	(0.277)	(0.298)	(1.922)	(0.207)
R-squared	0.391	0.259		0.262	0.234	
Number of observations	662	662	233	331	331	138
Panel C: SP, case and market/district fixed effects						
Is a private provider	1.246***	5.999**	0.611*	1.485***	9.504***	0.829***
	(0.424)	(2.891)	(0.327)	(0.316)	(2.062)	(0.205)
Has MBBS	-0.156	3.285	0.043			
	(0.638)	(2.589)	(0.257)			
Has some qualification	-0.131	2.518	0.157			
	(0.443)	(1.813)	(0.151)			
Age of provider	-0.004	-0.046	0.000	0.004	-0.066	0.004
	(0.013)	(0.059)	(0.008)	(0.019)	(0.089)	(0.101)
Gender of provider (1=Male)	0.653	-0.949	0.212	-0.070	-1.343	-0.288
	(0.428)	(4.207)	(0.327)	(0.437)	(3.306)	(0.309)
Patient load during visit	-0.096	-0.144	0.082**	-0.097**	-0.225	0.013
	(0.061)	(0.481)	(0.040)	(0.041)	(0.457)	(0.517)
R-squared	0.399	0.259		0.278	0.234	
Number of observations	638	638	221	302	302	126

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample, robust standard errors clustered at the provider level are in parentheses. All regressions include a constant. Observations are at the SP-provider interaction level, except in IRT score where each observation is a composite provider level score across all cases. Market fixed effects are used for the representative sample, and district fixed effects for the dual practice sample.

Table 4: Diagnosis in the public and private sectors (unstable angina and asthma cases only)

	(1)	(2)	(3)	(4)	(5)	(6)
	Representative sample			Dual practice sample		
	Gave diagnosis	Correct diagnosis (conditional)	Correct diagnosis (unconditional)	Gave diagnosis	Correct diagnosis (conditional)	Correct diagnosis (unconditional)
Panel A: SP and case fixed effects						
Is a private provider	0.168***	-0.014	0.016	0.095	-0.041	0.023
	(0.052)	(0.057)	(0.022)	(0.066)	(0.102)	(0.049)
R-squared	0.130	0.121	0.075	0.130	0.113	0.055
Number of observations	440	178	440	201	88	201
Mean of public	0.263	0.150	0.039	0.382	0.385	0.147
Mean of private	0.431	0.135	0.058	0.495	0.388	0.192
Mean of sample	0.418	0.135	0.057	0.438	0.386	0.169
Panel B: SP, case and market/district fixed effects						
Is a private provider	0.188***	-0.019	0.023	0.092	-0.056	0.025
	(0.061)	(0.071)	(0.027)	(0.068)	(0.107)	(0.049)
R-squared	0.218	0.301	0.145	0.150	0.175	0.067
Number of observations	440	178	440	201	88	201
Panel C: SP, case and market/district fixed effects						
Is a private provider	0.149**	-0.046	0.031	0.084	0.017	0.044
	(0.067)	(0.095)	(0.032)	(0.071)	(0.120)	(0.055)
Has MBBS	-0.092	0.108	0.008			
	(0.125)	(0.131)	(0.030)			
Has some qualification	0.023	-0.010	-0.012			
	(0.058)	(0.073)	(0.025)			
Age of provider	-0.002	-0.005*	-0.002	0.002	-0.001	0.000
	(0.002)	(0.003)	(0.001)	(0.004)	(0.009)	(0.004)
Gender of provider (1=Male)	-0.089	0.272***	0.079***	-0.125	-0.052	-0.086
	(0.134)	(0.092)	(0.030)	(0.095)	(0.174)	(0.071)
Patient load during visit	-0.003	-0.017*	-0.005	-0.017	-0.003	-0.005
	(0.008)	(0.009)	(0.005)	(0.020)	(0.033)	(0.011)
R-squared	0.222	0.362	0.159	0.185	0.217	0.097
Number of observations	423	173	423	183	80	183

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample, robust standard errors clustered at the provider level are in parentheses. All regressions include a constant. Observations are at the SP-provider interaction level. Market fixed effects are used for the representative sample, and district fixed effects for the dual practice sample.

Table 5: Treatment in the public and private sectors

(unstable angina and asthma cases only)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Representative sample						Dual practice sample					
	Correct treatment	Palliative treatment	Unnecessary treatment	Correct treatment only	Antibiotic	Number of medicines (Dispensed and/or Prescribed)	Correct treatment	Palliative treatment	Unnecessary treatment	Correct treatment only	Antibiotic	Number of medicines (Dispensed and/or Prescribed)
	(1=Yes)	(1=Yes)	(1=Yes)	(1=Yes)	(1=Yes)	(Continuous variable)	(1=Yes)	(1=Yes)	(1=Yes)	(1=Yes)	(1=Yes)	(Continuous variable)
Panel A: SP and case fixed effects												
Is a private provider	0.052	-0.038	0.061	-0.008	0.016	0.972***	0.151**	-0.126**	-0.021	0.019	-0.141**	0.002
	(0.045)	(0.056)	(0.072)	(0.023)	(0.062)	(0.279)	(0.061)	(0.057)	(0.054)	(0.026)	(0.067)	(0.200)
R-squared	0.260	0.215	0.066	0.044	0.079	0.087	0.274	0.309	0.108	0.025	0.120	0.127
Number of observations	440	440	440	440	440	440	201	201	201	201	201	201
Mean of public	0.211	0.526	0.737	0.026	0.263	2.092	0.373	0.637	0.833	0.020	0.490	2.833
Mean of private	0.270	0.496	0.808	0.017	0.279	3.097	0.566	0.465	0.838	0.040	0.374	2.919
Mean of sample	0.266	0.498	0.802	0.018	0.278	3.021	0.468	0.552	0.836	0.030	0.433	2.876
Panel B: SP, case and market/district fixed effects												
Is a private provider	0.051	0.040	0.095	-0.020	0.086	0.894***	0.156**	-0.127**	-0.022	0.018	-0.139**	-0.002
	(0.051)	(0.059)	(0.079)	(0.027)	(0.084)	(0.264)	(0.062)	(0.058)	(0.053)	(0.026)	(0.067)	(0.198)
R-squared	0.384	0.350	0.233	0.255	0.239	0.289	0.299	0.315	0.167	0.039	0.135	0.155
Number of observations	440	440	440	440	440	440	201	201	201	201	201	201
Panel C: SP, case and market/district fixed effects												
Is a private provider	0.101*	0.060	0.066	-0.005	0.112*	0.638**	0.181***	-0.106*	-0.021	0.018	-0.122*	-0.001
	(0.056)	(0.066)	(0.076)	(0.027)	(0.067)	(0.310)	(0.067)	(0.060)	(0.062)	(0.029)	(0.071)	(0.215)
Has MBBS	0.309***	0.246**	-0.132	0.106**	0.267***	-0.397						
	(0.074)	(0.100)	(0.084)	(0.047)	(0.075)	(0.515)						
Has some qualification	0.088**	0.086	0.029	-0.001	0.099	-0.116						
	(0.039)	(0.061)	(0.061)	(0.015)	(0.062)	(0.274)						
Age of provider	-0.001	-0.001	-0.002	-0.000	-0.000	-0.012	-0.002	-0.007	0.001	-0.002	-0.001	-0.019*
	(0.002)	(0.002)	(0.002)	(0.000)	(0.003)	(0.011)	(0.004)	(0.006)	(0.003)	(0.001)	(0.005)	(0.012)
Gender of provider (1=Male)	0.133	-0.118	-0.068	0.001	-0.029	-0.128	0.049	0.097	0.111	0.007	0.152	0.286
	(0.106)	(0.160)	(0.112)	(0.031)	(0.099)	(0.327)	(0.115)	(0.092)	(0.091)	(0.033)	(0.103)	(0.330)
Patient load during visit	-0.008	-0.017	0.007	-0.001	-0.008	0.009	0.004	0.004	0.013	-0.004	-0.000	0.074**
	(0.008)	(0.012)	(0.006)	(0.001)	(0.006)	(0.026)	(0.015)	(0.013)	(0.018)	(0.003)	(0.017)	(0.037)
R-squared	0.406	0.370	0.253	0.278	0.272	0.293	0.279	0.318	0.180	0.053	0.164	0.180
Number of observations	423	423	423	423	423	423	183	183	183	183	183	183

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample, robust standard errors clustered at the provider level are in parentheses. All regressions include a constant. Observations are at the SP-provider interaction level. Market fixed effects are used for the representative sample, and district fixed effects for the dual practice sample. In columns (6) and (12) the dependent variable is total number of medicines recommended to the patient (dispensed and/or prescribed). Medicines in the public sector are almost always dispensed at the clinic, whereas medicines in the private sector are both dispensed at the clinic and prescribed to be purchased elsewhere.

Table 6: Correlates of price charged (private interactions)

	(1)	(2)	(3)	(4)	(5)	(6)
	Fees in Rs.					
	Representative sample		**Dual practice sample**		**Pooled sample**	
	Binary regressions	Multiple regression	Binary regressions	Multiple regression	Binary regressions	Multiple regression
Time spent with SP (minutes)	1.763***	0.771	2.498***	2.017***	1.502***	0.805**
	(0.454)	(0.475)	(0.587)	(0.679)	(0.361)	(0.390)
Percentage of checklist items	0.411***	0.368***	0.355***	0.061	0.394***	0.309***
	(0.091)	(0.101)	(0.100)	(0.124)	(0.073)	(0.093)
Correct diagnosis (unconditional)	-3.749	-2.137	6.353	5.459	2.674	2.803
	(4.212)	(2.122)	(9.363)	(9.076)	(4.670)	(4.175)
Correct treatment	7.065***	0.050	6.301	1.508	7.633***	1.458
	(1.789)	(2.892)	(4.016)	(4.754)	(1.872)	(2.305)
Palliative treatment	8.036***	5.581***	11.748***	7.798*	8.124***	6.252***
	(2.056)	(2.036)	(4.344)	(4.663)	(1.811)	(1.863)
Unnecessary treatment	14.039***	4.030	15.220***	3.145	14.355***	5.545*
	(2.395)	(3.341)	(5.056)	(6.233)	(2.129)	(2.864)
Number of medicines dispensed	4.774***	4.215***	9.247***	11.513***	4.080***	3.937***
	(1.656)	(1.379)	(2.997)	(3.765)	(1.371)	(1.409)
Number of medicines prescribed	-0.202	-1.188	3.650**	3.891	0.926	-1.020
	(1.129)	(0.881)	(1.845)	(2.672)	(0.861)	(1.067)
Referred/Asked to see child	-19.161***	-13.301***	-10.082**	-3.638	-16.857***	-14.151***
	(4.115)	(3.636)	(4.722)	(4.495)	(3.356)	(3.229)
Has MBBS	24.325***	28.416***			14.516***	22.133***
	(6.644)	(7.997)			(4.605)	(4.195)
Has some qualification	4.444	5.399**			2.313	6.022***
	(3.276)	(2.139)			(2.929)	(2.197)
Patient load during visit	0.736	0.441	0.276	0.029	0.503	0.149
	(0.665)	(0.333)	(0.863)	(0.876)	(0.602)	(0.510)
Age of provider	-0.150	-0.103	0.233	0.226	-0.095	-0.018
	(0.144)	(0.091)	(0.231)	(0.214)	(0.119)	(0.083)
Gender of provider (1=Male)	-8.164**	-4.923	-1.101	-3.713	-7.474**	-3.098
	(3.497)	(4.969)	(4.845)	(5.460)	(2.918)	(4.069)
Constant		10.526		-11.589		3.386
		(6.561)		(12.095)		(5.913)
R2		0.393		0.466		0.361
Number of observations		543		152		695
Mean price charged		27.327		33.125		28.699
SD		26.079		28.580		26.851

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample and pooled sample, robust standard errors clustered at the location/market level are in parentheses. Observations are at the SP-provider interaction level. Interpretation of coefficents in "Binary regressions" needs caution. Each coefficient represents a separate regression of prices on the row variable and SP, case and district fixed effects. Multiple regressions include SP, case and district fixed effects. The pooled sample (Columns 5 and 6) combine the representative and dual practice samples.

Table 7: Wages in the public sector (public observations only)

	(1)	(2)	(3)	(4)
	Log of Monthly Salary (pooled sample)		**Desirability index** (PHC/CHC sample)	
	Binary regressions	Multiple regression	Binary regressions	Multiple regression
Percentage of checklist items	0.002	-0.001	0.004	0.003
	(0.003)	(0.002)	(0.009)	(0.009)
Time spent with SP (minutes)	-0.051**	-0.012	-0.061	-0.080
	(0.026)	(0.014)	(0.074)	(0.077)
Correct Treatment	0.055	-0.090*	-0.304	-0.132
	(0.066)	(0.048)	(0.237)	(0.202)
Has MBBS	1.055***	1.283***		
	(0.168)	(0.175)		
Has some qualification	-0.092	0.849***		
	(0.367)	(0.300)		
Age of provider	0.012**	0.019***	0.052***	0.062**
	(0.006)	(0.006)	(0.019)	(0.024)
Gender of provider (1=Male)	0.112	0.126	-0.530	-0.846
	(0.189)	(0.106)	(0.509)	(0.739)
Born in same district	-0.389***	0.015	-0.180	0.101
	(0.147)	(0.081)	(0.449)	(0.432)
Is a dual provider	0.582***	0.149*	0.076	-0.135
	(0.136)	(0.086)	(0.402)	(0.527)
Constant		8.044***		-1.470
		(0.316)		(1.198)
R2		0.625		0.165
Number of observations		301		182

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. Robust standard errors are in parentheses. The pooled sample (Columns 1 and 2) combine the representative and dual practice samples. The desirability index is a constructed using principal component analysis of proximity to several amenities (paved road, bus stop, railway station, Internet, post-office and bank), availability of infrastructure (stethoscope, spyghamometer, torchlight, weighing scale, hand washing facility, drinking water, staff toilet, patient toilet, fridge, sterilizers, electric connection, electric supply, power generator, telephone, computer, IV drip, cots/beds, disposable syringes), and PHC size (number of staff and number of patients). In binary regressions columns, each coefficient represents a separate regression of prices on the row variable, a constant and district fixed effects. Multiple regressions include district fixed effects.

