Quality and Accountability in Health: Audit Evidence from Primary Care Clinics in India

Jishnu Das

Alaka Holla

Michael Kremer

Karthik Muralidharan[1]

Abstract

Private primary care providers routinely account for more than 50 percent of first-contacts in low-income settings, rising to 80 percent in countries like India. The majority of these providers operate in single provider clinics with little regulatory oversight or government subsidies. No patients have health insurance beyond the free care that they can access in the public sector. Doctors in public clinics are paid a fixed salary and are typically more qualified than private providers who often have no formal medical qualifications. Nevertheless, private fee-charging providers account for over 70% of the market share for primary care. In this paper, we provide evidence on the quality and accountability of healthcare provision in rural India using a unique and original set of audit studies, where unannounced standardized patients were presented to a representative sample of rural public and private primary care providers in the Indian state of Madhya Pradesh. The three main findings from our audit studies suggest that that customer accountability in an unregulated, unsubsidized and uninsured private market elicits greater provider effort relative to the administrative accountability in the public sector. First, we find across all audit studies that public providers spent less time with patients, completed fewer items on a checklist of essential history and examination items, and were either no different or worse in their treatment and diagnostic accuracy. Second, we identify a sample of public sectors doctors who also have a private practice, and find that the *same doctor* exerts significantly higher effort on the *same medical case* when seeing a patient in private practice relative to in their public practice. Finally, we find a strong positive correlation between the prices charged to the standardized patients and the quality of care received, suggesting that the market rewards quality, which in turn provide incentives for better service delivery among private providers of healthcare. However, hedonic pricing in the private market also suggests that financial constraints may prevent the poor from accessing quality care.

[1] Das: World Bank and Center for Policy Research (CPR); jdas1@worldbank.org
Holla: World Bank; aholla@worldbank.org
Kremer: Harvard University, NBER, and BREAD; mkremer@fas.harvard.edu
Muralidharan: UC San Diego, NBER, and BREAD; kamurali@ucsd.edu

I. Introduction

Along the continuum of services that should be provided privately versus through the public sector, it is widely argued that the provision of health care falls in the latter category. Two compelling theoretical arguments bolster this assertion. First, the welfare function of governments is likely to differ from that of the consumers: As an emergency physician may point out, using customer satisfaction among narcotic addicted patients who are denied medication is a surefire way to wrongly assess the capabilities of a doctor. In such situations, private provision of health care—which depends on customer feedback through demand—will likely lead to inefficient provision (Prendergast, 2003). Further, the market for health care is, in many ways, similar to that for credence goods (see for instance, Dulleck and Kerschbamer 2006). That is, even as consumers may know *what* they purchased, they may not know whether it was really required. A new mother may know that she was operated upon, but in most cases will not have the expertise to judge whether a Caesarean section was really required. There are strong reasons to believe that such markets will be subject both to under and over-provision of care (Gruber and Owens 1996, Dulleck, Kershcbamer and Sutter 2011, Schneider 2012, Wolinsky 1993).

Yet, most people in low-income countries continue using the private sector for health care, even when public sector options are widely available. In India, the focus of this paper, the private sector accounts for 80 percent of first contacts for primary care and our nationwide data show that there is only a small difference in the share of the private sector in villages with and without public sector clinics. These, and similar other data, raise the possibility that consumers value attributes of health services that are uncorrelated to quality; one such example is a perceived "demand for medication and injections" in low-income countries. If so, the presence of the private sector can lower quality in the market and increase costs for consumers.

But an alternate view is also possible. With little effective regulation and poor administrative accountability, public sector care may be no better, or even worse, than existing private sector alternatives. For instance, research over the last decade has demonstrated widespread absenteeism (Chaudhury and Hammer 2005, Chadhury and others 2006) and very low effort among public sector providers (Das and Hammer, 2007 and Das, Hammer and Leonard 2008). Further, it is not necessarily the case that markets for credence goods will *always* produce inefficient outcomes—whether or not it does so depends critically on the structure of price markups for different services (Dulleck and Kerschbamer 2006). This very different view offers an alternate guide to consumer behavior: The choice of the private sector reflects accurate valuations of quality among consumers.

Given the implications for health policy deriving from these two very different pictures of the world, it is surprising that there is currently little evidence comparing quality in the public and private sector in a manner that appropriately adjusts for patient selection and allows for market (rather than administratively determined) pricing of services. This paper provides such evidence.

We examine quality in what is essentially an unregulated and unsubsidized market for private primary care with fully market determined prices. In our study population, there is no medical insurance other than the implicit insurance of a tax-funded public health delivery system. The use of the private health care is high and accounts for above 80 percent of first contacts and the majority of providers in the private sector has no medical training. Entry into primary care is free and basic arbitrage conditions are satisfied: Private sector providers earn roughly the wage of a secondary-school graduate in the labor market. In contrast, the public sector is (theoretically) staffed by fully trained doctors who provide free care at the point of service. The public sector also provides free medicines to patients.

We contrast quality among these private sector providers with those in the public system, which relies entirely on administrative and peer mechanisms, insofar as they exist and are enforced in our context. By systematically measuring quality and prices in the market for primary health care, this paper provides evidence on the promises and pitfalls of customer accountability, and contrasts it with the accountability (or the lack thereof) in the public provision of health care.

The specifics are as follows.

We conducted two rounds of audit studies using "standardized patients" among providers practicing in the primary public and private sectors in five districts of rural Madhya Pradesh—one of the poorest states in India with low rates of literacy/numeracy and poor health outcomes. In order to verify that our audit results were similar to those among actual consumer populations, the audit measures were then complemented with direct observations of provider-patient interactions, where a surveyor sat for a full day with the provider, recording various aspects of every interaction.

For the audit studies we trained 22 Standardized Patients (SPs) to depict uncomplicated presentations of unstable angina in a middle-aged adult, asthma in a young adult and dysentery in a child, who was sleeping at home. The training was extensive and exceeded 150 hours for each of the SPs. We then sent the standardized patients to a sample of providers practicing in the public and the private sector in the health markets of 60 villages; we had earlier enumerated households and providers in each of these markets and the villages were chosen by random sampling with some restrictions (more on this below). We highlight that there were no a priori restrictions on the qualification of providers in the private sector; consequently, close to 50 percent of the sample of private sector providers does not have a high-school degree.

In the first round, we asked the SPs to seek care from the sampled clinics regardless of the identity of the person providing care. Given frequent doctor absences from public sector clinics, this round

was designed to assess the care that a patient would receive from the public clinic on an average visit. As we will show below, in 64 percent of cases, the SP was seen by a person who was not the doctor, nor a staff member with formal medical training.

In the second round, we asked the SP to wait till the doctor came to the clinic. For operational purposes, this was implemented by having one surveyor sit in a tea-shop close to the clinic who would then call the SP when the doctor arrived; in one case, this was 6 days later. The results from this round allow us to compare and contrast the quality of care provided by "real" doctors relative to the care that a patient would receive on an "average visit". Previous studies (see Banerjee, Deaton and Duflo, 2004) document that there is no pattern of doctor absences, so whether patients in real life see the doctor or someone else can be thought of as luck of the draw.

In this round, the SPs also visited the *private sector* clinic of the public providers who had been sampled. For 80 percent of the public sample in this round, we could identify such private sector clinics, pointing to the widespread prevalence of dual-practice in the region. The second round thus allows us to look at the relative performance of the same provider in the public and private sector; the former subject to administrative accountability and peer monitoring and the latter only to customer accountability in a reasonably competitive market (see below).

Across the audits we collected three measures of quality and the price charged. These three measures of quality are: (A) completion of a medically necessary checklist of questions and examinations that a provider is expected to perform for a patient arriving with the given set of symptoms; (B) the diagnosis given by the care provider and; (C) the treatment given by the provider. Thus, for instance, in the case of unstable angina, a "good" doctor may ask about (for instance) the nature of the chest-pain (whether radiating or not, when it was first experienced, the severity) and attempt to complete the vital signs (blood-pressure, pulse, temperature and respiratory rate). She may then utter a

provisional diagnosis, and would certainly (A) dispense or prescribe an Aspirin, (B) refer the patient to a hospital and/or (C) refer the patient for an Electro-cardiogram. For much of our analysis, for reasons discussed below, we will focus on checklist completion as the appropriate measure of quality in our context.

In both rounds, providers in the private sector spent more time with patients and completed more checklist items. In our preferred specifications, the coefficient on checklist completion ranged from 7.9 to 9.06 percentage points in favor of the private sector. Given a mean checklist completion rate of 22.4 percent, this reflects a 36 to 40 percent or 0.5 standard-deviation difference. Particularly noteworthy is that the public-private difference in the second round is very similar to that in the first, where 64 percent of the care providers in the public clinics were not medically trained staff. This is because we are unable to detect any difference in checklist completion depending on whether the SP saw a "real" doctor (Round 2) or somebody else in the public clinic (Round 1). In fact, with district and SP fixed-effects, the difference between the real doctor and an untrained staff in the public clinic is zero and statistically insignificant at all conventional confidence levels. Of course, as the results from Round 2 with provider fixed-effects demonstrate, this is not because the capabilities of the "real" doctors are like that of the untrained staff, but because the "real" doctors behave completely differently in their public and private practices.

We find greater evidence of *under provision* in the public sector, with lower rates of correct treatment. The basic comparison of treatment in public versus private sector, which compares all public versus all private providers, shows a 7.5 percentage point premium for the private sector. When we include provider fixed-effects the premium increases further to 8.7 percentage points. This premium, on the very low base of correct treatment rates of 39 percent reflect close to a 25 percent increase in the chance of correct treatment. Particularly startling are the results for unstable angina: Doctors with

dual practices audited in the private sector were 32 percentage points less likely to give the correct treatment relative to those in the public sector—a difference of almost 100 percent.

Finally, we are unable to detect any evidence of *over-provision* in the private sector. The rates of unnecessary or harmful treatment are just as high in both the public and the private sector, with a mean of 75 percent (!). There are no statistically significant differences between the two sectors for all cases combined, and for individual cases there is no clear pattern.

That care is higher quality in the private sector suggests customer accountability in the market for health care rewards quality. Direct evidence on prices in the private sector confirm that higher quality is rewarded in the market (prices in the public sector are nominal and fixed—and we don't observe any requests for under-the-table compensation). Moving from the 5th to the 95th percentile of the checklist distribution increases prices by more than 100 percent. These results reflect more than observed characteristics of the provider that are correlated with quality, as there is no change in the coefficient relating prices to checklist completion when we include a host of controls for the qualification of the doctor, the doctor's patient load at the time of the SSP visit and an index of equipment in the facility. There is some evidence that more time with patients is part of this story as every additional minute result in higher prices, but checklist completion continues to be reflected in overall prices with time as an additional control.

Several ancillary results help make progress towards ruling out alternate explanations. First, on process variables that we could track in direct observations with patients (such as time spent, total questions asked, total examinations performed), the audit studies closely resemble the observational study. This helps rule out the kinds of differences noted previously between audit and "real" consumer results (see for instance, Ayres and Siegelman 1995and Golderg 1996).

Second, differences in case-loads and physical equipment do not explain the difference in quality between the two sectors. Public sector clinics have more of each of the different types of physical equipment necessary for these cases and controls for case-load do not alter the coefficients (in fact, most of the clinics have tremendous excess capacity).

Third, the public-private sector differences are not the direct impact of dual practice per se (for instance, the provider in the public clinic refers the patient to his private clinic). We find similar patterns among providers with and without a dual practice, and similar behavior among providers with geographically proximate versus geographically distant dual practices, where the distance to the dual practice decreases the possibility of market-overlap.

Fourth, the results are not specific to the low literacy rural region of Madhya Pradesh. We repeat Audit 2 in the urban setting of Delhi, where incomes and education-levels are the highest in the country and broadly replicate our findings.

The welfare consequences of these differences, in part, depend on the ability of patients to triage themselves into different kinds of care. For instance, if people who experience chest-pain are fully able to triage into "muscle strain" and "something serious", they may well visit the public sector for the former and the private sector for the latter. If all patients with chest pain in the public sector come with a muscle strain, it may be optimal for public sector providers to dispense symptomatic relief and move to the next patient. Theoretically, the question is what we should expect the provider to do for an "out-of-equilibrium" patient. To assess whether such patients would truly be "out-of-equilibrium" (that is, they never visit the public sector) we also provide evidence that at best 50 percent of the households in our sample would be able to triage themselves into public or private care with the symptoms related to these cases. We provide further robustness for our results by including a rich set of controls for the types of patients that the provider usually sees; these were

collected through an exit-survey on the same day that we completed the observation of clinical interactions. Again, we find that these controls have no discernable impact on our basic audit results.

