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

Jishnu Das
Alaka Holla
Aakash Mohpal
Karthik Muralidharan[†]

27 February 2015

Abstract: We present among the first direct evidence on the quality of healthcare in low-income settings using a unique and original set of audit studies, where *standardized patients* were presented to a nearly-representative sample of rural public and private primary care providers in the Indian state of Madhya Pradesh. We report three main findings. First, private providers are mostly unqualified but spent more time with patients, and completed more items on a checklist of essential history and examination items than public providers, while being no different in their diagnostic and treatment accuracy. Second, we identify the private practices of qualified public sector doctors, and show that the *same doctors* exert higher effort and are more likely to provide correct treatment in their private practices. Third, we find a strong positive correlation between provider effort and prices charged in the private sector, whereas we find no correlation between effort and wages in the public sector. Our results suggest that market-based accountability in the unregulated private sector may be providing better incentives for provider effort than administrative accountability in the public sector in this setting. While the overall quality of care is low in both the public and private sectors, 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.

JEL Codes: D40, H10, H41, I11
Keywords: Standardized patients, quality of healthcare, healthcare markets, audit studies, India

[†] Das: World Bank and Center for Policy Research (CPR); jdas1@worldbank.org
Holla: World Bank; aholla@worldbank.org
Mohpal: University of Michigan; amohpal@umich.edu
Muralidharan: UC San Diego, NBER, and BREAD; kamurali@ucsd.edu

We are especially grateful to Michael Kremer for his involvement as a collaborator in the early stages of this project and for subsequent discussions. We thank Brian Chan, Veena Das, Diana Tabak and Manoj Mohanan for their contribution to designing and implementing the SP methodology. 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. 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, and it is widely believed therefore that the unregulated market-based delivery of healthcare is socially undesirable (Arrow 1963). In addition to the problems of asymmetric information, if optimal care requires the potential denial of services to customers (such as steroids or antibiotics), market-based healthcare may be over-responsive to demand, leading to inefficient provision (Prendergast 2003). Partly as a result of these beliefs, the default policy approach to delivering healthcare for the poor in most low-income countries has been 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, a majority of households in low-income countries choose to visit fee-charging private health care providers in the private sector, with the market share of private providers in India being over 70%.[1] In India, the focus of our study, private providers often operate under little de facto regulation and most have no formal medical training or qualifications (Rohde and Viswanathan 1995, Banerjee et al. 2004, CPR 2011). While the high use of the private sector could, in part, reflect the absence of alternative public care, this cannot be the only explanation. For instance, in our data from rural India, the private sector share is 80% even in markets with a qualified public doctor offering free care through public clinics with more than 60% of visits to private providers with no formal qualifications.

The high market share of unqualified private healthcare providers raises fundamental questions about the functioning of healthcare markets in these 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 on comparable medical cases – both

[1] While precise numbers vary by location and year of study, the market share of private providers in low-income countries often exceeds 50%. For instance, data from the Demographic and Health surveys show that 50% of households seeking pediatric outpatient care in Africa and 70-80% in India visit the private sector with little variation over the 20 years that these surveys have been collected (DHS 2005, Selvaraj and Karan 2009, Harding 2009, Grepin 2014, Shivakumar 2014). The World Health Surveys provide a more complete (but somewhat outdated) picture by including adult morbidity and here, the numbers vary from 30% in Sub-Saharan Africa to between 70 and 80% in India (Wagstaff, 2013).

with and without adjusting for the differences in composition of providers across sectors? 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 measures of 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 in low-income settings.[2]

This paper presents among the first direct measures of condition-specific metrics of the quality of care in low-income countries using data from a unique audit study conducted in rural areas of the Indian state of Madhya Pradesh (MP). Specifically, standardized 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. The SPs then made over 1,100 unannounced visits to a near-representative sample of public and private providers of primary healthcare services in rural MP and recorded detailed data on condition-specific metrics of care for each interaction.[3] These metrics include the providers' adherence to a checklist of questions and examinations deemed essential for making a correct diagnosis, their likelihood of pronouncing a correct diagnosis, and the appropriateness of any prescribed treatments. For brevity, we refer to these metrics as the "quality of care".

We present results from two sets of comparisons. First, we sent SPs to a near-representative sample of public and private health facilities in rural MP, and we use this data to compare the representative patient experience across public and private clinics. These differences reflect variation in both provider composition (including knowledge and intrinsic motivation) and effort 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 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

[2] Banerjee et al. (2004), Chaudhury et al. (2006), and Das and Hammer (2007) highlight the problem of low doctor effort in the public sector (high absence, low time spent with patients) and low training in the private sector. There is also a large medical literature on quality of care in the public and private sector. However, Coarasa, Das and Gummerson (2014) examine 182 cited studies in two systematic reviews of this 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 of the measurement protocols in section 3.

sample" and compares the quality of care across the public and private practices of the *same* doctors on the *same* set of cases.[4]

We report three main findings. First, while the majority of private providers in the representative sample have no medical qualifications, we find that they exert significantly more effort than public providers and perform no worse on diagnosis and treatment. Private providers spent 1.5 minutes more with patients (62% more), and completed 7.4 percentage point more items on a checklist of essential history and examination items (47% 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), or to offer treatments that were not clinically warranted (provided by 70 percent of public providers).

Second, in the dual practice sample we find that the same doctors spent more time with SPs, complete more items on the checklist, and are also more likely to offer a correct treatment in their private practice relative to their own public practice, while being no more likely to offer an unnecessary treatment. 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).

Our third set of results examines the link between financial compensation and quality of care in the public and private sectors. Our SPs paid whatever providers charged them, which gives us data on the prices charged for each completed consultation. We find a positive correlation between the fees charged 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. Thus, the market appears to reward provider effort as well as the provision of medication (both correct ones as well as unnecessary ones). In the public sector, SPs were provided free or nominally-priced care. Since there is no variation in prices, we examine the correlation between doctors' wages and quality of care. In contrast to the private sector, we find no correlation between wages in the public sector and any measure of provider quality.

[4] We used 22 SPs across the three cases, and the dual sample features the same case across both public and private practice, but with different (randomly-assigned) SPs presenting the case. All analysis includes SP fixed effects.

Finally, while healthcare in public facilities is notionally free to the consumer, it is not free to the taxpayer. We calculate the per-patient cost in the public sector, and estimate that this is four times higher than the fees charged by private providers in our sample.[5] Thus, the unregulated private market for healthcare in rural MP, which is mainly staffed by unqualified providers, appears to deliver higher provider effort and comparable overall quality of care, at a much lower cost per patient interaction.

Our contributions to the existing literature are three fold. On the methodological front, as is well recognized, the fundamental problem of inferring product quality in credence goods markets is that the optimal action is consumer specific and inefficiencies could include under-treatment, over-treatment or both. Our use of unannounced SPs constitutes a significant improvement in the measurement of healthcare quality in low-income settings, where the literature to date has used either medical vignettes or structured self-reported questionnaires to measure provider knowledge and/or observations of medical practices or prescription analysis to code provider-patient actions.[6] There are four advantages to our approach. 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 real observations, we would observe only the presenting symptoms and would have to speculate about the true underlying illness.[7] Third, we are able to observe prices charged for *completed* transactions, which allows us to study to what extent the unregulated market rewards quality, which improves upon audit studies in other settings that

[5] The estimates of cost per patient in the public sector are conservative because they only include the wage costs and do not include the cost of facilities, equipment, or medicines (whereas fees in the private sector should be all-inclusive). However, the calculations do require an assumption that the types of cases seen are comparable across public and private primary healthcare providers.

[6] Examples include Leonard and Masatu (2006, 2007), and Das and Hammer (2007). However, as acknowledged by these studies, both methods have limitations. Measures of provider *knowledge* can be very different from provider *practice* (it is typically an upper bound) which should caution against using vignettes as measures of actual quality of care delivered to patients (Rethans et al 1991). In participant observations, researchers never know the real illness, observed providers may act differently under observation (Leonard and Masatu 2006), and differential sorting of patients and illnesses across providers can potentially confound true differences in provider quality.

[7] Note that the vignettes approach allows for this (because the actual illness is known) – but is limited by the fact that the vignettes measure theoretical knowledge of the provider and not actual practice.

obtain price quotes but do not complete the purchase.[8] Finally, Hawthorne effects are not a concern in the SP context because providers do not know that they are being observed.

Substantively, the methodological advances in measurement combined with our ability to observe the same doctor across public and private practices, allow us to better isolate and highlight differences in quality of care across the public and private sectors.[9] Further, the positive correlation between measures of provider effort such as time spent and checklist completion and prices charged in the private sector combined with the lack of correlation with wages in the public sector suggests that compensation for provider effort may be an important mechanism for the higher quality of care in the private practice of public doctors. This is in line with a literature from the United States that uses variation in administrative pricing to uncover incentive effects in the provision of care (for instance, Clemens and Gottlieb 2014).

Further, we are able to study how *market* rather than administrative prices behave in an unregulated setting. The SP methodology allows us to detect *both* under and over-treatment in the provision of care. Our first observation is that both under and over-treatment are widespread in the public and private provision of care: 70% of our SPs received care that can be characterized as unnecessary in both the public and private sector, and only 2.6% received care that was the correct treatment and nothing more. Critically, there is a positive correlation between checklist completion/correct treatment and price as well as unnecessary treatments and price. This suggests that in unregulated markets for healthcare, market prices do reflect some information on the quality of care , but also that patients cannot evaluate whether they are being over-treated and charged for medically 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

[8] For instance, Ayres and Siegelman (1995) showed that African-American auditors received higher first-offers when negotiating for a car relative to white auditors. However, Goldberg (1996) showed that this was because the variance of the offer distribution of the African-American population was higher, so that the final sales prices in the observational data were the same, and higher first-price offers reflected optimal bargaining strategies on the part of the seller. 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.

[9] Our approach is similar to that in the literature testing for moral hazard in agricultural labor markets by comparing worker effort and output under different contractual arrangements (Shaban 1987; Foster and Rosenzweig 1994). Our key advance is in applying this approach to a credence good setting where output is harder to measure and there is substantial direct provision of the good by the public sector.

mechanics and cab rides for tourists (Dulleck and Kerschbamer 2006, Gruber and Owings 1996, Dulleck, Kerschbamer and Sutter 2011, Schneider 2012, Wolinsky 1993).

Finally, as is well recognized, demonstrating inefficiencies in market provision does not automatically imply that public provision will do better if government failures are as or more prevalent than market failures. If governance and administrative accountability in the delivery of healthcare through the public sector is poor and difficult to fix (Banerjee et al 2008, Hanna and Dhaliwal 2014), customer-driven accountability through the market may present a legitimate alternative in spite of its many theoretical (and empirical) weaknesses. 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, our measures of healthcare quality, and discusses sampling and representativeness; section 4 presents the main results and discusses robustness; section 5 discusses pricing and cost effectiveness; section 6 concludes with a discussion of policy implications and caveats.