Table 8: Real patients in the public and private sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Representative sample					Dual sample				
	Time spent (mins)	Total questions	Physical examination	Dispensed/ prescribed medicines	Number of medicines	Time spent (mins)	Total questions	Physical examination	Dispensed/ prescribed medicines	Number of medicines
Panel A: no patient or provider controls, and no fixed effects										
Is a private provider	1.456***	0.799***	0.371***	-0.026**	0.500***	1.894***	1.154***	0.143**	-0.007	-0.017
	(0.323)	(0.180)	(0.108)	(0.011)	(0.121)	(0.570)	(0.319)	(0.064)	(0.010)	(0.134)
R-squared	0.054	0.030	0.103	0.003	0.017	0.115	0.082	0.017	0.001	0.000
Number of observations	1,137	1,137	1,133	1,138	1,138	1,085	1,083	1,082	1,091	1,091
Mean of public	2.378	2.994	0.473	0.994	2.319	1.499	3.284	0.678	0.991	3.190
Mean of private	3.833	3.793	0.844	0.968	2.819	3.393	4.439	0.821	0.983	3.169
Mean of sample	3.621	3.676	0.790	0.972	2.746	1.899	3.527	0.708	0.989	3.185
Number of public providers	29	29	29	29	29	51	51	51	51	51
Number of private providers	169	169	169	169	169	40	40	41	41	41
Panel B: no patient or provider controls, and market/district fixed effects										
Is a private provider	1.626***	0.630***	0.503***	-0.016	0.674***	1.910***	1.155***	0.154**	-0.008	-0.013
	(0.490)	(0.170)	(0.112)	(0.014)	(0.167)	(0.561)	(0.315)	(0.061)	(0.009)	(0.139)
R-squared	0.163	0.162	0.218	0.090	0.167	0.120	0.101	0.074	0.006	0.016
Number of observations	1,137	1,137	1,133	1,138	1,138	1,085	1,083	1,082	1,091	1,091
Panel C: including patient and provider controls, and market/district fixed effects										
Is a private provider	1.190***	0.654***	0.522***	0.009	0.602***	1.570***	0.561**	0.072	-0.016	-0.016
	(0.313)	(0.246)	(0.085)	(0.014)	(0.145)	(0.547)	(0.255)	(0.060)	(0.012)	(0.166)
Has MBBS degree	-0.466	0.373*	0.159**	-0.025	-0.337					
	(0.462)	(0.217)	(0.079)	(0.016)	(0.206)					
Has some qualification	0.334	0.027	0.011	-0.035**	-0.178					
	(0.378)	(0.153)	(0.052)	(0.015)	(0.146)					
Age of Provider	-0.025**	0.008*	0.001	0.001	0.007	-0.003	-0.012	-0.002	-0.000	-0.017**
	(0.011)	(0.005)	(0.002)	(0.001)	(0.006)	(0.015)	(0.009)	(0.003)	(0.000)	(0.008)
Gender of Provider (1=Male)	-1.337*	-0.744	0.009	0.008	-0.016	-0.495	-0.040	-0.103	0.007	0.034
	(0.705)	(0.729)	(0.090)	(0.018)	(0.209)	(0.369)	(0.374)	(0.099)	(0.013)	(0.257)
Patient can easily dress themselves	-0.058	-0.307	0.009	-0.036	0.083	-0.679	-0.807**	-0.115	-0.002	-0.319
	(0.475)	(0.601)	(0.118)	(0.023)	(0.413)	(0.516)	(0.315)	(0.075)	(0.031)	(0.203)
Patient can easily do light household work	0.722***	0.023	0.027	0.024	-0.191	-0.267	0.505***	-0.009	0.026*	0.057
	(0.215)	(0.216)	(0.036)	(0.020)	(0.184)	(0.318)	(0.182)	(0.047)	(0.015)	(0.139)
Patient can easily lift a 5kg bucket and walk for 100m	-0.433	0.179	-0.025	0.005	-0.233**	0.149	-0.253	-0.036	-0.017**	0.103
	(0.266)	(0.181)	(0.032)	(0.010)	(0.112)	(0.197)	(0.170)	(0.050)	(0.008)	(0.118)
Patient can easily walk 200-300m	0.416	-0.304	0.081*	-0.027*	0.019	0.554*	0.245	0.097	0.013	0.121
	(0.262)	(0.225)	(0.048)	(0.016)	(0.150)	(0.285)	(0.155)	(0.064)	(0.008)	(0.154)
R-squared	0.303	0.331	0.348	0.119	0.306	0.168	0.356	0.197	0.042	0.155
Number of observations	835	835	833	835	835	809	808	807	810	810

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample, robust standard errors clustered at the provider level are in parentheses. Observations are patient-provider interactions, and the sample has been limited to the SP sample. The regressions in Panel C include controls for patients' characteristics and patients' presenting symptoms. Controls for patients' characteristics include: whether patient has no education, number of questions asked by patient, and patients' asset index. Controls for patients' presenting symptoms include: number of days patient has been sick, patients' ease in performing activities of daily living, and indicators for a number of presenting symptoms (fever, cold, diarrhea, weakness, injury, vomiting, dermatological problem, pregnancy, and pain). In columns (5) and (10) the dependent variable is total number of medicines recommended to the patient (dispensed and/or prescribed).

Appendices (Text, Figures, Tables)

A Mapping and Sampling of Providers

A.1 Mapping of Providers: Representative Sample

We first randomly selected five districts in the state of Madhya Pradesh, stratified by region and an index of health outcomes. In each district, we sampled 20 villages by probability proportional to size (PPS). Because of the rural focus of the study, we restricted the sampling frame to villages with populations under 5,000. The sample of villages is thus representative of rural Madhya Pradesh.

In each sampled village, we conducted at least three participatory resource assessments in different locations within the village and obtained a list of all the healthcare providers that households' sought primary care services from. These lists were used primarily to identify the geographical locations that households sought care from. For instance, households may seek care from providers within the village, but also on the nearest highway. If 5 percent or more of households reported visiting a provider in an outside location, we identified that location as a "cluster village" and considered it a part of the "healthcare market" for the sampled village. Fifty-five sampled villages have one cluster village, 13 villages have two, and one village has three. The remaining 31 villages have no cluster villages (i.e. less than 5 percent of primary healthcare visits were to a location outside the village). For our sample as a whole, we identified 184 locations, including the 100 sampled villages.

Surveyors then visited each location and administered a *provider census* to all healthcare providers in the location - regardless of whether they had been mentioned in the participatory resource assessments. The provider census collects details on the providers' demographics, and practice and clinic characteristics.

After the provider census in the villages, we administered a short *household census* and obtained information on household demographics and healthcare seeking behavior. For each household member, we asked about incidence of any illness in the past one month, if they sought medical attention for that illness, and (if yes), the name and address of the provider they visited (regardless of the location of the provider). Surveyors mapped the household visits to the providers lists; this is the mapping we use to compute the fraction of visits to public and private providers and providers with different qualifications in Table 1. In instances where households reported visiting providers not already on the list, we probed for the providers' names, addresses and practice details and added the providers to our listing and census exercise. We verified through this exercise that we had covered providers for at least 95 percent of all households in each village. This exhaustive mapping process ensured that we mapped the complete "healthcare market" where households in our sampled villages sought primary care services.

A.2 Sampling of Providers for SP visits: Representative Sample

To make the exercise tractable, we conducted the SP work in three out of the five districts in our sample. Although SPs were recruited from the local community, they needed plausible reasons for their presence in the village (which they were not from), and the typical narrative was that they were traveling and/or passing through the village. In order to minimize SP detection, we excluded 5 remote markets (as assessed by road access) from the possible 60 markets, where, after consultation with field staff, we believed that a traveling excuse might not be plausible.

We sampled providers for the SPs to visit from a smaller set of "eligible providers" than what we had mapped. All public nurses and midwives (ANMs), community health workers (ASHA), and day-care center workers (Anganwadi) were excluded from the sample as they primarily provide preventive care such as vaccinations. We also excluded mobile and itinerant providers, chemists, and pharmacists from the sample. Finally, we excluded 55 providers with whom we could not complete the provider census prior to sampling (typically due to the unavailability of provider, we were able to conduct the census with only 17 of these providers in subsequent rounds). These restrictions remove an additional 7 markets from our study, because there were no eligible providers in these markets. We also drop two other markets because they share a cluster with other sampled villages and do not have eligible providers inside the village. Our study in the representative sample therefore covers 46 markets in 3 districts of Madhya Pradesh (see Table A.1). Based on the eligibility criteria defined above, these 46 markets have 649 eligible providers (130 public and 519 private) from which we sample.

In each market we randomly sampled up to two eligible providers in each public clinic and up to six private providers in each market.[1] In the private sector, we sampled one provider per clinic. We also sampled all MBBS providers in both public and private sectors. Since the unit of analysis for the representative sample is the clinic and not the provider, this sampling procedure was equivalent to sampling private clinics with simple random sampling (after sampling all private MBBS doctors), and sampling larger public clinics (those with two or more eligible providers) twice. We sampled a total of 247 providers of which 45 are public providers and 202 are private providers (Appendix Table A.1).

[1]One market in Gwalior district was an exception to this rule. In the cluster village of a particular market, we found 113 providers. In this market, we relaxed our sampling protocol and sampled 20 private providers.

A.3 Completion of SPs: Representative Sample

Based on the sampling methodology described above, we sampled 247 provider-clinics for the SP work in the representative sample. Since SPs sought care from whoever was practicing at the time of the visit, the relevant unit here is the provider-clinic. The sampled providers belong to 235 clinics, and the total number of unique providers sampled is 242 (5 sampled providers practice from multiple clinics and we treat these as different provider-clinic combinations for sampling). Of the 247, SPs completed at least one case in 224 provider-clinics for a completion rate of 91 percent.

At the case-level, SPs saw providers who were not originally sampled but were mapped in the first round in 27 interactions. Furthermore, for 18 interactions (corresponding to 8 public and 2 private clinics) SPs saw providers that we had not mapped and we do not know the identity of the provider. These were most likely staff present in the clinic who are not licensed to provide care, but who do so when the doctor is absent. The discrepancy between whom we sampled and whom we actually saw does not affect interpretation of our results in Panels A and B of Tables 3-5, but it does in Panel C, where we include controls for provider characteristics. Panels A and B present results without provider controls, so whether or not we have background data on the provider is irrelevant, because we know which market they were practicing in and whether they were public or private. This is why the public-private difference here should be interpreted as the difference in random visits to providers' clinics rather than providers. In Panel C, we present results including provider controls. Here, for 27 interactions where we saw providers we did not sample but mapped (and interviewed during the provider census), we use their background information. The 18 observations where we do not know the provider at all are dropped from the estimation sample.

A.4 Mapping of Providers: Dual Practice Sample

We obtained a list of all Primary Health Centers (PHCs) and Community Health Centers (CHCs) from the Ministry of Health of Madhya Pradesh. Excluding PHCs/CHCs which were mapped as part of the representative sample, we mapped 200 more facilities in this round. Of these 200 facilities, 40 did not have an MBBS provider posted (see Appendix Table A.2). In the remaining 160 PHCs/CHCs we located 216 providers (some providers were mapped to multiple facilities). Our field team then undertook detailed field work to find out if these providers operated private practices and, if yes, to locate their private practices. We were able to locate a private practice for 132 of the 216 providers (61.1 percent) (this is the sample we call the "dual practice sample"). After the mapping, we administered the provider census to all providers. To the extent possible, the census was administered in the private clinic of the provider.

A.5 Sampling of Providers: Dual Practice Sample

We sampled one MBBS doctor from every PHC/CHC with preference for one with a dual practice when there were multiple MBBS doctors in the clinic. In cases where a provider was posted to multiple public facilities, and where there were no additional MBBS providers in these facilities, we randomly sampled the provider from one of the multiple facilities they were posted to. With this sampling strategy, we sampled from 139 of the 160 facilities we could have sampled from. Of the 139 providers, 91 operated private practices (65.5 percent, see Table A.2).

A.6 Completion of SPs: Dual Practice Sample

SPs completed interactions with 116 of the 139 providers sampled. The main reason for non-completion is that providers were absent or were away on "long leave" in the 6-month phase between the listing and the SP work. We made up to 3 (and in one case 4) attempts to complete the SP-case interaction, and were forced to stop trying at that point due to the heightened risk of detection. Of the 48 providers without private facilities, SPs completed interactions with 32 providers (66.7 percent). Of the 91 providers with private practices, SPs were able to complete at least one interaction with 84 providers (92.3 percent, either public or private, Panel B2 of Table A.2).

The number of dual practice doctors sampled is 91, with 227 cases allocated to the public clinics and 228 to private clinics (we randomly assigned the unstable angina case to either the public or a private clinic). Completion rate in the dual practice sample varies by sector due to high absence rate of doctors in the public clinics (see Panel C2 of Table A.2). Of the 91 public doctors, we successfully completed at least one case with 78 percent. In the private sector, we completed at least one case with 92.3 percent. At the case level, completion rates for public and private doctors in the dual sample was 74 percent and 90 percent respectively. The number of dual practice providers for whom we have at least one observation in both their public and private practice is 70. We discuss the robustness of our results to differential non-completion of SP cases across public and private clinics in Appendix D.1.

B Standardized Patient Data Collection and Notes

B.1 Description of Tracer Conditions and Relevance for India

SPs presented either a case of unstable angina, asthma, or dysentery of an absent child.

- Unstable Angina: A 45-year-old male complains of chest pain the previous night. Appropriate history taking would reveal classic signs (radiating, crushing pain) and risk factors (smoking, untreated diabetes, and family history of cardiac illness) of unstable angina or an imminent myocardial infarction.

- Asthma: A 25-year-old male or female presents with difficulty breathing the night before the visit. When questioned appropriately, the SP reveals that the episode lasted for 10 to 15 minutes and involved a "whistling" sound (wheezing) and that he or she has had similar episodes before, often triggered by house cleaning and cooking smoke. The SP also reports a family history of similar symptoms.

- Dysentery: A 26-year-old father of a 2-year-old complains that his child has diarrhea and requests medicines. When probed, the SP reveals details of their water source and sanitation habits, in addition to the presence of fever and the frequency and quality of the child's stools.

For all cases, checklists of recommended history questions and examinations were developed together with an advisory committee and SPs were trained to recall the questions asked and examinations performed. These were then recorded during a debriefing with a field supervisor using a structured questionnaire within an hour of the interaction. In a recent study, we test the reliability of recall by comparing audio recordings with recall and find a very high correlation of 0.63 ($p<.001$) (Das et al., 2015*a*).

B.2 Relevance of Cases

Incidence of cardiovascular and respiratory diseases has been increasing, and diarrheal disease kills more than 200,000 children per year in India (Black et al., 2010; Patel et al., 2011) The Indian government's National Rural Health Mission (NRHM) has developed triage, management, and treatment protocols for unstable angina, asthma, and dysentery in public clinics, suggesting clear guidelines for patients presenting with any of these conditions (Jindal et al., 2005). The cases were also chosen to minimize risk to standardized patients since they could not portray any symptoms of infection given the documented high propensity to administer medicines intravenously with unsterilized needles and to use thermometers that have not been appropriately disinfected (Banerjee, Deaton and Duflo, 2004).

B.3 SP Recruitment, Script Development and Training

A total of 15 individuals were selected from an initial group of 45 who were extensively screened and trained for 3 weeks. The age and sex of recruited SPs corresponded to the relevant tracer conditions. For instance, angina was depicted by male SPs between 40 and 50 years old.