These results provide some grounds for optimism regarding the potential of customer accountability where government accountability is poor. On the face of it, a lightly regulated private sector with many different types of health care providers provides better care than that provided in public clinics with (supposedly) fully-trained providers. However the findings do not espouse an unregulated private sector approach to health care. We discuss three reasons in the conclusion: (A) that the public sector provides a location subsidy for highly qualified doctors to move to rural areas; (B) that the private market works the way it does only because there are no financial subsidies for care—one such financial subsidies are brought in to enhance equity, administrative pricing is required and; (C) that much could be done to better design health care markets that use the strengths and weaknesses of either sector once we recognize that that market for health care even for illiterate farmers in Madhya Pradesh provides a semblance of allocative efficiency.

The remainder of the paper is as follows. Section 2 provides the background, methods and data collection. Section 3 provides four figures highlighting the main results of the study. Section 4 presents the empirical specifications and results and Section 5 concludes.

II. Context and Data

Health markets in India differ fundamentally from heavily regulated markets in most middle and high-income economies. Here, we describe the market for primary care services and present empirical patterns that help characterize supply and demand and the basic organization of healthcare markets in rural India.

The vast majority of households in rural India do not have formal medical insurance. Although a federally funded insurance product for hospital care was recently introduced, it was not available in the state that we worked in, and it did not extend to primary care. Consequently, patients in our population can choose among a variety of public (government funded and managed) and private options. In the public sector, patients can obtain primary care in hospitals, in Community Health Centers (CHCs), in Public Health Centers (PHCs) and in sub-centers. These were organized around a triage model, where simple cases would be seen in sub-centers and primary health care centers and more complicated cases would be treated in community health centers with inpatient care typically available at the district hospital[2]. Given the triage system, there are many more sub-centers and primary health care centers than community health centers and district hospitals. According to official guidelines, there is supposed to be 1 primary health care center for every 10,000 population, 1 community health center for every 30,000 and one district hospital for every 100,000 population.

These institutions should be staffed by qualified health personnel and health assistants. Most doctors hold a Bachelor of Medicine and Bachelor of Surgery (MBBS) degree, the equivalent of an MD in the United States, and they all receive a fixed salary from the government. All consultations are either free or nominally priced (ranging from Rs. 2 to 5 or $0.03 to $0.09), and care is available on a walk-in basis. Nevertheless, a number of studies have documented severe problems with accountability in public clinics, usually measured through doctor absence. Nationwide, doctor absences averaged 40 percent on any given day in 2003 and XX percent in 2011 (Chaudhury and others, 2006 and 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. One study that tried

[2]Some kinds of inpatient care (particularly labor and delivery) is also provided at Community health centers.

to put in time-stamp machines and tie wages of nurses to showing up to work failed because the nurses broke the machines (Banerjee, Dufo and Glennerster, 2008).

Perhaps as a consequence, in India, like in many resource-poor settings, the private sector accounts for most primary care visits (Montagu and other citations here). In India, despite a host of *de jure* regulations, the *de facto* reality is that anyone can offer care in this sector, where qualifications range from MBBS degrees to no medical training whatsoever and where clinics can be independent structures that are almost identical to PHCs, small one-room shops, the provider's residence, or even the back of a motorcycle for providers that make home visits. Average fees for such visits in our sample are Rs.28 ($0.50) and usually include the cost of medicines. While providers operating without a medical license are not legal and face the threat of an occasional raid, as the data will show, they have come to be the dominant source of care in these markets.

Study Setting

Our data come from five districts in Madhya Pradesh, one of the poorest states in India with low overall levels of literacy and high rates of childhood mortality. The five study districts were chosen randomly following an accepted stratification of the state into five socio-cultural zones. Within every district, we sampled 20 villages randomly, subject only to the constraint that they lay outside a 5 kilometer buffer around major towns and cities. The restriction was put in place because we wished to sample health markets for all sampled villages and survey every provider in the relevant health markets; including towns and cities would have blown up the sample beyond the scope of our envisaged study.

To quantify the use of different kinds of providers, we conducted provider and household censuses in each of these 100 representative villages. In each of these villages, we asked people to name the locations where they sought care, which often included nearby villages or small markets on nearby

national and state highways. We then surveyed *all* providers in these locations, regardless of whether or not the providers themselves had been mentioned in the sample villages. We used this process to define the geographic boundaries of the effective healthcare market for households in our sample villages. During our household census, we also asked for the identity of any providers they had visited in the previous 30 days, allowing us to match households to specific providers in our sample and ensure that we did not miss providers in the provider census.

The data from the provider and household censuses reveal three main trends. First, the supply of providers is high once the entire market is considered (Table 1, Panel A). The average household can access 11 providers offering primary care services; excluding nurses and midwives in the public sector reduces this total to 10 per village. Half of them (5.4) operate privately and report no medical training at all. There is on average less than one public MBBS doctor available in these markets, and even less of their private sector counterparts. There are actually more unqualified providers dispensing care as doctors in public clinics (1.7) than qualified doctors, consistent with the high rates of absence among official health workers documented in India (Chaudhury and others, 2006).[3] In these markets, we also find providers with alternative qualifications—mostly in indigenous systems of medicine– who are allowed to offer primary services but are expected to provide treatments consistent with their training, although in practice, they are just as likely to dispense or prescribe allopathic medicine as MBBS doctors (see Das and others 2012 and Das and Hammer 2007).[4]

A couple of additional observations help contextualize these large numbers of providers in the health markets of our sampled villages. First, the averages could reflect some villages that are close to larger towns and whose providers would thus be counted in the health market for the village. For

[3] It is important to note that these providers freely told us about their lack of qualifications. To the extent that providers might want to hide this, we might expect the unqualified share in the public sector to be even higher.
[4] These include training in systems such as Ayurveda, Yoga and Naturopathy, Unani, Siddi, and Homeopathic Medicine.

instance, there is one village in our sample, located next to a town with 123 providers. Excluding this town from the analysis, or alternatively, using the median instead of mean number of providers, knocks out most of the MBBS providers in the private sector from the choice sets of our households—the median village in our sample does not have an MBBS provider in their choice set. Therefore, the effective options for the majority of households in our villages are a large number of unqualified or semi-qualified providers in the private sector, and trained providers in the public sector. Second, these numbers are not particular to the state we worked in. In parallel work across 1800 villages in 19 Indian states, we find very similar patterns and document an average of 6.2 health care providers for every rural Indian village.

These providers are mostly middle-aged men, half of whom have completed 12 or more years of education (Table 2, Panel A). These practices have been open for an average of 13 years (Table 2, Panel B) and receive around 15 patients per day. Most practices (89 percent) dispense medicines in the clinic itself and are fully equipped with the infrastructure and medical devices required for routine examinations, such as stethoscopes and blood pressure cuffs. Notably, only 15 percent of public providers in the representative sample self-reported a second (private) practice. However, our second audit study confirms that this is a gross underestimate of the trend in dual practice. By using chemists and other providers located in the same markets, we were ultimately able to locate private practices for 74 percent of public providers in Round 2 (Table 2, Panel A).

This profusion of providers is consistent with the high use of primary health services in our sampled villages (Table 1, Panel B). Nearly half of households reported seeking primary care services at least once in the 30 days prior to the survey. They overwhelmingly chose the private sector, which

accounted for 91 percent of all reported visits.[5] Even when they had access to at least one public sector provider practicing as a doctor in their own village, households still visited the private sector 79 percent of the time. Doctors with an MBBS degree accounted for only 3 percent of all reported visits, whereas unqualified providers captured 76 percent.

i. Two audit studies

We conducted two rounds of data collection using an audit framework with 22 standardized patients. The recruitment of the standardized patients, the case selection and the measures of quality are discussed below; here we describe the basic sampling strategy.

Round 1 of the audit study was conducted in three of the five districts in our study. We first eliminated extremely remote sample villages that could not be accessed by a road. We wanted to avoid SP detection since providers in such settings might expect to know all of their patients and since it would be difficult to come up with an excuse for passing through a village with limited connectivity. We also excluded community health workers, midwives, and providers that only made home visits. Among the remaining eligible providers, every public provider and every private provider with an MBBS degree were automatically sampled, as were the private practices of public providers. For each eligible public provider, we also sampled the closest private provider. The remaining providers were randomly selected until the number of private providers sampled per village reached a total of 6. To avoid detection, we limited the maximum number of providers who saw a standardized patient to two per clinic for public clinics and to one for private clinics.

[5] This private sector share is higher than what we find in the latest Demographic and Health Survey administered in India (NHFS 2005-2006). The DHS relies on the household to correctly classify the providers as public or private. We, however, asked households to name the providers they visited, and we obtained the providers' sector from the survey we administered among providers.

We randomly assigned the 22 standardized patients to providers to ensure that there was no correlation between patient characteristics and the attributes of providers/clinics. Each clinic received 3 patients in total – unstable angina, asthma, and proxy-dysentery (where a parent presents on behalf of an absent child) – unless we were sampling the second clinic of a dual-practice provider; in these cases (given the rarity of the case and therefore high likelihood of detection), the provider did not receive a second unstable angina case.

In this round of the study, we treated the clinic as our unit of observation, and standardized patients received care from whoever was acting like the doctor on that day. Following this strategy, for 63 percent of interactions, a provider with no medical training was substituting for the MBBS provider assigned to that clinic, leaving us with 11 public sector MBBS doctors in our sample.

We then conducted a second audit (Round 2) in all 5 districts of Madhya Pradesh in our sample with 116 doctors drawn from the universe of public MBBS doctors working in all villages in those districts. Since most of these doctors practiced privately as well, we also sent standardized patients to their 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. Standardized patients portrayed a dummy case (e.g. headache) if they encountered a substitute, and we sent in other standardized patients on our subsequent attempts. For this second round audit study, providers received two of our three cases—asthma and dysentery in a child sleeping at home—in both their public and private clinics. The third case—unstable angina—was too rare for us to send to both the public and private practices. Consequently, we randomized the providers into two

groups, one group receiving the unstable angina patient in his/her private practice and the other in the public.

Appendix 3 presents several checks on the sample, comparing (A) characteristics of the universe of providers to those sampled in the first audit; (B) characteristics of providers with and without dual practices and; (C) comparing providers who received the angina patient in the public versus the private practice. A couple of differences emerge. First, our sampled providers for Audit 1 in the private sector report more equipment in their clinics and more training than the universe; this difference comes from the elimination of extremely remote villages from our sample. Second, the sample of providers with dual practice has practiced in the current location for 2 more years (with a mean of 5.17 years for the single practice providers) and report seeing half the patients (15 per day) that the single practice sample reports. Finally, there are no differences in provider characteristics and availability of medical equipment for those who received the angina patient in the public versus private sector practices.

III. Audit Methodology

Our methodology builds on previous work measuring the quality of medical care in resource poor settings (Das and Hammer, 2007 and Das, Hammer and Leonard 2008, Leonard, Melkiory and Vialou 2007). Since measures of process quality typically available in OECD countries (like patient charts) are absent in resource poor settings, these studies advanced two different techniques to measure the quality of medical advice.[6,7] They measured "knowledge" by administering medical vignettes—essentially tests of knowledge for standardized case presentations—to multiple doctors

[6] Private providers don't keep any patient information, and even when public providers record details like names and symptoms of patients, they are notoriously incomplete and incorrect.

[7] Further, standard measures of quality in low-income countries focused on measures of medical equipment, infrastructure and the availability of medicines. These `structural measures' may have little to do with the quality of medical advice; for instance, the availability of medicines is a measure of the fiscal subsidy to a clinic rather than a measure of quality.

and they measured "practice" by observing doctor patient clinical interactions. Beyond their finding of very low levels of knowledge and practice, the main result was a difference between knowledge and practice, characterized by a large gap between what doctors tell you they would do in a vignette and what they actually do faced with a similar patient (Rethans, 1991, Das and Hammer 2007 and Leonard and others 2007).

However, comparisons across doctors using direct clinical observations raise the possibility of (A) omitted variable bias due to unobserved patient characteristics and illness severity and (B) Hawthorne effects because doctors knew they were being observed. In addition, observers never knew the `real' illnesses that patients were presenting with. Given the brevity of interactions (in the public sector, doctors spent 90 seconds with patients), these studies were unable to assess the accuracy of diagnosis and treatments. Therefore, to measure the quality of clinical interactions among primary care providers and make valid comparisons across sectors and qualifications, we employed unannounced standardized patients in an audit framework. We describe below three key features of the standardized patient approach as applied in our context—the recruitment of standardized patients, the selection of cases and the measures of quality; further medical details are provided in Das and others (2012).