2. Context

2.1. Healthcare in Rural India

Healthcare in India is delivered by both publicly and privately-run clinics and hospitals, and patients can choose among them. In the public sector, patients can obtain primary care on a walk-in basis in District hospitals, Community Health Centers (CHCs), Public Health Centers (PHCs), and sub-centers.[10] 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. Sub-centers are supposed to be staffed with qualified nurses and to be visited by doctors on a fixed rotation. Most doctors hold a Bachelor of Medicine and Bachelor of Surgery (MBBS) degree, the rough

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

equivalent of an MD in the United States, and they all receive a fixed salary from the government, with no variable compensation based on either patient load or quality of care.[11]

Consultations in public clinics are free or very nominally priced. Patients are also supposed to receive free medication, to the extent that it is available. Although a federally-funded insurance program for inpatient hospital care has been recently introduced, it did not extend to primary care and was not available in our study site during the period of our field work. Thus, the tax-funded public system of care was the only source of (implicit) public insurance in the system at the time of this study.

Although public facilities are theoretically accountable to administrative norms and procedures (documented in the *Civil Service Codes* for each state), both the perceptions of staff members and process measures of effort suggest severe deficiencies. Nationwide, doctor absences averaged 43 percent on any given day in 2003 and 40 percent in 2010 (Muralidharan et al. 2011; Centre for Policy Research 2011). These absences do not occur on predictable days or hours (Banerjee, Deaton and Duflo, 2004) and they are not easy to address.[12] 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 even disciplinary proceedings due to grave negligence (La Forgia et al, 2014).

While official policy documents of the Government mainly focus on improving the public system of primary healthcare (Planning Commission 2013), data from household surveys show that the fee-charging private sector accounts for over 70% of primary care visits (DHS 2005; Selvaraj and Karan 2009; CPR 2011; Shivakumar 2014). Barriers to entry for private healthcare providers are low. Provider qualifications range from MBBS degrees to no medical training whatsoever, and clinics can range from well-equipped structures almost identical to PHCs 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

[11] India also recognizes medical degrees from alternative schools of medicine including the BAMS (Bachelors in Ayurvedic Medical Sciences), the BHMS (Bachelor of Homeopathic Medical Sciences) and the BUMS (Bachelor of Unani Medical Sciences). However, providers with these qualifications are only licensed to prescribe medication in line with their training and are not given prescription rights on allopathic medicine. They also are not typically posted in the frontline healthcare system of PHCs, CHCs, and district hospitals that prescribe allopathic medicine.
[12] An experiment that tried time-stamp machines to link wages of nurses to their attendance did not generate any lasting impact over a 2-year period (Banerjee, Duflo and Glennerster, 2008).

medicines. While providers operating without a medical license are not legal and face the threat of an occasional raid, they have come to be the dominant source of care in these markets (as the data below will show). Regulations are seldom enforced and the court system for malpractice is severely underdeveloped (Jesani 1996). Further, attempts to shut down unqualified private providers typically fail because the providers are popular among local communities and often restart their practices (Rohde and Vishwanathan 1995).

2.1. Market Sampling and Summary Statistics

Our study was carried out in the Indian state of Madhya Pradesh (MP), which is 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 region and an index of health outcomes. We then conducted a household *census* in these villages, where we asked respondents to name all the providers from whom they sought primary care in the past thirty days and their locations (this list often included providers outside the sampled village, but in market clusters on the nearest main road). We then surveyed *all* providers in 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.

Table 1 (columns 1-3) present summary statistics based on the provider census (Panel A) and the household census (Panel B) in these markets. The census reveals a large supply of primary care providers (an average of 11 per market) and a high incidence of usage, with 46% of households reporting visiting a primary care provider in the 30 days prior to the survey. Our data show that rural residents frequently travel outside their own village for healthcare and therefore are able to access nearly three times as many healthcare providers relative to those in their own village. Comparing columns 2 and 3 also highlights the importance of considering not just the sampled village, but the effective market for healthcare services. There is less than one MBBS doctor per market on average, and almost none of them are in the sampled villages themselves and are only available outside the village. Overall access to MBBS doctors is very limited in rural MP, with only 4% of total patient interactions being with an MBBS doctor (Panel B).

The majority of the providers are private (7 out of 11 or 64%), and they account for 89% of household visits to a health provider in the month prior to the survey. Strikingly, 48% of all providers and 77% of all private providers (5.4 per village) have no formal training, but account for 77% of household visits. The share of visits to private providers is 88% when there is a public provider in the market, and is 83% even when there is a public MBBS doctor in the same market. The corresponding numbers for private *unqualified* providers is 74% when there is a public provider in the same market, and 60% when there is a public MBBS doctor in the same market.[13] These data nuance the view that the widespread use of private (and unqualified) healthcare providers is mainly driven by the lack of access to qualified public providers (Rao et al. 2011).

3. Measuring Healthcare Quality Using Standardized Patients

3.1. The Standardized Patient (SP) Methodology

Used routinely in training and evaluation of medical students in high-income countries, including the United States, standardized patients (SPs) are highly-trained 'fake patients' who present symptoms of an illness to a physician (like any other normal patient) to evaluate the quality of care received by a typical patient who would see that physician (Peabody et al. 2004; Rethans et al. 1991; Hutchinson et al. 1998; Ozuah, and Reznik 2007). They are coached to present not only the initial symptoms but to be able to answer an extensive range of questions corresponding to those that the physician may ask as part of history taking. We follow the same method (adapted to local conditions) and send SPs on unannounced visits to physicians during the course of a normal working day to measure the quality of healthcare received in rural India.

Survey enumerators were recruited from the districts where the study was conducted, and 22 of them were retained as standardized patients (SPs) for the study. Using a team comprising of a professional standardized patient trainer, two medical doctors, and a medical anthropologist who was familiar with the local forms of presenting symptoms and illnesses, the 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

[13] Note that even public facilities have a substantial number of 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 be the main healthcare providers in public clinics and also prescribe medication (given high doctor absence rates).

the mother or father of the child (see Das et al. 2012 for more details on the SP protocols).[14] The SPs visited sampled providers, who did not know they were receiving standardized patients and who therefore should have treated them as new patients.[15] After the interaction, SPs were debriefed within one hour, using 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 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 counties and minimized risk to SPs that could arise from unsafe invasive examinations, such as a blood test with an unsterilized needle.

Finally, we also picked cases where the role of suitable medical advice was important because real patients would be unlikely to be able to categorize them 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.[16] 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 health

[14] Das et al (2012) is co-authored with the medical experts mentioned above and validates the methodology of standardized patients with peer review from health specialists. The study presents summary statistics on overall quality of care in this setting. The current paper focuses on the economics of unregulated healthcare markets and does not seek to re-establish the validity of the SP methodology beyond describing the protocols and the measured outcomes. For more details on case presentations, please see all instruments posted on www.healthandeducationinindia.org

[15] The research ethics board of Innovations for Poverty Action approved this deception design. We describe sampling and representativeness in the next section.

[16] 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).

care provider would be 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 uses two different provider samples. First, we use the census of healthcare providers described earlier to construct a near representative sample of public and private healthcare providers in rural MP in three of the five sampled districts. 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 expect to know all of their patients, the SPs had to justify their presence in the area for example, through work 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 of a total of 46 villages across three districts.

Table 1 shows summary statistics from the provider and household censuses for the sampled and representative villages. Sampled villages have more providers on average than the representative sample of villages summarized in columns 1-3, which is not surprising since they are less remote. However, while the scale of sampled villages is larger, there is no difference in the fraction of different types of providers in the market and in the composition of provider types across the frame and the sample.

Since SPs visited clinics to obtain primary care, we excluded community health workers, midwives, and providers that only made home visits. We sampled providers in all public clinics (up to two providers per clinic), and sampled a maximum of six private providers in each market.[17] Data from this 'representative sample' allow us to compare the care provided across typical public and private clinics in rural MP (all our estimates are re-weighted by the inverse of the sampling probabilities to provide population representative averages). However, these differences reflect a combination of varying composition of providers (including their knowledge or professionalism), as well as the effect of practicing in the private sector.

[17] 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 1 for details on sampling.

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. We then identified the private practices of these doctors (we found a private practice for 61% of these doctors). We sampled and successfully administered SP visits to 118 public MBBS doctors, of whom 88 (75%) also had a private practice. Our 'dual sample' therefore consists of 88 MBBS doctors who also have a private practice, to whom SPs are presented 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. Appendix 1 and Tables A1 and A2 provide further details on the sampling and construction of both the representative and the dual sample of providers.

Table 2 (columns 1-3) provides summary statistics for the representative sample of providers. The providers are mostly middle-aged men and over 40% have not completed 12 or more years of education (Table 2, Panel A). Providers' practices have been open for 13-15 years, and they self-report receiving 28 and 16 patients per day averaged across private and public providers respectively. Most practices (82% of private and 100% 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. The main differences between public and private providers in the representative sample is that the former are more likely to have an MBBS degrees (26% vs. 8%), while the latter charge much higher consultation fees (an average of Rs. 3.7 in public clinics versus Rs. 48 in private clinics).

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

We randomly assigned three SPs to each sampled provider 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.[18] Since the rarity of unstable angina could raise suspicions if providers saw two travelers presenting the same case (even if the visits were separated by a few weeks as they typically were for the other two cases), 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 A4 (after we define measures of quality).

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 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 A3).[19] 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 2007 for details). We report both measures in our analysis.

[18] Since we had 22 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 while visiting 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.

[19] 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 than what the Indian government's own guidelines recommend. If we use the more extensive checklist, this would deflate the checklist adherence further below the low numbers that we document, but does not affect the relative performance of public and private providers (which is the focus of this paper).

Second, we examine diagnoses – whether or not the provider uttered any diagnosis to the patient and the accuracy of the diagnosis. We consider a diagnosis incorrect when it cannot even be considered partially correct – for example, a provider tells an asthma patient that she has a gastrointestinal problem or an unstable angina patient that the weather is causing his ailment. Our definitions of correct and incorrect diagnoses are presented in Table A3 - Panel B.