Scripts were developed under the guidance of a medical anthropologist with active SP participation that described the social and family contexts of the patient if a provider were to ask questions about these details. Joint script development and SP training ensured that the clinical symptoms and case history reflected the social and cultural milieu of which the SP was assumed to be a member and, second, the presentation of symptoms and answers to history were consistent with biomedical facts about the disease. SPs were trained to present symptoms and answer questions pertaining to case history that were medically correct. For example, all opening statements and questions pertaining to the type of cough and its duration were standardized. SPs were also trained to distinguish between questions to which answers could be improvised but had to be appropriate to the social role of the SP and answers that had to be given using local idioms but in a standardized format without any alterations.

All SPs underwent rigorous training for 100-150 hours that started with a focus on the cases and the development of scripts and proceeded to memorization and appropriate role-playing, as well as techniques to perfect recall of the questions asked and examinations completed during the interaction. Following the training, SPs visited doctors who were working with our team to provide feedback on their presentation and depiction of the cases. Finally, dry runs were completed with unannounced visits to consented providers to help build the confidence of the SPs and take them through a number of "real-life" situations. Field work started once protocols were in place for the variety of these experiences.

With consent from the Institutional Review Board at Harvard University, the study was first piloted in Delhi with 64 consented providers who had been previously informed that they would be visited by an SP within the next 6 months (see Das et al. (2012)). In the pilot phase of the study, a total of 248 out of a potential 256 SP interactions were completed. Within a month of the SP visit, field-workers visited the consented providers to inquire if they had been visited by an SP. In cases where the provider felt that an SP visit had occurred, we elicited the sex, approximate age and symptoms of the SP. We could confirm a match between the providers' suspicions and the actual SP sent to the provider in only 2 cases for a detection rate of less than 1 percent.

The Institutional Review Board of Innovations for Poverty Action and the Central and

State governments in India granted permission for the overall study. To minimize detection in rural Madhya Pradesh, where providers are more likely to recognize their entire patient population, the study proceeded as an audit, and providers were not aware that they were being visited by standardized patients. The Institutional Review Board at Innovations for Poverty Action granted clearance for this deception design. Clearance was granted because the risks to providers and their patients were minimal, whereas accurate measures of provider practice were nonexistent. The expected length of clinical interactions, patient loads, and levels of provider anxiety induced by the cases were thought to be small, and standardized patients had to pay providers whatever they charged. The waiver of consent is consistent with the principle that where the research subject provides a public service to other customers, the public have a right to know about the quality of the service provided (Norris, 2002).

B.4 Categorizing Treatment in SP Interactions

In rural Madhya Pradesh, as in much of India, providers often dispense medicines in the clinic rather than prescribe them for purchase from external chemists (some do both). Our field staff recorded names of all dispensed/prescribed medicines in SP exit interviews and used multiple resources to classify medicines as accurately as possible. Field staff were given a list of commonly used drugs in India along with their medical classification, and the CIMS Drug Information System (in print), which they used to record exact medicine names and classes. For drugs that were not immediately confirmed, they consulted local chemists and pharmacists and obtained correct names to the extent possible.

To construct our main treatment variables - correct, palliative, and unnecessary/harmful treatment - we obtained from a panel of doctors in the United States and India a full list of correct and palliative treatments/medicines for each case. These include nitrates, aspirin, clopidogrel, anti-platelet agents, blood thinners, beta blocker, morphine, other pain control, ACE inhibitor, and vasodilator for unstable angina; ORS, electrolytes, antibiotics, and zinc for dysentery; and inhaled-corticosteroids, leukotriene inhibitors, cromones, inhaled-anticholinergics, and oral-corticosteroids for asthma (see Table A.5).

After medicine coding in the field, the authors and members of the ISERDD team in Delhi verified the codes assigned to all medicines and recoded if them when necessary. To further ensure the coding was correct, we used a third party, a pharmaceutical consulting firm in Delhi, to independently verify our classification of medicines.

Medicine coding is relatively straightforward in instances where providers prescribe and SPs receive a written prescription. In cases where providers dispense, it was easier to obtain names when medicines came with packaging than when they did not. In the 1,123 complete SP interactions, SPs were recommended a total of 2,772 medicines corresponding to 969 unique medicines (by medicine names, ignoring unlabeled ones). We are unable to classify 14.18 percent of the all 2,772 medicines because they were unlabeled (providers dispensed them as loose samples or in crushed powder form). We are further unable to classify 3.64 percent of medicines (93 unique medicines by name) because we could not match them to secondary information sources. SPs received at least one unclassifiable medicine in 268 interactions (23.9 percent of all interactions). However, in 211 of these interactions (18.8 percent), SPs received classifiable medicines along with the unclassifiable medicines. In only 57 interactions (5.1 percent) were all medicines unclassifiable.

We construct our main treatment variables - correct treatment, palliative treatment and unnecessary treatment - after completing the medicine coding process described above. For each interaction, we determine if any recommended medicines fall into correct, palliative and/or unnecessary treatments, treating all unlabeled and unidentifiable medicines as unnecessary. It is possible that the unlabeled and unidentifiable medicines are really correct or palliative treatment. However, the likelihood that the provider dispenses an unclassifiable medicine is decreasing in other measures of provider quality from the SP study. We are therefore confident that such medicines are more likely unnecessary treatments than not. Our results are also robust to excluding interactions that include unclassifiable medicines.

C Theoretical Appendix

C.1 Problem Setup

A patient visits a provider endowed with a level of medical knowledge K, and presents a set of symptoms (this would correspond to the opening line of the SP script). The patient has a true illness denoted by n^{true}. Patients with different underlying illnesses may experience and present similar symptoms. In other words, given a set of symptoms, there is a distribution of n^{true} associated with the symptoms (we assume for analytical tractability that this distribution can be expressed on a single-dimensional real line, with n^{true} being a point on this line). A provider's job is to identify the true state of the patient and perform adequate treatments. The provider-patient transaction is modeled as a two-stage process: consultation and treatment. A subscript i for the i^{th} provider is used when there is a need to emphasize heterogeneity among providers, but is suppressed otherwise for notational simplicity.

C.2 Consultation Stage

The patient visits a provider and describes her symptoms based on which the provider forms a prior belief about the true illness that follows a normal distribution:

$$n^{prior} \sim N\left(\nu, \frac{1}{\alpha}\right)$$

The true illness of the patient (n^{true}) is unobserved to both the patient and the provider, and the prior belief can be thought of as the provider's belief about the distribution of illnesses in the region which cause the given symptoms. The provider exerts costly effort e to learn about n^{true}. We can interpret e as the number of checklist items completed by the provider or time spent with the patient. The provider draws a noisy signal $s \sim N(n^{true}, \frac{1}{\beta})$ by exerting e where $\beta = eK$. Thus, the marginal return to effort in terms of increased signal accuracy is higher when the provider's medical knowledge K is higher.[2] We assume a quadratic cost-of-effort function, with the cost of effort being equal to e^2.

The patient can observe the amount of effort expended (e) but cannot observe the signal (s) drawn by the provider as a result of the effort. Given s, the provider updates his belief about n^{true}. The posterior belief of the true state is given by:

$$n^{post} \sim N\left(\underbrace{\frac{\alpha\nu}{\alpha+\beta} + \frac{\beta s}{\alpha+\beta}}_{\equiv\mu}, \frac{1}{\alpha+\beta}\right)$$

where μ is the posterior mean. This is the result of standard Bayesian normal updating, and hence, a separate proof is omitted. Note that $n^{post} \to n^{true}$ as $\beta \to \infty$.

C.3 Treatment Stage

In the second stage, the provider makes treatment choices based on the posterior belief about the true state. The choice of treatments is expressed as an interval $[\mu - n, \mu + n]$, which maps into the empirical observation that most providers in our setting provide multiple medications. A wider range of treatments has a higher probability of covering the true illness and curing the patient of the current ailment but also increases long-term health costs.[3] The

[2]Note that the marginal return to e on signal accuracy diminishes as e becomes larger as illustrated in Figure A.6 (Panel B). Also, as in Rosenzweig (1995) a doctor with more knowledge may also have a more accurate prior to begin with, in addition to learning more with additional effort. We abstract away from this point to focus on deriving predictions for effort, treatment, and health outcomes for the same doctor across public and private practices. This corresponds to our dual sample.

[3]This assumption can reflect multiple channels, including adverse reactions to unnecessary drugs, the building of resistance to drugs that are not needed now but may be useful in future, or by the potential for

long-term health cost of excessive medication is modeled as $h(n) = n^2$

Let F_e denote the cumulative density function of the posterior belief given some level of effort e. Given K, the shape of the posterior belief is governed by e (e and β are used interchangeably when K is fixed). The probability that the interval $[\mu - n, \mu + n]$ includes n^{true} is denoted by $P_e(n)$ where $P_e(n) = F_e(\mu+n) - F_e(\mu-n)$. The patient's expected health outcome given n (which is a function of e) is H, which is given by $H(e, n) = P_e(n) - n^2$.

Note that for each individual patient, the interval either includes the true state or not with probability of $P_e(n)$ and $1 - P_e(n)$. Thus the optimal outcome for a patient is to receive only the correct treatment, and not receive any additional unnecessary treatments, and we can think of a high-quality provider as someone who provides this outcome, enabled by a precise posterior distribution of the true illness.

In practice, providers will choose effort and treatments to maximize their own utility, which may not be aligned with those of patients. We model provider utility as having two components. First, providers care about curing their patients and overall patient health. This can be attributed partly to altruism, intrinsic motivation to do the right thing, training and professionalism (Hippocratic oath), peer pressure and monitoring, and the liability and malpractice regime. We capture all of these factors with the parameter ϕ, which should be thought about as representing the extent to which providers value patient health in their own utility function in a setting without high-powered financial incentives. Thus, a higher ϕ represents greater alignment between provider and patient utility.

Second, providers also care about financial rewards, which in turn depends on how they are compensated. Under market pricing, providers can charge a consultation fee (τe) that is a function of a piece rate τ_i (determined by their qualifications and reputation) and effort expended (which is observable to patients), and a dispensing fee that increases linearly with the number of medicines provided (this is consistent with the correlates of market prices reported in Table 6).

They also have an incentive for improving patient health because this helps build their reputation and raises future demand (which we can think of as an increase in their consulting piece rate over time). However, patients can observe whether they were "cured" more easily than the costs of excessive medication, and this creates an incentive to over-treat because over-treatment increases the probability of spanning the true illness and providing a correct treatment. We denote the observed health outcome as $H^o(e, n)$, and true health as $H(e, n)$. Note that the idea that there is wedge between what patients consider to be optimal treatment and what a medical professional would consider optimal can be motivated in several

adverse interactions between drugs.

ways including differences in observability as well as by present-biased patients.[4]

C.4 Providers' Optimization Problem with and without Market Incentives

Denote the maximized utility of providers in the consultation stage and treatment stage by V_1 and V_2 respectively. Without market incentives, providers have low-powered incentives and maximize their utility:

$$V_1 = \max_e \left\{ -e^2 + V_2(e) \right\} \tag{1}$$
$$V_2(e) = \max_n \{\phi H(e,n)\}$$

where ϕ governs the extent to which providers care about patients' health without high-powered incentives.

In a market environment, providers face market incentives in addition to low-powered incentives. Now, a provider i charges a piece rate τ_i per unit of effort as a consultation fee and also charges p per unit of n for the treatment (we can think of p as the profit margin on medicines dispensed or the commission on medicines prescribed). Providers also care about their reputation in the market, which is determined by the health outcomes of their patients. Health outcomes are not fully observed in the market because the long-term health cost of excessive treatment is not as easily observed as the immediate relief of symptoms. Instead, reputation is based on the observed health outcome H^o, which is given by $H^o(e,n) = P_e(n) - \gamma_o n^2$ where $0 < \gamma_o < 1$, and δ, which is a parameter that governs the extent to which providers care about their reputation in the market. When there are market incentives, providers maximize their utility given by:

$$V_1(\tau_i) = \max_e \left\{ -e^2 + \tau_i e + V_2(e) \right\} \tag{2}$$
$$V_2(e) = \max_n \{\phi H(e,n) + \delta H^o(e,n) + np\}$$

[4] In an earlier working paper version (Das et al. (2015*b*)), we incorporate a third channel that providers care about - which is responding to patient-driven demand. Patients have their own expectation about proper treatment, and providers may satisfy patients by meeting their demand for medication in order to avoid a communication cost of explaining to patients that they do not need the treatment that they seek. We drop this extension here because our data does not allow us to contribute any empirical insights regarding this channel. We also assume that private providers have dynamic incentives to acquire a positive reputation, but we do not endogenize these market incentives since a static framework is adequate to interpret our empirical findings. A theoretical extension where we provide one potential way of endogenizing market incentives is available on request but is also omitted here because our data do not allow us to study the dynamics of reputation and price setting.

The first order conditions without market incentives are given by:

$$\phi f_e(\mu + n(e)) \frac{n(e)K}{\sqrt{\alpha + eK}} = 2e \tag{3}$$

$$f_e(\mu + n) = n \tag{4}$$

where f_e is the probability density function of the posterior belief given e. The term $f_e(\mu+n)$ captures the marginal benefit of increasing n through the higher probability of spanning the correct treatment, and the right hand side is the marginal cost of increasing n through the higher health cost of excessive treatment. In the absence of market incentives, note that providers choose n which maximizes H at any given e.

The first order condition in the consultation stage with market incentives is given by:

$$\tau_i + (\phi + \delta)\, f_e(\mu + n(e)) \frac{n(e)K}{\sqrt{\alpha + eK}} = 2e \tag{5}$$

and the first order condition in the treatment stage is given by:

$$f_e(\mu + n) + \frac{p}{2\,(\phi + \delta)} = \left(\frac{\phi + \gamma_o \delta}{\phi + \delta}\right) n \tag{6}$$

It is easy to see from (4) and (6) that given e, providers choose larger n when there are market incentives. Because there is a pecuniary benefit from n and also because the cost of excessive n is not fully observed in the market ($\gamma_o < 1$), given e, the marginal benefit of n is always greater and the marginal cost is always smaller with market incentives. Thus, providers choose excessive n where H is decreasing in n instead of where H is maximized. This means that by slightly decreasing n, the health outcome can be improved.

Whether market incentives induce higher effort depends on the relative size of the rewards for e and n in the market. As long as the rewards for n are not so large so as to dominate those for e, providers choose higher e with market incentives (see Das et al. (2015*b*) for the proof). Since our empirical results find that private providers always exert more effort (in both the representative and dual samples) and we also find a robust positive relationship between prices charged and effort expended, it appears that the τ in our setting is high enough to induce additional effort from providers facing market incentives.

C.5 Market Incentives and Health Outcomes

However, while provider effort may be higher under market incentives, the impact of market incentives on health outcomes is ambiguous and will depend on parameter values. In particular, when ϕ is very low, it is possible that health outcomes under market incentives are better; however, as ϕ increases, health outcomes without market incentives may be better.

Figures A.6 and A.7 in the illustrate this mechanism. Panel (A) in Figure A.6 illustrates a case where market incentives induce higher effort. MB_{with} and $MB_{without}$ are the left hand side of (5) and (3) with respect to e. MC_{with} and $MC_{without}$ are the right hand side of (5) and (3) with respect to e. The terms e^*_{with} and $e^*_{without}$ are the optimal levels of effort with and without market incentives, respectively, for small and large ϕ values. The rewards for effort in the market are sufficiently large in this case that e^*_{with} is larger than $e^*_{without}$. With larger ϕ the optimal choice of e is higher.