Standardized patients

Standardized patients are trained personnel who present a standard case to multiple providers, similar to audit approaches in studies of discrimination and credence goods (REF). In our case, to make their appearance, manner, and answers to unanticipated questions conform closely to providers' expectations of their regular patient population, the 22 standardized patients were recruited from local communities (outside sampled locations). During approximately 150 hours of training, they were carefully coached by a professional standardized patient trainer, doctors, and an

anthropologist to consistently portray the emotional, physical, and psychosocial aspects of the case and to accurately recall interactions with providers. All standardized patients were also thoroughly trained to make plausible excuses to avoid thermometers, needles, and pelvic exams and to hide medicines that doctors requested them to ingest in the clinic. After the standardized patient presented to the provider, he/she was debriefed within 1 hour using a structured questionnaire that detailed the questions and examinations that the provider completed or recommended, the treatments provided and the diagnosis given. For example, they had to remember whether or not the provider asked if the chest pain was radiating down the patient's arm in the unstable angina case or whether the provider listened to the patient's chest with a stethoscope. Standardized patients paid the total fees charged by providers, who did not know that they were receiving standardized patients and thus should have treated them as new patients.

Cases

Standardized patients presented one of three cases: unstable angina, asthma, and dysentery of an absent child. For unstable angina, a 45-year-old male complains of chest pain the previous night. Appropriate history taking would reveal classic signs (radiating, crushing pain) and risk factors (smoking, untreated diabetes, and family history of cardiac illness) of unstable angina or an imminent myocardial infarction. The asthma case features a 25-year-old male or female standardized patient presenting with difficulty breathing the night before the visit. When questioned appropriately, the standardized patient reveals that the episode lasted for 10 to 15 minutes and involved a "whistling" sound (wheezing) and that he or she has had similar episodes before, often triggered by house cleaning and cooking smoke. The standardized patient also reports a family history of similar symptoms. For the dysentery case, a 26-year-old father or mother of a 2-year-old child complains that his or her child has diarrhea and requests medicines. When probed, the standardized reveals

details of their water source and sanitation habits, in addition to the presence of fever and the frequency and quality of the child's stools.

These cases are relevant in the Indian context and in many middle- and low-income settings. Incidence of cardiovascular and respiratory diseases has been increasing, and diarrheal disease kills more than 200,000 children per year in India. 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 cases were also chosen to minimize risk to standardized patients since they could not portray any symptoms of infection given the documented high propensity to administer medicines intravenously with unsterilized needles and to use thermometers that have not been appropriately disinfected.

In choosing these cases, we also kept in mind the differential diagnosis and the ability or rural clinics to provide care. In particular, we were interested in presentations that "real" patients would not be immediately able to categorize as `life threatening' or `potentially non-harmful' and therefore sort into appropriate clinics. For instance, our presentation of unstable angina is chest pain which, even in countries like the United States is often mistaken by patients as arising from exertion or muscle strain. Similarly, wheezing and shortness of breath in asthma can also arise from short-term allergies to environmental contaminants. Finally, for any child with diarrhea, the key contribution of a health care provider is to assess whether the symptoms reflect a bacterial or viral etiology and the degree of dehydration. Depending on the provider's assessment, the treatment could range from sending the child home with appropriate instructions to keep the child hydrated (viral diarrhea, not severe dehydration) to immediate administration of intravenous (IV) fluids with antibiotics (dysentery with severe dehydration).

In our case, the appropriate treatment for unstable angina would be immediate referral to a hospital (with or without explicit recommendation for an electro-cardiogram) and the administration of aspirin. For asthma, the provider should recommend an inhaler (widely available in local pharmacies) or an appropriate steroid and for the child with dysentery, the provider should assess the degree of dehydration (not severe) and try to deduce whether a bacterial etiology is present using appropriate questioning. Given that children cannot be used as standardized patients, we coded the treatment as correct if the provider recommended rehydration using Oral Rehydration Salts (ORS).

Measures of Quality

The literature on credence goods identified both *over* and *under* provision as potential inefficient outcomes in such markets. However, the classic presentation of credence goods does not account for provider effort, which previous empirical research has identified as a severe problem in the salaried public sector. To flexibly account for both over and under-provision and differential provider effort and knowledge in the public and private sectors, we use three measures of quality—adherence to a medically necessary checklist of recommended questions and examinations, the nature of the treatment and the nature of the diagnosis.

Adherence to the medically necessary checklist is a continuous measure that records the questions and examinations that the provider completed or recommended and compares them to what is medically required (Appendix Table A1 details each item). Items that are medically required conform both to the Indian government's own guidelines on treatment in these cases and to the advice from our medical advisory panel. In fact, the checklist that we use tends to be more parsimonious than what the Indian government's own guidelines recommend; using the more extensive (and more correct) checklist would deflate the percentage completed further below the low numbers we document below.

We also evaluate whether providers' treatment protocols are *correct*, *helpful*, and/or *unnecessary or harmful*. We use a lenient definition of *correct treatment* for each case and allow providers to be both correct and incorrect at the same time. Take the case of unstable angina, for example. We classify providers as correct if they recommended aspirin, clopidogrel, or other anti-platelet agents or if they referred patients to a hospital. If providers took any one of these actions, we marked him as being correct. This does not mean, however, that the same provider could not have also prescribed or dispensed an unnecessary or harmful treatment, such as antibiotics or psychiatric medicine. The provider could have also given a medicine that could be considered helpful for unstable angina, such as pain medication, but alone would not be considered sufficient for ensuring the safety of the patient. Appendix Table A2 lists the correct and incorrect treatments for each case.

For the final component, we also recorded whether or not a provider *uttered any diagnosis*, right or wrong, to the SP. Because we knew what illness our "patients" were presenting, we can assess whether the provider *uttered a correct diagnosis* and whether a provider *uttered an incorrect 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. Appendix Table A2 lists the correct and incorrect diagnoses that we used for each case. We note though that in close to 50 percent of the cases, the diagnosis is missing because the patient did not receive one. [8]

These three measures of quality are distinct, but closely related. Figure 1 for instance, shows the link between consultation time—a measure of provider effort—and checklist adherence (left) as well as

[8] In the pilot phase of the project, we noted this problem and attempted to correct it by randomizing the standardized patients into two groups, where the second group, as they were leaving the clinic, would turn around and ask the provider if he could tell him what was wrong. But our efforts were in vain. In most cases, the provider would get upset when the standardized patient asked the question and retorted with phrases indicating that the patient would not understand, since he/she was not the doctor. In our data, randomization into this group has no discernable impact on the likelihood of receiving a diagnosis.

the link between checklist adherence and the correctness of treatment. Both figures plot the kernel density for the dependent variable (consultation time and checklist adherence) on the right axis and the non-parametric plot between the dependent variable and checklist adherence/correctness of treatment on the left axis.

As is immediately clear, the more time a provider spent with the patient, the greater the compliance with the checklist. Although the average consultation time in our sample was only 3.1 minutes (identical to that documented in Das and Hammer, 2007 for doctors in Delhi), increasing consultation time from 0 to 10 minutes increased checklist adherence from 0 to just under 40 percent. As we may expect, there is substantial concavity in this relationship; the first 5 minutes really matter, but additional time from (say) 20 to 25 minutes has a much smaller (but still positive impact). Also clear is that greater adherence to the checklist increases the likelihood of correct treatment. Again, perhaps the most striking feature of these data is the low overall checklist adherence—even for our parsimonious list of items, average completion was just 21.1 percent.[9] At the same time, providers who did comply with the checklist significantly increased the likelihood of correct treatment, a relationship that is basically linear over the relevant portions of the density. Appendix Figure 1 shows a similar figure for the likelihood of giving a right diagnosis versus checklist adherence, and again, overall rates of correct diagnosis are low (topping out at 25 percent), but increased checklist adherence always leads to greater rates of correct diagnosis.

Despite the close link between checklist adherence and the correctness of treatment and diagnosis, there is a subtle distinction among them. What we would ideally like is to rank doctors on the basis of their posterior beliefs following the patient presentation. This would include all the potential states of the world (the different diseases) and an associated probability with each. Then, a second

[9] Thus, the majority of providers missed the most elementary questions such as checking for radiating pain in a patient presenting with chest pain (16 percent did so); these basic results are further discussed in Das and others (2012).

mapping would link this posterior belief to the treatment. The way that quality data like ours are typically collected, this posterior assessment is missing. Instead, what we observe is a particular treatment and diagnosis. In the case of the child with diarrhea for instance, we may observe that the provider diagnosed the patient with viral diarrhea and gave ORS, but we cannot tell whether he/she was able to rule out other differential diagnosis. The treatment would be `correct' if our case was indeed one of viral diarrhea, but would be `incorrect' if the case was bacterial diarrhea that required anti-infectives. The checklist adherence provides a measure that is arguably more in line with the posterior assessments we have in mind. For instance, the checklist items for unstable angina are constructed so that we can assess whether the doctor completed the questions and examinations not only to *rule-in* unstable angina, but also to *rule-out* competing explanations such as a stomach ulcer or muscle strain. In an ideal world, the adherence to checklist would be similar to observations of multiple treatments given to multiple standardized patients presenting with chest pain—one with unstable angina, another with muscle strain and a third with stomach ulcer.

i. Identifying impact of incentives in public and private sectors

These two audits will allow us to measure differences in the quality of care that patients received from providers in the public and private sector. In Round 1, we can measure this difference in a sample of providers representative of rural providers in Madhya Pradesh. Since we deployed multiple standardized patients in each market, we can net out differences that might arise, for example, if primary health care centers (PHCs) are located in areas with low demand for quality. Since standardized patients saw multiple providers within a market, we can also control for standardized patient fixed effects to account for any residual correlation between treatment and characteristics of the standardized patient. Formally, we can estimate,

$$q_{i(scp)m} = \beta_0 + \beta_1 public_{ip} + \beta_2 qualification_p + \beta_3 patient\ load_i + \delta_s + \delta_c + \delta_m + \varepsilon_{i(scp)m} \quad (1)$$

where we regress the quality *q* we observe in interaction *i* between a standardized patient *s* presenting case *c* and a provider *p* in market *m* and control for providers' qualifications and the number of patients waiting in the clinic at the time of the interaction. Since different cases may require different lines of history-taking and examination or pose different challenges for providers, we also include a set of indicators δ_c for the cases presented by the SP and a full set of SP fixed effects, δ_s, to capture any variation observed across SPs.

In this model, β_1 jointly estimates the effect of being in the private sector and differences in provider characteristics across the two sectors. For instance, providers in the private sector are typically medically untrained, may have different levels of altruism and may have different risk preferences. The data from Round 2 allow us to eliminate potential differences due to differing provider characteristics since standardized patients visited the *same* doctor in his/her public and private sector clinic. We thus estimate a public-private quality difference net of provider fixed effects, δ_p:

$$q_{i(scp)v} = \beta_0 + \beta_1 public_{ip} + \beta_2 patient\ load_i + \delta_p + \delta_c + \varepsilon_{i(scp)v} \quad (2)$$

In this case, we estimate the average difference between the public and private practices of the same provider, except for the case of unstable angina, where providers randomly received the standardized patient in either their public or private practice (see Appendix Table X confirms that there are no observable differences among these providers). In our robustness section, we discuss and rule-out the possibility that our estimate reflects deliberately lower effort in the public sector among doctors with dual job holdings.

IV. Results

Quality in the Public and Private Sectors: Figures and Mean Comparisons

Two pictures and a table present the basic results from our exercise. Given that we used an audit framework and that providers were randomized into receiving the audit patient in his/her public or private clinics, subsequent regressions help identify point-estimates, provide standard-errors and rule-out alternative explanations for our results.

Figures 2 and 3 follow the same basic format: We plot kernel densities of the standardized score across all three audit patients, where the standardized score represents standard-deviation differences in checklist adherence between different groups of providers, aggregated across all three cases. Figure 2 highlights the much lower quality, as evinced through checklist adherence, in the public clinics relative to private qualified (recall that very few of these are MBBS providers) and private unqualified (the bulk) providers. As is immediately clear, checklist adherence in the public sector is significantly lower than in the private sector—the distribution for the public sector has a mean and mode that is far to the left of the private sector distributions, with a right skew. Neither do we find large differences between qualified and unqualified providers in the private sector. A likely explanation for this latter finding is that even the qualified providers in our sample are not MBBS providers and that the quality of medical education for alternate degrees may be very poor.