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, helpful, and/or unnecessary /harmful, based on inputs from our panel of doctors, pharmacists, and a pharmaceutical company (Table A3 - Panel C lists the specific treatments that fall into each category). Since providers can dispense or prescribe multiple medicines, we classify each medicine as correct, helpful, or unnecessary/harmful and thus allow their 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 being presented and that would relieve/mitigate the underlying condition. Helpful treatments are palliatives 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 helpful. 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 could take a longer view with antibiotic resistance in mind and consider it as ultimately harmful. However, none of the treatments we observed were directly contra-indicated as having immediately negative effects, and hence most of these represent unnecessary treatments as opposed to directly harmful ones.[20]

There are several advantages to using the checklist completion rate as a measure of provider quality. First, it has the highest statistical power because it is a continuous measure of quality,

[20] If the overall quality of care was 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.

whereas both diagnosis and treatment are binary. Second, there is no censoring unlike in the case of a diagnosis (which is not pronounced in the majority of cases made) and interpretation is less subjective than coding the 'correctness' of a treatment (see details below). Finally, it is possible that a provider may have gotten lucky in pronouncing a correct diagnosis or providing a correct treatment for the specific case that we had in mind, but may have made a mistake if the true underlying ailment was different (with the *same* presenting symptoms). In other words, it is the checklist that allows a provider to identify a specific diagnosis and treatment *regardless* of what the illness is for the presenting symptoms.

On the other hand, patient welfare depends mostly on being treated correctly by the provider, and not on checklist completion per se (or the diagnosis pronounced) and so we also focus on the rate of provision of correct treatment for the case that was presented. However, even after classifying all medicines as correct, helpful, and unnecessary, there are two challenges in coding the "correctness" of a treatment. The first is: "How should we interpret a referral?" 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.[21] Since we do 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 these as correct, our results are unchanged.

A second challenge in classifying treatments is that in the proxy 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.

While checklist completion rate and provision of a correct treatment are different outcomes, Figure 1 shows that the checklist completion rate is a very good predictor of the probability of

[21] Field notes from the SPs suggest that this often happened in public clinics where the qualified doctor was absent and where the officiating provider heard the case presented and told the SP to go somewhere else, as he was not able to treat him. By necessity, this is coded as a “referral” in our data.

providing a correct treatment, and the two measures of quality are highly correlated.[22] Further, checklist completion is strongly positively correlated with the time spent by the provider with the patient, which suggest that the checklist items were appropriate since providers do more items when they spend more time with the patient (Figure 2).

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 of providers we estimate:

$$q_{i(scp)m} = \beta_0 + \beta_1 Private_{ip} + \beta_2 X_p + \delta_s + \delta_c + \delta_m + \varepsilon_{i(scp)m} \quad (1)$$

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 term of interest. Since we pool cases and SPs and there may be systematic differences across them, all our 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 between public and private providers.

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.[23] To isolate the impact of private sector practice, we re-estimate (1) in the dual

[22] As Das et al. (2012) point out this also suggests that providers believed the SP presentations. In a Bayesian model, the provider, faced by a patient with chest pain, would assign subjective probabilities to differential diagnoses. If further questioning led the provider to believe that the patient did not have a serious condition, we should find that the likelihood of correct treatment declines with questioning. We find precisely the opposite. The providers who got it wrong did not ask many questions or undertake examinations; providers who questioned the patient and conducted exams were more likely to conclude that the SP had the condition that they presented with.

[23] A further caveat in interpreting β_1 is that it equally weights all providers and does not account for differential market shares of the providers in the representative sample. Given differentiated characteristics of providers

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[24]; then we add district fixed effects (since the sample for the dual practice sample was drawn from a district-level frame of public MBBS doctors and not at the market level as was done 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 (1) in the representative sample. Our outcome variable is 'provider effort', measured by time spent and checklist completion. While the results are very 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. Base levels of effort among representative public providers are 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 spend 1.5 minutes more per patient and complete 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 score 0.61 standard deviations higher.

Columns 4-6 repeat the analysis in the dual sample, and we find the same pattern. 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%) in 35% less time (0.9 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' double time spent with patients, completed

including access and price, there is no obvious way to weight them, which is why we present mean effects in a representative sample.

[24] 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. However, since the case was randomized across the public and private practices of the doctor and assignment was balanced on measures of quality on the other case (see Table A4), 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 (2) with provider fixed effects and the results are unchanged (but are then mostly driven by variation in the asthma case).

50 percent more checklist items, and scored 0.73 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 Vermeesch (2012) look at checklist completion as a result of the introduction of performance pay in Rwanda. They find that performance pay increased checklist completion by 0.09 standard-deviations; we find that the difference in checklist completion across public and private practices of the same doctor is seven times larger.

These differences are seen clearly in Figures 3-5. Figure 3A plots the IRT-score distribution (based on checklist completion) of public and private providers in the representative sample, Figure 4A does so for the dual sample, and Figure 5A pools all four samples together. Figures 3B, 4B, and 5B plot the cumulative distribution functions (CDF) for the same comparisons. Distributions of checklist completion for private providers first-order stochastically dominates that of the public providers (Figure 3B), and the corresponding distribution for the private practices of public providers also first-order stochastically dominates that of their public practices (Figure 4B). Finally Figures 5A and 5B show that 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).

Focusing on individual checklist items (Table A5) on each 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. In addition to β_1, Table 3 (columns 1-3) also shows that there is no correlation between the possession of any formal medical qualification and checklist completion, suggesting that formal qualifications may be a poor predictor of actual quality of care received.[25] There is also suggestive evidence that doctors with a higher patient load may be better ones (they have higher IRT scores) but spend less time with patients.

[25] We also find no correlation between possession of an MBBS degree and checklist completion, but note that there are very few MBBS doctors in the representative sample (see Table A1 for details on sample composition). Also, as noted earlier, MBBS doctors complete more checklist items per unit time, but spend less time with patients.

4.3. Diagnosis

Table 4 follows 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 who 15 percent offer a correct one. The unconditional probability of a correct diagnosis was 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 significantly higher in the dual practice sample (15 percent vs. 4 percent), which is not surprising since these comprise of trained MBBS doctors. Even among these doctors, however, there is no significant difference in the likelihood of a correct diagnosis between their public and private practices. Overall, the summary statistics, our price regressions (seen later), and our detailed field work suggest that pronouncing a correct diagnosis is not seen by providers (and the market) as being essential in this setting.

4.4. Treatment

Table 5 reports on several outcomes related to the treatment offered, coded as discussed in section 3.3. The mean probability of receiving at least one correct treatment from a representative public provider was 27 percent. However, these providers offered non-indicated treatments at much higher rates, with a 66 percent probability of providing a helpful treatment, and a 70 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 only 2.6 percent. Finally, we also report two measures that could proxy for over-treatment – the rates of antibiotic prescriptions and polypharmacy (provision of multiple medications).[26] Antibiotics were prescribed or dispensed in 26 percent of interactions (though they were not indicated), and 70 percent of treatment protocols could be characterized by polypharmacy

In the representative sample, we do not find a significant difference between public and private providers on the probability of providing a correct, helpful, or unnecessary treatment;

[26] The WHO considers poly-pharmacy a useful proxy for over-treatment.

however, point estimates suggest that private providers have a higher probability of providing both correct and unnecessary treatments. Consistent with higher rates of treatments being offered in the private sector, private providers in the representative sample have a 16.5 percentage point higher rate of poly-pharmacy.

In the dual practice sample, treatment outcomes are significantly better in the private practices of public doctors (relative to their performance in their public practices). The rate of correct treatment is 39 percent higher (14.8 percentage points on a base of 38 percent), and the rate of antibiotic provision is 25 percent lower (12.1 percentage points on a base of 48 percent). These results are robust to the inclusion of controls and alternative definitions of correct treatment (see Table A6).

4.5. Knowledge and Effort of Public and Private Providers

The results in Tables 3 and 5 suggest that the typical private provider exerts significantly greater effort than his public counterpart, but that this greater effort does not lead to better treatment outcomes. How should we interpret this result?[27] One possible explanation is that the private providers exert greater effort, but have lower medical knowledge (note that the majority are unqualified). Since quality of care is a function of both provider knowledge and effort, it appears that the private providers make up for lower knowledge with higher effort, leading to comparable overall levels of treatment accuracy. To examine this possibility further, we use the 'discrimination' parameter of each checklist item (as estimated by the IRT-model), to classify the individual items into terciles of low, medium, and high discrimination items. As discussed earlier, higher "discrimination" items are those that are more effective at distinguishing provider quality. The terciles for each individual item are indicated in Table A5, and we see for instance that in the unstable angina case, asking whether the pain was radiating was a high discrimination item compared to asking when the pain started, which was a low discrimination item.[28]

Table 6 reports the same specifications as in Table 3, but compares public and private providers on the fraction of checklist items completed by the level of item discrimination. In

[27] Note that the most obvious explanation for this result (that checklist completion is not a good predictor of better treatment) is not true given the strong positive correlations between the two seen in Figure 1.
[28] 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.

general, providers are much 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 high-discrimination checklist items. However, in the dual sample doctors are significantly more likely to complete both low and high-discrimination items in their private practice. These results suggest that while the representative private provider does exert more effort, this effort may be directed in areas where marginal productivity is lower (which is consistent with lower knowledge). The results also highlight the importance of using the dual sample for isolating the effect of incentives since differences across private and public providers in the representative sample confound variation in provider composition, knowledge, and effort.

4.6. Real patients

While the use of SPs to measure quality of healthcare has several advantages over the next best method of participant observations (POs) where doctors are observed during the course of a typical work day, one limitation with the SP method is that it can only be conducted on a small number of cases, which may not adequately represent the range of illnesses that patients come with in this setting. Further, we may worry that the audit methodology may represent "off equilibrium" situations in the market that do not extend to its general functioning.[29] We therefore supplemented our data collection (after completing the SP modules) by conducting day-long observations in provider clinics to code actual provider-patient interactions. We conduct these observations both in the representative and in the dual sample observing the provider in both private and public practices. While we cannot code the accuracy of the diagnosis and treatment from these observations (since we do not observe the underlying illness), we record several observable characteristics of each patient interaction based on over 1000 interactions in both samples.

Table 7 reports results using a similar specification as in (1) but with the data from real patient interactions, and uncovers similar patterns. Private providers spend more time with patients, ask more questions, and are more likely to conduct a physical exam. But they are less likely to dispense medicine on premises and are more likely to prescribe them instead. We find

[29] Note that while the SP methodology allows us to compare the same cases across public and private providers, it is possible that it may be off the equilibrium path for a patient with a serious condition to visit one type of provider.

the same results in both the representative and in the dual samples, and the results are robust to including market fixed effects and patient controls (based on conducting exit interviews with the patients as they were leaving the clinics). Thus, our results from observing real patients corroborate those from the SPs, with significantly higher levels of effort among private providers using both methods.