Panel (B) traces posterior variance $\frac{1}{\alpha+\beta}$, the inverse of posterior precision, as a function of e holding K constant. The y-intercept, $\frac{1}{\alpha}$, is the posterior variance when $e = 0$. The term $\frac{1}{\alpha+\beta}$ decreases with e at diminishing rates because $\beta = eK$. When ϕ is small, a difference in e is translated into a substantial difference in $\frac{1}{\alpha+\beta}$. When ϕ is large, the marginal effect of effort on $\frac{1}{\alpha+\beta}$ is small.

Panel (C) illustrates the optimal level of treatment with and without market incentives, n^*_{with} and $n^*_{without}$, when the posterior variance with market incentives is substantially smaller than that without market incentives. MB_{with} and $MB_{without}$ are the left hand side of (6) and (4) with respect to n. MC_{with} and $MC_{without}$ are the right hand side of (6) and (4) with respect to n. The slope of MC_{with} is smaller than one because the health cost of excessive treatment is not fully observed, and hence, penalties for additional treatment in the market are weaker than what providers would impose on themselves under low-powered incentives. p, the unit price of n, is added to MB_{with}, so MB_{with} asymptotes to $\frac{p}{2(\phi+\delta)}$ rather than to 0. When the posterior variance with market incentives is substantially smaller than that without incentives, the optimal level of n with market incentives can be smaller in spite of incentives for excessive treatment. Panel (D) illustrates the optimal level of treatment when the posterior variance with market incentives is only slightly smaller than that without market incentives. In this case, the effects of market incentives on excessive treatment dominate, and the optimal level of n is larger with market incentives.

Figure A.7 illustrates the health outcome produced with and without market incentives with different values of ϕ. H increases with ϕ because e increases with ϕ, and n is invariant to ϕ given e when there are no market incentives and decreases with ϕ when there are market incentives. At low levels of e, a small difference in e is translated into a substantial difference in the posterior precision. Although market incentives induce excessive n, the effect of higher posterior precision on the health outcome dominates the offsetting effect of excessive n. However, as ϕ increases, e under both environment increases, and the marginal effect of e on the posterior precision, and hence on the health outcome, becomes smaller. At sufficiently high levels of e, higher e with market incentives generates a difference in the posterior precision that is too small to offset the effect of excessive n. Thus, when ϕ is high, the health outcome without market incentives is higher.

This may be typical in high-income countries with better oversight of medical training and practice, which is the context where Arrow (1963) is implicitly set. However, in settings with very low ϕ as seen in India and other low-income countries - exemplified by high doctor absence rates (Chaudhury et al., 2006) - it is possible that market incentives may lead to better outcomes.[5] Thus, an important goal of our theoretical framework is to illustrate how ideas about the optimal organization of healthcare that may have been developed in high-income settings may not apply equally to low-income settings with weak state capacity for running a well-functioning public health system.

D Differential Case Completion and Patient Sorting Across Sectors

D.1 Differential Case Completion Across Public and Private Sectors

As we mention in the text, in the dual sample, SPs were more likely to complete an interaction with MBBS doctors in their private clinics than in their public clinics due to the higher absence rates of doctors in their public sector practices. The differential completion rates could bias our estimates comparing the quality of care across public and private practices of the same doctor (the problem is exactly analogous to differential attrition from treatment and control groups in a randomized experiment). If doctors who are more absent in their public practice also provide poorer care when they are present, our estimates of the public-private differences would represent a lower bound of the true differences. Conversely, if doctors who are more absent from public clinics provide better care when they are present, our estimates will be inflated.

Our data allows us to directly test for the likely direction of this bias, because we can compare effort and treatment outcomes by whether or not the case was completed in the

[5] See Muralidharan and Sundararaman (2011) for an adaptation of the multi-tasking framework of Holmstrom and Milgrom (1991) and Baker (1992) that yields similar insights in the context of performance-linked pay for teachers (showing that outcomes could improve under performance pay if the default level of teacher effort was low, but could worsen if the default level was high). A key difference in our context is that the high-powered incentives do not come from administratively set performance-linked bonuses, but market rewards for effort and reputation.

first attempt in each sector or whether additional visits were needed. Panel A of Table A.13 reports means of effort and treatment outcomes by number of attempts and by sector. Panel B presents these differences in a regression format including case and SP fixed effects. In the private sector, we find no difference in either the IRT-score for checklist completion, or the likelihood of providing a correct treatment as a function of whether SPs managed to complete the case in the first attempt or made additional visits to do so. However, in their public practices, doctors who were not found on the first visit had significantly lower IRT scores and likelihood of providing a correct treatment. Thus, doctors who are more absent in the public sector are likely those who exert lower effort even when they are present. The coefficient on the interaction between "public practice" and "completed in first attempt" in Panel B formalizes this and shows that public doctors who were present on the first visit had significantly higher IRT scores and likelihood of providing a correct treatment.

To account for potential bias from differential non-completion of cases across public and private practices in the dual sample, we present re-weighted results in Table A.14. In each sector, we impute missing values (where cases were not completed due to doctor absence) with the average of outcome variables for those providers with whom cases were completed after multiple visits (the averages are calculated separately across public and private practices) and re-estimate equation 1. Panel A presents the original estimates (corresponding to Tables 3 and 5) and in Panel B we report the re-weighted estimates. For each effort measure, the re-weighted estimates are larger than the original estimate (although they are not statistically different). Results are similar for the correct treatment outcome - the re-weighted estimate is 20.3 percentage points, which is larger than the original estimate of 15.1 percentage points. Overall, these results suggest that differential case completion across sectors attenuates our main results on effort and correct treatment, and that the estimates presented in the paper are likely to be a lower bound of the true differences in quality of care across public and private practices of the same provider.

D.2 Differential Patient Sorting Across Public and Private Sectors

As discussed in Section 7.2, a further consideration in interpreting our results is the issue of statistical discrimination. Specifically, while the use of SPs allow us to control for differential case mix across public and private providers by presenting the same case in both settings, it is possible that the cases presented may have been off the equilibrium path for either or both public and private clinics in this setting. Even if the presented cases map well into the overall morbidity patterns and care seeking behavior of the population, it is possible that patients choose to visit different provider types (public or private) for different types of conditions. Patients may choose public facilities for more serious conditions, or vice versa.

If there are large systematic differences in the type of patient and case that is presented to public or private clinics, the quality of care differences we record across public and private clinics may partly reflect statistical discrimination.

Note that this is a very difficult problem to address in general because observing real provider-patient interactions precludes the concern of off-equilibrium behavior, but we cannot code the quality of care accurately because we do not know the underlying ailment. On the other hand, the SP method allows for better measurement of quality of care, but may represent an off-equilibrium interaction. But, it is challenging to solve both problems simultaneously. This is why we present results from both approaches in the main text and show that the results are consistent across Tables 3, 5 (SP) and Table 8 (real patient observations).

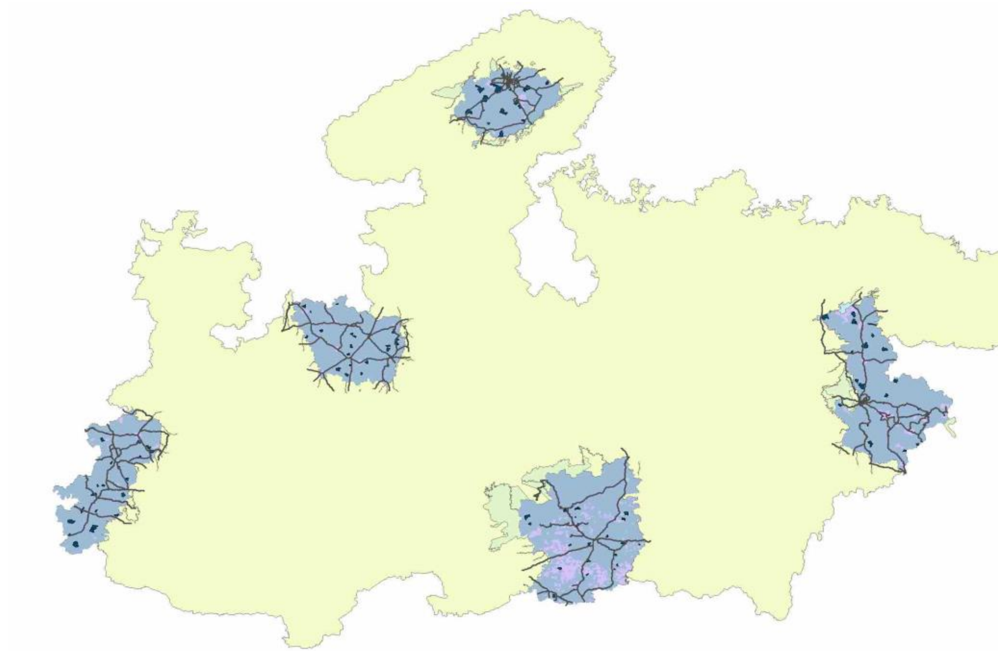
Here, we provide evidence against differential patient sorting using more data. In addition to observing real patient interactions (as described in section 7.1), we conducted patient exit interviews immediately after their provider interactions, where we asked patients the reasons for their visit, including a list of symptoms, their morbidity levels (measured by their ease of conducting activities of daily living), and other background and demographic questions. In Table A.17 we present estimates of differences in patient characteristics across public and private clinics. For the representative sample, for each outcome variable (rows of the table), Columns (1) and (2) present means in the public and private sectors respectively, and Columns (3) and (4) present coefficients from regressing the outcome variable on a private indicator with and without market fixed effects. Columns (5)-(8) repeat the same exercise for the dual sample except that we use district fixed effects instead of market fixed effects.

Overall, we see that for almost all illness symptoms, patients are equally likely to go to either a public or private provider (in both representative and dual practice samples). Out of the 18 patient and case characteristics comparisons presented in columns 4 and 8, we find significant differences in only two. Similarly, we find no difference in average morbidity among patients visiting public and private clinics (as measured by activities of daily living).

Where we do find some difference is in patient affluence and education (especially in the dual sample), which is not surprising because MBBS providers charge a higher fee. However, as we discussed in the main text, the optimal *initial* effort and treatment in the cases we chose should not depend on the patients ability to pay for *follow up* treatments (for instance in the angina case, the patient could be given an aspirin and referred to a public hospital, which would have been coded as a correct treatment). Overall, the similarity in the type and intensity of symptoms presented across public and private clinics suggest that differential patient sorting across case type is not likely to affect our results.

Appendix Figures and Tables

Panel A: Location of Madhya Pradesh in India



Panel B: Sampled Districts of Madhya Pradesh – Chhindwara, Gwalior, Jhabua Rajgarh, and Shahdol