Figure 3 disaggregates the patterns further by using Audit 2 to separate public MBBS providers and private MBBS providers, the majority of whom are public sector doctors in their private clinics. There are a couple of noteworthy features. First, there is a significant difference between the public unqualified providers who gave care to 63 percent of the SPs in Audit 1 and the doctors in the public sector. Public sector doctors come out looking significantly better than the public sector unqualified providers, though still to the left of unqualified providers in the private sector. Second, the same doctors in their private sector clinics provide higher quality care. There is a full shift of the

distribution to the right, and in fact in their private sector care, the public doctors provide the highest quality care among our entire sample. The simple means are telling: Private sector doctors without any medical qualifications completed 21.6 percent of all checklist items, relative to 15.7 percent for public sector providers without qualifications. Doctors in the public sector completed 18.8 percent of all checklist items, but the same doctors in their private clinics completed 28.1 percent. Observed in their private clinics, these doctors were the best in the entire system. But, observed in their public clinics, the same doctors were the worst in the entire system.

Table 3 compares treatments among different kinds of providers. We follow the credence good literature and differentiate between over and under-treatment. An SP was "under treated" if he or she did not receive a treatment that included the right course of action for her presentation. This includes cases where the SP received the correct treatment only, but also cases where in addition to the correct treatment the SP may have received helpful treatments (for instance, symptomatic pain relief in the unstable angina case) or even incorrect treatments. Alternatively, over-treatment includes all cases where an SP was given incorrect treatment only, but also cases where the incorrect treatment was coupled with other drugs that were helpful or even correct. Thus, an SP who was correctly referred to the hospital with unstable angina but also received an antibiotic and a pain killer would, strictly speaking, not have been under-treated but *would* have been over-treated due to the antibiotic.

For those not familiar with the Wild East of medical care in India, the numbers will be startling. Across all three cases, the correct treatment *only* was given in less than 5 percent of all cases across all types of providers. At least one incorrect treatment was given in more than 50 percent of SP presentations. The numbers look slightly better for "correct treatment plus", a category that includes interactions were a correct treatment was coupled with a helpful or incorrect treatment, but even

here, the fractions range from 25 percent (the untrained public providers) to 44 percent (public sector providers in their private clinics). In contrast, nearly one fifth of the SPs received *only* an incorrect treatment, with a high of 28 percent among the untrained public providers. The basic reality is that less than a half of patients with these cases will receive any correct treatment in the rural medical marketplace, and close to a fourth will leave the clinic with only incorrect treatments for their ailments.

The variation across types of providers follows the observed patterns for checklist adherence. Untrained providers in the public sector come out looking the worst with a 25-26 percent likelihood of giving the correct treatment relative to 39 to 44 percent among trained providers. Across all types of providers, incorrect treatments are just as high—and substantially *higher* for the doctors (74 to 77 percent relative to 52 to 58 percent). Finally, there is a 10 percent difference in the likelihood of providing the correct treatment for MBBS doctors in their private relative to their public clinics (for those with a dual practice) and virtually no difference in treatment patterns between doctors with and without a dual practice in their public clinics.

The case-wise breakup further shows large differences in the likelihood of correct treatment for unstable angina for dual practice doctors depending on the sector of care, coupled with a discount for dysentery. For unstable angina, being in the private sector increases the likelihood of correct treatment by almost 100 percent (from 22 to 41 percent) and for asthma by 15 percent (from 58 to 68 percent). For dysentery, the likelihood of correct treatment declines in the private clinic, though incorrect treatment is again equally high across all categories.

These basic comparisons of means and densities presages the results from our regressions: Consistently we will find higher adherence to checklist in the private sector (with and without provider fixed-effects). On treatments, we find public sector doctors do better than those in the

private sector, but this is primarily due to a difference in their qualifications. Controlling with qualifications reverses the difference, and under-treatment is higher in the public sector with no accompanying evidence of higher over-treatment.

Quality in the Public and Private Sectors: Regressions

The regression analogs to these figures are presented in Tables 3, 4 and 5, Table 3 looks at pure measures of effort (consultation time), adherence to checklist and the standardized score across all cases for checklist adherence. Table 4 looks for evidence of under-treatment in the public and private sectors. We look at both the provision of information—whether the provider offered a diagnosis and whether the provider offered a correct diagnosis—and the correctness of treatment as defined above. Finally, Table 5 looks at evidence of over-treatment, where we focus on the likelihood of providing an incorrect diagnosis and of providing an unnecessary or harmful treatment. For each of these tables, we include a full set of location (village) fixed-effects and a full set of indicator variables for the standardized patients. We note that as a check of the reliability of the audit methodology, the joint test that all SP fixed-effects are jointly insignificant is never rejected in these specifications.

The basic messages are as follows. First, providers in the public sector spend a lot less time than in the private sector, ranging from 1.12 minutes for the dual sample to 1.5 minutes for the full sample. This may not seem like much, but the average consultation time in our sample is 3.12 minutes (Table 3). Therefore, in the public sector, the typical time spent with a patient is around 90 seconds. Checklist adherence is lower by 8 to 9 percentage points in the public sector, which represents a 40 to 50 percent discount relative to the private sector, which is 0.6 standard-deviation on the normalized score. These results are virtually identical with and without provider fixed-effects and

patient load at the time of the visit has little discernable impact—not surprising because with an average of 10 to 15 patients per day, these providers operate under enormous excess capacity.

Second, there is strong evidence of under-treatment in the public sector. Public sector providers are less likely to give any diagnosis to their patients, but on average are just as likely to have given a correct diagnosis (which only 6 percent our SPs received). They are 7.5 – 8.7 percentage points less likely to give the correct treatment. On a base of 39 percent, this represents a 20 percent discount. Again, patient load has no discernable impact.

Third, mirroring the descriptive tables, public-private differences in time spent, adherence to checklist and the likelihood of providing a diagnosis are equally strong across all three cases, but the patterns for correct treatment are case dependent (Table 6). Specifically, with and without provider fixed-effects there is a strong discount in the public sector for correctness of treatment in the asthma case; given a mean rate of 38 percent, the discount of 16 to 21 percentage points is substantially high. For unstable angina, there is a small (but statistically insignificant) premium in the public sector of 8 percentage points when we compare all public and private providers. However, in the dual sample, the signs are reversed and the public sector discount is again large and significant: SPs seen in the private sector practice were almost 100 percent more likely to be correctly treated. We ascribe this difference between the dual sample and the full sample as reflecting differences in the knowledge and competence of the providers in the market. Finally, we find no discernable impacts of public sector practice on treatment for dysentery. In both the public and private sector, providers were equally likely to give antibiotics (helpful, but not necessary) and equally unlikely to prescribe an oral rehydration solution (necessary).

Prices in the Private Sector

As the preceding results show, both adherence to checklists and the rate of correct treatments are higher in the private sector, especially when we focus on provider fixed-effects. The higher quality in the private sector suggests that price incentives, which are the only form of accountability in the private sector, reward quality, and they reward the "right" type of quality. Our final results look at the direct relationship between the prices that the SPs were charged and our measures of quality. Formally, we regress the prices charged to our SPs on these various dimensions of quality $\boldsymbol{q}$

$$p_{i(scp)m} = \beta_0 + \beta_1 qualifications_p + \beta_2 patient\ load_i + \gamma \sum \boldsymbol{q} + \delta_c + \delta_s + \delta_m + \varepsilon_{i(scp)m} \quad (3)$$

after limiting the sample to private providers since we see only 3 prices in the public sector (Rs. 0, Rs. 2, and Rs. 5), corresponding to the official schedule of fees that can be charged in public clinics. Given the number of providers in these markets, it should not be unreasonable to assume that both providers and patients are price-takers, and because our SPs paid whatever price the provider charged, the γ coefficients should capture the returns that providers can earn by providing that dimension of quality during the interaction. We highlight that the audit nature of the evidence implies that the estimated price-quality relationship is not contaminated by the sorting of patients to providers.

Presaging the results relating price to quality, Figure 4 shows the non-parametric relationship between the fees charged to the SPs and adherence to the checklist for the private sector as a whole and for only the dual sample—public providers observed in their private practice. As is clear, there is a basic price-quality hedonic relationship in this market, with higher quality resulting in higher prices. The marginal return to quality is high: Moving from the bottom quintile of checklist adherence to the top increases prices from Rs.20 to Rs.48, or close to 150 percent. Also of independent interest is that the returns to quality for the "dual" sample—public providers in their private practice are

identical to that of the general private sector; this is a basic arbitrage condition that we would expect holds when the market for health care functions independent of quality signals such as the presence of a public sector practice.

Table 7 then shows the estimated returns to quality as measured through checklist adherence in the private sector. Because checklist adherence may be correlated to other clinic attributes that are easily observed (is the clinic clean, does the provider have the necessary equipment), which in and of themselves may not result in higher quality, in successive specifications we include a richer set of observed characteristics. Specifically, Column 1 shows the bivariate correlation; Column 2 then includes SP dummies and location fixed-effects; Column 3 introduces provider characteristics—including the tenure in the current location as a proxy for reputation and Column 4 introduces clinic characteristics. Finally, Column 5 investigates whether what is being rewarded is consultation time, instead of checklist adherence. Recall that in Figure 1, we found that consultation time and checklist adherence were strongly correlated; not surprisingly, the more time a provider spends with a patient, the more they do. However, for the purposes of the price-quality relationship, if what is captured in Figure 4 is this basic relationship, it should lend pause in interpreting the relationship as a reward to quality, rather than just effort. In all regressions, we flexibly account for intra-cluster correlations using corrected standard errors.

Across all specifications, we find a strong and positive return to checklist adherence on prices. With no controls, every additional percent completed increases fees by Rs.0.60, and this coefficient does not change with the inclusion of location and SP fixed-effects. Adding in provider controls decreases the estimated effect somewhat to 0.49, in part due to a distinct premium for MBBS providers in the market. The estimated coefficient further declines to 0.38 when we add in independent controls for consultation length (which is also rewarded with an additional Rs.1.65 per

minute; any other aspects of the interaction (including controls for the regular patient body or the specific types of infrastructure in the clinic) have no further impact on this coefficient. We view this as strong evidence that customer accountability in this virtually unregulated market creates strong incentives for medically appropriate quality of care.

V. Robustness

Several ancillary results help interpret the results on public-private differences. These additional results can be broadly grouped into three categories. The first is whether the audit methodology reflects similar patterns with real results; the second is whether the focus on dual practice biases our results towards finding positive effects for the private sector because public sector providers deliberately decrease quality and the third is whether these results hold beyond the particular context of rural Madhya Pradesh.

Audit versus real patients

In one celebrated example, the results from an audit study of discrimination in the car sale market were very different from those in observational data (Ayres and Siegelman 1995 and Golderg 1996). In that case, Goldberg (1996) shows that the audit study elicited the first price offer, but not the completed sale. Because the two populations studied differ in the overall distribution of valuations, the optimal bargaining solution called for a higher initial price offer, but an equal final price. In our particular case, this is less of a problem since we observe a "completed sale", but the basic critique, which applies to all audit studies, is that the standardization of patients does not fully replicate what would happen if the same *body of patients* arrived in the public and private sectors. In our particular case, it may be, for instance, that households deliberately triage into the public or private sectors depending on the severity of their illness. Consequently, patients with chest pain reporting to the public sector are more likely to suffer from muscle strain and in the private sector from unstable

angina. There is no clear solution to this problem, since any scheme that tries to reproduce the same patient distributions in both sectors will fundamentally alter the observed pricing and accountability schemes—whose impact we are interested in estimating to begin with. Furthermore, in this study the cases were deliberately chosen to reflect conditions that could either be "serious" or require palliative care and it is unlikely that households would be able to fully triage on the basis of the symptoms alone. In fact, given the low rates of correct diagnosis and treatment, it is difficult to believe that the most educated population in the region (the providers) were less likely than households themselves to parse the symptoms appropriately.

Nevertheless, observations with real patients can help provide further insights into patient sorting and the potential impact on our results. To look at this, we observed interactions between the providers in our samples and real patients in their clinics for a day, recording various salient aspects of each interaction. We then interviewed the same patients as they left the clinic with a short questionnaire on their sickness, their overall health status, asset ownership, education and the distance they had travelled.

Table 8 first shows various aspects of the interaction between doctors and their real patients, with and without provider fixed-effects. Because observers were not medically trained and we never knew what illnesses the patients came with, we use coarser measures of quality—the time spent, the number of questions asked, whether the doctor examined the patient and whether they gave medicines. In both the overall and the dual sample, the public sector always does worse—and by large numbers. For instance, the public sector spent 1.11 minute less per patient, relative to a mean of 3.7 minutes in the private sector. The average interaction with real patients lasted 90 seconds—just as with our standardized patients. They asked an average of 3.2 questions and only in 64 percent

of interactions were a single examination completed. These results are virtually identical with provider fixed-effects.