4.7. Robustness

Our main results pool data across cases to maximize power. For completeness, we also show the results from in Tables 3-5 by case (Table A6). The superior performance of private providers on time spent 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 11 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, such as treating all 'referrals' as a 'correct' treatment (Table A7 shows the specific treatments offered by case, including referral frequency; Table A8 shows that the results 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 fraction of cases where the provider did not provide these and instead asked to see the child (see Table A7). Since there is a possibility that the 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 A9). We also show the core results with controls for clinic-level infrastructure and facilities (Table A10), and all the results continue to hold as in Tables 3-5, suggesting that the results are not being driven by differences in facilities and infrastructure across public and private clinics.

5. Results – Pricing and Cost Effectiveness

A unique aspect of our study is that we are able to observe prices charged for a completed transaction in an audit study, while simultaneously having objective measures of quality. Since, we also observe the wages paid in the public sector, and have corresponding measures of quality of care in the public sector as well, we can compare the extent to which market and state provision compensates providers for quality of care.

5.1. Correlates of Prices Charged among Private Providers

Figure 6 shows a strong positive correlation between price charged and checklist completion, suggesting that the market rewards effort. Table 8 (column 1) shows binary correlations between prices charged and our various metrics of healthcare quality, while columns 2-8 show the correlates of prices charged with measures of quality in multiple regressions. We see in column 1 of Table 8 that the market rewards several measures of quality of care including time spent, checklist completion rates, and provision of a correct treatment. On the other hand, we find no price premium for pronouncing a correct diagnosis, and also find a price penalty for referrals, which may reduce provider incentives to refer patients away in cases where it is optimal to do so. Finally, we find a price premium for dispensing medicines as well as for 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. The provision of a correct treatment is no longer independently rewarded, but this is likely due to the high correlation between the provision of a correct treatment and the checklist completion rate (as seen in Figure 1), and also with the incidence of dispensing medicines (both of which continue to be significant in the multiple regressions as well.

The correlates of pricing observed in Table 8 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. On the other hand, the results are consistent with evidence from other 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.[30]

A plausible interpretation of our results is that "observable" markers of quality to individual patients (such as time spent, checklist completion, and provision of medication) are rewarded by the market. Since these observable markers are also correlated with the provision of correct treatment, it appears that the market rewards better quality of care, while at the same time providing incentives for unnecessary treatments. Finally, while patients may not be able to assess the "quality" of care in an individual interaction with a provider, the "market" may be able to do this better over time.

5.2. Correlates of Public Sector Wages

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, correct diagnosis, or correct treatment (Table 9). On the contrary the correlations suggest that better paid providers are more likely to give out an unnecessary treatment. Thus, while the administrative structure of wages in the public sector reward qualifications and seniority, we find no correlation between wages and the quality of care provided.[31]

5.3. 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 10 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 facilities, equipment, medicines or administration), whereas the fees charged are the only

[30] However, unlike the existing credence good literature, we are also able to observe over-provision in the public sector and find that both public and private providers have similarly high levels of provision of unnecessary treatments (Table 5).

[31] These results are similar to those found in publicly-provided education in India, where teacher salaries increase with qualifications and seniority, but are not correlated with teachers' effectiveness at raising test scores (see Muralidharan 2013). It is theoretically possible for public doctors to be given non-monetary rewards for performance, such as more desirable postings. In practice, these are typically obtained by more senior doctors, and we find no correlation between seniority and quality of care provided. We also find no correlation between quality of care provided and measures of "posting quality" such as proximity to facilities (results available on request).

source of revenue for private providers, and hence will cover all operating costs. It also assumes that all patients shown in the official records of the PHC/CHC were true patients. Finally, as is standard in comparative cost effectiveness analysis of this sort, we assume that there is a comparable case mix for primary-health visit across public and private facilities.

Thus, even though private providers charge much more per consultation than public ones (as seen in Table 1), the fees charged by private providers of around Rs. 50/consultation are around a fifth of the cost per patient interaction in the public sector. The relative social cost per patient interaction is likely to be even higher in the public sector because the calculation above does not include the cost of the greater number of unsuccessful trips that patients make to see a doctor in the public sector. In our own case, SPs made more unsuccessful trips to see a doctor in the public practice relative to the private practice (1.64 attempts per case versus 1.39) and had a significantly lower rate of successfully seeing a public MBBS doctor in his public practice relative to his private practice (75% versus 93%).

6. Discussion and Conclusion

Using an audit methodology, we present the first set of results on the quality of public and privately provided healthcare in a low-income country that features a de facto unregulated private sector. Comparisons of representative public and provider samples suggest that patients in our setting have few good options for healthcare — public or private. Private sector providers, the majority of whom have no formal medical training, spend more time with patients and are more likely to adhere to a checklist of recommended case-specific questions and examinations, but their effectiveness appears to be ultimately limited by their low level of medical knowledge. Public sector clinics, though theoretically staffed by qualified providers, are characterized by lower provider effort. Posts are vacant and doctors are frequently absent, so that even in a public sector clinic, the patient often sees a provider without formal training. The lower effort (compared to the private sector), appears to offset the benefit of having 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. At least on the basis of these data, there is little evidence that patients are harming their health more by going to the private relative to the public sector, and the price paid could well reflect patient demand for provider effort (including more reliable presence at the clinic).

Comparing the *same* provider in the public and private sector allow us to isolate the effect of customer accountability in the private sector and compare it with administrative accountability in the public sector. The first appears to perform better on all counts. Adherence to checklists and correct treatment rates are higher in the provider's private clinic, and rates of incorrect treatments are identical in both sectors.

Better treatment according to medical guidelines is consistent with the hedonic price-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. Where customer accountability *does* fail is in its ability to control the extent of unnecessary medication. Patients frequently receive treatments that they didn't need, and they pay for them. Surprisingly, however, the rate of provision of unnecessary medication is equally high in the public clinics. Finally, our best estimates of cost per patient interaction suggest that the public healthcare system in India spends over four times more but does not deliver better outcomes than the private sector.

Our results are directly relevant for Indian and global health policy debates, which have been hampered by a lack of empirical evidence on the quality of clinical interactions in the public and private sector. Under the status quo, considerable attention has been focused on inadequate access to publicly-provided healthcare and the need to increase spending on the public healthcare system (Reddy et al. 2011; Shivakumar et al. 2011; Planning Commission 2013). Our results suggest that enthusiasm for the public sector as the primary source of healthcare in resource poor settings has to be tempered by the extent to which administrative accountability is enforced in the system. More broadly, the quality of healthcare depends both on provider knowledge and effort, and there are likely positive returns to investing in improved incentives for effort in the public system of healthcare delivery (where providers are more qualified) or increased training and credentialing among private healthcare providers, who have better incentives for effort.[32]

[32]Our results should not necessarily be interpreted as recommending a fee-for-service model of compensating healthcare providers to provide them with incentives for effort. There is considerable evidence from the US that such a model can induce over-treatment (Clemens and Gottlieb 2014), which may be especially undesirable in a low-income setting. On the other hand, it is worth noting that the status quo public healthcare system in India

However, 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; Planning Commission 2013; Shivakumar 2014). This viewpoint is often justified by ad hoc assumptions that patients—particularly those who are poor and illiterate—are unable to make accurate decisions regarding their health care. While certainly possible, such an assertion would have to be backed by empirical evidence on patient demand and quality of care. Our paper is one of the first attempts to do so, and expanding this methodology to other conditions and settings will allow for a richer understanding of the functioning of medical systems in settings with low resources and administrative capacity.

provides negative incentives to doctors for exerting effort, since greater effort is likely to lead to an increased load of patients with no increase in compensation.

References (to be updated)

Arrow, Kenneth J. 1963. "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review* 53 (2): 941-73.

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

Banerjee, Abhijit V., Angus Deaton, and Esther Duflo. 2004. "Wealth, Health, and Health Services in Rural Rajasthan." *American Economic Review* 94 (2): 326-30.

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 European Economic Association* 6 (2-3): 487-500.

Center for Policy Research. 2011. "Mapping Medical Providers in Rural India: Four Key Trends." Policy Brief, Center for Policy Research, New Delhi, India.

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

Das, Jishnu, Alaka Holla, Veena Das, Manoj Mohanan, Diana Tabak, and Brian Chan. 2012. "In Urban and Rural India, A Standardized Patient Study Showed Low Levels of Provider Training and Huge Quality Gaps." *Health Affairs* 31 (12): 2774-84.

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

DHS. 2007. "National Family Health Survey (NFHS-3) 2005-2006: Key Findings." Available at http://dhsprogram.com/publications/publication-SR128-Summary-Reports-Key-Findings.cfm

Dulleck, Uwe, and Rudolf Kerschbamer. 2006. "On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods." *Journal of Economic Literature*, 44 (1): 5-42.

Dulleck, Uwe, Rudolf Kerschbamer, and Matthias Sutter. 2011. "The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition."*American Economic Review*, 101 (2): 526-55.

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

Foster, Andrew D., and Mark R. Rosenzweig. 1994 "Information, Learning, and Wage Rates in Rural Labor Markets." *Journal of Human Resources* 28 (4): 759-790.

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.

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

Hutchinson, Brian, Christel A. Woodward, Geoffrey R. Norman, Julia Abelson, Judy A. Brown. 1998. "Provision of preventive care to standardized patients." *Canadian Medical Association Journal* 158 (2): 185-193.

Jesani, Amar. 1996. “Laws and Health Care Providers: A Study of Legalization and Legal Aspects of Health Care Delivery.” CEHAT, Bombay

Leonard, Kenneth L., and Melkiory C. Masatu. 2005. “The Use of Direct Clinician Observation and Vignettes for Health Services Quality Evaluation in Developing Countries.” *Social Science and Medicine* 61 (9): 1944-51.

Leonard, Kenneth L., and Melkiory C. Masatu. 2006. “Outpatient Process Quality Evaluation and the Hawthorne Effect.” *Social Science and Medicine* 63 (9): 2330-40.

Leonard, Kenneth L., and Melkiory C. Masatu. 2007. “Variations in the Quality of Care Accessible to Rural Communities in Tanzania.” *Health Affairs* 26 (3): 280-92.

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

Ozuah, Philip O., and Marina Reznik. 2007. "Residents' asthma communication skills in announced versus unannounced standardized patient exercises." *Ambulatory Pediatrrics* 7 (6): 445-448.

Peabody, John W, Jeff Luck, Peter Glassman, Sharad Jain, Joyce Hansen, Maureen Spell, and M Lee. 2004. “Measuring the Quality of Physician Practice by Using Clinical Vignettes: A Prospective Validation Study.” *Annals of Internal Medicine* 141: 771–80.