Figure A.1

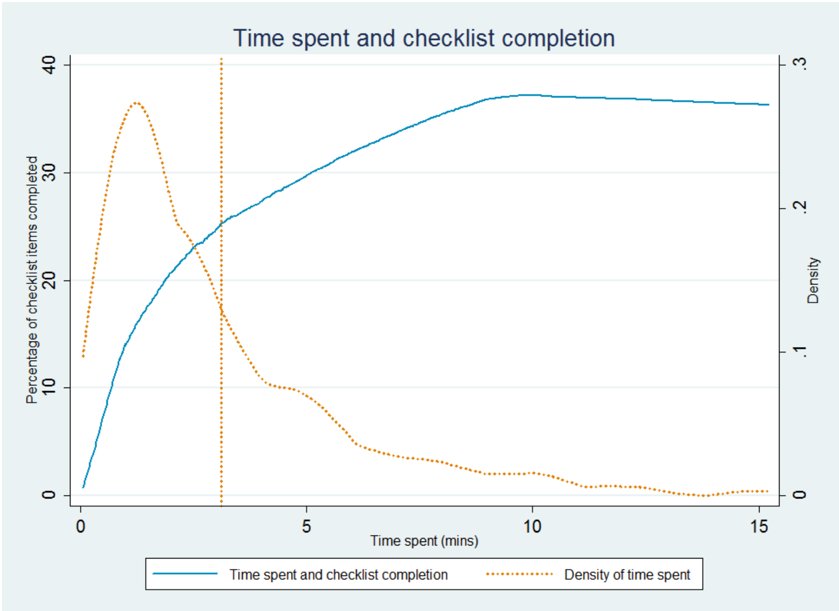


Figure A.2

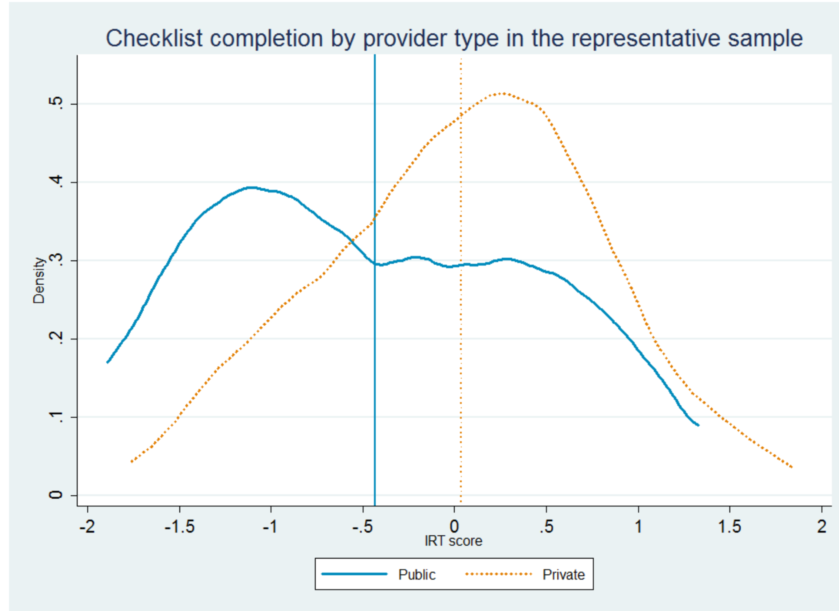


Figure A.3

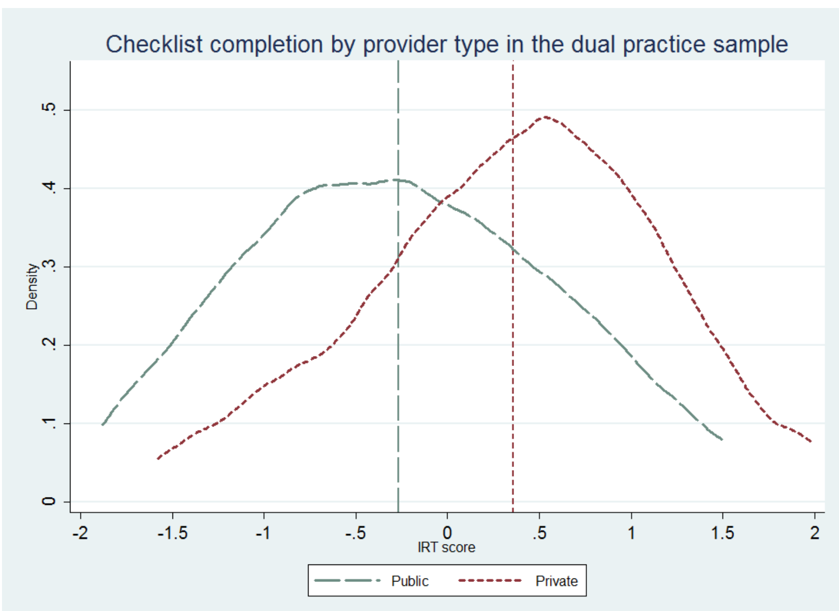


Figure A.4

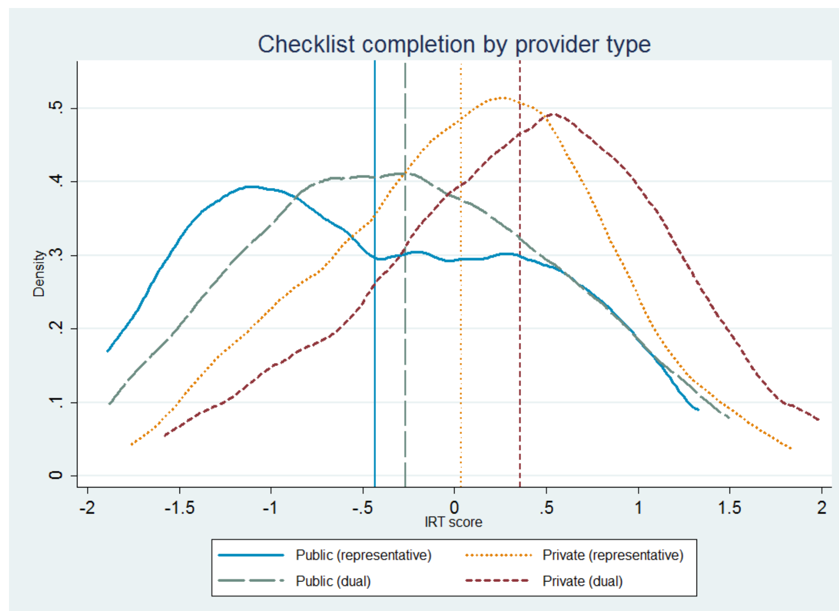
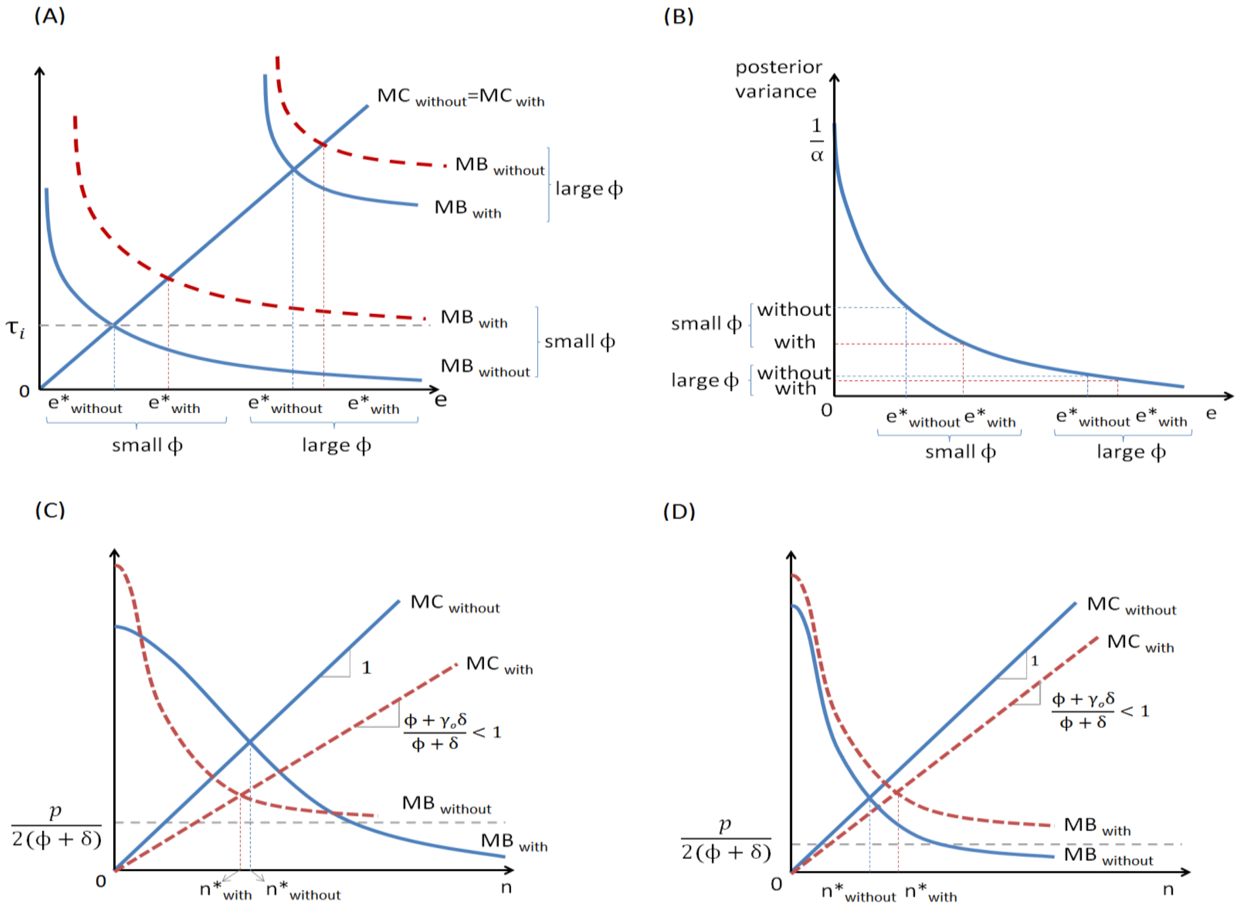


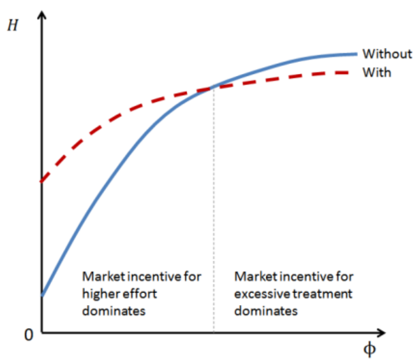
Figure A.5

Figure A.6: Optimal choice of effort and treatment with high and low ϕ with and without market incentives



Notes: In panel (A), $\mathrm{MB_{with}}$ and $\mathrm{MC_{with}}$ are the marginal benefit and the cost of e with market incentives, and $\mathrm{MB_{without}}$ and $\mathrm{MC_{without}}$ are those without market incentives. e^*_{with} and e^*_{without} are optimal levels of effort with and without market incentives for small and large ϕ values. In panel (B), The graph traces the posterior variance $\frac{1}{\alpha+\beta}$ with e on the x-axis. The y-axis intercept $\frac{1}{\alpha}$ is the posterior variance when $e = 0$. In panel (C) and panel (D), $\mathrm{MB_{with}}$ and $\mathrm{MC_{with}}$ are the marginal benefit and the cost of n with market incentives, and $\mathrm{MB_{without}}$ and $\mathrm{MC_{without}}$ are those without market incentives. n^*_{with} and n^*_{without} are optimal levels of treatment with and without market incentives for small and large ϕ values. Panel (C) and panel (D) compares the optimal level of treatment with and without market incentives when the posterior variance with market incentives is substantially smaller than that without market incentives and when the two posterior beliefs are similar.

Figure A.7: Health outcome with and without market incentives with varying ϕ



Notes: The graph illustrates the health outcome produced with and without market incentives with different values of ϕ. The y-axis is the health outcome H and x-axis is the magnitude of low-powered incentive, ϕ. The solid line traces H without market incentives and the dotted line traces H with market incentives.

Table A.1: Sampling and completion of SPs in the representative sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Markets	Number of providers			Number of MBBS providers		
		Total	Public	Private	Total	Public	Private
Panel A: Sampling and completion by market							
Total eligible	60	719	144	575	51	23	28
Markets selected for SP	46	649	130	519	50	23	27
Reasons for not sampling market							
Remote market	5						
No eligible provider	7						
Common cluster market, no provider within village	2						
Sampled for SPs		247	45	202	28	12	16
Not sampled for SPs	14	472	99	373	23	11	12
Completed SPs	46	224	36	188	23	9	14
Panel B: Sampling and completion by sector							
Public Sector		(Number of providers with whom SPs were completed)					
At least 1 public provider sampled	22	151	36	115	20	9	11
At least 1 public provider completed	20	141	36	105	20	9	11
At least 1 public MBBS provider sampled	10	98	21	77	18	8	10
At least 1 public MBBS provider completed	9	87	19	68	18	9	9
Private Sector							
At least 1 private provider sampled	44	218	30	188	22	8	14
At least 1 private provider completed	44	218	30	188	22	8	14
At least 1 private MBBS provider sampled	8	68	5	63	16	2	14
At least 1 private MBBS provider completed	7	67	5	62	16	2	14
Private and Public Sector							
Markets with at least 1 public and 1 private provider sampled	20	145	30	115	19	8	11
Markets with at least 1 public and 1 private provider completed	18	135	30	105	19	8	11

Notes: In 5 markets where SP work was over completed, the SP saw a provider other than a sampled provider

Table A.2: Mapping, sampling and completion in the dual practice sample

	(1) Number of Facilities		(2)	(3) Providers	(4)	(5)	(6) Cases	(7)
			Number of providers	Percentage of total	Percentage of sampled	Number of cases	Percentage of total	Percentage of sampled
Panel A: Mapping								
Total	200	Total	216					
without doctors	40	without private clinics	84	38.9%				
with doctors	160	with private clinics	132	61.1%				
Panel B1: Sampling								
Total	139	Total	139			599		
		without private clinics	48	34.5%		144	24.0%	
		with private clinics	91	65.5%		455	76.0%	
Panel B2: Completion								
Total	116	Total*	116		83.5%	460		76.8%
		without private clinics*	32		66.7%	87		60.4%
		with private clinics*	84		92.3%	373		82.0%
Panel C1: Sampling in dual practice sample								
Total	81	Provider-clinics	182			455		
		in public clinics	91	50.0%		227	49.9%	
		in private clinics	91	50.0%		228	50.1%	
Panel C2: Completion in dual practice sample								
Total	81	Provider-clinics*	155		85.2%	373		82.0%
		in public clinics*	71		78.0%	168		74.0%
		in private clinics*	84		92.3%	205		89.9%

Notes: * counts all providers with whom at least one case was completed. Reasons for not completing SP surveys include transfer of provider or an inability to find the provider for an interview. In these cases our field staff typically made three (in some cases four) attempts to complete a case. During fieldwork we replaced five sampled providers with other providers. In two cases, it was because the provider was on sick leave, two cases because provider had been transferred and one case because provider had gone on training.

Table A.3: Characteristics of Private Providers in the Representative Sample

	(1)	(2)	(3)	(4)
	Total	MBBS Providers	Providers with alternative qualifications	Unqualified providers
Number of Providers	772	40	192	540
Qualification details				
Duration of degree (months)	22.6	57.5	47.9	11.3
Did an internship as part of degree	0.244	0.900	0.625	0.059
Duration of internship (months, conditional)	2.7	12.3	8.7	0.7
Additional training				
Received additional training	0.793	0.325	0.688	0.864
Duration (months, conditional)	29.9	19.6	22.6	32.2
Trained by practising physician or learned by observation	0.224	0.125	0.307	0.203
Duration (months, conditional)	24.8	14.4	21.5	27.0
Trained as a compounder	0.198	0.025	0.063	0.258
Duration (months, conditional)	43.3	60.0	36.5	43.8
Trained at another institution or hospital	0.240	0.175	0.265	0.236
Duration (months, conditional)	19.3	17.6	17.4	20.2
Training other providers				
Has trained other providers	0.1082	0.0769	0.1780	0.0857

Notes: The MBBS degree is equivalent to the MD degree in the United States and stands for "Bachelor of Medicine & Bachelor of Surgery." Providers in the MBBS category includes all providers with only MBBS degrees and those with an MBBS and a specialization degree. Providers in the "Providers with alternative qualifications" includes the following degrees: Bachelor of Ayurvedic Medicine and Surgery (BAMS), BIMS, Bachelor of Unani Medicine and Surgery (BUMS), Bachelor of Homoeopathic Medicine and Surgery and Diploma of Homeopathic Medicine and Surgery (BHMS/DHMS), Diploma in Homeopathy and Biochemistry (DHB), Bachelor of Electro Homeopathic Medicine and Surgery (BEHMS/BEMS), Bachelor of Science in Nursing and Master of Science in Nursing (BSc Nursing/MSc Nursing. Providers in "Unqualified providers" includes Rural Medical Practitioners (RMP), providers with unverifiable degrees, and providers with no formal training. The majority of providers in this category are providers with no formal training.

Table A.4: Randomization balance for dual sample providers' assignment of Unstable Angina cases

	(1)	(2)	(3)	(4)	(5)	(7)	(8)	(9)	(10)
	Asthma outcomes							**Dysentery outcomes**	
	Time spent (mins)	Percent checklist completed	Gave diagnosis	Correct diagnosis	Correct treatment	Palliative treatment	Unnecessary treatment	Time spent (mins)	Percent checklist completed
Is private	1.497***	13.190***	0.181	0.077	0.131	-0.230**	-0.017	0.302	9.109**
	(0.483)	(3.292)	(0.118)	(0.099)	(0.113)	(0.117)	(0.075)	(0.241)	(4.119)
Received Unstable Angina in private	0.433	5.441	0.100	0.075	-0.194	-0.079	0.094	0.205	-0.862
	(0.518)	(3.534)	(0.127)	(0.106)	(0.121)	(0.126)	(0.080)	(0.255)	(4.356)
(Is private) x (Received Unstable Angina in private)	0.143	-2.996	-0.214	-0.094	0.044	0.131	-0.051	0.268	-0.604
	(0.719)	(4.898)	(0.176)	(0.147)	(0.168)	(0.174)	(0.111)	(0.354)	(6.053)
Constant	1.644***	13.687***	0.307***	0.150**	0.639***	0.487***	0.873***	0.783***	17.088***
	(0.347)	(2.367)	(0.085)	(0.071)	(0.081)	(0.084)	(0.054)	(0.172)	(2.941)

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. Robust standard errors are in parentheses. All regressions include district fixed effects.