In Appendix Table A5, we return to the SP sample and now add in a rich set of "real patient" controls from the exit survey. These include where the patient came from, Activities of Daily Living (a measure of health status) and the detailed list of symptoms. Again, there is no change in our estimated coefficients on consultation time, adherence to checklist and likelihood of giving a diagnosis. The coefficient on correct treatment decreases, but is statistically indistinguishable from our previous estimate.

Dual Practice

A second concern is that our results reflect the particular nature of our sample—providers with dual practices could systematically decrease quality in the public sector to increase demand for their services in the private sector. We first note that if this were the case, it could only reflect the sensitivity of provider behavior to price incentives. If providers did not care about incentives, they should have no reason to decrease their quality in the public sector. Second, there are two potential margins through which dual practice could affect provider behavior—they could either decrease effort on the extensive margin by not showing up to work, or they could decrease effort on the intensive margin by providing less effort when at work. The degree to which they will do so depends on market segmentation—if the private and public clinics for instance, are in different locations, we would expect the provider to operate on the extensive rather than the intensive margin. Our standardized patients, for instance, were never referred to the private clinic by the public sector doctor.

To look at this, we undertook several comparisons. First, we compared the practice of providers with and without a dual practice, noting that potential selection into dual practice remains a concern.

We find (Table A3) that providers with a dual practice spend less time with their patients and report lower checklist adherence, but are just as likely to provide correct treatments and utter a diagnosis. Critically, providers with dual practice were 4 percentage points *less* likely to refer the standardized patient, although the result is not statistically significant. We also looked for potential heterogeneity by market segmentation, using the fact that some dual practice providers have their public and private clinics in the same location and others choose different villages and towns. Again, we were unable to find any differences in their behavior depending on the location of their practice (results not shown).

External Validity

Our final set of results assess whether the public-private difference noted here is particular to the rural, largely illiterate population that we studied. For instance, it could be that in urban areas, monitoring the public sector is easier through administrative means. Table 9 presents results from the pilot phase of the project, which was conducted in the richest urban Indian state—Delhi. Here, we present data from 231 interactions for the same cases, noting that private sector doctors were purposively sampled and do not have the same competence or qualifications as those in the public sector. We find a much higher discount in the public sector in terms of time and adherence to the checklist. Yet, the public sector does better in terms of correctness of treatment and the likelihood of giving incorrect treatments. Thus, lower effort in the public sector is robust to the setting of the study, but the correctness of treatments is not. Whether this reflects the fact that the providers in the public and private samples are inherently different in their medical training (the private sample was less qualified and less competent) or higher administrative accountability remains an open question.

VI. Conclusion

Our audit study of public and private care in the health market revealed higher quality in the private sector—both in terms of adherence to a medically required checklist and the correctness of treatment. Further, prices in the private sector were strongly correlated to quality suggesting that customer accountability, even in a setting where the majority of the population is illiterate, creates incentives for quality. These results add to our understanding of accountability in the market for credence goods, and are, to our knowledge, the first evidence using an audit methodology in a largely unregulated medical market.

References

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

Ayres, Ian. "Fair Driving: Gender and Race Discrimination in Retail Car Negotiations." *Harvard Law Review* 104 (February 1991): 817-72.

Banerjee, A., Duflo, E. and Rachel Glennerster. 2008. "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System." *Journal of the European Economic Association*, Vol.(6):487-500.

Banerjee,A, Deaton, A, Duflo, E. "Health care delivery in rural Rajasthan." *Economic and Political Weekly*. 2004. 39 (9):944–949.

Canice Prendergast, 2003. "The Limits of Bureaucratic Efficiency," *Journal of Political Economy*, University of Chicago Press, vol. 111(5), pages 929-958, October.

Centre for Policy Research, The MAQARI Team. "Mapping Medical Providers in Rural India: Four Key Trends". Centre for Policy Research, New Delhi. 2011. Available at: http://cprindia.org/sites/default/files/policy%20brief_1.pdf

Chaudhury, Nazmul and Jeffrey Hammer. 2005. "Ghost Doctors: Absenteeism in Bangladeshi Health Facilities." *World Bank Economic Review*, 18(3): 423-41.

Chaudhury, Nazmul., Jeffrey Hammer, Michael Kremer, K. Muralidharan and F. Halsey Rogers. 2006. "Missing in Action: Teachers and Health Workers in Developing Countries." *Journal of Economic Perspectives* 20(1) Spring.

Das J, Hammer J, Leonard K. 2008. The quality of medical advice in low-income countries. *Journal of Economic Perspectives*. Spring; 22(2):93-114.

Das, J, Hammer, J. Money for Nothing: The Dire Straits of Medical Practice in Delhi, India. 2007. *Journal of Development Economics* 83(1): 1—36.

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

Dulleck U, Kerschbamer R. 2006. "On Doctors, Mechanics and Computer Specialists: The Economics of Credence Goods," *Journal of Economic Literature* p5-42

Dulleck U., Kerschbamer R., Sutter M. "The Economics of Credence Goods: On the Role of Liability, Verifiability, Reputation and Competition". *American Economic Review* 2011; 101:526-555.

Goldberg, Pinelopi Koujianou. 1996. "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey." *Journal of Political Economy*, Vol. 104, No. 3 (Jun., 1996), pp. 622-654

Gruber, Jonathan and Maria Owings. 1996. "Physician Financial Incentives and Cesarean Section Delivery.", *The RAND Journal of Economics* , Vol. 27, No. 1 (Spring, 1996), pp. 99-123

Leonard, KL., Melkiory, CM, Vialou, A. 2007. "Getting doctors to do their best: the roles of ability and motivation in health care." *Journal of Human Resources*, 42 (3): 682-700.

Maestad, O, Ole, F. N. 2010. "Overworked? On the relationship between workload and health worker performance." *Journal of Health Economics,* 29(5): 686-698.

Mcguire T. "Physician Agency". In: Culyer A. J., Newhouse J. P., editors. Handbook of Health Economics. Vol. 1. Elsevier; 2000. p. 461-536.

Rethans, J. J., Sturmans, F, Drop,R, Van der Vleuten, CPM, Hobus, P. 1991. "Does competence of general practitioners predict their performance? Comparison between examination setting and actual practice." *British Medical Journal* 303 (6814): 1377–80.

Schneider H. 2012. "Agency Problems and Reputation in Expert Services: Evidence from Auto Repair". *Journal of Industrial Economics,* 60:406-433.

Wolinsky A. "Competition in a Market for Informed Experts' Services". 1993. *RAND Journal of Economics,* 24:380-398.