Planning Commission of India. 2013. “Health, Nutrition, and Family Welfare.” Available at http://planningcommission.nic.in/sectors/health.php?sectors=hea

Rao, Mohan, Krishna D. Rao, A. K. Shiva Kumar, Mirai Chatterjee, Thiagarajan Sundararaman. 2011. “Human Resources for Health in India.” *The Lancet* 377 (9765): 587-598.

Rethans, Jan-Joos, Ferd Sturmans, Riet Drop, Cees van der Vleuten, and Pie Hobus. 1991. "Does competence of general practitioners predict their performance? Comparison between examination setting and actual practice." *British Medical Journal* 303 (6814): 1377-80.

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

Schneider, Henry. 2012. "Agency Problems and Reputation in Expert Services: Evidence from Auto Repair." *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): 983-920.

Wolinsky, Asher. 1993. "Competition in a Market for Informed Experts' Services." *The RAND Journal of Economics* 24 (3): 380–398.

World Bank. 2003. "World Development Report 2004." The World Bank, Washington D.C.

Appendix 1: Mapping and Sampling of Providers in the Representative Sample

A. Mapping of Providers

We first sampled five districts in the state of Madhya Pradesh, stratified by region and an index of health outcomes (the sampled districts were Chhindwara, Gwalior, Jhabua, Rajgarh, and Shahdol). Within each district, we randomly sampled 20 villages using probability proportional to size (PPS) sampling after restricting the frame of villages to those with a population below 5.000. Thus the sample of villages is representative of the state of Madhya Pradesh.

In each of the 100 sampled villages, we first conducted detailed Participatory Resource Assessments (PRA) with groups of villagers to obtain a list of all providers households sought primary care services from. To obtain a comprehensive list of providers households regularly accessed, at least three PRAs were conducted in each village and names and addresses of providers both within and outside the village were recorded. Surveyors then compiled a list of all providers by location (inside village and outside village) and visited every provider inside the village and verified the listing. We also administered a Provider Census to obtain basic demographic, practice characteristics and clinic characteristics details from the provider. During the verification, providers were also asked to list all other providers practicing within the village and any new provider was added to the list and a census was administered to the new provider.

We then administered a short household census to every household in the sample village which collected information on household demographics and health care seeking behavior. We asked household heads to list all members of the household and report the incidence of any primary illnesses in the past one month. For all reported morbidities in the past month, we then asked households to report if they sought care and to identify the name and practice location of the provider they received care from. During this process surveyors used the list of providers they had obtained from PRAs to aid households in identifying providers they has visited. When households reported visiting providers outside of the sampled village, we collected providers' name, address and village/town name and practice details to the extent possible.

For each sampled village, after the completion of all household surveys, we analyzed information from the household survey in the field and listed all locations (villages and towns) other than the sampled village where households reported they visited in the last one month. If 5 percent or more households reported visiting a provider in an outside location, we identified that location as a "cluster village" and considered it a part of the "health care market" for the sampled village.

Once all cluster villages for a sampled village were identified, our surveyors visited those villages and mapped all providers practicing in the cluster locations, regardless of whether households reported visiting them in the past month or not. Once again, PRAs were conducted in multiple locations of each cluster village and a provider list was generated. Surveyors then visited each provider on the list, administered a census and also asked providers to list any other

provider practicing in that location. This exhaustive mapping process ensures that we mapped the universe of providers that households in our sampled villages sought primary care services from and had access to.

B. Sampling of Providers for SP visits

The SPs needed to have a plausible reason for being in the village (which they were not from), and the typical narrative was that they were traveling and passing by the village. We therefore dropped remote villages from the sample where SPs were administered because the SPs would have had a higher risk of detection in these locations. We dropped two districts (Jhabua and Shahdol) where several of the villages were remote, and restricted our initial sampling frame for SP visits to the 60 sampled villages in the three districts of Chhindwara, Gwalior, and Rajgarh (for the representative sample).

The universe of providers for SP work is smaller than what we mapped. Four broad categories of providers in our data were excluded from SP work. First, we excluded all nurses, midwives, community health workers, day-care center workers, chemists and pharmacists both in the public and private sector. Although in our surveys, household directly report that they seek primary care from these provider types, government regulations forbid these health workers from providing primary care. Second, we exclude all mobile and itinerant providers from the sampling universe. Administering SPs to these providers would have been extremely difficult. Third, we exclude all providers with whom we could not complete a Provider Census during the mapping process. Finally, we also exclude providers who practice from home and do not have a signboard (in the absence of basic information on the provider and/or a publicly visible signboard identifying a provider, the SP would have risked detection by visiting such a provider). Overall, we had a total of 719 eligible providers in the 60 sampled villages, of whom 144 were public and 575 were private (Appendix Table A1).

Excluding the abovementioned provider types also rules out 8 markets from our sample. This is because these markets did not have a single eligible provider type. Furthermore, we excluded 4 markets because these were extremely small, remote and lacked accessibility. Administering SPs in these markets would have involved a high risk of detection. These four markets were identified as such at the suggestion of our field staff and their remoteness verified using GPS locations of households and providers. Finally, two villages – one that shares a common cluster with another village and one that has another sampled village as its cluster – did not have any eligible providers within the sampled village. Thus the unique number of markets we sample from is 46. Restricted to these 46 markets, we had 649 providers in our sampling universe; 130 of those are public providers and 519 are private providers.

Once the list of markets to be sampled and the sampling frame was established, we randomly sampled up to two eligible providers in each public clinic and up to six private providers in each market. One market in Gwalior district was an exception to this rule. In the cluster village for

this particular market, we found 113 providers practicing in the cluster village. In this market, we relaxed our sampling methodology to sample up to 20 private providers. Following the sampling methodology described above, we sampled a total of 247 providers (Appendix Table A1). The breakdown between public and private is 45 and 202 respectively. There were 28 providers with MBBS degrees in the sample, 12 are public and 16 are private.

C. Completion of SPs

We managed to complete at least one SP interaction with 224 providers we had mapped. Of these, 214 providers are one of the 247 providers we sampled. Ten are providers who we had mapped but did not sample. These happened to be practicing in the clinic at the time of SP visit and the SP was completed with them. In 10 additional clinics (18 observations) our SPs saw providers other than the one mapped and sampled, whose identity we could not verify (these providers were not a part of our mapping exercise). Seventeen of these 18 instances occurred in the public sector and one in the private sector.

Appendix 1B: Mapping and Sampling of Providers in the Dual Practice Sample

A. Mapping of Providers

We obtained a list of all Primary Health Centers (PHCs) and Community Health Centers (CHCs) from the Ministry of Health of Madhya Pradesh. After excluding the PHCs/CHCs which were covered as part of the representative sample, we mapped 200 facilities in this process. Of these 200 facilities, 40 did not have a MBBS provider posted (see Appendix Table A2). In the remaining 160 PHCs/CHCs we located 216 unique providers (some providers are mapped to multiple facilities). Our field team then embarked on detailed field work to find out if the providers operated private practices and if so, to locate their private practices. We found that 132 of the 216 providers (61.1 percent) operated a private practice (dual practice providers). Once the mapping exercise was complete, we administered our Provider Census to all mapped providers. To the extent possible, the census was administered in the private clinic of the provider.

B. Sampling of Providers

We sampled one provider from every PHC/CHC with preference for a dual practice provider. Often a provider is posted to multiple public facilities, and in cases where there were no additional provider 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 143 of the 160 facilities we could have sampled from. The total number of providers is 143, and of these, 94 operated private practices.

C. Completion of SPs

We managed to complete at least one SP observation with 118 of the 143 providers sampled. Of the 49 providers without private facilities, we completed an SP with 30 providers. Of the 94

providers with private practices, we were able to get at least one observation with 88 providers (either public or private). The number of dual practice providers for whom we have at least one observation in both public and private is 69. As discussed in the text, and seen in Table A10, it was more difficult to complete a case in the public practice of the public MBBS doctors (because of high absence rates), and we had a lower completion rate in the public practices of these doctors than in the private practices (75% versus 93%).

Table 1: Health market attributes

	(1)	(2)	(3)	(4)	(5)	(6)
	Madhya Pradesh (5 districts, 100 markets)			SP Sample Villages (3 districts, 46 markets)		
	Total	Inside village	Outside village	Total	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 provider in last 30 days	0.46			0.58		
	(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 (conditional on public availability)	0.88	0.89	0.83	0.95	0.96	0.91
	(0.33)	(0.31)	(0.38)	(0.22)	(0.20)	(0.28)
Fraction of visits to private sector (conditional on public MBBS availability)	0.83	0.84	0.79	0.93	0.98	0.90
	(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 (conditional on public availability)	0.74	0.82	0.54	0.81	0.86	0.64
	(0.44)	(0.38)	(0.50)	(0.39)	(0.35)	(0.48)
Fraction of visits to unqualified providers (conditional on public MBBS availability)	0.60	0.77	0.38	0.66	0.81	0.39
	(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 month	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 providers (5 districts)				Dual Practice sample (5 districts)		
	Public	Private	p-value of (1)-(2)	All public	Public without dual practice	Public with dual practice	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	0.84	0.85	0.87
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				
No medical training	0.61	0.68	0.42	0.00	0.00	0.00				
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 RMP, other degrees (which could not be verified) and providers who self-reported no formal training. In the public sector of the representative sample, there are 22 with no formal qualifications, 4 with RMP and 1 who reported other degree. In the private sector, there are 128 with no formal qualfication, 46 with RMP and 10 who reported other 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: Provider 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***	0.512**	1.471***	8.888***	0.729***
	(0.250)	(2.488)	(0.211)	(0.267)	(1.762)	(0.178)
R-squared	0.305	0.160		0.237	0.219	
Number of observations	662	662	233	331	331	138
Mean of public	2.388	15.287		1.562	17.677	
Mean of private	3.703	22.302		2.965	28.223	
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***	0.608**	1.475***	8.882***	0.729***
	(0.244)	(1.948)	(0.273)	(0.259)	(1.762)	(0.180)
R-squared	0.391	0.259		0.258	0.233	
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**	0.500*	1.452***	9.414***	0.770***
	(0.319)	(2.338)	(0.301)	(0.268)	(1.827)	(0.190)
Has MBBS	-0.156	3.285	0.043			
	(0.568)	(2.940)	(0.257)			
Has some qualification	-0.131	2.518	0.157			
	(0.299)	(1.716)	(0.151)			
Age of provider	-0.004	-0.046	0.000	0.005	-0.064	0.004
	(0.012)	(0.071)	(0.008)	(0.015)	(0.102)	(0.101)
Gender of provider (1=Male)	0.653	-0.949	0.212	-0.077	-1.383	-0.288
	(0.544)	(3.529)	(0.327)	(0.386)	(2.639)	(0.309)
Patient load during visit	-0.096*	-0.144	0.082**	-0.106*	-0.283	0.013
	(0.052)	(0.554)	(0.040)	(0.062)	(0.424)	(0.517)
R-squared	0.399	0.259		0.275	0.233	
Number of observations	638	638	221	302	302	126