Table A.5: Checklist items, diagnoses and treatments

	(1) Unstable angina	(2) Asthma	(3) Dysentery
Panel A: Checklist Items			
History questions	where is the pain, when started, severity of pain, radiation, previous similar, since when, shortness of breath, sweating, beedi-cigarette, family history	current breathing probes, cough, expectoration probes, previous breathing problems, since when problems, shortness constant of episodic, what triggers, fever, chest pain, weight loss, beedi-cigarette, family history	age of child, qualities of school, frequency, quantity of stool, urination, child active/playful, fever, abdominal pain, vomitting, source of water, what has child eaten, child taking fluids
Examinations	pulse, bp, auscultation (front or back), temperature attempt, ecg in/outside clinic	pulse, bp, auscultation (front or back), temperature attempt	
Panel B: Diagnosis			
Correct	Heart attack, angina, myocardial infarction, attack	Asthma, asthma attack	Dysentery, bacteria
Incorrect	Blood pressure problem, gastrointestinal problem, muscle problem, the weather, injury, nerve pull, lack of blood, swelling in chest, pain from drinking cold water, heavy work, bad blood, decaying lungs, chest congestion	Blood pressure problem, gastrointensinal problem, heart problem, the weather, cough in chest, thyroid problem, weakness, lack of blood, infection in windpipe, pregnancy, allergy	Weather, heat in liver, acidity, diarrhea
Panel C: Treatment			
Correct	Aspirin, clopidogrel/other anti-platelet agents, do an ECG.	Bronchodilators, theophylline, inhaled or oral corticosteroids, leukotriene inhibitors, cromones, inhaled anticholinergics	ORS, rehydration
Palliative	Nitroglycerin, blood thinners, betablockers, ACE inhibitors, vasodilators, other cardiac medication, morphine, other pain medication, referral or referral for an ECG.	Anti-allergy medication	Antibiotics,zinc
Unncessary or harmful	Antibiotics, oral rehydration salts, oral electrolyte solution, zinc, steroids, inhaler, bronchodilators, theophylline, inhaled corticosteroids, leukotriene inhibitors, cromones, inhaled anti-cholinergics, oral cortico-steroids, other anti-asthmatic medication, anti-allergy medication, psychiatric medication.	Aspirin, clopidogrel, anti-platelet agents, blood thinners, betablockers, ACE inhibitors, vasodilators, other cardiac medication, morphine, other pain medication, oral rehydration salts, oral electrolyte solution, zinc, antibiotics, anti-ulcer medication, psychiatric medication	Aspirin, clopidogrel, anti-platelet agents, blood thinners, betablockers, ACE inhibitors, vasodilators, other cardiac medication, morphine, other pain medication, steroids, inhaler, bronchodilators, theophylline, inhaled corticosteroids, leukotriene inhibitors, cromones, inhaled anti-cholinergics, oral cortico-steroids, other anti-asthmatic medication, anti-allergy medication, psychiatric medication

Notes: See Appendix B for coding of treatments

Table A.6: List of checklist items used in the treatment of SPs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Item discrimination tercile	**Representative sample**				**Dual practice sample**			
		All	Public	Private	Difference (4)-(3)	All	Public	Private	Difference (9)-(8)
Panel A: Unstable Angina									
History questions									
where is the pain	high	0.659	0.486	0.694	0.208***	0.582	0.514	0.667	0.153
when started	low	0.369	0.270	0.389	0.119*	0.149	0.162	0.133	-0.029
doing when began	high	0.074	0.054	0.078	0.024	0.119	0.081	0.167	0.086
severity of pain	low	0.258	0.162	0.278	0.116*	0.284	0.162	0.433	0.271***
radiation	high	0.143	0.108	0.150	0.042	0.299	0.216	0.400	0.184*
previous similar	medium	0.392	0.270	0.417	0.146**	0.328	0.270	0.400	0.130
since when	low	0.263	0.216	0.272	0.056	0.209	0.108	0.333	0.225**
quality of pain	high	0.115	0.108	0.117	0.009	0.179	0.108	0.267	0.159**
pain changes	low	0.060	0.054	0.061	0.007	0.104	0.054	0.167	0.113*
shortness of breath	medium	0.138	0.081	0.150	0.069	0.045	0.054	0.033	-0.021
nausea	medium	0.295	0.297	0.294	-0.003	0.209	0.054	0.400	0.346***
sweating	high	0.290	0.270	0.294	0.024	0.313	0.189	0.467	0.277***
beedi-cigarette	low	0.069	0.054	0.072	0.018	0.134	0.081	0.200	0.119*
family history	high	0.014	0.000	0.017	0.017	0.045	0.000	0.100	0.100**
Examination questions									
pulse	low	0.392	0.243	0.422	0.179**	0.537	0.432	0.667	0.234**
bp	medium	0.313	0.135	0.350	0.215***	0.373	0.216	0.567	0.350***
auscultation (either front or back)	low	0.447	0.189	0.500	0.311***	0.522	0.432	0.633	0.201*
temperature attempt	medium	0.134	0.108	0.139	0.031	0.134	0.054	0.233	0.179**
ecg in/outside clinic	medium	0.230	0.243	0.228	-0.015	0.313	0.270	0.367	0.096
Number of observations		217	37	180		67	37	30	
Panel B: Asthma									
History questions									
current breathing probes	medium	0.601	0.385	0.647	0.262***	0.552	0.431	0.667	0.236***
cough	low	0.677	0.590	0.696	0.106	0.575	0.462	0.681	0.220***
expectoration probes	low	0.148	0.077	0.163	0.086*	0.045	0.015	0.072	0.057*
previous breathing problems	high	0.439	0.333	0.462	0.129*	0.410	0.277	0.536	0.259***
previous episode probes	medium	0.184	0.128	0.196	0.067	0.201	0.123	0.275	0.152**
since when problems	medium	0.475	0.385	0.495	0.110	0.328	0.231	0.420	0.190***
how often happens	high	0.108	0.128	0.103	-0.025	0.067	0.046	0.087	0.041
shortness constant or episodic	low	0.103	0.051	0.114	0.063	0.090	0.046	0.130	0.084**
what triggers	medium	0.117	0.077	0.125	0.048	0.164	0.092	0.232	0.140**
how long lasts	high	0.067	0.077	0.065	-0.012	0.052	0.015	0.087	0.072**
childhood illness	medium	0.027	0.000	0.033	0.033	0.030	0.015	0.043	0.028
age	high	0.170	0.308	0.141	-0.166***	0.537	0.585	0.493	-0.092
fever	low	0.309	0.231	0.326	0.095	0.306	0.215	0.391	0.176**
chest pain	low	0.336	0.154	0.375	0.221***	0.231	0.169	0.290	0.121**
weight loss	high	0.000	0.000	0.000	0.000	0.015	0.015	0.014	-0.001
night sweats	high	0.054	0.051	0.054	0.003	0.067	0.046	0.087	0.041
beedi-cigarette	high	0.018	0.026	0.016	-0.009	0.045	0.015	0.072	0.057*
family history	medium	0.022	0.000	0.027	0.027	0.037	0.031	0.043	0.013
Examination questions									
pulse	low	0.502	0.256	0.554	0.298***	0.388	0.308	0.464	0.156**
bp	medium	0.278	0.205	0.293	0.088	0.239	0.108	0.362	0.255***
auscultation (either front or back)	low	0.516	0.333	0.554	0.221***	0.649	0.492	0.797	0.305***
temp attempt	low	0.166	0.103	0.179	0.077	0.082	0.077	0.087	0.010
Number of observations		223	39	184		134	65	69	

(continued on next page)

Table A.6 continued

	(1) Item discrimination tercile	(2) Representative sample All	(3) Representative sample Public	(4) Representative sample Private	Difference (3)-(2)	(5) Dual practice sample All	(6) Dual practice sample Public	(7) Dual practice sample Private	Difference (6)-(5)
Panel C: Dysentery									
History questions									
age of child	low	0.919	0.795	0.945	0.150***	0.930	0.921	0.939	0.019
qualities of stool	low	0.167	0.077	0.186	0.109**	0.271	0.159	0.379	0.220***
frequency	medium	0.288	0.179	0.311	0.132**	0.372	0.270	0.470	0.200***
quantity of stool	high	0.050	0.000	0.060	0.060*	0.031	0.016	0.045	0.030
urination	high	0.018	0.000	0.022	0.022	0.008	0.016	0.000	-0.016
active/playful	high	0.032	0.026	0.033	0.007	0.000	0.000	0.000	0.000
fever	medium	0.171	0.077	0.191	0.114**	0.295	0.222	0.364	0.141**
abdominal pain	low	0.113	0.077	0.120	0.043	0.256	0.222	0.288	0.066
vomiting	low	0.216	0.077	0.246	0.169***	0.295	0.254	0.333	0.079
source of water	high	0.023	0.000	0.027	0.027	0.016	0.000	0.030	0.030*
what has eaten	medium	0.050	0.000	0.060	0.060*	0.093	0.032	0.152	0.120***
taking fluids	medium	0.023	0.000	0.027	0.027	0.062	0.048	0.076	0.028
Number of observations		222	39	183		130	63	67	

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%.

Table A.7: Effort in the public and private sectors by checklist item discrimination terciles

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome variable: Percentage of recommended type of checklist items					
	Representative sample			Dual practice sample		
	Low discrimination	Medium discrimination	High discrimination	Low discrimination	Medium discrimination	High discrimination
Panel A: SP and case fixed effects						
Is a private provider	10.982***	7.085**	1.760	10.650***	11.728***	5.288***
	(3.281)	(2.875)	(2.143)	(2.583)	(2.616)	(1.766)
R-squared	0.144	0.175	0.238	0.280	0.235	0.319
Number of observations	662	662	662	330	330	330
Mean of public	21.770	13.975	10.197	28.225	14.690	10.072
Mean of private	32.966	21.322	12.235	41.288	28.874	15.245
Mean of sample	32.108	20.759	12.079	34.756	21.782	12.659
Panel B: SP, case and market/district fixed effects						
Is a private provider	11.290***	8.597***	1.594	10.705***	11.733***	5.226***
	(3.549)	(3.141)	(2.540)	(2.577)	(2.607)	(1.762)
R-squared	0.253	0.256	0.300	0.302	0.247	0.323
Number of observations	662	662	662	330	330	330
Panel C: SP, case and market/district fixed effects						
Is a private provider	8.538**	7.317**	1.657	11.879***	12.550***	4.660**
	(3.717)	(3.382)	(2.876)	(2.823)	(2.729)	(1.854)
Has MBBS	2.548	5.175*	2.307			
	(4.091)	(2.978)	(1.850)			
Has some qualification	2.300	4.764**	0.721			
	(2.017)	(2.208)	(2.063)			
Age of provider	-0.151*	-0.009	0.044	-0.072	-0.138	-0.043
	(0.078)	(0.090)	(0.062)	(0.141)	(0.114)	(0.099)
Gender of provider (1=Male)	1.009	-1.353	-2.369	2.822	-2.740	-3.631
	(6.644)	(3.138)	(3.958)	(4.328)	(3.696)	(3.465)
Patient load during visit	-0.041	-0.396	0.050	-0.428	-0.126	-0.182
	(0.622)	(0.430)	(0.518)	(0.454)	(0.676)	(0.449)
R-squared	0.254	0.262	0.301	0.291	0.252	0.331
Number of observations	638	638	638	301	301	301

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual practice sample, robust standard errors clustered at the provider level are in parentheses. Observations are at the SP-provider interaction level. Checklist item discrimination parameters are estimated using the IRT methodology. The classification of items into terciles of difficulty is done within each case, but the results are robust to classifying the items jointly across all cases. Market fixed effects are used for the representative sample, and district fixed effects for the dual practice sample.

Table A.8: Effort, diagnosis and treatment by case

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Effort		**Diagnosis**			**Treatment**					
	Time spent	Checklist	Gave diagnosis	Correct diagnosis (conditional)	Correct diagnosis (unconditional)	Correct treatment	Palliative treatment	Unnecessary treatment	Correct treatment Only	Antibiotic	Number of medicines
Panel A1: Unstable angina, representative sample, with SP fixed effects											
Is a private provider	1.101***	7.890**	0.112	0.033	0.011	0.021	-0.070	0.083	-0.026	0.024	0.782***
	(0.302)	(3.860)	(0.076)	(0.067)	(0.028)	(0.031)	(0.096)	(0.092)	(0.027)	(0.053)	(0.233)
R-squared	0.083	0.138	0.016	0.155	0.082	0.033	0.021	0.056	0.016	0.030	0.043
Number of observations	217	217	217	102	217	217	217	217	217	217	217
Mean of public	2.592	17.354	0.378	0.071	0.027	0.027	0.784	0.730	0.027	0.135	2.054
Panel A2: Unstable angina, dual practice sample, with SP fixed effects											
Is a private provider	3.370***	13.640**	0.184*	0.186	0.144*	0.286***	-0.007	0.052		-0.053	0.447
	(1.027)	(5.380)	(0.109)	(0.183)	(0.076)	(0.094)	(0.081)	(0.130)		(0.110)	(0.362)
R-squared	0.225	0.116	0.337	0.141	0.153	0.182	0.063	0.054		0.073	0.175
Number of observations	61	61	61	29	61	61	61	61	61	61	61
Mean of public	1.954	18.341	0.394	0.077	0.030	0.030	0.909	0.667	0.000	0.273	2.242
Panel B1: Asthma, representative sample, with SP fixed effects											
Is a private provider	1.952***	6.015*	0.224***	-0.123	0.021	0.082	-0.008	0.040	0.010	0.009	1.158***
	(0.449)	(3.548)	(0.084)	(0.134)	(0.034)	(0.088)	(0.082)	(0.078)	(0.037)	(0.104)	(0.372)
R-squared	0.200	0.172	0.209	0.065	0.067	0.043	0.029	0.076	0.038	0.019	0.095
Number of observations	223	223	223	76	223	223	223	223	223	223	223
Mean of public	3.301	17.716	0.154	0.333	0.051	0.385	0.282	0.744	0.026	0.385	2.128
Panel B2: Asthma, dual practice sample, with SP fixed effects											
Is a private provider	1.431***	11.970***	0.044	-0.078	-0.009	0.128	-0.151*	-0.054	0.025	-0.165*	-0.224
	(0.362)	(2.361)	(0.085)	(0.149)	(0.071)	(0.084)	(0.078)	(0.055)	(0.045)	(0.089)	(0.202)
R-squared	0.202	0.228	0.091	0.102	0.060	0.132	0.111	0.111	0.044	0.101	0.122
Number of observations	122	122	122	51	122	122	122	122	122	122	122
Mean of public	1.875	16.102	0.373	0.545	0.203	0.525	0.458	0.915	0.034	0.593	3.119
Panel C1: Dysentery, representative sample, with SP fixed effects											
Is a private provider	0.846***	7.088**									
	(0.219)	(2.850)									
R-squared	0.091	0.108									
Number of observations	222	222									
Mean of public	1.281	10.897									
Panel C2: Dysentery, dual practice sample, with SP fixed effects											
Is a private provider	0.395**	5.279**									
	(0.173)	(2.468)									
R-squared	0.095	0.340									
Number of observations	119	119									
Mean of public	0.879	16.228									

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual practice sample, robust standard errors clustered at the provider level are in parentheses. All regressions include a constant and controls for provider qualifications, age, gender, and patient load. Observations are standardized provider-patient interactions. In column (11) the dependent variable is total number of medicines recommended to the patient (dispensed and/or prescribed).