Tables and Figures

Figure 1a

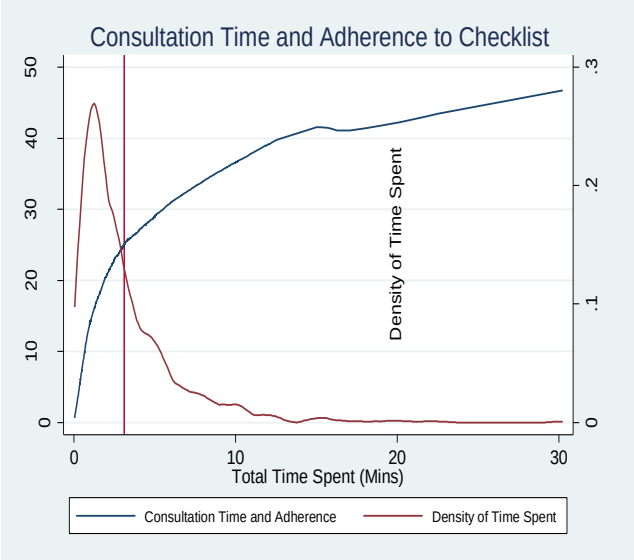


Figure 1b

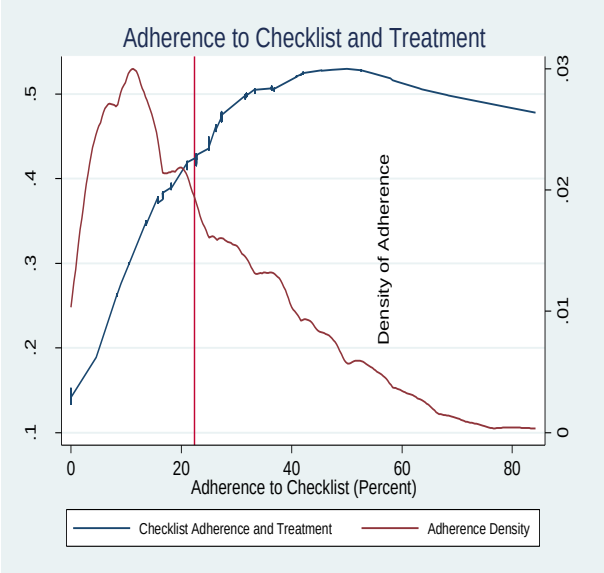


Figure 2: Adherence to Checklists

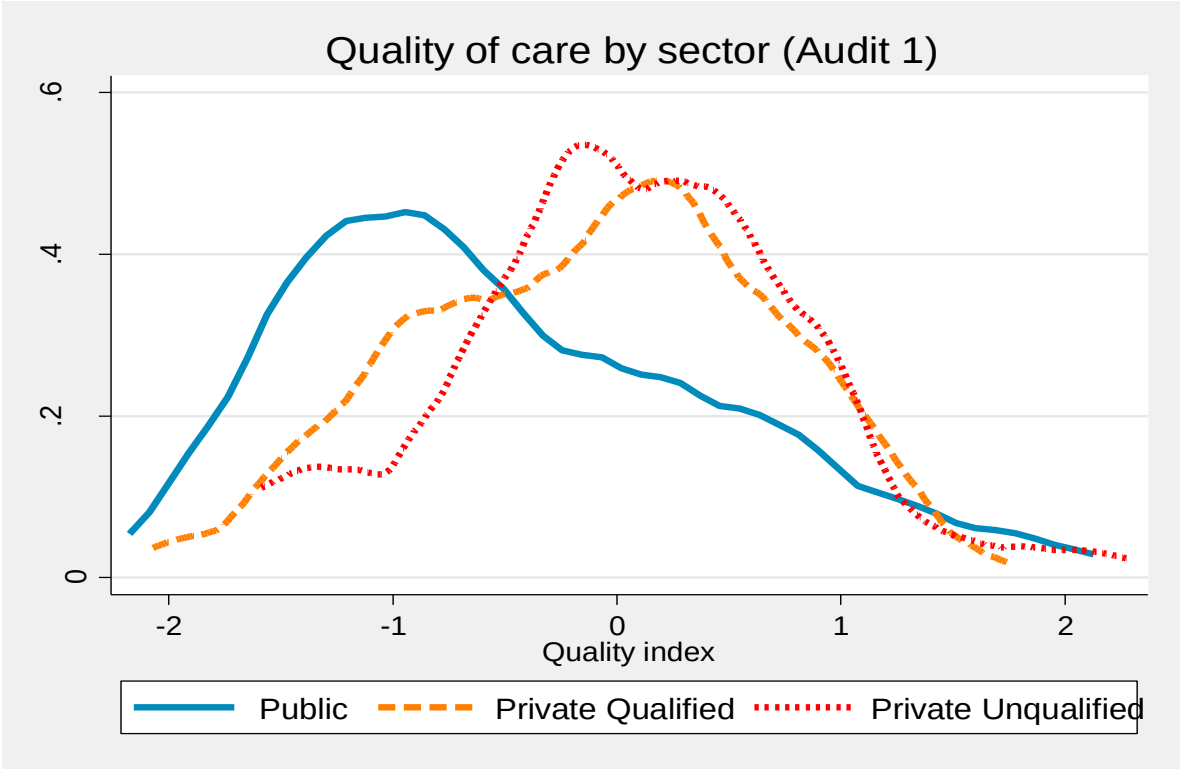


Figure 3: Adherence to Checklist and Dual Practice

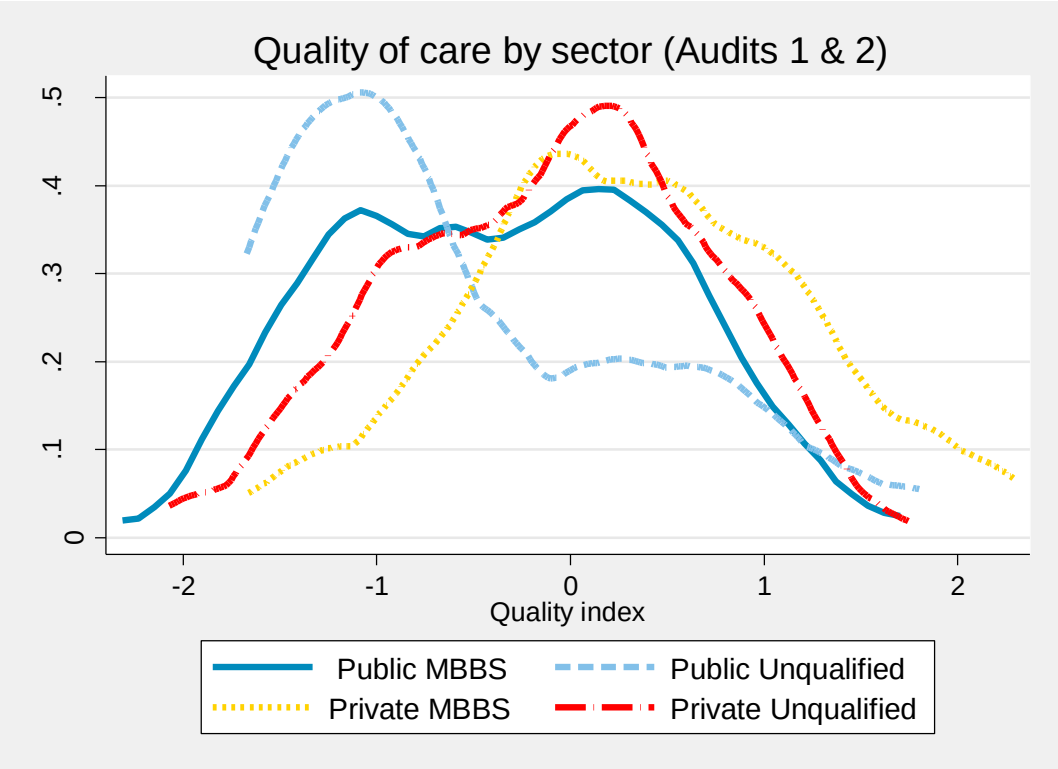


Figure 4

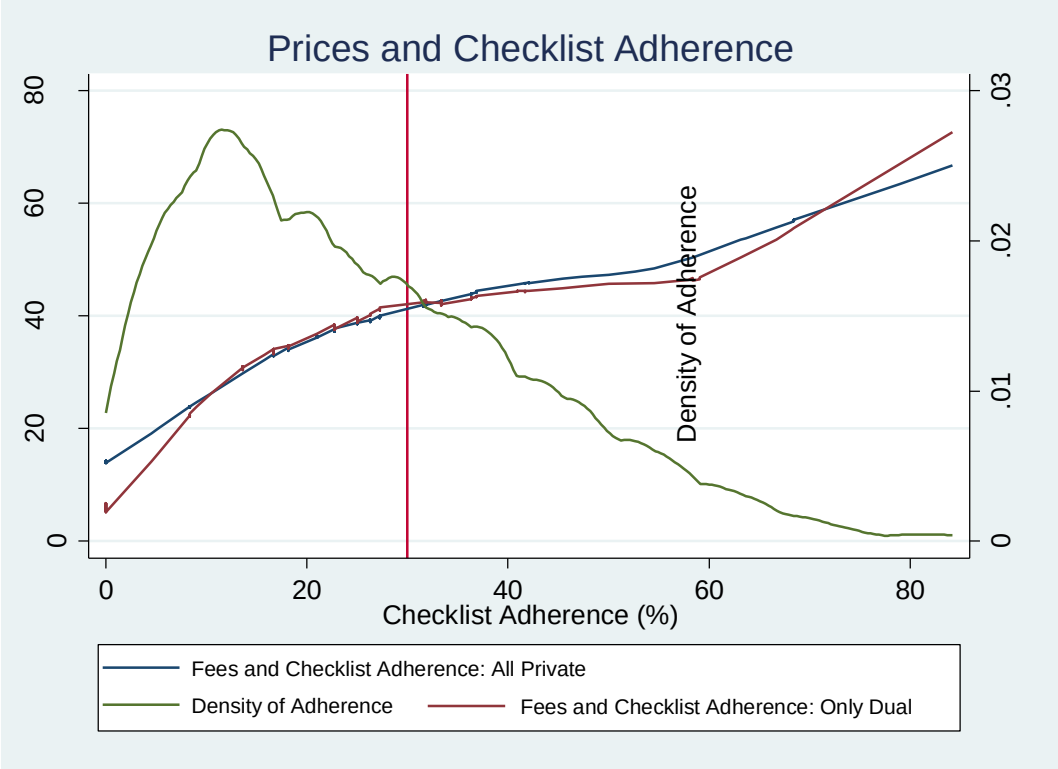

Prices and Checklist Adherence
0
20
40
60
80
0
.01
.02
.03
Density of Adherence
0
20
40
60
80
Checklist Adherence (%)
Fees and Checklist Adherence: All Private
Density of Adherence
Fees and Checklist Adherence: Only Dual

Table 1: Health market attributes

	(1)	(2)	(3)
	Total	**Inside village**	**Outside village**
Panel A: Number of providers			
Total	11.05	3.06	7.99
	(1.25)	(0.37)	(1.29)
Public MBBS	0.50	0.05	0.45
	(0.11)	(0.02)	(0.10)
Public alternative qualification	0.22	0.07	0.15
	(0.05)	(0.03)	(0.04)
Public paramedical	1.58	1.13	0.45
	(0.19)	(0.15)	(0.13)
Public unqualified	1.70	0.67	1.03
	(0.17)	(0.10)	(0.15)
Private MBBS	0.42	0.00	0.42
	(0.16)	0.00	(0.16)
Private alternative qualification	1.92	0.23	1.69
	(0.36)	(0.07)	(0.37)
Private unqualified	5.40	1.81	3.59
	(0.60)	(0.22)	(0.61)
Panel B: Composition of demand			
Population (2001 Census of India)	3885	1354	2531
	(385.46)	(103.56)	(378.58)
Probability provider contact in last 30 days	0.46		
	(0.00)		
provider was inside/outside village		0.65	0.35
		(0.00)	(0.00)
Distance traveled to visited provider (km)	1.66	0.40	3.92
	(0.02)	(0.01)	(0.03)
Probability visit was to private sector	0.91		
	(0.00)		
Probability visit was to private sector in villages with at least 1 public doctor	0.79		
	(0.01)		
Probability visit was to MBBS doctor	0.03		
	(0.00)		
Probability visit was to unqualified doctor	0.76		
	(0.00)		
Number of villages	100		
Number of households	23306		
Number of reported household-visits	18632		

Note: Standard errors 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. Villages "outside village" 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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Audit 1			**Audit 2**			**Audits 1 and 2: Dual**		
	Public	**Private**	**p-value of (1)-(2)**	**Public**	**Private**	**p-value of (4)-(5)**	**Public**	**Private**	**p-value of (7)-(8)**
Panel A: Provider characteristics									
Age	47.46	43.35	0.05	44.60	45.55	0.55	45.31		
Male	0.89	0.95	0.11	0.88	0.85	0.67	0.85		
More than 12 years of basic education	0.54	0.50	0.67	0.61	0.67	0.43	0.68		
MBBS degree	0.26	0.07	0.00	1.00	1.00	0.00	0.97		
No medical training	0.63	0.68	0.53	0.00	0.00	0.00	0.01		
Operates more than 1 practice (*Self-reported*)	0.09	0.02	0.02			0.00			
Operates more than 1 practice (*Verified*)				0.74	1.00		1.00		
Panel B: Practice characteristics									
Provider tenure in current location	15.82	13.70	0.25	6.12	7.64	0.13	7.02	6.80	0.84
Dispenses medicines in clinic	1.00	0.81	0.00	0.49	0.36	0.08	0.35	0.35	0.00
Fees	3.65	51.09	0.00	3.75	58.57	0.00	3.90	55.90	0.00
Patients per day	27.76	15.74	0.00	19.17	16.91	0.32	15.31	14.62	0.74
Electricity	0.94	0.95	0.96	0.99	1.00	0.39	1.00	1.00	0.00
Stethoscope	0.97	0.94	0.45	1.00	1.00	0.00	1.00	1.00	0.00
Blood pressure cuff	0.83	0.75	0.30	1.00	1.00	0.00	1.00	1.00	0.00
Thermometer	0.94	0.92	0.61	0.97	0.97	0.90	0.98	0.97	0.53
Weighing scale	0.86	0.52	0.00	0.93	0.81	0.03	0.94	0.77	0.01
Handwashing facility	0.89	0.81	0.27	0.83	0.80	0.63	0.86	0.79	0.29
Number of providers	38	188		101	83		71	71	

Notes: Unit of observation is a provider. The *dual sample* consists of providers who received a standardized patient in both their public and private practices. The provider mapping and complete provider census yielded information about whether or not a provider operates more than practice. Audit 1 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 Audit 2, however, the existence of additional medical practices was *verified* by repeated observation. Means for fees have been calculated from direct observations of clinical interactions. All other variables derive from a survey administered during the census of providers.

Table 3: Treatment Patterns for All and Specific Cases

FULL	Private Untrained	Private Other	Private MBBS	Public MBBS with Private Practice in Public	Public Untrained	Public MBBS
Correct Treatment ONLY	0.04	0.05	0.04	0.04	0.05	0.05
Correct Treatment and Incorrect Treatment	0.02	0.01	0.03	0.00	0.03	0.00
Correct Treatment and Helpful Treatment	0.03	0.01	0.04	0.01	0.00	0.01
Correct Treatment and Helpful Treatment and Incorrect treatment	0.16	0.20	0.33	0.35	0.17	0.33
Correct Treatment Plus	0.26	0.26	0.44	0.40	0.25	0.39
Helpful Treatment only	0.14	0.17	0.13	0.17	0.14	0.16
Helpful treatment and incorrect treatment	0.14	0.19	0.18	0.20	0.11	0.23
Incorrect Treatment Only	0.19	0.17	0.20	0.22	0.28	0.20
All unlabeled	0.06	0.05	0.00	0.00	0.05	0.00
At least one medicine unlabeled	0.19	0.15	0.02	0.00	0.18	0.00
Incorrect Treatment Plus	0.52	0.56	0.74	0.77	0.58	0.76

UNSTABLE ANGINA	Private Untrained	Private Other	Private MBBS	Public MBBS with Private Practice in Public	Public Untrained	Public MBBS
Correct Treatment ONLY	0.13	0.11	0.15	0.11	0.15	0.14
Correct Treatment and Incorrect Treatment	0.01	0.00	0.00	0.00	0.00	0.00
Correct Treatment and Helpful Treatment	0.02	0.00	0.02	0.00	0.00	0.00
Correct Treatment and Helpful Treatment and Incorrect treatment	0.14	0.11	0.24	0.11	0.15	0.12
Correct Treatment Plus	0.29	0.23	0.41	0.22	0.30	0.26
Helpful Treatment only	0.03	0.14	0.04	0.16	0.05	0.14
Helpful treatment and incorrect treatment	0.28	0.32	0.48	0.49	0.30	0.47
Incorrect Treatment Only	0.04	0.09	0.04	0.14	0.05	0.12
All unlabeled	0.11	0.07	0.00	0.00	0.05	0.00
At least one medicine unlabeled	0.24	0.16	0.04	0.00	0.25	0.01
Incorrect Treatment Plus	0.47	0.52	0.76	0.73	0.50	0.71

DYSENTERY	Private Untrained	Private Other	Private MBBS	Public MBBS with Private Practice in Public	Public Untrained	Public MBBS
Correct Treatment ONLY	0.00	0.00	0.00	0.02	0.00	0.01
Correct Treatment and Incorrect Treatment	0.03	0.02	0.01	0.00	0.00	0.00
Correct Treatment and Helpful Treatment	0.00	0.00	0.00	0.00	0.00	0.00
Correct Treatment and Helpful Treatment and Incorrect treatment	0.06	0.02	0.20	0.31	0.09	0.30
Correct Treatment Plus	0.10	0.05	0.21	0.33	0.09	0.31
Helpful Treatment only	0.40	0.39	0.32	0.34	0.36	0.34
Helpful treatment and incorrect treatment	0.10	0.18	0.08	0.09	0.00	0.10
Incorrect Treatment Only	0.23	0.18	0.34	0.22	0.45	0.24
All unlabeled	0.03	0.02	0.00	0.00	0.05	0.00
At least one medicine unlabeled	0.13	0.18	0.00	0.00	0.05	0.00
Incorrect Treatment Plus	0.43	0.41	0.63	0.63	0.55	0.64