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. Observations are standardized provider-patient interactions, except in IRT score column where each observation is a provider. The score is computed using all cases, plausible values scores are used. Market fixed effects are used for the representative sample, and district fixed effects for 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	-0.050	0.018
	(0.052)	(0.057)	(0.022)	(0.068)	(0.105)	(0.053)
R-squared	0.130	0.121	0.075	0.130	0.114	0.054
Number of observations	440	178	440	201	88	201
Mean of public	0.263	0.150	0.039	0.380	0.395	0.150
Mean of private	0.431	0.135	0.058	0.495	0.380	0.188
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	0.089	-0.067	0.018
	(0.072)	(0.093)	(0.031)	(0.069)	(0.109)	(0.054)
R-squared	0.218	0.301	0.145	0.149	0.176	0.066
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	0.083	0.005	0.037
	(0.081)	(0.111)	(0.035)	(0.072)	(0.122)	(0.058)
Has MBBS	-0.092	0.108	0.008			
	(0.093)	(0.134)	(0.039)			
Has some qualification	0.023	-0.010	-0.012			
	(0.074)	(0.075)	(0.028)			
Age of provider	-0.002	-0.005*	-0.002	0.002	-0.001	0.000
	(0.003)	(0.003)	(0.001)	(0.004)	(0.008)	(0.003)
Gender of provider (1=Male)	-0.089	0.272*	0.079*	-0.125	-0.054	-0.086
	(0.126)	(0.145)	(0.041)	(0.109)	(0.182)	(0.079)
Patient load during visit	-0.003	-0.017	-0.005	-0.018	-0.004	-0.005
	(0.014)	(0.011)	(0.004)	(0.018)	(0.034)	(0.013)
R-squared	0.222	0.362	0.159	0.185	0.217	0.096
Number of observations	423	173	423	183	80	183

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. Observations are standardized provider-patient interactions. Market fixed effects are used for the representative sample, and district fixed effects for 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	Helpful treatment	Unnecessary treatment	Correct treatment Only	Antibiotic	Poly-pharmacy	Correct treatment	Helpful treatment	Unnecessary treatment	Correct treatment Only	Antibiotic	Poly-pharmacy
Panel A: SP and case fixed effects												
Is a private provider	0.068	0.014	0.056	-0.020	0.016	0.130*	0.147**	0.029	-0.031	-0.009	-0.119*	0.075
	(0.056)	(0.055)	(0.074)	(0.021)	(0.062)	(0.068)	(0.064)	(0.063)	(0.054)	(0.024)	(0.068)	(0.048)
R-squared	0.302	0.051	0.070	0.029	0.079	0.054	0.271	0.041	0.075	0.018	0.114	0.138
Number of observations	334	365	392	440	440	440	199	200	201	201	201	201
Mean of public	0.267	0.662	0.696	0.026	0.263	0.697	0.380	0.730	0.820	0.030	0.480	0.800
Mean of private	0.361	0.695	0.771	0.006	0.279	0.837	0.566	0.760	0.812	0.020	0.386	0.901
Mean of sample	0.353	0.692	0.765	0.007	0.278	0.827	0.472	0.745	0.816	0.025	0.433	0.851
Panel B: SP, case and market/district fixed effects												
Is a private provider	0.026	-0.001	0.104	-0.022	0.086	0.165**	0.148**	0.028	-0.031	-0.010	-0.121*	0.076
	(0.071)	(0.075)	(0.076)	(0.024)	(0.069)	(0.069)	(0.064)	(0.062)	(0.054)	(0.025)	(0.068)	(0.048)
R-squared	0.450	0.261	0.265	0.061	0.239	0.219	0.294	0.090	0.118	0.067	0.130	0.177
Number of observations	334	365	392	440	440	440	199	200	201	201	201	201
Panel C: SP, case and market/district fixed effects												
Is a private provider	0.067	0.023	0.099	-0.027	0.112	0.137*	0.170**	0.051	-0.014	-0.024	-0.100	0.093*
	(0.087)	(0.091)	(0.081)	(0.024)	(0.080)	(0.075)	(0.069)	(0.066)	(0.058)	(0.024)	(0.071)	(0.053)
Has MBBS	0.207**	0.062	-0.090	0.051	0.267***	-0.057						
	(0.096)	(0.118)	(0.096)	(0.033)	(0.086)	(0.090)						
Has some qualification	0.052	0.057	-0.045	0.028	0.099	-0.025						
	(0.073)	(0.076)	(0.069)	(0.022)	(0.063)	(0.054)						
Age of provider	-0.002	-0.004	-0.001	-0.001	-0.000	-0.003	-0.002	-0.003	-0.001	-0.002*	-0.001	-0.004
	(0.003)	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)	(0.004)	(0.004)	(0.003)	(0.001)	(0.004)	(0.003)
Gender of provider (1=Male)	0.148	-0.228**	-0.093	0.012	-0.029	-0.128*	0.037	0.201*	0.177*	-0.048	0.152	0.158*
	(0.128)	(0.110)	(0.107)	(0.010)	(0.132)	(0.075)	(0.100)	(0.106)	(0.091)	(0.046)	(0.100)	(0.082)
Patient load during visit	-0.015	-0.004	-0.003	-0.000	-0.008	-0.006	0.001	0.016	0.023*	-0.003	0.002	0.026***
	(0.018)	(0.018)	(0.012)	(0.001)	(0.008)	(0.011)	(0.015)	(0.011)	(0.013)	(0.004)	(0.016)	(0.009)
R-squared	0.460	0.279	0.284	0.092	0.272	0.240	0.273	0.133	0.162	0.122	0.159	0.219
Number of observations	321	350	378	423	423	423	181	182	183	183	183	183

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. Observations are standardized provider-patient interactions. Market fixed effects are used for the representative sample, and district fixed effects for dual practice sample.

Table 6: Provider 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	10.519***	11.745***	5.122***
	(3.281)	(2.875)	(2.143)	(2.404)	(2.360)	(1.746)
R-squared	0.144	0.175	0.238	0.279	0.236	0.318
Number of observations	662	662	662	330	330	330
Mean of public	21.770	13.975	10.197	28.254	14.592	10.020
Mean of private	32.966	21.322	12.235	41.104	28.800	15.234
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	10.594***	11.709***	5.077***
	(2.609)	(2.535)	(1.969)	(2.358)	(2.381)	(1.751)
R-squared	0.253	0.256	0.300	0.301	0.247	0.322
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	11.786***	12.518***	4.523**
	(3.030)	(3.092)	(2.381)	(2.480)	(2.466)	(1.794)
Has MBBS	2.548	5.175	2.307			
	(3.949)	(3.449)	(2.916)			
Has some qualification	2.300	4.764*	0.721			
	(2.563)	(2.447)	(1.831)			
Age of provider	-0.151	-0.009	0.044	-0.069	-0.135	-0.042
	(0.102)	(0.094)	(0.077)	(0.139)	(0.138)	(0.101)
Gender of provider (1=Male)	1.009	-1.353	-2.369	2.773	-2.792	-3.651
	(4.541)	(5.383)	(3.361)	(3.586)	(3.565)	(2.593)
Patient load during visit	-0.041	-0.396	0.050	-0.501	-0.203	-0.211
	(0.736)	(0.557)	(0.528)	(0.576)	(0.572)	(0.416)
R-squared	0.254	0.262	0.301	0.290	0.252	0.330
Number of observations	638	638	638	301	301	301

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. Robust standard errors clustered at the market level are in parenthesis. 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 dual practice sample.

Table 7: Real patients in the public and private sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Audit 1				Dual sample			
	Time spent (mins)	Total questions	Physical examination	Dispensed medicines	Time spent (mins)	Total questions	Physical examination	Dispensed medicines
Panel A: no patient controls or fixed effects								
Is a private provider	1.456***	0.799***	0.371***	-0.241***	1.894***	1.154***	0.143**	-0.560***
	(0.323)	(0.180)	(0.108)	(0.082)	(0.569)	(0.318)	(0.063)	(0.097)
R-squared	0.054	0.030	0.103	0.029	0.115	0.082	0.017	0.398
Number of observations	1,137	1,137	1,133	1,138	1,085	1,083	1,082	1,090
Mean of public	2.378	2.994	0.473	0.765	1.499	3.284	0.678	0.963
Mean of private	3.833	3.793	0.844	0.524	3.393	4.439	0.821	0.403
Mean of sample	3.621	3.676	0.790	0.559	1.899	3.527	0.708	0.844
Number of public providers	29	29	29	29	51	51	51	51
Number of private providers	169	169	169	169	40	40	41	41
Panel B: including patient controls and market/district fixed effects								
Is a private provider	1.192***	0.615**	0.510***	-0.380***	1.464***	0.494***	0.080**	-0.523***
	(0.325)	(0.248)	(0.086)	(0.114)	(0.300)	(0.128)	(0.039)	(0.035)
Has MBBS degree	-0.427	0.416*	0.155**	-0.407***				
	(0.516)	(0.227)	(0.079)	(0.132)				
Has some qualification	0.349	0.021	0.004	-0.057				
	(0.388)	(0.143)	(0.050)	(0.073)				
Age of Provider	-0.027**	0.006	0.001	0.000	-0.004	-0.011**	-0.001	0.005***
	(0.010)	(0.005)	(0.002)	(0.002)	(0.006)	(0.005)	(0.002)	(0.001)
Gender of Provider (1=Male)	-1.433**	-0.780	0.016	-0.134	-0.362*	0.056	-0.102*	0.046
	(0.688)	(0.756)	(0.087)	(0.097)	(0.202)	(0.203)	(0.054)	(0.030)
R-squared	0.295	0.331	0.368	0.339	0.167	0.354	0.195	0.498
Number of observations	835	835	833	835	812	811	810	813

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. Robust standard errors clustered at the village level in parenthesis. Observations are patient-provider interactions, and the sample has been limited to the SP sample. All regressions 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 panel B, mean of public is from markets that have both (public and private) types of providers. In the fixed effects regressions, market fixed effects are used in the representative sample and district fixed effects in the dual practice sample.