Table A.9: Summary of treatment by case

	(1)	(2)	(3)	(4)	(5)	(6)
	Representative sample			Dual practice sample		
	Public	Private	Difference (2)-(1)	Public	Private	Difference (5)-(4)
Panel A: Unstable Angina						
Correct treatment	0.03	0.06	0.03	0.03	0.30	0.27***
Correct treatment (alternate)	0.46	0.37	-0.09	0.41	0.63	0.23**
Palliative treatment	0.78	0.71	-0.07	0.92	0.90	-0.02
Unnecessary treatment	0.73	0.80	0.07	0.68	0.73	0.06
Aspirin	0.03	0.04	0.02	0.03	0.23	0.21***
Anti-platelet agents	0.03	0.01	-0.02	0.00	0.03	0.03
Referred	0.30	0.24	-0.05	0.22	0.33	0.12
ECG	0.24	0.23	-0.02	0.27	0.37	0.10
ECG & Referred	0.11	0.12	0.01	0.08	0.17	0.09
Antibiotic	0.14	0.17	0.03	0.30	0.20	-0.10
Number of observations	37	180		37	30	
Panel B: Asthma						
Correct treatment	0.38	0.50	0.12*	0.57	0.68	0.11*
Palliative treatment	0.28	0.29	0.01	0.48	0.28	-0.20***
Unnecessary treatment	0.74	0.83	0.09*	0.92	0.88	-0.04
Bronchodilators	0.33	0.36	0.03	0.51	0.59	0.09
Theophylline	0.13	0.22	0.09*	0.31	0.32	0.01
Oral Corticosteroids	0.15	0.31	0.16**	0.15	0.25	0.09*
Antibiotic	0.38	0.40	0.02	0.60	0.45	-0.15**
Number of observations	39	184		65	69	
Panel C: Dysentery						
Correct treatment	0.08	0.13	0.05	0.33	0.22	-0.11*
Palliative treatment	0.44	0.61	0.18**	0.75	0.61	-0.13*
Unnecessary treatment	0.28	0.56	0.28***	0.35	0.40	0.05
ORS	0.05	0.12	0.07	0.33	0.21	-0.12*
Asked to see child	0.33	0.14	-0.20***	0.27	0.42	0.15**
Antibiotic	0.44	0.61	0.18**	0.75	0.61	-0.13*
Number of observations	39	183		63	67	

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. In Unstable Angina, alternate definition for correct treatment codes referrals and referrals for ECG as correct. In the dysentery case, note the large and significant differences in "asked to see the child" across public and private providers in the representative and dual samples. If we were to assume the same rate of correct treatment by public and private providers in the cases where they "asked to see the child" as in the cases where a treatment was provided, then the differences in correct treatment are no longer significant in either sample. If we carry out a bounding exercise, the differences are still not significant, and the standard errors are too wide for meaningful inference. This is why we exclude the dysentery case in our pooled analysis of treatment across cases.

Table A.10: Robustness of treatment results with alternative definition for correct treatment for unstable angina

	(1)	(2)	(3)	(4)
	All (compare with table 4)		**Unstable angina only (compare with table A8)**	
	Representative sample	Dual practice sample	Representative sample	Dual practice sample
	Correct treatment	Correct treatment	Correct treatment	Correct treatment
Panel A: SP fixed effects				
Is a private provider	-0.014	0.138**	-0.112	0.232*
	(0.063)	(0.069)	(0.088)	(0.120)
R-squared	0.075	0.091	0.092	0.081
Number of observations	440	201	217	67
Mean of public	0.421	0.510	0.459	0.405
Mean of private	0.421	0.667	0.360	0.633
Mean of sample	0.421	0.587	0.367	0.507
Panel B: SP and market/district fixed effects				
Is a private provider	0.001	0.142**	-0.065	0.210*
	(0.069)	(0.061)	(0.118)	(0.118)
R-squared	0.196	0.101	0.298	0.192
Number of observations	440	201	217	67
Panel C: SP and market/district fixed effects				
Is a private provider	-0.009	0.150**	-0.203	0.197
	(0.070)	(0.065)	(0.141)	(0.125)
Has MBBS	0.340***		0.233	
	(0.081)		(0.147)	
Has some qualification	0.164***		0.139	
	(0.057)		(0.095)	
Age of provider	0.000	-0.005	0.002	-0.006
	(0.004)	(0.004)	(0.004)	(0.006)
Gender of provider (1=Male)	0.256	0.007	0.334**	-0.167
	(0.158)	(0.107)	(0.170)	(0.161)
Patient load during visit	-0.030***	-0.003	-0.022**	-0.030
	(0.008)	(0.018)	(0.010)	(0.022)
R-squared	0.244	0.112	0.352	0.242
Number of observations	423	183	208	61

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual practice sample, robust standard errors clustered at the provider level are in parentheses. All regressions include a constant. Observations are at the SP-provider interaction level. Columns (1) and (2) also include case fixed effects. Market fixed effects are used for the representative sample, and district fixed effects for the dual practice sample. Alternative definition for Unstable Angina adds "referral" and "referral for ECG" as correct treatment.

Table A.11: Robustness of provider effort results to exclusion of dysentery cases

	(1)	(2)	(3)	(4)	(5)	(6)
	Representative sample			Dual practice sample		
	Time Spent (mins)	Percentage of checklist items	IRT score	Time Spent (mins)	Percentage of checklist items	IRT score
Panel A: SP and case fixed effects						
Is a private provider	1.531***	6.942**	0.551**	2.261***	12.421***	0.755***
	(0.306)	(3.307)	(0.212)	(0.449)	(2.414)	(0.206)
R-squared	0.225	0.152		0.177	0.157	0.031
Number of observations	440	440	233	201	201	199
Mean of public	2.956	17.540		1.960	17.553	
Mean of private	4.548	24.335		4.094	30.378	
Mean of sample	4.427	23.820		3.011	23.870	
Panel B: SP, case and market/district fixed effects						
Is a private provider	1.907***	7.593**	0.668**	2.269***	12.361***	0.759***
	(0.453)	(3.829)	(0.277)	(0.450)	(2.418)	(0.207)
R-squared	0.341	0.278		0.201	0.166	
Number of observations	440	440	233	201	201	138
Panel C: SP, case and market/district fixed effects						
Is a private provider	1.654***	6.087	0.611*	2.132***	12.433***	0.829***
	(0.579)	(4.409)	(0.327)	(0.464)	(2.738)	(0.210)
Has MBBS	-0.062	6.415*	0.124			
	(0.987)	(3.792)	(0.369)			
Has some qualification	-0.159	2.737*	0.176			
	(0.567)	(1.648)	(0.200)			
Age of provider	-0.002	0.027	0.005	0.017	-0.012	-0.002
	(0.017)	(0.105)	(0.008)	(0.029)	(0.145)	(0.009)
Gender of provider (1=Male)	1.460***	2.136	0.164	-0.332	-3.055	-0.085
	(0.554)	(5.474)	(0.410)	(0.675)	(4.828)	(0.342)
Patient load during visit	-0.188***	-0.333	0.021	-0.107*	0.087	0.001
	(0.056)	(0.342)	(0.039)	(0.061)	(0.675)	(0.041)
R-squared	0.357	0.283		0.224	0.171	
Number of observations	423	423	221	183	183	126

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual practice sample, robust standard errors clustered at the provider level are in parentheses. All regressions include a constant. Observations are at the SP-provider interaction level, except in IRT score where each observation is a composite provider level score across all cases. Market fixed effects are used for the representative sample, and district fixed effects for the dual practice sample.

Table A.12: Robustness of results to inclusion of facilties controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Effort			**Diagnosis**			**Treatment**					
	Time spent	Checklist	IRT Score	Gave diagnosis	Correct diagnosis (conditional)	Correct diagnosis (unconditional)	Correct treatment	Palliative treatment	Unnecessary treatment	Correct treatment only	Antibiotic	Number of medicines
Panel A: Representative sample, with SP, case and market fixed effects												
Is a private provider	1.207***	7.826***	0.731**	0.197***	-0.023	0.039	0.143***	0.082	0.115	-0.009	0.153**	0.861***
	(0.454)	(2.608)	(0.333)	(0.064)	(0.102)	(0.034)	(0.055)	(0.069)	(0.082)	(0.027)	(0.075)	(0.318)
Facilities index	0.012	1.679**	0.120	0.051***	0.014	0.010	0.034**	0.026	0.038*	-0.001	0.029	0.203**
	(0.124)	(0.676)	(0.078)	(0.015)	(0.033)	(0.010)	(0.016)	(0.018)	(0.023)	(0.003)	(0.024)	(0.095)
R-squared	0.356	0.265		0.233	0.362	0.161	0.410	0.379	0.267	0.280	0.275	0.313
Number of observations	634	634	220	420	171	420	420	420	420	420	420	420
Panel B: Dual practice sample, with SP, case and district fixed effects												
Is a private provider	1.233***	9.087***	0.875***	0.039	-0.035	0.001	0.183**	-0.134*	-0.014	0.023	-0.154**	-0.108
	(0.284)	(2.090)	(0.235)	(0.079)	(0.129)	(0.068)	(0.075)	(0.075)	(0.063)	(0.028)	(0.077)	(0.216)
Facilities index	-0.205	-0.963	0.029	-0.038	-0.029	-0.028	-0.063*	-0.017	0.001	0.001	-0.039	-0.256**
	(0.187)	(1.315)	(0.121)	(0.039)	(0.073)	(0.036)	(0.036)	(0.040)	(0.033)	(0.014)	(0.050)	(0.126)
R-squared	0.322	0.243	0.081	0.220	0.199	0.091	0.320	0.306	0.158	0.052	0.146	0.198
Number of observations	272	272	272	164	73	164	164	164	164	164	164	164

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual practice sample, robust standard errors clustered at the provider level are in parentheses. Observations are at the SP-provider interaction level. All regressions include a constant and controls for provider qualifications, age, gender, and patient load. Market fixed effects are used for the representative sample, and district fixed effects for the dual practice sample. Columns (1)-(3) include all cases and can be compared with Table 3. The remaining columns include Unstable Angina and Asthma cases only - compare Columns (4)-(6) with Table 4; and Columns (7)-(12) with Table 5. In column (12) the dependent variable is the total number of medicines recommended to the patient (dispensed and/or prescribed). Note that the reason for not including the controls for an index of Facility quality in the main results in Tables 3-5 is that we are missing data on the facility index for around 4% of the representative sample and 18% of the dual sample. However, as we see here, the results are robust to including the facility controls.

Table A.13: Differential case completion in the dual practice sample

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
		Effort				Treatment						
		Fraction of cases	Time spent	Checklist	IRT Score	*Fraction of cases*	Correct treatment	Palliative treatment	Unnecessary treatment	Correct treatment only	Antibiotic	Number of medicines
Panel A: Summary Statistics												
Is a public provider	Completed in first attempt	0.586	1.574	18.291	-0.361	0.574	0.423	0.615	0.833	0.026	0.423	2.782
	Completed in later attempt	0.154	1.509	15.347	-0.758	0.191	0.208	0.708	0.833	0.000	0.708	3.000
	Not completed	0.260				0.235						
	Difference (first - later)		0.065	2.944	0.397*		0.215**	-0.093	0.000	0.026	-0.285***	-0.218
Is a private provider	Completed in first attempt	0.719	3.000	28.804	0.362	0.417	0.553	0.421	0.803	0.053	0.355	2.803
	Completed in later attempt	0.180	2.919	26.383	0.550	0.123	0.609	0.609	0.957	0.000	0.435	3.304
	Not completed	0.101				0.061						
	Difference (first - later)		0.081	2.421	-0.187		-0.056	-0.188*	-0.154**	0.053	-0.080	-0.502**
Panel B: Differential completion												
Is a public provider			-1.583***	-10.971***	-1.194**		-0.381***	0.072	-0.107	0.002	0.305**	-0.212
			(0.503)	(3.717)	(0.465)		(0.103)	(0.109)	(0.075)	(0.006)	(0.136)	(0.340)
Completed in first attempt			0.165	0.862	-0.170		-0.095	-0.146	-0.155**	0.049**	-0.074	-0.474*
			(0.526)	(3.311)	(0.247)		(0.087)	(0.101)	(0.065)	(0.025)	(0.132)	(0.257)
Is a public provider x Completed in first attempt			0.081	2.172	0.385		0.291**	0.067	0.152	-0.028	-0.222	0.202
			(0.560)	(4.326)	(0.514)		(0.120)	(0.122)	(0.102)	(0.037)	(0.163)	(0.441)
R-squared			0.239	0.215	0.244		0.281	0.316	0.093	0.033	0.145	0.105
Number of observations			331	331	331		201	201	201	201	201	201

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. In Panel A, significance stars are for t-tests comparing completion in first attempt vs. completion in later attempt. The columns "fraction of cases" is different for effort and treatment variables because the former treats all cases while the latter considers only unstable angina and asthma cases. In Panel B, robust standard errors clustered at the provider level are in parentheses. Observations are at the SP-provider interaction level except in Column (4) where it is at the provider level. All regressions include a constant, and SP and case fixed effects. In column (11) the dependent variable is the total number of medicines recommended to the patient (dispensed and/or prescribed).

Table A.14: Reweighted estimates for differential case completion in the dual sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Effort			Treatment					
	Time spent	Checklist	IRT Score	Correct treatment	Palliative treatment	Unnecessary treatment	Correct treatment only	Antibiotic	Number of medicines
Panel A: Original estimates									
Is a private provider	1.507***	8.977***	0.755***	0.151**	-0.126**	-0.021	0.019	-0.141**	0.002
	(0.298)	(1.935)	(0.207)	(0.061)	(0.057)	(0.054)	(0.026)	(0.067)	(0.200)
R-squared	0.241	0.220		0.274	0.309	0.108	0.025	0.120	0.127
Number of observations	331	331	138	201	201	201	201	201	201
Panel B: Reweighted estimates									
Is a private provider	1.575***	10.236***	0.894***	0.203***	-0.135***	0.041	0.015	-0.126**	0.149
	(0.217)	(1.457)	(0.160)	(0.046)	(0.044)	(0.040)	(0.019)	(0.052)	(0.158)
R-squared	0.250	0.207		0.239	0.276	0.052	0.018	0.100	0.063
Number of observations	455	455	182	273	273	273	273	273	273

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. Robust standard errors clustered at the provider level are in parentheses. Panel A replicates original results (corresponding to Tables 3 and 5) to facilitate comparison. The effort regressions use all cases while the treatment regressions use only the unstable angina and asthma cases. Observations are at the SP-provider interaction level except in Column (3) where it is at the provider level. All regressions include a constant, and SP and case fixed effects. In Panel B, the used SP and case fixed effects are those for assigned SP and case. In column (9) the dependent variable is the total number of medicines recommended to the patient (dispensed and/or prescribed).