ASTHMA	Private Untrained	Private Other	Private MBBS	Public MBBS with Private Practice in Public	Public Untrained	Public MBBS
Correct Treatment ONLY	0.00	0.02	0.02	0.03	0.00	0.03
Correct Treatment and Incorrect Treatment	0.03	0.00	0.06	0.00	0.09	0.00
Correct Treatment and Helpful Treatment	0.07	0.02	0.09	0.02	0.00	0.03
Correct Treatment and Helpful Treatment and Incorrect treatment	0.28	0.48	0.51	0.53	0.26	0.51
Correct Treatment Plus	0.39	0.52	0.68	0.58	0.35	0.57
Helpful Treatment only	0.00	0.00	0.00	0.00	0.00	0.00
Helpful treatment and incorrect treatment	0.06	0.07	0.12	0.14	0.04	0.19
Incorrect Treatment Only	0.30	0.23	0.14	0.27	0.30	0.22
All unlabeled	0.05	0.07	0.00	0.00	0.04	0.00
At least one medicine unlabeled	0.21	0.11	0.04	0.00	0.26	0.00
Incorrect Treatment Plus	0.67	0.77	0.84	0.94	0.70	0.92

Table 4: Provider effort in the public and private sectors

	(1)	(2)	(3)
Dependent variable:	Time spent	Percentage of checklist items	IRT score
Panel A: Full sample (Village and SP fixed effects), N=1,105			
Public	-1.53***	-7.87***	-0.59***
	(0.25)	(1.71)	(0.16)
Unstable angina	1.55***	-24.4***	-1.68***
	(0.36)	(2.06)	(0.14)
Dysentery	-1.86***	-4.79***	0.011
	(0.20)	(1.16)	(0.053)
MBBS	-0.16	2.55	-0.022
	(0.68)	(2.10)	(0.13)
Other qualification	0.047	3.79***	-0.011
	(0.39)	(1.05)	(0.17)
Patient load during visit	-0.091**	-0.27	0.0097
	(0.035)	(0.22)	(0.0090)
Panel B: Dual-practice sample (Provider fixed effects), N=344			
Public	-1.12***	-9.06***	-0.68***
	(0.23)	(2.07)	(0.13)
Unstable angina	0.76	2.82	-0.058
	(0.46)	(2.27)	(0.068)
Dysentery	-1.70***	-2.08	0.00047
	(0.22)	(1.77)	(0.026)
Patient load during visit	-0.16**	-0.64	-0.019
	(0.075)	(0.66)	(0.031)
Mean	3.12	22.38	-0.07
SD	3.09	16.82	0.86
Mean: Public	2.03	18.34	-0.35
SD: Public	1.86	15.80	0.89
Mean: Private	3.65	24.53	0.07
SD: Private	3.43	16.96	0.82

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. Observations are standardized patient-provider interactions. The *dual sample* consists of providers visited separately in both their public and private clinics.

Table 5: Under-treatment in the public and private sectors

Dependent variable:	(1) Uttered a diagnosis	(2) Uttered correct diagnosis	(3) Correct treatment	(4) Any helpful treatment
Panel A: Full sample (Village and SP fixed effects), N=1,105				
Public	-0.13***	-0.027	-0.075**	-0.068
	(0.038)	(0.020)	(0.037)	(0.042)
Unstable angina	-0.85***	-0.97***	-0.44***	0.33***
	(0.053)	(0.024)	(0.058)	(0.058)
Dysentery	-0.18***	-0.073***	-0.41***	0.045
	(0.040)	(0.020)	(0.046)	(0.072)
MBBS	-0.070	0.0017	0.088	-0.0012
	(0.064)	(0.013)	(0.097)	(0.082)
Other qualification	0.018	-0.0014	0.027	0.14**
	(0.051)	(0.016)	(0.055)	(0.062)
Patient load during visit	-0.0018	-0.0029	-0.0098	0.0013
	(0.0050)	(0.0026)	(0.0093)	(0.0052)
Panel B: Dual-practice sample (Provider fixed effects), N=3				
Public	-0.11**	-0.0068	-0.087*	-0.0099
	(0.052)	(0.031)	(0.044)	(0.045)
Unstable angina	0.11	-0.087	-0.24***	0.089
	(0.078)	(0.053)	(0.079)	(0.067)
Dysentery	-0.25***	-0.20***	-0.35***	-0.043
	(0.057)	(0.038)	(0.062)	(0.056)
Patient load during visit	0.00099	-0.014	-0.010	0.0075
	(0.017)	(0.0086)	(0.015)	(0.0089)
Mean	0.32	0.06	0.39	0.70
SD	0.47	0.24	0.49	0.46
Mean: Public	0.27	0.07	0.38	0.71
SD: Public	0.44	0.25	0.49	0.46
Mean: Private	0.35	0.06	0.40	0.70
SD: Private	0.48	0.23	0.49	0.46

Notes: *** p<0.01, ** p<0.05, * p<0.1. Observations are standardized patient-provider interactions. The *dual sample* consists of providers visited separately in both their public and private clinics.

Table 6: Over-treatment in the public and private sectors

Dependent variable:	(1) Uttered incorrect diagnosis	(2) Unnecessary or harmful treatment
Panel A: Full sample (Village and SP fixed effects), N=1,105		
Public	-0.052*	0.028
	(0.030)	(0.042)
Unstable angina	0.063	-0.28***
	(0.045)	(0.054)
Dysentery	-0.080**	-0.31***
	(0.033)	(0.039)
MBBS	-0.14***	-0.029
	(0.036)	(0.059)
Other qualification	-0.033	-0.050
	(0.041)	(0.040)
Patient load during visit	0.0032	0.0050
	(0.0044)	(0.0051)
Panel B: Dual-practice sample (Provider fixed effects), N=344		
Public	-0.011	0.045
	(0.036)	(0.043)
Unstable angina	0.12*	-0.16***
	(0.066)	(0.053)
Dysentery	-0.047	-0.28***
	(0.038)	(0.058)
Patient load during visit	0.016	0.018
	(0.011)	(0.014)
Mean	0.15	0.76
SD	0.35	0.43
Mean: Public	0.12	0.76
SD: Public	0.33	0.43
Mean: Private	0.16	0.75
SD: Private	0.37	0.43

Notes: *** p<0.01, ** p<0.05, * p<0.1. Observations are standardized patient-provider interactions. The *dual sample* consists of providers visited separately in both their public and private clinics.

Table 7: Quality in the public and private sectors, by case

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Time spent	Percentage of checklist items	Uttered a diagnosis	Uttered correct diagnosis	Uttered incorrect diagnosis	Correct treatment	Unnecessary or harmful treatment	Any helpful treatment
Panel A: Asthma								
SP and Village FEs (Full sample)	-2.15***	-7.56***	-0.15**	-0.019	-0.072	-0.21***	0.059	-0.077
	(0.39)	(2.41)	(0.075)	(0.044)	(0.049)	(0.064)	(0.044)	(0.070)
Provider FEs (Dual sample)	-1.37***	-9.97***	-0.091	-0.0097	-0.031	-0.16**	0.091*	-0.083
	(0.37)	(2.48)	(0.083)	(0.067)	(0.052)	(0.073)	(0.052)	(0.077)
Mean of dependent variable	4.29	23.80	0.36	0.11	0.14	0.60	0.92	0.68
SD of dependent variable	3.66	14.72	0.48	0.32	0.35	0.49	0.28	0.47
Mean: Public	2.57	18.75	0.30	0.15	0.08	0.54	0.94	0.69
SD: Public	2.11	14.00	0.46	0.36	0.27	0.50	0.24	0.47
Mean: Private	5.15	26.53	0.40	0.10	0.17	0.64	0.91	0.68
SD: Private	3.97	14.39	0.49	0.30	0.38	0.48	0.29	0.47
Panel B: Dysentery								
SP and Village FEs (Full sample)	-0.46***	-5.54**	-0.11**	-0.00057	-0.035	0.042	0.080	-0.046
	(0.15)	(2.19)	(0.048)	(0.011)	(0.037)	(0.055)	(0.074)	(0.083)
Provider FEs (Dual sample)	-0.36*	-7.12**	-0.099		-0.039	0.097	0.031	0.10
	(0.18)	(3.47)	(0.066)		(0.053)	(0.063)	(0.074)	(0.072)
Mean of dependent variable	1.58	19.03	0.17	0.01	0.08	0.20	0.60	0.68
SD of dependent variable	1.44	15.53	0.37	0.09	0.27	0.40	0.49	0.47
Mean: Public	1.10	16.67	0.12	0.01	0.06	0.26	0.62	0.70
SD: Public	0.80	14.02	0.33	0.09	0.24	0.44	0.49	0.46
Mean: Private	1.81	20.35	0.20	0.01	0.09	0.17	0.59	0.67
SD: Private	1.60	16.15	0.40	0.09	0.29	0.37	0.49	0.47
Panel C: Unstable angina								
SP and Village FEs (Full sample)	-2.62***	-13.7***	-0.10	-0.062***	-0.11	0.081	-0.16	-0.078
	(0.65)	(4.78)	(0.087)	(0.023)	(0.097)	(0.12)	(0.100)	(0.085)
Dual sample	-3.10***	-15.4***	-0.15	-0.12	0.14	-0.32***	-0.016	-0.059
	(0.98)	(5.17)	(0.12)	(0.078)	(0.11)	(0.11)	(0.10)	(0.096)
Mean of dependent variable	3.54	24.69	0.46	0.06	0.24	0.38	0.75	0.77
SD of dependent variable	3.03	19.88	0.50	0.23	0.43	0.49	0.43	0.42
Mean: Public	2.54	19.97	0.41	0.04	0.25	0.32	0.72	0.74
SD: Public	2.03	19.69	0.49	0.20	0.43	0.47	0.45	0.44
Mean: Private	4.04	27.10	0.48	0.06	0.24	0.42	0.77	0.78
SD: Private	3.32	19.69	0.50	0.24	0.43	0.50	0.42	0.41

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. Observations are standardized patient-provider interactions. This table displays the coefficient on *public* with specifications identical to the specifications in Table 3-5 aside from the controls for case. The *dual sample* consists of providers visited separately in both their public and private clinics. For the *unstable angina* case, providers with dual practices did not get two separate visits; rather an unstable angina standardized patient visited only one randomly selected practice. Thus, for this case, the *public* coefficient is measuring the difference between dual practice providers visited in their public practice and those that were visited in their private practice.

Table 8: Prices and Quality					
	1	2	3	4	5
	No Controls	Location and SSP Dummies	Adding Provider Controls	Adding Facility Index	Adding Consultation Length
Checklist Adherence	0.597	0.57	0.496	0.49	0.379
	[0.073]**	[0.066]**	[0.070]**	[0.067]**	[0.069]**
MBBS Degree			13.173	12.609	12.853
			[8.257]	[8.346]	[9.299]
No Qualification			-6.714	-5.467	-6.227
			[3.773]	[3.656]	[3.488]
Experience			0.006	-0.016	-0.038
			[0.105]	[0.113]	[0.113]
Facilities Index				1.72	1.834
				[1.371]	[1.341]
Consultation Time					23.613
					[9.236]*
Constant	21.458	34.532	32.308	32.782	1.657
	[1.796]**	[7.026]**	[9.701]**	[9.715]**	[0.588]**
Observations	754	754	543	540	540
R-squared	0.1	0.3	0.3	0.31	0.33
Robust standard errors in brackets					
* significant at 5%; ** significant at 1%					

Table 9: Real patients in the public and private sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Village Fixed Effects				Provider Fixed Effects (Dual sample)			
	Time spent	Total questions	Examined patient	Gave medicines	Time spent	Total questions	Examined patient	Gave medicines
Public	-1.11***	-0.85***	-0.32***	0.42***	-1.77***	-0.90***	-0.17***	0.48***
	(0.28)	(0.20)	(0.077)	(0.079)	(0.50)	(0.29)	(0.065)	(0.095)
MBBS	-0.40	0.50**	0.16**	-0.43***				
	(0.50)	(0.21)	(0.076)	(0.14)				
Some qualification	0.57	0.087	0.072	-0.0025				
	(0.37)	(0.11)	(0.052)	(0.078)				
Number of observations	2,528	2,527	2,521	2,454	1,010	1,008	1,009	973
Mean dependent variable	2.68	3.59	0.75	0.78				
SD dependent variable	2.43	1.62	0.44	0.42				
Mean: Public	1.64	3.22	0.64	0.97				
SD: Public	1.11	1.38	0.48	0.18				
Mean: Private	3.74	3.95	0.85	0.59				
SD: Private	2.90	1.75	0.36	0.49				

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. Observations are patient-provider interactions, and the sample has been limited to the SP sample. Controls for patients' SES: whether patient has no education. Controls for patients' presenting symptoms indude: indicators for a number of presenting symptoms (fever, cold, diarrhea, weakness, injury, vomiting, dermatological problem, pregnancy, and pain).