Table 8: Prices charged and correct treatment (pooled sample, private interactions, unstable angina and asthma only)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Binary regressions (Fees in Rs.)	**Multiple regressions** (Fees in Rs.)					
Time spent with SP (minutes)	1.088***	0.645	0.635	0.564	0.356	0.693*	0.704*
	(0.343)	(0.450)	(0.452)	(0.490)	(0.442)	(0.384)	(0.386)
Percentage of checklist items	0.367***	0.316***	0.323***	0.270**	0.294**	0.196*	0.218**
	(0.089)	(0.105)	(0.109)	(0.119)	(0.126)	(0.105)	(0.095)
Correct diagnosis (unconditional)	3.972		-1.832	-0.391	-0.522	1.220	0.727
	(4.959)		(4.484)	(4.776)	(5.039)	(4.858)	(4.796)
Correct treatment	10.266***			5.279*	4.157	0.312	-0.231
	(2.423)			(3.048)	(2.894)	(3.124)	(3.123)
Helpful treatment	7.605***			-0.133	-3.715	-5.108*	-5.570**
	(2.686)			(2.460)	(2.884)	(2.679)	(2.806)
Unnecessary treatment	16.647***			13.587***	8.651**	5.940**	5.579*
	(2.820)			(2.812)	(4.048)	(3.030)	(3.146)
Dispensed medicine	8.592***				7.000**	17.527***	18.923***
	(2.566)				(2.854)	(3.015)	(3.018)
Total medicines	4.776***				2.327	2.419*	2.755**
	(0.929)				(1.461)	(1.257)	(1.292)
Gave/prescribed antibiotic	3.900*				-1.637	-0.715	0.227
	(2.110)				(2.841)	(2.438)	(2.435)
Has MBBS	17.286***					32.863***	33.125***
	(3.717)					(5.810)	(5.292)
Has some qualification	0.210					9.396***	8.766**
	(3.311)					(3.415)	(3.461)
Patient load during visit	0.618						0.692
	(0.608)						(0.490)
Age of provider	0.022						0.090
	(0.171)						(0.114)
Gender of provider (1=Male)	-7.027						-7.684
	(4.600)						(5.210)
Referred in Unstable Angina	-18.377***						
	(4.836)						
Constant		20.757***	20.753***	8.684***	7.400**	-5.144	-3.651
		(1.995)	(1.995)	(3.335)	(3.345)	(3.146)	(9.213)
R-squared		0.096	0.096	0.147	0.169	0.346	0.378
Number of observations		490	490	384	384	384	373
Mean price charged		32.214	32.214	31.177	31.177	31.177	31.367
SD		27.637	27.637	27.736	27.736	27.736	27.959

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. Robust standard errors clustered at the market level are in parenthesis. Observations are standardized provider-patient interactions. Interpretation of coefficents in "Binary regressions" needs caution. Each coefficient represents a separate regression of prices on the row variable. Multiple regressions include district fixed effects. For the dependent variable "Referred in Unstable Angina", results from a bivariate regression is shown. This variable is only applicable to unstable angina cases and hence does not enter the multivariate regressions.

Table 9. Wages in the public sector (pooled sample, public observations only)

	(1)	(2)	(3)
	Binary regressions (Log of monthly salary)	**Multiple regressions** (Log of monthly salary)	
Percentage of checklist items	0.002	0.004	-0.000
	(0.003)	(0.003)	(0.002)
Time spent with SP (minutes)	-0.049*	-0.083***	-0.022
	(0.025)	(0.029)	(0.016)
Correct Treatment	0.011	-0.103	-0.123**
	(0.059)	(0.063)	(0.058)
Helpful Treatment	0.116	0.030	0.036
	(0.099)	(0.111)	(0.066)
Wrong Treatment	0.176**	0.221***	0.083
	(0.072)	(0.080)	(0.068)
Provider gave a diagnosis	0.102	0.019	0.025
	(0.084)	(0.113)	(0.090)
Gave correct diagnosis	0.060	0.030	-0.019
	(0.119)	(0.129)	(0.106)
Has MBBS	1.056***		1.334***
	(0.168)		(0.212)
Has some qualification	-0.094		0.875***
	(0.367)		(0.331)
Age of provider	0.011**		0.018***
	(0.006)		(0.006)
Gender of provider (1=Male)	0.114		0.099
	(0.188)		(0.103)
Born in same district	-0.384***		0.037
	(0.146)		(0.081)
Is a dual provider	0.578***		0.155*
	(0.135)		(0.085)
Constant		10.027***	7.961***
		(0.104)	(0.357)
R2		0.097	0.611
Number of observations		318	288

Notes: *** Significant at 1%, ** Significant at 5%, * Significant at 10%. Robust standard errors clustered at the market level are in parenthesis. Observations are standardized provider-patient interactions. All multiple regressions include district fixed effects.

Table 10: Cost in the public sector

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 cost 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. Wages data were collected in the year 2010, note that we use 2010 wage data to compute cost per patient in 2008 and 2009. Wages in 2008 and 2009 could have been lower.

Appendix Tables

Table A1: 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 the 5 markets where SP work was over completed, the SP saw a provider other than a sampled provider

Table A2: Mapping, Sampling and Completion in the Dual Practice Sample

Facilities	(1) Number of facilities	(2) Percentage	Providers	(3) Number of providers	(4) Percentage
Panel A: Mapping					
Total	200		Total	216	
without doctors	40	20.0%	without private facilities	84	38.9%
with doctors	160	80.0%	with private facilities	132	61.1%
Panel B: Sampling					
Total	143		Total	143	
			without private facilities	49	34.3%
			with private facilties	94	65.7%
Panel C: Completed (at least one case)					
Total	102		Total	118	
			without private facilities	30	25.4%
			with private facilities	88	74.6%

Notes: Reasons for not completing SP surveys include transferred and provider not found. In almost all cases our field staff made at least three 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 A3: 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
Helpful	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:

Table A4: Randomization balance for dual sample providers' assignment of Unstable Angina cases

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Asthma outcomes						Dysentery outcomes	
	Time spent (mins)	Percent checklist completed	Correct treatment	Helpful treatment	Gave diagnosis	Correct diagnosis	Time spent (mins)	Percent checklist completed
Is private	1.497***	13.190***	0.131	-0.045	0.181	0.077	0.302	9.109**
	(0.483)	(3.292)	(0.113)	(0.084)	(0.118)	(0.099)	(0.241)	(4.119)
Received Unstable Angina in private	0.433	5.441	-0.194	0.125	0.100	0.075	0.205	-0.862
	(0.518)	(3.534)	(0.121)	(0.091)	(0.127)	(0.106)	(0.255)	(4.356)
(Is private) x (Received Unstable Angina in private)	0.143	-2.996	0.061	-0.056	-0.214	-0.094	0.268	-0.604
	(0.719)	(4.898)	(0.169)	(0.125)	(0.176)	(0.147)	(0.354)	(6.053)
Constant	1.644***	13.687***	0.640***	0.843***	0.307***	0.150**	0.783***	17.088***
	(0.347)	(2.367)	(0.081)	(0.061)	(0.085)	(0.071)	(0.172)	(2.941)

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

Table A5: List of checklist items used in the treatment of SPs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Item discrimination tercile	Representative sample			Dual practice sample		
		Public	Private	Difference (3)-(2)	Public	Private	Difference (6)-(5)
Panel A: Unstable Angina							
History questions							
where is the pain	high	0.486	0.694	0.208***	0.528	0.645	0.117
when started	low	0.270	0.389	0.119*	0.167	0.129	-0.038
doing when began	high	0.054	0.078	0.024	0.083	0.161	0.078
severity of pain	low	0.162	0.278	0.116*	0.167	0.419	0.253**
radiation	high	0.108	0.150	0.042	0.222	0.387	0.165*
previous similar	medium	0.270	0.417	0.146**	0.278	0.387	0.109
since when	low	0.216	0.272	0.056	0.111	0.323	0.211**
quality of pain	high	0.108	0.117	0.009	0.111	0.258	0.147*
pain changes	low	0.054	0.061	0.007	0.056	0.161	0.106*
shortness of breath	medium	0.081	0.150	0.069	0.056	0.032	-0.023
nausea	medium	0.297	0.294	-0.003	0.056	0.387	0.332***
sweating	high	0.270	0.294	0.024	0.194	0.452	0.257**
beedi-cigarette	low	0.054	0.072	0.018	0.083	0.194	0.110*
family history	high	0.000	0.017	0.017	0.000	0.097	0.097**
Examination questions							
pulse	low	0.243	0.422	0.179**	0.417	0.677	0.261**
bp	medium	0.135	0.350	0.215***	0.222	0.548	0.326***
auscultation (either front or back)	low	0.189	0.500	0.311***	0.444	0.613	0.168*
temperature attempt	medium	0.108	0.139	0.031	0.028	0.258	0.230***
ecg in/outside clinic	medium	0.243	0.228	-0.015	0.278	0.355	0.077
Number of observations		37	180		36	31	
Panel B: Asthma							
History questions							
current breathing probes	medium	0.385	0.647	0.262***	0.422	0.671	0.250***
cough	low	0.590	0.696	0.106	0.453	0.686	0.233***
expectoration probes	low	0.077	0.163	0.086*	0.016	0.071	0.056*
previous breathing problems	high	0.333	0.462	0.129*	0.266	0.543	0.277***
previous episode probes	medium	0.128	0.196	0.067	0.109	0.286	0.176***
since when problems	medium	0.385	0.495	0.110	0.234	0.414	0.180**
how often happens	high	0.128	0.103	-0.025	0.047	0.086	0.039
shortness constant or episodic	low	0.051	0.114	0.063	0.047	0.129	0.082**
what triggers	medium	0.077	0.125	0.048	0.094	0.229	0.135**
how long lasts	high	0.077	0.065	-0.012	0.016	0.086	0.070**
childhood illness	medium	0.000	0.033	0.033	0.016	0.043	0.027
age	high	0.308	0.141	-0.166***	0.578	0.500	-0.078
fever	low	0.231	0.326	0.095	0.219	0.386	0.167**
chest pain	low	0.154	0.375	0.221***	0.172	0.286	0.114*
weight loss	high	0.000	0.000	0.000	0.016	0.014	-0.001
night sweats	high	0.051	0.054	0.003	0.047	0.086	0.039
beedi-cigarette	high	0.026	0.016	-0.009	0.016	0.071	0.056*
family history	medium	0.000	0.027	0.027	0.031	0.043	0.012
Examination questions							
pulse	low	0.256	0.554	0.298***	0.313	0.457	0.145**
bp	medium	0.205	0.293	0.088	0.109	0.357	0.248***
auscultation (either front or back)	low	0.333	0.554	0.221***	0.484	0.800	0.316***
temp attempt	low	0.103	0.179	0.077	0.063	0.100	0.038
Number of observations		39	184		64	70	

(continued on next page)

Table A5 continued

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

Notes:

Table A6: 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	Helpful treatment	Unnecessary treatment	Correct treatment Only	Antibiotic	Poly-pharmacy
Panel A1: Unstable angina, representative sample, with SP fixed effects											
Is a private provider	1.101***	7.890**	0.112	0.033	0.011	0.036	-0.008	0.079	-0.016	0.024	0.139
	(0.303)	(3.418)	(0.093)	(0.074)	(0.030)	(0.047)	(0.091)	(0.100)	(0.030)	(0.062)	(0.089)
R-squared	0.083	0.138	0.016	0.155	0.082	0.038	0.032	0.063	0.029	0.030	0.046
Number of observations	217	217	217	102	217	150	181	187	217	217	217
Panel A2: Unstable angina, dual practice sample, with SP fixed effects											
Is a private provider	3.232***	12.856**	0.214**	0.164	0.142*	0.292***	-0.055	0.099	-0.027	-0.013	0.167
	(1.012)	(5.367)	(0.109)	(0.153)	(0.076)	(0.095)	(0.105)	(0.115)	(0.027)	(0.112)	(0.106)
R-squared	0.214	0.105	0.349	0.141	0.153	0.189	0.123	0.139	0.140	0.070	0.212
Number of observations	61	61	61	29	61	60	61	61	61	61	61
Panel B1: Asthma, representative sample, with SP fixed effects											
Is a private provider	1.952***	6.015**	0.224***	-0.123	0.021	0.096	0.037	0.036	-0.025	0.009	0.121
	(0.475)	(2.940)	(0.071)	(0.201)	(0.043)	(0.108)	(0.101)	(0.091)	(0.028)	(0.094)	(0.087)
R-squared	0.200	0.172	0.209	0.065	0.067	0.042	0.029	0.071	0.028	0.019	0.058
Number of observations	223	223	223	76	223	184	184	205	223	223	223
Panel B2: Asthma, dual practice sample, with SP fixed effects											
Is a private provider	1.418***	12.167***	0.033	-0.078	-0.015	0.113	0.109	-0.083	-0.022	-0.151*	0.046
	(0.374)	(2.545)	(0.092)	(0.158)	(0.080)	(0.090)	(0.084)	(0.065)	(0.033)	(0.092)	(0.058)
R-squared	0.201	0.235	0.090	0.102	0.060	0.120	0.112	0.096	0.069	0.097	0.108
Number of observations	122	122	122	51	122	121	121	122	122	122	122
Panel C1: Dysentery, representative sample, with SP fixed effects											
Is a private provider	0.846***	7.088***									
	(0.231)	(2.052)									
R-squared	0.091	0.108									
Number of observations	222	222									
Panel B2: Dysentery, dual practice sample, with SP fixed effects											
Is a private provider	0.395**	5.279**									
	(0.181)	(2.569)									
R-squared	0.095	0.340									
Number of observations	119	119									

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 controls for provider qualifications, age, gender, and patient load. Observations are standardized provider-patient interactions.

Table A7: 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	0.04	0.08	0.05	0.03	0.30	0.27***
Correct treatment (alternate)	0.55	0.48	-0.07	0.42	0.61	0.20*
Aspirin	0.03	0.04	0.02	0.03	0.23	0.20***
Anti-platelet agents	0.03	0.01	-0.02	0.00	0.03	0.03
Referred	0.30	0.24	-0.05	0.22	0.32	0.10
ECG	0.24	0.23	-0.02	0.28	0.35	0.08
ECG & Referred	0.11	0.12	0.01	0.08	0.16	0.08
Antibiotic	0.14	0.17	0.03	0.28	0.23	-0.05
Unnecessary treatment	0.66	0.74	0.09	0.67	0.77	0.11
Number of observations	37	180		36	31	
Panel B: Asthma						
Correct treatment	0.47	0.61	0.14*	0.58	0.68	0.10
Bronchodilators	0.33	0.36	0.03	0.52	0.59	0.07
Theophylline	0.13	0.22	0.09*	0.31	0.31	0.00
Oral Corticosteroids	0.15	0.31	0.16**	0.16	0.24	0.09
Antibiotic	0.38	0.40	0.02	0.59	0.46	-0.14*
Unnecessary treatment	0.73	0.82	0.09	0.91	0.83	-0.08*
Number of observations	39	184		64	70	
Panel C: Dysentery						
Correct treatment	0.08	0.14	0.06	0.33	0.22	-0.11*
ORS	0.08	0.13	0.05	0.33	0.22	-0.11*
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*
Unnecessary treatment	0.11	0.41	0.30***	0.43	0.33	-0.10
Number of observations	39	183		63	67	

Notes: In Unstable Angina, alternate definition for correct treatment codes referrals and referrals for ECG as correct

Table A8: Robustness of treatment results with alternative definition for unstable angina

	(1)	(2)	(3)	(4)
	All (compare with table 5)		Unstable angina only (compare with table 6)	
	Representative sample	Dual practice sample	Representative sample	Dual sample
	Correct treatment	Correct treatment	Correct treatment	Correct treatment
Panel A: SP fixed effects				
Is a private provider	0.009	0.125*	-0.079	0.206*
	(0.062)	(0.070)	(0.085)	(0.121)
R-squared	0.085	0.085	0.120	0.070
Number of observations	350	200	166	67
Mean of public	0.508	0.520	0.548	0.417
Mean of private	0.536	0.660	0.479	0.613
Mean of sample	0.533	0.590	0.485	0.507
Panel B: SP and market/district fixed effects				
Is a private provider	-0.032	0.127*		
	(0.066)	(0.070)		
R-squared	0.234	0.094		
Number of observations	350	200		
Panel C: SP and market/district fixed effects				
Is a private provider	-0.065	0.132*		
	(0.079)	(0.075)		
Has MBBS	0.122			
	(0.095)			
Has some qualification	0.087			
	(0.073)			
Age of provider	0.001	-0.005		
	(0.004)	(0.004)		
Gender of provider (1=Male)	0.253	-0.008		
	(0.198)	(0.112)		
Patient load during visit	-0.042***	-0.006		
	(0.011)	(0.018)		
R-squared	0.263	0.105		
Number of observations	336	182		

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 controls for provider qualifications, age, gender, and patient load. Observations are standardized provider-patient interactions. Dual sample refers to providers who operate both public and private clinics. Columns (1) and (2) also include case fixed effects. Market fixed effects are used for the representative sample, and district fixed effects for dual practice sample. Alternative definition for Unstable Angina adds "referral" and "referral for ECG" as correct treatment.

Table A9: 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**	0.471**	2.200***	12.267***	0.696***
	(0.306)	(3.307)	(0.209)	(0.419)	(2.376)	(0.187)
R-squared	0.225	0.152		0.170	0.154	
Number of observations	440	440	233	201	201	138
Mean of public	2.956	17.540		1.970	17.481	
Mean of private	4.548	24.335		4.043	30.196	
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***	0.491*	2.201***	12.271***	0.696***
	(0.367)	(2.727)	(0.292)	(0.406)	(2.392)	(0.189)
R-squared	0.341	0.278		0.194	0.164	
Number of observations	440	440	233	201	201	
Panel C: SP, case and market/district fixed effects						
Is a private provider	1.654***	6.087*	0.405	2.061***	12.243***	0.711***
	(0.481)	(3.354)	(0.347)	(0.421)	(2.514)	(0.199)
Has MBBS	-0.062	6.415*	0.206			
	(0.847)	(3.805)	(0.254)			
Has some qualification	-0.159	2.737	0.119			
	(0.435)	(2.159)	(0.156)			
Age of provider	-0.002	0.027	0.003	0.018	-0.004	0.003
	(0.017)	(0.088)	(0.008)	(0.024)	(0.141)	(0.010)
Gender of provider (1=Male)	1.460*	2.136	0.060	-0.343	-3.130	-0.186
	(0.789)	(4.284)	(0.325)	(0.605)	(3.609)	(0.315)
Patient load during visit	-0.188***	-0.333	0.034	-0.132	-0.054	-0.033
	(0.059)	(0.609)	(0.044)	(0.103)	(0.614)	(0.038)
R-squared	0.357	0.283		0.218	0.170	
Number of observations	423	423	221	183	183	126

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. Observations are standardized provider-patient interactions, except in IRT score column where each observation is a provider. The score is computed using all cases, plausible values scores are used. Market fixed effects are used for the representative sample, and district fixed effects for dual practice sample.

Table A10: 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 (uncondition	Correct treatment	Helpful treatment	Unnecessary treatment	Correct treatment Only	Antibiotic	Poly-pharmacy
Panel A: Representative sample, with SP, case and market fixed effects												
Is a private provider	1.207***	7.826***	0.635**	0.197**	-0.023	0.039	0.123	0.103	0.150*	-0.031	0.153*	0.185**
	(0.363)	(2.494)	(0.305)	(0.085)	(0.126)	(0.038)	(0.090)	(0.095)	(0.084)	(0.025)	(0.081)	(0.077)
Facilities index	0.012	1.679***	0.129*	0.051**	0.014	0.010	0.047*	0.063**	0.036*	-0.003	0.029	0.038**
	(0.112)	(0.600)	(0.072)	(0.023)	(0.033)	(0.011)	(0.024)	(0.026)	(0.021)	(0.002)	(0.021)	(0.018)
R-squared	0.356	0.265		0.233	0.362	0.161	0.469	0.303	0.295	0.094	0.275	0.256
Number of observations	634	634	220	420	171	420	318	347	375	420	420	420
Panel B: Dual practice sample, with SP, case and market fixed effects												
Is a private provider	1.218***	9.185***	0.751***	0.030	-0.035	-0.004	0.172**	0.006	-0.027	0.000	-0.146*	0.059
	(0.259)	(1.927)	(0.213)	(0.081)	(0.135)	(0.069)	(0.072)	(0.073)	(0.061)	(0.021)	(0.080)	(0.059)
Facilities index	-0.206	-0.928	-0.033	-0.040	-0.029	-0.029	-0.074*	-0.072*	-0.007	0.005	-0.038	-0.070**
	(0.157)	(1.148)	(0.111)	(0.041)	(0.079)	(0.037)	(0.039)	(0.040)	(0.036)	(0.005)	(0.045)	(0.031)
R-squared	0.321	0.244		0.219	0.199	0.092	0.318	0.133	0.128	0.065	0.144	0.213
Number of observations	272	272	114	164	73	164	162	163	164	164	164	164

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 controls for provider qualifications, age, gender, and patient load. Observations are standardized provider-patient interactions. Dual sample refers to providers who operate both public and private clinics. In panel A, mean of public is from markets that have both (public and private) types of providers. Market fixed effects are used for the representative sample, and district fixed effects for 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. Columns (4)-(6) with Table 4; and Columns (7)-(12) with Table 5