Table A.15: Correlates of price charged

(private interactions, excludes cases where all medicines are unidentifiable)

	(1)	(2)	(3)	(4)	(5)	(6)
	Fees in Rs.					
	Representative sample		**Dual practice sample**		**Pooled sample**	
	Binary regressions	Multiple regression	Binary regressions	Multiple regression	Binary regressions	Multiple regression
Time spent with SP (minutes)	1.720***	0.618	2.625***	2.279***	1.484***	0.709*
	(0.476)	(0.477)	(0.587)	(0.692)	(0.377)	(0.401)
Percentage of checklist items	0.397***	0.339***	0.364***	0.055	0.386***	0.291***
	(0.089)	(0.096)	(0.100)	(0.129)	(0.071)	(0.084)
Correct diagnosis (unconditional)	-4.269	-3.647*	7.504	5.494	2.690	2.685
	(3.978)	(1.993)	(9.350)	(9.046)	(4.658)	(4.148)
Correct treatment	6.199***	-1.564	7.744*	4.475	7.306***	0.602
	(1.757)	(2.919)	(4.145)	(4.967)	(1.934)	(2.404)
Palliative treatment	7.711***	2.198	10.435**	7.757	7.796***	3.542**
	(1.810)	(1.722)	(4.242)	(4.873)	(1.743)	(1.726)
Unnecessary treatment	15.794***	3.147	14.973***	5.137	15.655***	4.888*
	(2.842)	(2.963)	(5.032)	(6.240)	(2.451)	(2.746)
Dispensed medicines	19.525***	16.400***	16.118***	12.371*	16.511***	15.688***
	(2.993)	(2.726)	(6.070)	(7.019)	(2.319)	(2.830)
Prescribed medicines	-2.931	-4.331	7.540	-2.854	0.071	-4.133
	(3.600)	(3.639)	(5.997)	(6.734)	(2.918)	(3.202)
Number of medicines	5.540***	1.630	5.863***	3.016	5.283***	1.111
	(0.842)	(1.394)	(1.783)	(2.987)	(0.787)	(1.348)
Referred/Asked to see child	-20.348***	-10.054***	-9.882**	-4.867	-17.533***	-11.860***
	(4.999)	(3.683)	(4.763)	(4.888)	(3.911)	(3.021)
Has MBBS	23.517***	27.905***			14.155***	23.516***
	(6.150)	(7.830)			(4.369)	(3.923)
Has some qualification	4.305	6.067***			2.127	6.952***
	(3.768)	(2.282)			(3.376)	(2.370)
Patient load during visit	1.017	0.867**	-0.073	-0.285	0.512	0.276
	(0.888)	(0.404)	(0.807)	(0.810)	(0.748)	(0.581)
Age of provider	-0.186	-0.111	0.267	0.248	-0.119	-0.018
	(0.155)	(0.100)	(0.239)	(0.218)	(0.126)	(0.089)
Gender of provider (1=Male)	-8.238**	-5.876	-1.284	-3.760	-7.475**	-3.810
	(3.518)	(4.543)	(4.882)	(5.580)	(2.961)	(3.919)
Constant		9.745		-11.295		2.234
		(7.179)		(11.810)		(6.345)
R2		0.446		0.444		0.398
Number of observations		495		154		649
Mean price charged		27.638		32.740		28.849
SD		26.557		28.592		27.118

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual sample and pooled sample, robust standard errors clustered at the location/market level are in parentheses. Observations are at the SP-provider interaction level. Interpretation of coefficents in "Binary regressions" needs caution. Each coefficient represents a separate regression of prices on the row variable and SP, case and district fixed effects. Multiple regressions include SP, case and district fixed effects. The pooled sample (Columns 5 and 6) combine the representative and dual practice samples.

Table A.16: Cost in the public sector

	(1)	(2)
Panel A: Staff per facility	**N**	**Average monthly wage (Rs.)**
Medical Officer in Charge/Medical Officer	1.92	Rs.32,245
GNM/ANM/VHN/LHV	3.24	Rs.16,305
MPW/MNA/Assistant/Compounder	1.43	Rs.16,657
Pharmacist/Chemist/Lab Assistant/Technician	0.8	Rs.16,571
Paramedic/other	6.08	Rs.13,387
All	13.47	Rs.17,315
Number of facilities	115	
Panel B: Visits to the public facilities per month		
Year 2008	111,039	
Year 2009	113,230	
Year 2010	111,473	
Panel C: Average per patient cost		
Year 2008	Rs.241.87	
Year 2009	Rs.237.66	
Year 2010	Rs.241.61	

Notes: We use an extremely conservative measure of per patient cost in the public sector facility. We assume that salary costs are the only cost in running a public health facility. Furthermore, we assume that every patient that visits the public health facility visits for a primary care visit, while people also visit public health facilities for preventative services such as vaccination. Wage data were collected in the year 2010, which we use to compute cost per patient in 2008 and 2009. Wages in 2008 and 2009 could have been lower.

Table A.17: Real patients' characteristics in the public and private sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Representative sample				Dual practice sample			
	Public	Private	Difference (coeff. on private)		Public	Private	Difference (coeff. on private)	
			no fixed effects	market fixed effects			no fixed effects	district fixed effects
Patient/Case Characteristics								
Number of symptoms	1.446	1.568	0.122**	0.092	2.075	2.113	0.038	0.026
			(0.057)	(0.081)			(0.095)	(0.101)
Fever	0.309	0.445	0.136***	0.135**	0.550	0.548	-0.002	0.012
			(0.034)	(0.054)			(0.043)	(0.043)
Cold	0.272	0.195	-0.077	-0.015	0.476	0.434	-0.042	-0.047
			(0.049)	(0.062)			(0.054)	(0.050)
Diarrhea	0.105	0.151	0.046	0.008	0.066	0.075	0.009	0.006
			(0.033)	(0.040)			(0.014)	(0.015)
Weakness	0.148	0.209	0.061*	0.047	0.182	0.176	-0.006	-0.016
			(0.034)	(0.047)			(0.029)	(0.031)
Injury	0.093	0.069	-0.023	-0.045	0.061	0.070	0.010	0.011
			(0.023)	(0.030)			(0.016)	(0.017)
Vomitting	0.031	0.116	0.085***	0.046*	0.056	0.057	0.001	0.001
			(0.019)	(0.025)			(0.018)	(0.018)
Dermatological	0.062	0.054	-0.007	0.016	0.086	0.070	-0.016	-0.017
			(0.024)	(0.023)			(0.021)	(0.022)
Pregnancy	0.037	0.010	-0.027	0.013	0.035	0.058	0.022	0.024
			(0.033)	(0.018)			(0.019)	(0.019)
Pain	0.426	0.346	-0.080	-0.127	0.648	0.659	0.011	-0.008
			(0.081)	(0.104)			(0.043)	(0.037)
Number of days sick	0.623	1.584	0.961	-2.264	1.570	1.742	0.172	-0.438
			(4.295)	(2.819)			(1.068)	(1.022)
Activities of Daily Living								
Can easily dress	1.000	0.983	-0.017***	-0.019*	0.957	0.938	-0.020	-0.018
			(0.006)	(0.009)			(0.023)	(0.023)
Can easily work	0.856	0.901	0.045	0.077	0.748	0.798	0.050	0.050
			(0.051)	(0.051)			(0.047)	(0.049)
Can easily lift	0.698	0.730	0.032	0.038	0.666	0.692	0.027	0.017
			(0.104)	(0.124)			(0.071)	(0.071)
Can easily walk	0.623	0.699	0.076	0.146	0.785	0.755	-0.030	-0.049
			(0.131)	(0.104)			(0.074)	(0.071)
Patient Background and Demographics								
New patient	0.944	0.850	-0.094**	-0.001	0.911	0.903	-0.008	-0.003
			(0.036)	(0.043)			(0.037)	(0.038)
Age	30.006	25.401	-4.605	-5.082	28.913	30.700	1.788	1.410
			(3.087)	(3.530)			(2.042)	(2.040)
Is Male	0.494	0.579	0.086	0.021	0.487	0.454	-0.033	-0.039
			(0.053)	(0.059)			(0.042)	(0.041)
Assets index	0.455	0.411	-0.044	-0.238	-0.077	1.006	1.084***	1.146***
			(0.423)	(0.442)			(0.220)	(0.211)
Has formal education	0.565	0.517	-0.048	-0.053	0.546	0.637	0.091**	0.087**
			(0.085)	(0.081)			(0.035)	(0.034)
No. of questions patient asked	0.369	0.478	0.109	0.387**	0.488	0.956	0.467***	0.472***
			(0.103)	(0.152)			(0.125)	(0.125)
Is from this village	0.759	0.529	-0.230***	-0.149**	0.538	0.582	0.045	0.036
			(0.060)	(0.063)			(0.049)	(0.051)
Came by foot	0.741	0.451	-0.290***	-0.158***	0.594	0.414	-0.180**	-0.186***
			(0.044)	(0.041)			(0.068)	(0.068)

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. For the representative sample, robust standard errors clustered at the market level are in parentheses. For the dual practice sample, robust standard errors clustered at the provider level are in parentheses. Data are from patient-exit surveys which we obtained by observing all providers for a full day of practice. Columns (3) and (7) present binary regression coefficients from estimating the relevant row variable on an indicator for private provider visit, and thus represent the mean difference of the row variable between the private and public sectors. Columns (4) and (8) repeat the exercise but add market fixed effects in the representative sample and district fixed effects in the dual sample.

Table A.18: Difference between dual and non-dual providers' treatment of SPs (public sample only)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Effort			Diagnosis			Treatment						
	Time spent	Checklist	IRT Score	Gave diagnosis	Correct diagnosis (conditional)	Correct diagnosis (unconditional)	Correct treatment	Palliative treatment	Unnecessary treatment	Correct treatment Only	Antibiotic	Number of medicines	Referred patient
Panel A: Dual practice sample, with SP, case and district fixed effects													
Is a dual provider	-0.950***	-5.673*	-0.281	-0.005	-0.001	0.002	-0.021	-0.014	-0.022	-0.018	-0.106	-0.209	-0.021
	(0.368)	(3.266)	(0.247)	(0.076)	(0.109)	(0.054)	(0.066)	(0.072)	(0.066)	(0.026)	(0.083)	(0.247)	(0.049)
R-squared	0.161	0.048		0.120	0.273	0.061	0.337	0.212	0.099	0.044	0.139	0.157	0.162
Number of observations	163	163	102	163	63	163	163	163	163	163	163	163	163
Mean of non-dual observations	2.883	23.653		0.393	0.292	0.115	0.311	0.689	0.836	0.033	0.557	2.934	0.131
Mean of dual observations	1.960	17.553		0.382	0.385	0.147	0.373	0.637	0.833	0.020	0.490	2.833	0.078
Mean of sample	2.306	19.836		0.387	0.349	0.135	0.350	0.656	0.834	0.025	0.515	2.871	0.098
Panel B: Dual practice sample, with SP, case and district fixed effects													
Is a dual provider	-0.911**	-6.300**	-0.376	-0.078	-0.156	-0.057	-0.033	0.010	-0.061	-0.013	-0.156*	-0.286	-0.058
	(0.421)	(3.129)	(0.251)	(0.088)	(0.142)	(0.063)	(0.082)	(0.086)	(0.075)	(0.028)	(0.092)	(0.307)	(0.054)
Age of provider	-0.032**	-0.122	0.000	-0.000	-0.003	-0.002	-0.003	-0.008*	0.001	-0.002	-0.007	-0.030**	0.000
	(0.015)	(0.142)	(0.011)	(0.004)	(0.008)	(0.003)	(0.004)	(0.004)	(0.003)	(0.001)	(0.004)	(0.014)	(0.002)
Gender of provider (1=Male)	0.024	-0.162	0.073	-0.035	-0.066	-0.046	0.021	0.164	0.150	-0.040	0.256**	0.464	-0.163*
	(0.650)	(4.820)	(0.465)	(0.132)	(0.184)	(0.089)	(0.132)	(0.121)	(0.110)	(0.051)	(0.115)	(0.400)	(0.086)
Patient load during visit	-0.015	1.475*	-0.005	0.020	-0.025	-0.002	-0.014	0.034**	0.007	-0.006	0.008	0.024	0.018
	(0.073)	(0.891)	(0.061)	(0.020)	(0.048)	(0.018)	(0.019)	(0.015)	(0.022)	(0.004)	(0.027)	(0.088)	(0.016)
R-squared	0.215	0.137		0.147	0.350	0.106	0.355	0.266	0.203	0.099	0.257	0.259	0.276
Number of observations	139	139	89	139	54	139	139	139	139	139	139	139	139

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. Robust standard errors clustered are in parentheses. All regressions include a constant. Observations are at the SP-provider interaction level. In column (13) the dependent variable is the total number of medicines recommended to the patient (dispensed and/or prescribed).

Table A.19: Robustness to alternative metrics for public-private comparison

(Representative Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Effort		Diagnosis			Treatment					
	Time spent	Checklist	Gave diagnosis	Correct diagnosis (conditional)	Correct diagnosis (unconditional)	Correct treatment	Palliative treatment	Unnecessary treatment	Correct treatment only	Antibiotic	Number of medicines
Panel A: Best public vs. best private (by correct treatment)											
Is a private provider	1.632***	11.288***	0.235***	0.033	0.079	0.162**	0.074	0.169	-0.014	0.143	1.147***
	(0.388)	(2.855)	(0.090)	(0.136)	(0.054)	(0.079)	(0.077)	(0.117)	(0.056)	(0.109)	(0.429)
R-squared	0.453	0.417	0.430	0.714	0.363	0.592	0.447	0.353	0.218	0.435	0.463
Number of observations	286	286	192	76	192	192	192	192	192	192	192
Mean of public	2.547	16.000	0.271	0.154	0.042	0.271	0.521	0.708	0.042	0.250	2.063
Mean of private	3.613	24.551	0.438	0.238	0.104	0.438	0.535	0.750	0.049	0.292	3.014
Mean of sample	3.352	22.458	0.396	0.224	0.089	0.396	0.531	0.740	0.047	0.281	2.776
Panel B: Best public vs. best private (by checklist items)											
Is a private provider	3.216***	16.987***	0.263**	0.119	0.079	0.141	0.034	0.167	-0.027	0.222	1.581***
	(0.916)	(5.003)	(0.116)	(0.160)	(0.056)	(0.095)	(0.104)	(0.139)	(0.028)	(0.156)	(0.503)
R-squared	0.586	0.501	0.610	0.823	0.487	0.616	0.699	0.468	0.540	0.473	0.674
Number of observations	191	191	129	63	129	129	129	129	129	129	129
Mean of public	2.481	18.832	0.333	0.133	0.044	0.200	0.556	0.689	0.022	0.178	1.800
Mean of private	4.708	30.269	0.571	0.146	0.083	0.286	0.595	0.845	0.012	0.310	3.381
Mean of sample	3.938	26.317	0.488	0.143	0.070	0.256	0.581	0.791	0.016	0.264	2.829

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. Robust standard errors clustered at the market level are in parenthesis. All regressions include a constant and SP, case, and market fixed effects. Observations are at the SP-provider interaction level. In column (11) the dependent variable is the total number of medicines recommended to the patient (dispensed and/or prescribed).