Table 10: Quality in the public and private sectors in urban Delhi

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Time spent	Percentage of checklist items	IRT score	Uttered diagnosis	Uttered correct diagnosis	Uttered incorrect diagnosis	Correct treatment	Incorrect treatment	Antibiotics	Referred patient
Public	-5.42***	-16.1***	-1.25***	-0.16	-0.097*	0.0045	0.28**	-0.18	-0.21	0.085*
	(0.47)	(2.23)	(0.068)	(0.086)	(0.041)	(0.048)	(0.083)	(0.15)	(0.12)	(0.041)
MBBS	0.80	6.14***	0.34***	-0.14	0.025	-0.10	0.037	0.17	0.17	-0.065
	(0.71)	(1.17)	(0.042)	(0.11)	(0.053)	(0.089)	(0.11)	(0.14)	(0.16)	(0.14)
Some qualification	2.24***	4.48	0.32***	-0.014	0.027	-0.058	-0.019	-0.13***	-0.068*	0.080
	(0.60)	(2.53)	(0.054)	(0.11)	(0.034)	(0.11)	(0.15)	(0.023)	(0.034)	(0.12)
Dysentery	0.12	6.75***	0.31***	-0.34***	0.044***	-0.14***	0.13***	-0.28***	0.025**	
	(0.093)	(0.22)	(0.017)	(0.0078)	(0.0038)	(0.0099)	(0.016)	(0.014)	(0.0084)	
Patient load during visit	0.00015	-0.0069	-0.0054	0.00058	-0.0011	-0.000079	-0.0063**	0.0020	0.0011	-0.0024
	(0.018)	(0.054)	(0.0047)	(0.0015)	(0.00078)	(0.0016)	(0.0025)	(0.0039)	(0.0023)	(0.0013)
Number of observations	229	231	231	231	231	231	231	231	231	180

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standardized patients in parentheses. Observations are standardized patient-provider interactions. All specifications indude standardized patient fixed effects.

Appendix Tables and Figures

Table A1. Checklist items			
	Asthma	**Dysentery**	**Unstable angina**
Questions about:	Cough, onset of problems, current breathing problem, previous breathing problems, age, fever, chest pain, previous attacks, expectoration, triggers, whether problem is constant or episodic, frequency of problems, duration of problem, family	Age of child, stool frequency, stool quality, stool quantity, vomiting, abdominal pain, what the child has eaten, whether child is active or playful, whether child has been drinking any liquids, household's water source, urination	Pain location, pain start time, previous similar pain, pain severity, first experience of this kind of pain, pain radiation, pain quality, sweating, nausea, shortness of breath, activity during onset of pain, smoking habit, family history of similar illness.
Exams:	Temperature attempt, blood pressure measurement, pulse measurement, auscultation of chest.	N/A (patient was not present)	Temperature attempt, Electro-cardio gram, blood pressure measurement, pulse measurement, auscultation of the chest.

Table A2. Diagnoses and treatments

		Asthma	Dysentery	Unstable angina
Diagnoses	Correct	Asthma, asthma attack	Dysentery, bacteria	Heart attack, angina, myocardial infarction, attack
	Incorrect	Blood pressure problem, gastrointensinal problem, heart problem, the weather, cough in chest, thyroid problem, weakness, lack of blood, infection in windpipe, pregnancy,	Weather, heat in liver, acidity	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.
Treatment	Correct	Bronchodilators, theophylline, inhaled or oral corticosteroids, leukotriene inhibitors, cromones, inhaled anticholinergics	ORS, rehydration	Aspirin, clopidogrel/other anti-platelet agents, referral.
	Helpful	Bronchodilators, theophylline, inhaled or oral corticosteroids, leukotriene inhibitors, cromones, inhaled anticholinergics	Antibiotics,zinc	Nitroglycerin, blood thinners, betablockers, ACE inhibitors, vasodilators, other cardiac medication, morphine, other pain medication, recommend or do an ECG.
	Unncessary or harmful	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	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.

Table A3. Sample checks

	Means						P-values		
	Universe	Audit 1	Public dual MBBS	Public non-dual MBBS	Received angina in public	Received angina in private	Universe = Audit 1	Public dual MBBS = Public non-dual MBBS	Unstable angina private = Unstable angina public
Provider and practice characteristics									
Age	43.14	43.98	45.10	42.62	45.51	45.17	0.36	0.26	0.89
Male	0.89	0.94	0.84	0.91	0.83	0.90	0.01	0.33	0.44
At least 12 years of basic education	0.85	0.94	0.91	0.85	0.90	0.94	0.00	0.31	0.57
MBBS degree	0.06	0.10	1.00	1.00	0.95	1.00	0.05	0.00	0.20
No medical training	0.73	0.67	0.00	0.00	0.03	0.00	0.09	0.00	0.37
Was born in taluk (county)	0.18	0.21	0.05	0.06	0.03	0.07	0.28	0.81	0.45
Provider tenure in current location	13.62	14.05	7.02	5.17	6.81	8.00	0.58	0.15	0.46
Dispenses medicines in clinic	0.86	0.84	0.35	0.91	0.36	0.34	0.57	0.00	0.88
Fees	43.31	43.54	3.90	2.89	3.90	3.89	0.90	0.00	0.97
Patients per day	15.90	17.76	15.31	31.11	16.03	17.66	0.12	0.00	0.69
Electricity	0.88	0.95	1.00	0.97	1.00	1.00	0.01	0.20	
Stethoscope	0.90	0.95	1.00	1.00	1.00	1.00	0.02	0.00	
Blood pressure cuff	0.71	0.77	1.00	1.00	1.00	1.00	0.09	0.00	
Thermometer	0.89	0.92	0.98	0.95	1.00	0.97	0.13	0.30	0.28
Weighing scale	0.54	0.58	0.94	0.90	0.97	0.90	0.40	0.46	0.25
Handwashing facility	0.78	0.83	0.86	0.79	0.86	0.87	0.15	0.39	0.87

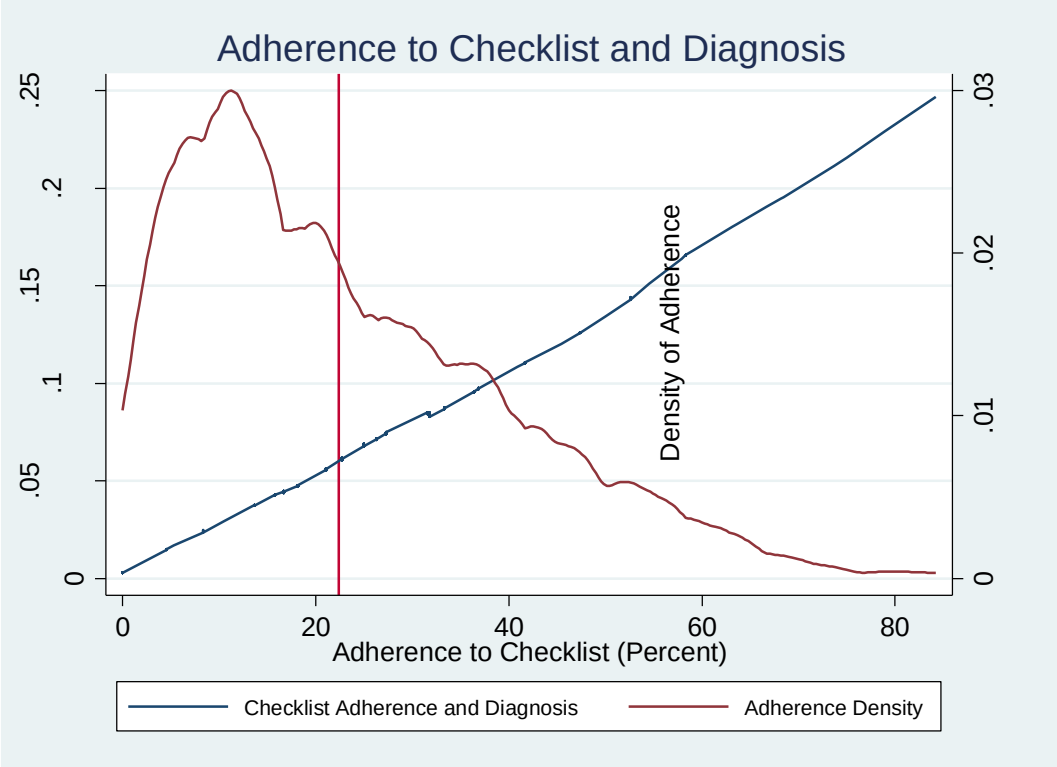
Table A3.1. Sample checks for Asthma and Dysentery Only

	Public dual MBBS	Public non-dual MBBS	Received angina in public	Received angina in private	Public dual MBBS = Public non-dual MBBS	Unstable angina private = Unstable angina public
Quality from audit studies						
Time spent	1.60	2.28	1.30	1.68	0.01	0.16
Percentage of checklist items	17.33	20.73	16.19	19.47	0.65	0.20
IRT score	-0.39	-0.20			0.42	
Utter diagnosis	0.27	0.29	0.21	0.22	0.78	0.83
Uttered correct diagnosis	0.09	0.06	0.07	0.12	0.78	0.32
Uttered incorrect diagnosis	0.15	0.13	0.08	0.10	0.90	0.69
Total medicines dispensed	0.04	0.05	0.01	0.10	0.36	0.18
Total medicines prescribed	2.28	2.43	2.32	2.29	0.66	0.92
Correct treatment	0.39	0.38	0.60	0.44	0.29	0.09
Wrong treatment	0.78	0.76	0.55	0.58	0.81	0.73
Antibiotic	0.56	0.59	0.65	0.66	0.85	0.98
Referred patient	0.10	0.14	0.03	0.00	0.76	0.37

Table A3.2. Sample checks

	Means						P-values		
	Universe	Audit 1	Public dual MBBS	Public non-dual MBBS	Received angina in public	Received angina in private	Universe = Audit 1	Public dual MBBS = Public non-dual MBBS	Unstable angina private = Unstable angina public
Characteristics of real patients									
Days sick	4.75	5.11	5.23	6.08	5.00	5.45	0.26	0.10	0.36
Number of questions patients pose	0.44	0.47	0.48	0.47	0.50	0.46	0.24	0.86	0.46
Has no education	0.47	0.47	0.44	0.46	0.47	0.41	0.91	0.56	0.06
Has electricity	0.94	0.94	0.93	0.97	0.93	0.94	0.76	0.04	0.47
Owns phone	0.96	0.97	0.93	0.87	0.59	0.72	0.16	0.01	0.00
Owns form of transport	0.81	0.83	0.75	0.66	0.67	0.83	0.06	0.00	0.00
Came by foot	0.44	0.48	0.57	0.64	0.58	0.57	0.02	0.06	0.81
Lives in the village	0.53	0.57	0.54	0.66	0.53	0.55	0.03	0.00	0.63
ADL: Dress	0.98	0.98	0.96	0.96	0.97	0.94	0.81	0.65	0.05
ADL: Light work	0.89	0.89	0.75	0.86	0.81	0.68	0.92	0.00	0.00
ADL: Carry	0.71	0.71	0.67	0.85	0.70	0.62	0.98	0.00	0.03
ADL: Walk	0.68	0.68	0.78	0.88	0.81	0.68	0.82	0.00	0.00
Fever	0.42	0.40	0.50	0.45	0.49	0.51	0.39	0.14	0.41
Cold	0.19	0.18	0.42	0.26	0.41	0.45	0.33	0.00	0.23
Pain	0.34	0.33	0.52	0.56	0.48	0.56	0.62	0.23	0.04
Diarrhea	0.14	0.15	0.06	0.06	0.06	0.06	0.35	0.98	0.67
Dermotological problem	0.05	0.05	0.07	0.12	0.07	0.07	0.81	0.01	0.94
Injury	0.07	0.08	0.06	0.04	0.06	0.07	0.63	0.13	0.47
Pregnancy	0.02	0.01	0.04	0.02	0.05	0.02	0.34	0.08	0.04
Vomitting	0.10	0.10	0.04	0.03	0.05	0.03	0.63	0.45	0.25
Weakness	0.13	0.13	0.07	0.09	0.07	0.07	0.74	0.40	0.85

Figure A1


Adherence to Checklist and Diagnosis
0
.05
.1
.15
.2
.25
0
.01
.02
.03
Density of Adherence
0
20
40
60
80
Adherence to Checklist (Percent)
Checklist Adherence and Diagnosis
Adherence Density